A NEW CLASS OF ROBUST POISSON RIDGE-TYPE ESTIMATOR: SIMULATION AND THEORY APPROACH

Albert, Seno Ofonmbuk*
Department of Statistics, University of Abuja
Olanrewaju Samuel Olayemi
Department of Statistics, University of Abuja
Oguntade, Emmanuel Segun
Department of Statistics, University of Abuja
Corresponding Author: senosuzie@gmail.com

ABSTRACT

This study evaluates the performance of various Poisson regression estimators under conditions of multicollinearity and data contamination (outliers) using simulated data. Specifically, the mean squared errors (MSEs) of traditional estimators Robust Poisson Maximum Likelihood Estimator (RPMLE), Robust Poisson Ridge (RPR), Robust Poisson Liu (RPL), and Robust Poisson Kibria–Lukman (RPK-L) were compared with three newly proposed estimators (prop1, prop1ME, and prop1HM) across different sample sizes (n = 200, 300, 400), correlation strengths (ρ = 0.8–0.99), and contamination levels (τ = 20%). Results consistently show that RPMLE yields the highest MSE, confirming its vulnerability to outliers and multicollinearity. While classical shrinkage methods perform better than RPMLE, their efficiency declines with increasing ρ and τ. In contrast, the proposed estimators demonstrate superior robustness, with prop1HM achieving the lowest MSE across nearly all designs. Notably, prop1HM maintains its performance even under high multicollinearity and small sample sizes, highlighting its adaptability. As contamination intensifies, the performance gap between the proposed and traditional estimators widens, emphasizing the resilience of the new methods. These findings suggest that prop1HM, in particular, is a reliable and efficient alternative for Poisson regression analysis in the presence of multicollinearity and outlier contamination.

Keywords: Poisson, Regression, Outlier, Multicollinearity, Estimate,

ISSN NUMBER: 1116-249X

INTRODUCTION

Count data frequently arise in fields such as epidemiology, econometrics, and engineering. One of the most commonly employed methods for modeling such data is the Poisson regression model (PRM), which assumes that the conditional mean and variance of the response variable are equal. Despite its theoretical appeal, the practical application of the PRM is often hampered by two prominent challenges: outliers and multicollinearity.

Outliers, which are observations that deviate markedly from the general pattern of the data, can severely distort maximum likelihood estimates, reduce model efficiency, and yield misleading inferential results. Conversely, multicollinearity a condition where predictor variables are highly correlated inflates the variance of coefficient estimates, leading to instability and unreliable inference (Cantoni & Ronchetti, 2001; Hall & Shen, 2010)

To address these issues, various statistical techniques have been proposed**.** Robust regression methods are commonly employed to mitigate the influence of outliers (Chen *et al*., 2014; Marazzi, 2021), whereas biased estimation techniques, such as ridge and Liu-type estimators, are often used to alleviate multicollinearity (Månsson & Shukur, 2011; Lukman *et al*., 2021). However, most existing methods focus on resolving either outliers or multicollinearity in isolation. In practice, datasets often suffer from both problems simultaneously, diminishing the effectiveness of methods designed to handle only one issue (Abonazel & Saber, 2020; Dawoud & Abonazel, 2022).

To bridge this methodological gap, several robust Poisson estimators have been introduced. Notable examples include those proposed by Cantoni and Ronchetti (2001), Hall and Shen (2010), Chen et al. (2014, 2018), Hosseinian and Morgenthaler (2011), and Abonazel and Saber (2020). Among them, the robust Poisson estimator by Hosseinian and Morgenthaler (2011) based on the weighted maximum likelihood method (WMLM) has gained significant traction for its efficacy in

ISSN NUMBER: 1116-249X

managing outliers in the response variable. However, integrated solutions that can effectively tackle both outliers and multicollinearity in the PRM are still scarce. Although some attempts have been made such as the Robust Poisson Ridge Regression (RPRR) estimator more comprehensive and efficient techniques are needed (Abonazel *et al*., 2022). This motivates the current study, which introduces novel robust Poisson two-parameter estimators designed specifically to handle both multicollinearity and outliers simultaneously. These estimators are evaluated using theoretical analysis and simulation studies

