JOURNAL OF STATISTICAL MODELING & ANALYTICS (JOSMA)
(ISSN: 2180-3102)
UNIVERSITI MALAYA

Robust Two-Parameter Poisson Regression Model Estimator

Ilesanmi Akanbi Ajibode[1*], Timothy Olabisi Olatayo[2] and Biodun Tajudeen Efuwape[3]

[1*]*Department of Mathematics and Statistics, School of Pure and Applied Sciences, The Federal Polytechnic, Ilaro, Ogun State, Nigeria*
[2,3]*Department of Statistics, Faculty of Science, Olabisi Onabanjo University, Ago Iwoye, Ogun State*

* *Corresponding author: ilesanmi.ajibode@federalpolyilaro.edu.ng*

RESEARCH ARTICLE

Received 28 May 2025
Accepted 3 Mar 2026
Published 9 May 2026

Abstract

Multicollinearity and outliers are some of the problems of parameter estimation in the Poisson regression analysis. The existing researches used the maximum likelihood estimator, Poisson Ridge estimator, transformed-M Ridge estimator, and the two-parameter estimator of Poisson regression models have taken the issue of multicollinearity and outliers with interest in the explanatory but not with the response variables. This paper suggested use of transform M-estimator assisting in transformation of the response variable, making it less sensitive to extreme and irregular observations and stabilizing variance. This is useful in instances where the response has skewness, or non-constant variance or other non-standard distributional properties. The transform M-estimator is also effective in enhancing efficiency in the Poisson regression as the relationship between means and variances is skewed. A two-parameter estimator is a biased regression estimator, which adds two additional tuning (biasing) parameters to the regression to enhance better estimation, particularly when there is multicollinearity among the predictor variables. The estimator is an extension of single-parameter estimators such as ridge or Liu estimators, is more flexible, and in many cases the estimator is more effective but has the drawback of the response variable being counted. Building on these strengths, the present study integrates the transform M-estimator with the two-parameter estimator to produce a more stable and reliable method suited for complex count-data as the response variable. The proposed estimator is able to reduce the effect of multicollinearity and outliers while dealing with response variables that the data are counted in nature exhibiting outliers. Simulation study and real-life data was used to validate the efficacy of the new estimator. The result of the finding shows that the introduced estimator (MT-TPE) performed better than existing estimators in the presence of outliers in the response variable and multicollinearity in the explanatory variables using the MSE. The ability of an estimator to maintain efficiency makes it a valuable tool in modeling, particularly in datasets commonly encountered in real-world scenarios. Thus, the estimator (MT-TPE) provided a robust method for mitigating the effects of multicollinearity and outliers in count data.

Keywords: Accident, Count data, Multicollinearity, Outlier, Transformed-M estimator.

1. Introduction

The Poisson regression model is one way to model count data where the response variable represents the frequency of occurrences of an event throughout a fixed interval/area (Schober, et al., 2021; Coxe, et al., 2009). It is called that because it was developed by Siméon Denis Poisson, a French mathematician who made basic advances in the understanding of probability and statistics (David, 1962). In many cases, Poisson regression is applied when a response variable is Poisson distributed-a distribution that expresses the probability of the frequency/rate of events occurring in a of time or space. The distribution is characterised by one parameter, denoted as (λ), which generally stands for the average rate of occurrence. Poisson regression usually assumes independence of predictor variables across time. However, it is very important to mention here that in such cases, predictors may independently vary over time and thus such variables will not be independent.

Multicollinearity is a statistical feature that occurs when, within a regression model, significant correlation exists among two or more predictors. Under these conditions, strong correlations may give way to a number of challenges in performing regression analysis. Where there is multicollinearity, one cannot visually identify the unique or distinctive contribution of the interrelated factors to the dependent variable. This may cause an inflation of the regression coefficients' standard errors, perhaps making it impossible to ascertain the significance of the coefficients. According to Aguinis (2004), slight changes in dataset would result in large swings in the estimation of the coefficients, which would lower the overall level of reliability of the model.

These issues have motivated numerous scholars to develop better estimators in place of the maximum likelihood estimator in dealing with multicollinearity, thereby making the Poisson regression analysis increasingly more accurate and more reliable in accommodating the negative effects of correlated predictor variables on the estimations of parameters and stability of the model. Månsson and Shukur, (2011); Rashad and Algamal, (2019); Qasim et al., (2020); Jadhav, (2021); Oranye et al., (2021); Lukman et al., (2021, 2022); Amin et al., (2021, 2022); Ertan and Akay, (2022, 2023) among others, have suggested several remedies for handling collinearity in Poisson regression models.

Besides the collinearity problem in Poisson regression models, it has been further compounded by the existence of outliers, especially in the y-direction. Extreme value is such that it is far away from the general pattern of a dataset and could influence the regression model overwhelmingly. Because these can distort the estimates of parameters and the fit of models, the predictions will not be reliable. Outliers can unduly affect estimates in Poisson regression, given that they may make the model dynamically unstable. Therefore, attention must be given to the impact of observation on outliers for validity and reliability in the final Poisson regression analysis result.

It is observed from the literature that the robust regression estimate outperforms the ML and Shrinkage estimators obtained from GLMs in estimating parameters when outliers are present in the data (Croux and Haesbroeck, 2003; Valdora and Yohai, 2014). Künsch et al. (1989) proposed conditionally unbiased bounded-influence estimators (CEs) which enjoy very high robustness, especially for a small fraction of contaminating outliers. Based on the framework of Poisson regression models, Valdora and Yohai (2014) obtained the transformed M-estimator.

Generally, multicollinearity and extreme values are combinations that present complex problems that may affect the reliability and interpretability parameters. If, for example, such a situation arises where the two phenomena coexist, then their combined impact can be loud (Olaluwoye, et al., 2025). Such multicollinearity could amplify the influence of outliers, making it difficult to affirm the true relationship between predictors and its response variable. These would skew the stability of the estimates of the parameters and further increase the issue of multicollinearity. Together, both of these phenomena tend to influence model predictions and make them unreliable, which further complicates the issue of influential variable identification. In that respect, critical knowledge of both multicollinearity and outliers are thus crucial in making regression modeling strong. The reason is that there is a need to have an approach that will consider the two when they co-exist so as to guarantee valid and accurate results.

The literature on methods that simultaneously address multicollinearity among explanatory variables and the existence of outliers—particularly in the count-based response variable—within Poisson regression models remains limited. Although existing two-parameter estimators are capable of reducing the effect of multicollinearity, they fall short when outliers are present in the response variable, making it inadequate for handling such combined challenges in Poisson regression estimation. Examples include Arum et al. (2022) and Lukman et al. (2023b). Notably, this current study attempts to add to the highly limited literature in this domain. Since then, individual research focused on either multicollinearity or outlier management in the studies of Poisson regression. How to combine the methodologies to handle effectively both phenomena simultaneously has seldom been discussed. Hence, the present study tries to fill the existing loophole through the development of a robust estimator for Poisson regression analysis. The study contributes to the estimation of stronger and more robust parameters when data are affected by both multicollinearity and outliers in the y-direction, furthering the methodology that could be available to researchers and practitioners for a variety of purposes in which count data is used. In addition, the ridge type M-estimator uses a single ridge penalty parameter which introduces uniform shrinkage of the regression parameter coefficients towards zero, hence reducing variance in the presence of multicollinearity and somehow introduces bias because of its inability to adapt to the underlying structure of the design matrix. However, the two-parameter transformed M-estimator incorporates additional tuning parameter that allows the estimator to adjust to strength of shrinkage, direction, and degree of bias correction. This helps to stabilize the estimation more effectively when the information matrix is ill-conditioned, which leads to improved handling of multicollinearity and more robust performance in infinite samples.

In Section 2, we briefly outline the Poisson regression model and then discuss the Maximum Likelihood method and Ridge estimator. After that, we briefly introduce M-estimators using transformations (MT) and our proposed methods. Continuing onward, Section 3 provides the results of a Monte Carlo study that examines the performance of the proposed methods versus other popular estimators within the context of Poisson regression. Section 4 extends our study on the real dataset and compares the proposed methods against the existing state-of-art methods based on mean squared error metrics. Finally, Section 5 presents our findings and conclusions.