METHODOLOGY

The Poisson regression model (PRM) is a standard framework for analyzing count data, assuming the response variable $\left(Y_i\right)$ follows a Poisson distribution with mean $\mu_i\left(\beta\right)=\exp\left(X_i\beta\right)$ where X_i denotes the i-th observation in the design matrix and β is a vector of unknown parameters. Estimation of these parameters is typically carried out using the Pseudo-Maximum Likelihood (PML) method, operationalized through the Iteratively Reweighted Least Squares (IWLS) algorithm. This approach ensures consistent and asymptotically efficient parameter estimates under the standard assumptions of the Poisson model. The PML estimator is defined as:

$$\beta_{PMLE}=\left(J\right)^{-1}X^TWs \tag{1}$$

Where $J=X^TWX$, s is an n – dimensional vector with the ith element $s=\log\hat{\mu}_i+\frac{\left(y_i-\hat{\mu}_i\right)}{\hat{\mu}_i}$ *and* $\hat{W}=diag\left[\hat{\mu}_i\right]$. The MLE is asymptotically normally distributed, with its covariance matrix derived from the inverse of the observed information matrix, which corresponds to the negative second derivative of the log-likelihood function and expressed as:

$$Cov\left(\hat{\beta}_{PMLE}\right)=\left[-E\left(\frac{\partial l}{\partial\beta_j\partial\beta_j^T}\right)\right]^{-1}=J^{-1} \tag{2}$$

ISSN NUMBER: 1116-249X

The mean square error can calculated as follows:

$$MSE\left(\hat{\beta}_{PMLE}\right)=\sum_{i=1}^{p}\frac{1}{\lambda_i} \tag{3}$$

Where λ_i is the ith eigen value of the matrix J. The robust Poisson estimator by Hosseinian and Morgenthaler (2011) based on the weighted maximum likelihood method (WMLM) has gained significant traction for its efficacy in managing outliers in the response variable (y). The weighted function $\left(W^*\right)$ of the WMLM depends on μ_i and two constants ($c1$ and $c2$):

$$W^* = \begin{cases} 1 if \ \frac{m}{c_1} < \mu_i < c_1 m; \\ \frac{c_1\mu_i}{m} if \ \mu_i < \frac{m}{c_1}; \\ 0 \quad otherwise \end{cases}$$

. Where m is known as the median of $\mu_i, c_1 = 2 \, and \, c_2 = 3$. Then, the WMLM estimation is defined as:

$$\sum_{i=1}^{n}\frac{\mu_i'}{\mu_i}W^*\left(y_i-\mu_i\right)x_i=0 \tag{4}$$

A Robust Poisson maximum likelihood (RPML) estimator can be obtained by applying the Newton–Raphson algorithm to the system of equations in Equation (4). For further methodological details, refer to Hosseinian (2009) and Hosseinian and Morgenthaler (2011).Then, the RPML estimator by Abonazel *et al* (2022) is defined as:

$$\hat{\beta}_{RPMLE}=\left(J\right)^{-1}X^T\hat{W}^*\hat{s}^*, \tag{5}$$

ISSN NUMBER: 1116-249X

Where $\mathcal{J} = X^T \hat{W}^* X, s$ is an n – dimensional vector with the ith element $\hat{s}^* = \log \hat{\mu}_i^* + \frac{\left(y_i^* - \hat{\mu}_i^*\right)}{\hat{\mu}_i^*}$ $and\, \hat{W}^* = diag\left[\hat{\mu}_i^*\right]$. The RPMLE follows, with its covariance matrix defined as the inverse of the observed information matrix, derived from the second derivative of the log-likelihood function.

$$Cov\left(\hat{\beta}_{RPMLE}\right) = \left[-E\left(\frac{\partial l}{\partial \beta_j \partial \beta_j^T}\right)\right]^{-1} = \mathcal{J}^{*-1} \tag{6}$$