2. Materials and Methods

2.1 Existing Methodology

The probability distribution of a response variable r_i that follows a Poisson distribution is expressed as follows:

$$f(r_i) = \frac{exp\,(-t_i)t_i^{r_i}}{r_i!}, r_i = 0,1,2, \dots \quad (1)$$

where $r_i > 0$. The Poisson distribution in (1) is characterized by equal mean and variance. The model is expressed in terms of the mean response. The mean response is related to linear predictors through a function h, as shown in equation (2).

$$h(r_i) = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p = x_i'\beta, \quad (2)$$

such that $h(.)$ is a monotone differentiable link function and a common example is the log link function. The likelihood function for equation (1) is defined as follows:

$$l(\beta) = \prod_{i=1}^{n} \frac{exp\,(-t_i)t_i^{r_i}}{r_i!} \quad (3)$$

The log-likelihood function from equation (3) is as follows:

$$\ln l(\beta) = \sum_{i=1}^{n} r_i \ln(t_i) - \sum_{i=1}^{n} t_i - \sum_{i=1}^{n} ln\,(r_i!) \quad (4)$$

The parameters β are estimated iteratively using the Fisher Scoring method as follows:

$$\beta^{c+1} = \beta^c + I^{-1}(\beta^c)S(\beta^c) \quad (5)$$

where the $S(\beta)$ represents the score vector, I represent the observed information matrix, and c represents the number of iteration until convergence. Hence, the Poisson MLE of β is:

$$\hat{\beta}^{PML} = \left(X'\widehat{H}X\right)^{-1} X'\widehat{H}\hat{v} \quad (6)$$

such that $\widehat{H} = diag[\hat{t}_i]$ and $\hat{v}$ is a vector with the i^{th} element equals $\hat{v} = log(\hat{t}_i) + \frac{r_i - \hat{t}_i}{\hat{t}_i}$.

In the context of MLE, the effectiveness of an estimator is often assessed using properties like bias, variance, and asymptotic efficiency. The MSE is a combination of its variance and squared bias.

$$MSEM\left(\hat{\beta}^{PML}\right) = \left(X'\widehat{H}X\right)^{-1} \quad (7)$$

The scalar mean squared error (SMSE) is the trace of the MSE matrix defined as follows:

$$SMSE\left(\hat{\beta}^{PML}\right) = trace\left(X'\widehat{H}X\right)^{-1} = \sum_{j=1}^{p+1} \frac{1}{e_j} \quad (8)$$

where e_j is the j[th] eigenvalue of the $X'\widehat{H}X$ matrix. As mentioned earlier, the presence of intercorrelation among predictors presented a challenge to the MLE. In response to this issue, regularization techniques, like ridge regression, have been embraced to circumvent multicollinearity in linear regression. To tackle this challenge more specifically in the PRM, Månsson and Shukur (2011) introduced the Poisson ridge regression estimator (PRR) as an alternative to MLE, offering a more efficient approach to estimation in the presence of multicollinearity.

$$\hat{\beta}^{PRR} = \left(X'\widehat{H}X + k_1 I\right)^{-1} X'\widehat{H}X\hat{\beta}^{PML} \quad (9)$$

where I is a $p \times p$ identity matrix, and the ridge parameter k is defined as:

$$k_1 = \frac{1}{Max(\hat{\gamma}_j^2)}, \quad (10)$$

where $\gamma = T'\hat{\beta}^{PML}$ and $T'X'\widehat{H}XT = E = diag\left(e_1, \ldots, e_{p+1}\right)$, $e_1 \geq e_2 \geq \ldots \geq e_{p+1}$, E is the matrix of eigenvalues of $X'\widehat{H}X$ and T is the matrix whose columns are the eigenvectors of $X'\widehat{H}X$.The scalar mean squared error is expressed as follows:

$$SMSE[\hat{\gamma}^{PRR}] = \sum_{j=1}^{p+1} \frac{e_j}{(e_j + k_1)^2} + k_1^2 \sum_{j=1}^{p+1} \frac{\gamma_j^2}{(e_j + k_1)^2} \quad (11)$$

Asar and Genc (2017) defined a two-parameter estimator to overcome collinearity in the Poisson regression model. The PTP is defined as:

$$\hat{\beta}^{PTP} = \left(X'\widehat{H}X + k_2 I\right)^{-1}\left(X'\widehat{H}X + k_2 d_1 I\right)\hat{\beta}^{PML} \quad (12)$$

where ridge parameter k is defined as:

$$k_2 = Max\left(\frac{e_j}{e_j\hat{\gamma}_j^2(1-d_1) - d_1}\right) \quad (13)$$

$$d_1 = Min\left(\frac{e_j\hat{\gamma}_j^2}{e_j\hat{\gamma}_j^2 + 1}\right). \quad (14)$$

The scalar mean square error is defined as follows:

$$SMSE[\hat{\gamma}^{PTP}] = \sum_{j=1}^{p+1} \frac{(e_j + k_2 d_1)^2}{e_j(e_j + k_2)^2} + k_2^2(d_1 - 1)^2 \sum_{j=1}^{p+1} \frac{\gamma_j^2}{(e_j + k_2)^2} \quad (15)$$

2.2 Transformed M-estimator (MT)

Robust regression estimators are adopted to account for outliers in the regression model. Examples of such estimators in linear regression include the M-estimator, the MM-estimator, least absolute deviation, least median squares estimator, the S-estimator, the least trimmed squared estimator, and others. Specifically, the M-estimator is considered highly effective in addressing outliers in the y-direction. Valdora and Yohai (2014) developed the conditionally unbiased bounded influence estimator and the transformed M-estimator to handle outliers in the Poisson regression model, respectively. Recent studies by Lukman et al. (2023b) and Arum et al. (2022) demonstrate that the transformed M-estimator consistently outperforms the conditionally unbiased bounded influence estimator. For this study, our attention will be specifically directed towards the transformed M-estimator (PMT).
Valdora and Yohai (2014) defined the M-estimator of β based on transformations as

$$\hat{\beta}_{MT} = arg \sum_{i=1}^{n} \rho(t(y_i) - m(g^{-1}(\beta' x))) \, . \tag{16}$$

ρ-function is tukey's bisquare function and m (.) is the mean function.

For the Poisson regression with $g(\phi) = log\,(\phi)$, the variance-stabilizing transformation for this family of distributions is $t(y) = \sqrt{y}$. PMT can be found in the R library "poissonMT". Further details on this estimator can be found in the literature (Valdora and Yohai, 2014; Arum et al., 2022; Lukman et al., 2023b).

2.3 Proposed Estimator

Lukman et al. (2023) developed the Poisson ridge-MT estimator by integrating the ridge estimator with the transformed M-estimator (MT). The Poisson Ridge-MT estimator of β (PR-MT) is defined as:

$$\hat{\beta}_{PMT}^{PRR} = \left(X'\hat{H}X + k_3 I\right)^{-1} X'\hat{H}X\hat{\beta}_{PMT} \tag{17}$$

where $k_3 = \frac{1}{max\left(\hat{\gamma}_{j,PMT}^2\right)}$. The scalar mean squared error is as follows:

$$SMSE[\hat{\gamma}_{PMT}^{PRR}] = \sum_{j=1}^{p+1} \frac{e_j}{(e_j+k_3)^2} + k_3^2 \sum_{j=1}^{p+1} \frac{\gamma_{j,PMT}^2}{(e_j+k_3)^2} \, . \tag{18}$$

k_3 is a data-driven ridge penalty. This makes shrinkage depend on the magnitude of the largest estimated Poisson Transformed M-estimator. Small values of k_3 when coefficients are large, then we have less shrinkage while large values whenever the coefficients are small, hence stabilization. e_j is the jth eigenvalue of the weighted cross-product matrix $X'\hat{H}X$.

$$X'\hat{H}X = Q \text{ diag } (e_1, e_2, \ldots\ldots e_{p+1})\, Q'$$

Q is the orthonormal eigenvector matrix of $X'\hat{H}X$ such that $Q'Q$ = I. The transformed coordinates of the Poisson transformed M-estimator is defined as $\hat{\gamma}_{PMT} = Q'\hat{\beta}_{PMT}$, then $\gamma_{j,PMT}$ is the jth coordinate of $\hat{\beta}_{PMT}$ in the eigen basis Q. This represents the contribution of the coefficient $\hat{\beta}_{PMT}$ along the direction associated with eigenvalue e_j.

The new estimator for this study was developed by pooling the two-parameter estimator together with the Poisson transformed M-estimator to form the Poisson Two-parameter Transformed M-estimator (MT-TPE), which is defined as follows:

$$\hat{\beta}_{MT}^{TPE} = \left(X'\hat{H}X + k_4 I\right)^{-1}\left(X'\hat{H}X + k_4 d_2 I\right)\hat{\beta}_{MT} \tag{19}$$

$$k_4 = Max\left(\frac{e_j}{e_j\hat{\gamma}_{j,MT}^2(1-d_2)-d_2}\right) \tag{20}$$

$$d_2 = Min\left(\frac{e_j\hat{\gamma}_{j,MT}^2}{e_j\hat{\gamma}_{j,MT}^2+1}\right). \tag{21}$$

The SMSE is defined as:

$$SMSE[\hat{\gamma}_{MT}^{TPE}] = \sum_{j=1}^{p+1}\frac{(e_j+k_4d_2)^2}{e_j(e_j+k_4)^2} + k_4^2(d_2-1)^2\sum_{j=1}^{p+1}\frac{\gamma_{j,MT}^2}{(e_j+k_4)^2} \tag{22}$$

From 19,

$$\hat{\beta}_{TME}^{PTTP} = (N + k_4 I)^{-1}(N + k_4 d_2 I)\hat{\beta}_{TME} \tag{23}$$

where,

N = $X'\hat{H}X$

I is the p x p identity matrix

$\hat{H}$ = diag $\left[\frac{\tau^2 r_i}{\zeta'' r_i}\right]$

τ is the derivative of ζ

r_i are the residuals of g(y_i) - $X_i\beta$

I is a p x p identity matrix

k_4 and $d_2$are the tuning parameters such that k_4 > 0 and 0<d_2<1.

$\hat{\beta}_{TME}$ is the transformed M-estimator

The expected value of $\hat{\beta}_{TME}^{PTTP}$, was derived by utilizing linearity of expectation. Assuming $\hat{\beta}_{TME}$ is an unbiased estimator of β (E [$\hat{\beta}_{TME}$] = β).

Then, the expected value is:

$$E[\hat{\beta}_{TME}^{PTTP}] = E[(N + k_4 I)^{-1}(N + k_4 d_2 I)\hat{\beta}_{TME}] \tag{24}$$

$$E[\hat{\beta}_{TME}^{PTTP}] = (N + k_4 I)^{-1}(N + k_4 d_2 I)E[\hat{\beta}_{TME}] \tag{25}$$

$$E[\hat{\beta}_{TME}^{PTTP}] = (N + k_4 I)^{-1}(N + k_4 d_2 I)\beta \tag{26}$$

From 23,

$$\hat{\beta}_{TME}^{PTTP} = (N + k_4 I)^{-1}(N + k_4 d_2 I)\hat{\beta}_{TME} \tag{27}$$

The variance of $\hat{\beta}_{TME}^{PTTP}$ is derived knowing that the estimator is related to $\hat{\beta}_{TME}$, which is a distribution around the true parameter β. Assuming normality, the variance of $\hat{\beta}_{TME}$ is approximately $(X'\hat{H}X)^{-1}$under large variance of $\hat{\beta}_{TME}^{PTTP}$. Then,

$$\text{Var}[\hat{\beta}_{TME}^{PTTP}] = (N + k_4 I)^{-1}(N + k_4 d_2 I)Var(\hat{\beta}_{TME})[(N + k_4 I)^{-1}]'[(N + k_4 d_2 I)]' \tag{28}$$

Hence,

$$\text{Var}\left[\hat{\beta}_{TME}^{PTTP}\right] = (N + k_4 I)^{-1}(N + k_4 d_2 I)(X'\hat{H}X)^{-1}[(N + k_4 I)^{-1}]'[(N + k_4 d_2 I)]' \tag{29}$$

The bias of $\hat{\beta}_{TME}^{PTTP}$ assuming the bias of $\hat{\beta}_{TME}$ is negligible. Then,

$$\text{Bias}\left[\hat{\beta}_{TME}^{PTTP}\right] = \text{E}\left[\hat{\beta}_{TME}^{PTTP}\right] - \beta \tag{30}$$

Also,

$$\text{Bias}\left[\hat{\beta}_{TME}^{PTTP}\right] = \text{E}\left[(N + k_4 I)^{-1}(N + k_4 d_2 I)\,\hat{\beta}_{TME}\right] - \beta \tag{31}$$

By linearity of expectation and assuming:

$$\text{E}\left[\,\hat{\beta}_{TME}\right] = \beta \tag{32}$$

$$\text{Bias}\left[\hat{\beta}_{TME}^{PTTP}\right] = [(N + k_4 I)^{-1}(N + k_4 d_2 I)\beta] - \beta \tag{33}$$

Then,

$$\text{Bias}\left[\hat{\beta}_{TME}^{PTTP}\right] = [(N + k_4 I)^{-1}(N + k_4 d_2 I) - I]\,\beta \tag{34}$$

The Mean Square Error (MSE), which connotes the expectation of the squared difference between the estimator and the true parameter, combines both the variance and squared bias. Then,

$$\text{MSE}\,[\hat{\beta}_{TME}^{PTTP}] = \text{E}\,[(\hat{\beta}_{TME}^{PTTP} - \beta)^2] \tag{35}$$

Hence,

$$\text{MSE}\,[\hat{\beta}_{TTME}^{PTTP}] = \text{Var}\,[\hat{\beta}_{TME}^{PTTP}] + [Bias[\hat{\beta}_{TME}^{PTTP}]]^2 \tag{36}$$

Using spectral decomposition of N, then, N = MTM', where M is the matrix whose columns are the eigen vectors of N and T is a diagonal matrix containing the eigenvalues of N. That is;

$$\text{T} = \text{diagonal}\,\{v_1, v_2, \ldots, v_{p+1}\} \tag{37}$$

The modification introduced by k_4 and d_2 can be expressed as:

$$[\text{N} + k_4 I] = \text{M}[\text{T} + k_4 I]M' \tag{38}$$

$$[\text{N} + k_4 d_2 I] = \text{M}[\text{T} + k_4 d_2 I]M' \tag{39}$$

The variance can then be expressed as:

$$\text{Var}[\hat{\beta}_{TME}^{PTTP}] = [N + k_4 I]^{-1}[\text{N} + k_4 d_2 I][X'\hat{H}X]^{-1}[N + k_4 d_2 I]'[[N + k_4 I]^{-1}]^1 \tag{40}$$

Substituting using spectral decomposition, the Var $[\hat{\beta}_{TME}^{PTTP}]$ can be expressed as:

$$\text{Var}\,[\hat{\beta}_{TME}^{PTTP}] = [M[T + k_4 I]M']^{-1}[\text{M}[\text{T} + k_4 d_2 I]M'][\text{M}T^{-1}M'][M[T + k_4 d_2 I]M']'[[M[T + k_4 I]M']^{-1}]' \tag{41}$$

This resulted to:

$$\text{Var}\,[\hat{\beta}_{TME}^{PTTP}] = M[T + k_4 I]]^{-1}[\text{T} + k_4 d_2 I]T^{-1}[T + k_4 d_2 I]\ \ [T + k_4 I]^{-1}]'M' \tag{42}$$

$$\gamma = M'\beta \tag{43}$$

Substituting spectral decomposition in bias;

$$\text{Bias}[\hat{\beta}_{TME}^{PTTP}] = [M[T + k_4 I]^{-1}(T + k_4 d_2 I)M' - MIM']\,\beta \tag{44}$$

$$\text{Bias}[\hat{\beta}_{TME}^{PTTP}] = M[[T + k_4 I]^{-1}(T + k_4 d_2 I) - I]\,M'\beta \tag{45}$$

$$\text{Bias}\,[\hat{\beta}_{TME}^{PTTP}]^2 = (k_4 d_2 - 1)^2 M[T + k_4 I]^{-1}\,\gamma\gamma'[T + k_4 I]\quad M' \tag{46}$$

Then,

$$\text{MSE}\,[\hat{\beta}_{TM}^{PTTP}] = \text{Var}\,[\hat{\beta}_{TM}^{PTTP}] = M[T + k_4 I]]^{-1}[\text{T} + k_4 d_2 I]T^{-1}[T + k_4 d_2 I]\ \ [T + k_4 I]^{-1}]'M' + [M[[T + k_4 I]^{-1}(T + k_4 d_2 I) - I]\,\gamma\gamma'[[T + k_4 I]^{-1}(T + k_4 d_2 I) - I]'M' \tag{47}$$

Then expressing the MSE in terms of sums (trace) over the individual components using the eigenvalues for it to be in a more explicit and simplified form. Then,

$$\text{MSE}\,[\hat{\beta}_{TME}^{PTTP}] = \text{Var}\,[\hat{\beta}_{TME}^{PTTP}] = M[T + k_4 I]]^{-1}[\text{T} + k_4 d_2 I]T^{-1}[T + k_4 d_2 I]\,[T + k_4 I]^{-1}]'M' + (k_4 d_2 - 1)^2 M[T + k_4 I]^{-1}\,\gamma\gamma'[T + k_4 I]\ \ ]\,M' \tag{48}$$

There exist the eigenvalues v_i of T. Then,

$$\text{MSE}\,[\hat{\beta}_{TME}^{PTTP}] = \sum_{i=1}^{p}\left[\frac{(v_i + k_4 d_2)^{\,2}}{(v_i + k_4)^{\,2}} + \frac{(k_4 d_2 - 1)^{\,2}\gamma_i^2}{(v_i + k_4)^{\,2}}\right] \tag{49}$$

γ_i^2's are the coefficients in the vector β projected into the basis of M. This provides a decomposition component-wise calculation of MSE based on each eigenvalue and the corresponding squared bias effect, trying back to the adjustment made by the parameters k_4 and d_2.

In order to estimate the tuning parameters k_4 and d_2, the values of the parameter k_4 and d_2 are obtained by differentiating equation 49 with respect to k_4 and d_2. Hence,

$$k_4 = \frac{v_i}{v_i\gamma_i^2(1 - d_2)^2 - d_2} \tag{50}$$

$$d_2 = \frac{v_i \gamma_i^2 - \frac{v_i}{k_4}}{[1 + v_i \gamma_i^2]} \quad (51)$$

In order to get the minimum of d_2, there is a need to consider the behaviour of the tuning parameter k_4 as it changes. Since k_4 is in the denominator, k_4 increasing will decrease the term $\frac{v_i}{k_4}$, then increasing d_2. Hence, the minimum of d_2 occurs when k_4 is at its minimum value, which is just a little above v_i. The estimator's performance is compared in the next two sections via simulation and real-life data analysis.

2.4 Simulation Design

In this section, a response variable that follows a Poisson distribution and correlated predictors was evaluated (Arum et al., 2022; Lukman et al., 2023b). The methodology employed to introduce multicollinearity is as follows:

$$x_{ij} = \sqrt{(1-\vartheta^2)}\xi_{ij} + \vartheta\xi_{i,p+1}, i = 1,2,\dots,n,\ j = 1,2,3,\dots,p, \quad (52)$$

where x_{ij} denotes the jth predictors in the ith row of the design matrix, ϑ denotes the level of multicollinearity such that ϑ= 0.8, 0.9, 0.99, and 0.999. These levels represented moderate, high, very high and near perfect collinearity respectively which helps to evaluate how collinear predictors affect the stability and precision of the estimators. ξ_{ij} denotes independent standard normal values. The study investigated models with three (3) to five (5) continuous predictors. The intercept values were taken as -1, 0, and 1; the three different intercept values were used in order to test how the estimator behaves under different baseline levels of the response variable (Lukman et al., 2022). Additionally, slope coefficients are selected such that $\sum_{j=1}^{p+1} \beta_j^2 = 1$ for the following sample sizes: 30, 50, 100, 200 and 500. These values were chosen to allow for a moderate sample size (100) representing typical study conditions while sample size (500) allows assessment of asymptotic behaviour and stability, providing insight into the estimator's consistency and robustness as data availability increases. Following Arum et al., (2022), the contamination formula is:

$$y_{[i]}=10*\max(y)+y_{[i]}, \quad (53)$$

This is a contamination formula that helps in magnifying the current maximum of the response variable and adding it back into the original observation. It helps in producing artificial large response values, simulating severe outliers' condition. This is essential for assessing how well an estimator can withstand severe deviations in the response variable.

The study investigated two percentages of outliers in the y-direction (20%). The estimator's performance was evaluated using the estimated mean squared error, defined as follows:

$$MSE(\hat{\beta}) = \frac{1}{1000}\sum_{j=1}^{1000}\left(\hat{\beta}_{ij} - \beta_i\right)^T\left(\hat{\beta}_{ij} - \beta_i\right) \quad (54)$$

where $\hat{\beta}_{ij}$ denotes the estimate of the ith parameter in jth replication and β_i represents the true parameter values. We repeat this experiment 1000 times to obtain robust results, which are meticulously presented in Tables 1 to 6 for thorough analysis and interpretation.

3. Results and Discussion

3.1 Results of the Simulation

The findings of the simulation (table 1-9) reveal that there is a significant variation in the performance of the estimators with different sample size, predictors, instances of high multicollinearity, and the presence of a response variable with outliers. The maximum likelihood estimator (MLE) has higher means squared errors (MSE), especially at smaller sample sizes and when there is a great deal of multicollinearity, and this is a manifestation of its sensitivity to both outliers and to the instability of correlated predictors. Ridge-type estimators minimize variance by regularizing MLE, but still, the MSE is larger than that of the robust two-parameter transformed M-estimator (MT-TPE) that unites the advantages of the two-parameter strategy with the robustness of the transformed M framework. In all the cases, the MSE is at its lowest with MT-TPE, signifying its greater capacity to deal with extreme observations and stabilisation of coefficient estimates at the same time. The greater the sample size and the greater the true β, the more this advantage is noted because more information enables the estimator to better differentiate between the underlying signal and the noise as well as reducing the effects of the outliers. When varying numbers of predictors are compared, it is revealed that as dimensionality grows, both MSE of MLE and Ridge1 changes increase, especially in the presence of strong multicollinearity, but at the same time, MT-TPE does not change its performance, indicating that it is not susceptible to the increasing correlation of predictors and dimensionality.

The effect of multicollinearity is also evident: at $\rho = 0.8$, the differences among estimators are modest, but as ρ increases to 0.9 and 0.99, MLE and Ridge1 show inflated MSE values, where a MT-TPE continues to yield stable and low MSE. The numerical values further illustrate these patterns; for example, for p=3, $\rho = 0.9$, $\beta_0 = -1$, and n = 30, MT-TPE achieves an MSE of 0.0423 compared to 0.0736 for MLE.

In general, these results confirm that MT-TPE effectively mitigates the destabilising effects of both outliers and multicollinearity, providing a reliable and efficient estimation method in complex count-data environments, while MLE and other standard estimators remain vulnerable to contamination and predictor correlation.

Table 1: Summary of the simulated data of the estimators for p=3 and $\rho = 0.8$ at 0.2 level of outlier

β_0	n	MLE	RIDGE1	MT-Ridge	TPE	MT-TPE
	30	0.002534474	0.002534234	0.002533065	0.002532967	0.002508534
	50	0.01372033	0.01369396	0.01362433	0.01360298	0.01309815
-1	100	0.005782858	0.005779182	0.005774443	0.005761798	0.005717805
	250	0.0006495846	0.0006495628	0.000649336	0.0006494594	0.0006461371
	500	0.0004261848	0.0004261749	0.0004260554	0.0004261069	0.0004245318
	30	0.01652434	0.01648869	0.01627384	0.0163236	0.01546505
	50	0.003007279	0.003006832	0.003002833	0.003005244	0.002987449
0	100	0.002231212	0.002230909	0.002225686	0.002229243	0.002221738
	250	0.000482586	0.0004825761	0.0004824316	0.0004825089	0.0004821331
	500	0.0001552805	0.0001552797	0.0001552643	0.0001552732	0.0001552297
	30	0.003538943	0.003538368	0.003528583	0.003532662	0.003522852
	50	0.002303648	0.002303424	0.002301228	0.002301767	0.002298368
+1	100	0.0006385427	0.0006385265	0.0006382657	0.0006384443	0.0006380961
	250	0.0001873388	0.0001873378	0.0001873183	0.000187337	0.0001873066
	500	2.845365e-05	2.845364e-05	2.845323e-05	2.845375e-05	2.845281e-05

Table 2: Summary of the simulated data of the estimators for $p = 4$ and $\rho = 0.8$ at 0.2 level of outlier

β_0	n	MLE	RIDGE1	MT-Ridge	TPE	MT-TPE
	30	0.02835322	0.02827542	0.0279205	0.02795451	0.02226891
	50	0.01966437	0.01962867	0.01963542	0.01946265	0.01872512
-1	100	0.001093501	0.001093466	0.001092991	0.001093189	0.001090482
	250	0.001249241	0.001249173	0.001248736	0.001248598	0.001240933
	500	0.0002666484	0.0002666463	0.0002666058	0.0002666253	0.0002663112
	30	0.007935053	0.007931679	0.007889767	0.007871102	0.007740559
	50	0.01966437	0.01962867	0.01963542	0.01946265	0.01872512
0	100	0.001250206	0.00125016	0.001249316	0.001249772	0.001246059
	250	0.0003503633	0.0003503604	0.0003502461	0.0003503332	0.0003501098
	500	9.898711e-05	9.898692e-05	9.897926e-05	9.898486e-05	9.896514e-05
	30	0.002167514	0.002167383	0.002166134	0.002167964	0.002162142
	50	0.001308575	0.00130853	0.00130761	0.001308012	0.001306894
+1	100	0.0002940834	0.0002940816	0.0002940489	0.0002940818	0.0002940035
	250	6.457816e-05	0.00006457811	6.457622e-05	6.457757e-05	6.457397e-05
	500	6.753933e-05	6.753926e-05	6.753705e-05	6.753862e-05	6.753465e-05

Table 3: Summary of the simulated data of the estimators for $p = 5$ and $\rho = 0.8$ at 0.2 level of outlier

β_0	n	MLE	RIDGE1	MT-Ridge	TPE	MT-TPE
	30	0.02688509	0.02681900	0.02643585	0.02658729	0.02526034
	50	0.007003451	0.007001322	0.006983861	0.006986212	0.006800115
-1	100	0.0008091451	0.0008091329	0.0008087941	0.0008089923	0.0008076444
	250	0.0005146195	0.0005146138	0.0005145434	0.0005145436	0.0005136034
	500	0.0002343132	0.0002343122	0.0002342974	0.0002342994	0.0002340661
	30	0.02645243	0.02639804	0.026285350	0.02604522	0.02474995
	50	0.003401424	0.003401118	0.003388534	0.003394779	0.003370758
0	100	0.0008091451	0.0008091329	0.0008087941	0.0008089923	0.0008076444
	250	3.874687e-05	3.874686e-05	3.874557e-05	3.874683e-05	3.874384e-05
	500	7.406186e-05	7.406179e-05	7.406176e-05	7.406092e-05	7.405416e-05
	30	0.005070327	0.005069613	0.005057956	0.005066397	0.005045899
	50	0.001043080	0.001043059	0.001042611	0.001042871	0.001041921
+1	100	0.0004769098	0.0004769066	0.0004768312	0.0004768622	0.0004766988
	250	3.784525e-05	3.784524e-05	3.784467e-05	3.784524e-05	3.784395e-05
	500	0.0000254271	2.542709e-05	2.542685e-05	2.542705e-05	2.542649e-05

Table 4: Summary of the simulated data of the estimators for $p = 3$ and $\rho = 0.9$ at 0.2 level of outlier

β_0	n	MLE	RIDGE1	MT-Ridge	TPE	MT-TPE
-1	30	0.07358132	0.0727447	0.06819571	0.07460531	0.04233877
	50	0.01135735	0.01134667	0.01107961	0.01125652	0.01089696
	100	0.005883798	0.00588058	0.005874611	0.005858191	0.005790166
	250	0.0006654659	0.000665447	0.0006653124	0.0006651587	0.00066251
	500	0.0003754402	0.0003754337	0.0003753572	0.0003753448	0.0003744207
0	30	0.006988402	0.006985366	0.006976007	0.006944733	0.00679034
	50	0.007217015	0.007212502	0.007185429	0.007169378	0.007077714
	100	0.001653811	0.001653694	0.001649658	0.00165183	0.001649611
	250	0.0002591952	0.0002591931	0.0002591559	0.0002591588	0.0002590239
	500	8.629929e-05	0.0000862991	8.629126e-05	8.629483e-05	8.628245e-05
+1	30	0.00113852	0.001138487	0.001137738	0.001137667	0.00113633
	50	0.003856892	0.003856227	0.00384776	0.003846256	0.00383351
	100	0.0003961609	0.0003961564	0.0003960623	0.0003961041	0.0003959729
	250	7.073109e-05	0.000070731	7.072722e-05	7.072872e-05	0.000070725
	500	7.772372e-05	7.772357e-05	7.771928e-05	7.772032e-05	7.771664e-05

Table 5: Summary of the simulated data of the estimators for $p = 4$ and $\rho = 0.9$ at 0.2 level of outlier

β_0	n	MLE	RIDGE1	MT-Ridge	TPE	MT-TPE
-1	30	0.002606321	0.002606164	0.002599877	0.002603841	0.002598874
	50	0.02334278	0.02329950	0.02314244	0.02302988	0.02103018
	100	0.003510148	0.003509682	0.003504445	0.003503667	0.003459681
	250	0.0003822403	0.0003822371	0.0003821693	0.0003821692	0.0003820691
	500	0.0002315230	0.0002315217	0.0002314955	0.0002314958	0.0002313379
0	30	0.01650520	0.01648945	0.01593725	0.01631996	0.0157367
	50	0.00441158	0.004411019	0.004401224	0.004400492	0.004335858
	100	0.001832753	0.001832652	0.001829656	0.001831032	0.001822818
	250	0.0001443624	0.0001443621	0.0001443396	0.0001443529	0.0001443307
	500	4.960107e-05	4.960104e-05	4.959939e-05	4.960002e-05	4.959484e-05
+1	30	0.004689425	0.004688815	0.004679709	0.004678446	0.004667139
	50	0.001055837	0.001055813	0.001055625	0.001055261	0.001054709
	100	0.0004561022	0.0004560987	0.0004560036	0.0004560082	0.0004558523
	250	4.213538e-05	4.213536e-05	4.213515e-05	4.213467e-05	4.213232e-05
	500	4.992732e-05	4.992729e-05	4.992609e-05	4.992624e-05	4.992467e-05

Table 6: Summary of the simulated data of the estimators for $p = 5$ and $\rho = 0.9$ at 0.2 level of outlier

β_0	n	MLE	RIDGE1	MT-Ridge	TPE	MT-TPE
-1	30	0.14734815	0.14403616	0.1433941	0.13780273	0.1167574
	50	0.004945677	0.004945027	0.004931531	0.004935199	0.004916128
	100	0.005939013	0.005937729	0.005922749	0.005923204	0.005878504
	250	0.0007274752	0.0007274652	0.0007271863	0.0007272531	0.0007246707
	500	0.0004654728	0.0004654681	0.0004654102	0.0004653836	0.000464532
0	30	0.004810844	0.004810382	0.004803857	0.004799352	0.004759629
	50	0.001697033	0.001696989	0.001695136	0.001695834	0.001693364
	100	.002105455	0.002105351	0.002100949	0.002103580	0.00209561
	250	0.0004241609	0.0004241581	0.0004239495	0.0004240889	0.0004236421
	500	0.0001268848	0.0001268846	0.0001268703	0.0001268787	0.0001268596
+1	30	0.006617432	0.006616004	0.00661138	0.006587429	0.00658614
	50	0.001108895	0.001108877	0.001108215	0.001108392	0.001107576
	100	0.0003040702	0.0003040692	0.0003040319	0.0003040378	0.0003039858
	250	7.223566e-05	7.223561e-05	7.223307e-05	7.223381e-05	7.222935e-05
	500	2.687182e-05	2.687181e-05	2.687181e-05	2.687157e-05	2.687115e-05

Table 7: Summary of the simulated data of the estimators for $p = 3$ and $\rho = 0.99$ at 0.2 level of outlier

β_0	n	MLE	RIDGE1	MT-Ridge	TPE	MT-TPE
-1	30	0.1951033	0.1907713	0.1906894	0.1653005	0.1523856
	50	0.16692205	0.16145827	0.1492007	0.14217875	0.1211016
	100	0.05212345	0.0518602	0.05074748	0.05161398	0.04400227
	250	0.003077234	0.003076854	0.003070393	0.003070955	0.003049226
	500	0.004880053	0.004878344	0.004859538	0.004857942	0.004722201
0	30	0.0815178	0.08090337	0.07614947	0.07547992	0.07070461
	50	0.03689775	0.03683210	0.03661245	0.03747151	0.03194901
	100	0.03717211	0.03705234	0.03572300	0.03652301	0.03358286
	250	0.003156653	0.003156227	0.003149846	0.003153232	0.003146206
	500	0.001452675	0.001452589	0.001449763	0.001450779	0.001446421
+1	30	0.0490708	0.04891634	0.04802042	0.05015392	0.0467922
	50	0.03376005	0.03368939	0.03323736	0.03414098	0.03300157
	100	0.006036703	0.006035583	0.006017863	0.006134425	0.005967314
	250	0.0002410299	0.0002410288	0.0002409843	0.0002409918	0.0002409492
	500	0.0004991241	0.0004991173	4.991239e-04	0.0004988834	0.000550114

Table 8: Summary of the simulated data of the estimators for $p = 4$ and $\rho = 0.99$ at 0.2 level of outlier

β_0	n	MLE	RIDGE1	MT-Ridge	TPE	MT-TPE
-1	30	0.4198819	0.4090720	0.38552534	0.5281544	0.29453766
	50	0.0626447517	0.0624641977	0.06175070	0.0617163322	0.05162347
	100	0.02115796	0.0211431	0.02097002	0.02143592	0.02087984
	250	0.007598569	0.007596592	0.007560765	0.007570433	0.007441784
	500	0.003942042	0.003941497	0.003932439	0.003933453	0.003887565
0	30	0.3133638	0.3089864	0.2991399	0.4640804	0.2433861
	50	0.006004936	0.006004356	0.005996267	0.005994433	0.005905879
	100	0.003318064	0.003317884	0.003313961	0.003316589	0.003304332
	250	0.0007920301	0.0007920211	0.0007916006	0.0007920151	0.0007903347
	500	6.540178e-05	6.540174e-05	6.539759e-05	6.539934e-05	6.538961e-05
+1	30	0.1260908	0.1255252	0.1201554	0.1436472	0.1191191
	50	0.02449153	0.02446903	0.02431152	0.02470685	0.02377550
	100	0.004070807	0.004070493	0.004060908	0.004081547	0.004053032
	250	0.0006179888	0.0006179838	0.0006177842	0.0006178221	0.0006175947
	500	0.0002180985	0.0002180979	0.0002180704	0.0002180742	0.0002180384

Table 9: Summary of the simulated data of the estimators for $p = 5$ and $\rho = 0.99$ at 0.2 level of outlier

β_0	n	MLE	RIDGE1	MT-Ridge	TPE	MT-TPE
-1	30	0.2351629	0.2333389	0.2298596	0.2738194	0.1644140
	50	0.3703659	0.3659753	0.33997626	0.5839340	0.27890673
	100	0.03557225	0.03553832	0.03549224	0.03543567	0.03452888
	250	0.002266638	0.002266569	2.266638e-03	0.002264608	0.00135008
	500	0.001249046	0.001249023	0.00124826	0.001248317	0.00124677
0	30	0.1249072	0.1246875	0.1231694	0.1635118	0.1239407
	50	0.00834909	0.008348179	0.008340567	0.008328493	0.008327629
	100	0.0007934455	0.0007934404	0.0007929569	0.0007933117	0.0007927572
	250	0.002228419	0.002228347	0.002226124	0.002226828	0.002220668
	500	0.0005133787	0.0005133758	0.0005130817	0.0005132561	0.0005129411
+1	30	0.06159215	0.06154444	0.06115958	0.08902041	0.06054266
	50	0.001941365	0.001941331	0.001939803	0.001944214	0.001938152
	100	0.03036582	0.03034339	0.03000693	0.03069241	0.02973167
	250	0.0002189877	0.0002189874	0.000218966	0.0002189707	0.0002189374
	500	0.0009842936	0.0009842808	0.0009838728	0.0009839541	0.0009833016

3.2 Results of the Accident Data

In addressing the challenges posed by multicollinearity and outliers within the dataset, Accident data from Lagos State for January to December, from 2013 to 2022 was analysed. The dataset includes a dependent variable (Y- total accident) and five independent variables (Xi_1 – Speed Violation, Xi_2 – Usage of Phone while Driving, Xi_3 - Overloading, Xi_4 – Mechanical Deficient Vehicle and Xi_5 – Route Violation). The existence of multicollinearity was established using correlation matrix (table 10). In addition, the response variable which happened to be a count data also exhibits presence of outliers which was confirmed using grubbs test as presented in table 11. Consequently, the conclusion is drawn that the model exhibits both multicollinearity and outliers in y-direction, making it a more suitable fit for this study.

Table 10: Correlation Matrix of the Accident Data

	X_1	X_2	X_3	X_4	X_5
X_1	1	.994**	.919**	.973**	.824**
X_2	.994**	1	.947**	.986**	.800**
X_3	.919**	.947**	1	.966**	.850**
X_4	.973**	.986**	.966**	1	.800**
X_5	.824**	.800**	.850**	.800**	1

Table 11: Grubbs Test for Outlier

Statistics	G	U	p-value
Value	4.22727	0.84857	0.000672
Decision	alternative hypothesis: highest value 276 is an outlier		

Table 12: Poisson regression estimates

	MLE	RIDGE1	MT-RIDGE	TPE	MT-TPE
Intercept	4.726562777	4.725842249	4.6414391676	4.710677181	**4.6265706579**
SE	0.053692195	0.053691673	0.057409421	0.053680330	0.057396905
Z-value	89.4286201	89.4181915	80.023291846	89.1911496	79.8215625
p-value	0.000000e+00	0.000000e+00	0.000000e+00	0.000000e+00	0.0000000000
X_1	-0.01019844	-0.01014913	-0.011185136	-0.009111443	**-0.0101677430**
SE	0.004166184	0.004166144	0.004460114	0.004165281	0.004459207
Z-value	-2.2616707	-2.2517560	-1.787355660	-2.0359203	-1.5951930
p-value	2.371775e-02	2.433770e-02	7.388002e-02	4.175836e-02	0.1106690864
X_2	0.014428137	0.014374318	0.0250833312	0.013241613	**0.0239728136**
SE	0.005426501	0.005426451	0.005828721	0.005425372	0.005827542
Z-value	2.3148513	2.3065653	3.951992752	2.1261498	3.7923388
p-value	2.062107e-02	2.107906e-02	7.750309e-05	3.349079e-02	0.0001492351
X_3	-0.003674270	-0.00366620	0.0035023814	-0.003496725	**0.0036685706**
SE	0.002585221	0.002585223	0.002883202	0.002585263	0.002883277
Z-value	-0.3304519	-0.3279591	2.226789479	-0.2738013	2.2730084
p-value	7.410585e-01	7.429426e-01	2.596135e-02	7.842373e-01	0.0230256740
X_4	0.013583302	0.013579074	-0.0187152899	0.013490500	**-0.0188022194**
SE	0.006508651	0.006508748	0.007186018	0.006510844	0.007188209
Z-value	3.8768111	3.8762637	-3.424888533	3.8644769	-3.4329553
p-value	1.058344e-04	1.060727e-04	6.150517e-04	1.113275e-04	0.0005970406
X_5	-0.001729168	-0.00171466	0.0003145788	-0.001409244	**0.0006140153**
SE	0.001865911	0.001865871	0.002014593	0.001865011	0.002013754
Z-value	-4.7210067	-4.7146304	-0.000318941	-4.5757455	0.1241648
p-value	2.346802e-06	2.421499e-06	9.997455e-01	4.745272e-06	0.9011847993
Wald (p-value)	0	0	0	0	**0**
MSE	0.0034881730	0.0034876480	0.0034870955	0.0037203270	**0.0034597819**

Table 12 presents Poisson regression results for accident data using five estimators; MLE, Ridge, MT-Ridge, TPE and MT-TPE. The result shows that the intercepts for the different estimators remained positive and significant across all the estimators ($p < 0.05$), indicating a strong baseline accident count. Examining the predictors, the coefficients maintained their expected signs, but magnitudes and significance vary depending on the estimator. This shows the impact of shrinkage and robustness on the parameter estimates.

The result for coefficients X_1 is negative across the estimators examined, suggesting that the predictors reduced accident occurrence. The MLE and Ridge estimators produce significant results ($p < 0.05$) with small standard errors, whereas MT-Ridge and MT-TPE have reduced significance and increased standard error. This affirmed that the robust estimators, especially MT-TPE, down-weight the influence of outliers leading to more conservative but reliable inference.

In terms of $X_2's$ all the estimators yielded positive coefficients, affirming that the factor increases expected accident counts. MT-Ridge and MT-TPE produce larger coefficients with small p-values, affirming the ability of robust methods to detect stronger effects when classical estimators may underestimate them as a result of multicollinearity or variance inflation. The higher standard errors for MT-based techniques are offset by the higher coefficient magnitudes, resulting in robust and statistically significant effects.

For X_3, classical estimator shows negative and non-significant effects while MT-Ridge and MT-TPE show positive and significant coefficients ($p < 0.05$). The shift suggests that extreme accident counts masked the true effect in non-robust models, and the MT-TPE estimator corrects for this, providing better parameter estimation.

For the explanatory variable X_4, the classical models yielded positive and significant coefficients ($p < 0.05$); however, MT-Ridge and MT-TPE reverse the sign while remaining significant. This affirms that robust estimation uncovers the underlying relationships obscured by the multicollinearity and influential observations, highlighting the importance of robust methods in accident modelling. There is an increase in standard errors for the MT-based estimators, however they remain reasonable relative to coefficient magnitude, preserving statistical reliability.

Lastly, X_5 exhibited negative and significant coefficients for classical estimators, which suggests a protective effect on accident occurrence. However, MT-ridge and MT-TPE, yielded coefficients close to zero and non-significant (p>0.05), indicating that the robust techniques appropriately reduced misleading effects caused by outliers or correlated predictors.

In general, the MT-TPE estimator is effective for accident data because of the count nature of the response variable which may exhibit extreme values and correlated predictors. The estimators combined the strength of two-parameter shrinkage and M-estimator to stabilize coefficient estimates and provide a more reliable inference. These two challenges were observed in the accident dataset, in which MT-TPE effectively down-weights influential values in the response variable and also shrinks unstable coefficients, resulting in meaningful Wald statistics and p-values that accurately shows predictor significance.

The simulation studies and the real-life data affirmed the theoretical with practical superiority of the MT-TPE estimator stability which enabled policymakers to identify risk factors thereby guiding targeted interventions in the areas of improved road design and traffic enforcement.

In the half-normal plot (figure 1), each of the points represents an ordered absolute residual from the fitted Poisson regression model plotted against the theoretical half-normal quantiles. The diagnostic is to visually detect potential outliers or influential observations by identifying points that deviate substantially from the reference pattern. Points 13 and 101 are labelled because they are the farthest from the red reference line, indicating that their residual values are disproportionately large relative to the rest of the data. These observations do not follow the expected half-normal trend and therefore stand out as extreme value since a half-normal plot assumes that well-behaved residuals should align roughly along the reference line.

The diagnostic plot supports the overall conclusion that the dataset contains response outliers-consistent with the contamination previously introduced and with the typical behaviour of accident data, which often include sporadic extreme counts. The presence of these outliers validates the use of robust estimators such as MT-TPE, which is specifically designed to down-weight unusually large residuals. Hence, figure 1 provides empirical justification for adopting robust estimation techniques in modelling accident counts.

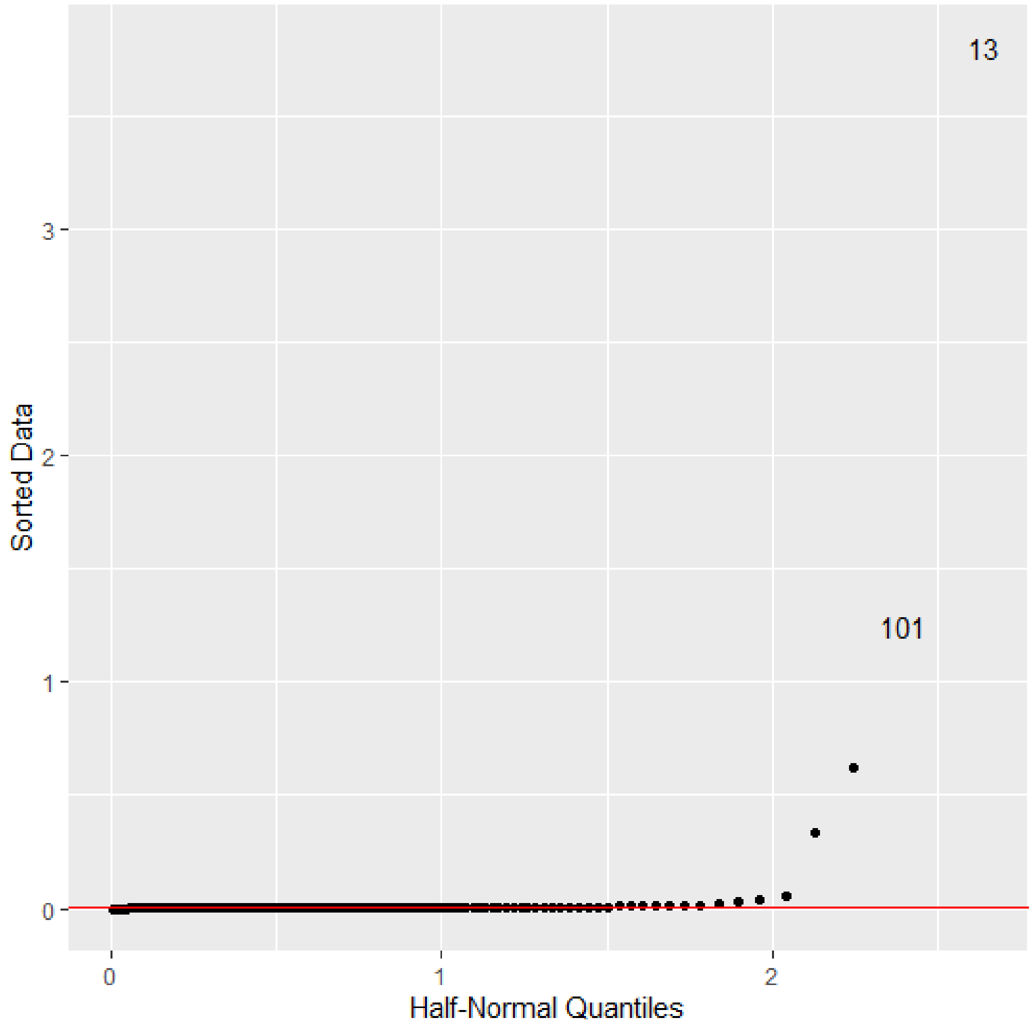

Figure 1: Diagnostic plot of the Accident data

4. Conclusion

The issues of correlated predictors and outliers in the response variable complications have frequently been ignored in statistical modeling, particularly in regression, even though increasing evidence suggests that they have a large effect on the precision of the estimated values. The recent work by Lukman et al. (2023) who developed a mixture of the ridge approach with powerful estimators and Arum et al. (2022) who creatively built a modified jackknife ridge approach by adjusting it to robustness has significantly contributed to the development of estimators that will not falter when the data are not favorable.

It is based on this that the current research extends this by substituting the traditional ridge component with a more adaptable two-parameters estimator and also incorporating it with the transformed M robustness framework. The result of this combination is the new estimator referred to as two-parameters transformed M-estimator that shows better performance across all simulation conditions, produced the lowest mean squared error and was always much better than the existing estimators, especially in conditions that were characterised by intense predictor dependence and outliers in responses. The results are supported by the analysis of real data with the help of the accident dataset where the estimator showed significant improvement in terms of stability and accuracy.

Generally, the findings supported the promise of MT-TPE as an effective substitute to analyse complex count data, providing a greater efficient estimator with better predictive accuracy. Future studies can further develop this research by adding adaptive tuning techniques, high dimensional variants in order to further show its practical and theoretical capacity.

5. Acknowledgements

The authors would like to express their gratitude to all those who contributed to the success of the research work.

6. References

Akay, K. U., & Ertan, E. (2022). A new improved Liu-type estimator for Poisson regression models. *Hacettepe Journal of Mathematics and Statistics*, *51*(5), 1484–1503.

Amin, M., Akram, M. N., & Amanullah, M. (2022). On the James-Stein estimator for the Poisson regression model. *Communications in Statistics - Simulation and Computation*, *51*(10), 5596–5608. https://doi.org/10.1080/03610918.2020.1775851

Amin, M., Akram, M. N., & Kibria, B. M. G. (2021). A new adjusted Liu estimator for the Poisson regression model. *Concurrency and Computation: Practice and Experience*, *33*(20), Article e6340. https://doi.org/10.1002/cpe.6340

Arum, K. C., Ugwuowo, F. I., & Oranye, H. E. (2022). Robust modified jackknife ridge estimator for the Poisson regression model with multicollinearity and outliers. *Scientific African*, *17*, Article e01386. https://doi.org/10.1016/j.sciaf.2022.e01386

Asar, Y., & Genç, A. (2018). A new two-parameter estimator for the Poisson regression model. *Iranian Journal of Science and Technology, Transactions A: Science*, *42*(2), 793–803. https://doi.org/10.1007/s40995-017-0174-4

Ayinde, K., Lukman, A. F., & Arowolo, O. (2015). Robust regression diagnostics of influential observations in linear regression model. *Open Journal of Statistics*, *5*(4), 273–283. https://doi.org/10.4236/ojs.2015.54030

Belsley, D. A., Kuh, E., & Welsch, R. E. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity*. John Wiley & Sons.

Bianco, A. M., & Yohai, V. J. (1996). Robust estimation in the logistic regression model. In H. Rieder (Ed.), *Robust statistics, data analysis, and computer intensive methods* (pp. 17–34). Springer.

Chatterjee, S., & Hadi, A. S. (1986). Influential observations, high leverage points, and outliers in linear regression. *Statistical Science*, *1*(3), 379–416.

Coxe, S., West, S. G., & Aiken, L. S. (2009). The analysis of count data: A gentle introduction to Poisson regression and its alternatives. *Journal of Personality Assessment*, *91*(2), 121–136. https://doi.org/10.1080/00223890802634175

Croux, C., & Haesbroeck, G. (2003). Implementing the Bianco and Yohai estimator for logistic regression. *Computational Statistics & Data Analysis*, *44*(1–2), 273–295. https://doi.org/10.1016/S0167-9473(03)00042-2

David, F. N. (1962). *Games, gods and gambling: The origins and history of probability and statistical ideas from the earliest times to the Newtonian era*. Griffin.

Ertan, E., & Akay, K. U. (2023). A new class of Poisson ridge-type estimator. *Scientific Reports*, *13*(1), Article 5940. https://doi.org/10.1038/s41598-023-32715-w

Fox, J. (2016). *Applied regression analysis and generalized linear models* (3rd ed.). SAGE Publications.

Huber, P. J. (1973). Robust regression: Asymptotics, conjectures and Monte Carlo. *Annals of Statistics*, *1*(5), 799–821.

Jadhav, N. H. (2021). A new linearized ridge Poisson estimator in the presence of multicollinearity. *Journal of Applied Statistics*, *48*(13-15), 2606–2621. https://doi.org/10.1080/02664763.2021.1887103

Künsch, H. R., Stefanski, L. A., & Carroll, R. J. (1989). Conditionally unbiased bounded-influence estimation in general regression models, with applications to generalized linear models. *Journal of the American Statistical Association*, *84*(406), 460–466.

Lukman, A. F., Adewuyi, E., Månsson, K., & Kibria, B. M. G. (2021). A new estimator for the multicollinear Poisson regression model: Simulation and application. *Scientific Reports*, *11*(1), Article 3732. https://doi.org/10.1038/s41598-021-82582-w

Lukman, A. F., Aladeitan, B., Ayinde, K., & Abonazel, M. R. (2022). Modified ridge-type for the Poisson regression model: Simulation and application. *Journal of Applied Statistics*, *49*(12), 3121–3134. https://doi.org/10.1080/02664763.2021.1889998

Lukman, A. F., Arashi, M., & Prokaj, V. (2023). Robust-stein estimator for overcoming outliers and multicollinearity. *Scientific Reports*, *13*(1), Article 5100. https://doi.org/10.1038/s41598-023-32219-7

Månsson, K., & Shukur, G. (2011). A Poisson ridge regression estimator. *Economic Modelling*, *28*(4), 1475–1481. https://doi.org/10.1016/j.econmod.2011.02.023

Månsson, K., Kibria, B. M. G., Sjolander, P., & Shukur, G. (2012). Improved Liu estimators for the Poisson regression model. *International Journal of Statistics and Probability*, *1*(1), 2–6.

Olaluwoye, O. T., Lukman, A. F., Alrasheedi, M. A., et al. (2025). Robust estimation methods for addressing multicollinearity and outliers in beta regression models. *Scientific Reports*, *15*, Article 11649. https://doi.org/10.1038/s41598-025-85553-7

Oranye, H. E., & Ugwuowo, F. I. (2021). Modified jackknife Kibria–Lukman estimator for the Poisson regression model. *Concurrency and Computation: Practice and Experience*, *34*(6), Article e6751. https://doi.org/10.1002/cpe.6751

Qasim, M., Kibria, B. M. G., Månsson, K., & Sjölander, P. (2020). A new Poisson Liu regression estimator: Method and application. *Journal of Applied Statistics*, *47*(12), 2258–2271. https://doi.org/10.1080/02664763.2019.1707485

Rashad, N. K., & Algamal, Z. Y. (2019). A new ridge estimator for the Poisson regression model. *Iranian Journal of Science and Technology, Transactions A: Science*, *43*(6), 2921–2928. https://doi.org/10.1007/s40995-019-00773-y

Rousseeuw, P. J., & Yohai, V. (1984). Robust regression by means of S estimators in robust and nonlinear time series analysis. In J. Franke, W. Härdle, & R. D. Martin (Eds.), *Lecture notes in statistics* (Vol. 26, pp. 256–274). Springer-Verlag.

Schober, P., & Vetter, T. R. (2021). Count data in medical research: Poisson regression and negative binomial regression. *Anesthesia & Analgesia*, *132*(5), 1378–1379. https://doi.org/10.1213/ANE.0000000000005398

Türkan, S., & Özel, G. (2016). A new modified jackknifed estimator for the Poisson regression model. *Journal of Applied Statistics*, *43*(10), 1892–1905. https://doi.org/10.1080/02664763.2015.1125861

Valdora, M., & Yohai, V. J. (2014). Robust estimation in generalized linear models. *Journal of Statistical Planning and Inference*, *146*, 31–48. https://doi.org/10.1016/j.jspi.2013.09.016