The mean square error can calculated as follows:

$$MSE\left(\hat{\beta}_{RPMLE}\right) = \sum_{i=1}^{p} \frac{1}{\lambda_i^*} \tag{7}$$

Where γ^* $and\, \lambda_j^*$ are the eigenvectors and Eigen values of $\mathcal{J}$, where $\mathcal{J} = \gamma^{*'} \Lambda^* \gamma^*$ $with\, \Lambda^* = diag\left(\lambda_j^*\right)$

Robust Poisson Ridge Regression

In a recent contribution, Abonazel and Dawoud (2022) introduced the Robust Poisson Ridge Regression (RPRR) estimator, designed to concurrently address the issues of outliers and multicollinearity within the Poisson Regression Model (PRM). The formulation of the RPRR estimator is defined as follows:

$$\hat{\beta}_{RPRR} = \left(\mathcal{J} + kI\right)^{-1} X^T \hat{W}^* \hat{s}^* \tag{8}$$

And the MSE of the RPRR is:

$$MSE\left(\hat{\beta}_{PRRE}\right) = \sum_{i=1}^{p+1} \frac{\lambda_i^*}{\left(\lambda_i^* + k\right)^2} + \sum_{i=1}^{p} \frac{k^2 \hat{\alpha}_i^{*2}}{\left(\lambda_i^* + k\right)^2} \tag{9}$$

ISSN NUMBER: 1116-249X

Robust Poisson Liu Regression

In a recent study, Dawoud *et al.* (2022) developed the Robust Poisson Ridge Regression (RPL) estimator, which is specifically designed to simultaneously mitigate the effects of outliers and multicollinearity within the Poisson Regression Model (PRM). The RPRR estimator is formally expressed as:

$$\hat{\beta}_{RPL} = \left(\mathcal{F} + I\right)^{-1}\left(\mathcal{F} + dI\right)\hat{\beta}_{RPMLE} \tag{10}$$

And the MSE of the RPL is:

$$MSE\left(\hat{\beta}_{PRRE}\right) = \sum_{i=1}^{p+1} \frac{\left(\lambda_i^* + d\right)}{\lambda_i^*\left(\lambda_i^* + 1\right)^2} + \left(1 - d\right)^2 \sum_{i=1}^{p+1} \frac{\hat{\alpha}_i^{*2}}{\left(\lambda_i^* + 1\right)^2} \tag{11}$$

Robust Poisson Kibra Lukman Regression

Dawoud et al. (2022) recently introduced the Robust Poisson K-L Ridge (RPKL) estimator as a methodological advancement aimed at addressing the dual challenges of outliers and multicollinearity within the Poisson Regression Model (PRM). The formulation of the RPKL estimator is given by:

$$\hat{\beta}_{RPKL} = \left(\mathcal{F} + kI\right)^{-1}\left(\mathcal{F} - kI\right)\hat{\beta}_{RPMLE} \tag{12}$$

And the MSE of the RPKL is:

$$MSE\left(\hat{\beta}_{PRKL}\right) = \sum_{i=1}^{p+1} \frac{\left(\lambda_i^* - k\right)}{\lambda_i^*\left(\lambda_i^* + k\right)^2} + 4k^2 \sum_{i=1}^{p+1} \frac{\hat{\alpha}_i^{*2}}{\left(\lambda_i^* + k\right)^2} \tag{13}$$

The proposed estimator

Given that the PL, PKL, and PMKL estimators have demonstrated greater efficiency than the Poisson ridge regression (PRR) estimator within the Poisson regression framework (PRM), robust counterparts of these estimators have been developed to address both multicollinearity and the

ISSN NUMBER: 1116-249X

presence of outliers, thereby extending one-parameter methods to more resilient alternatives. Recently, Albert *et al*. (2025) introduced a novel two-parameter estimator within the PRM context, which has shown superior performance compared to existing estimators such as PMLE, PRR, PL, and PKL. Building on this advancement, we propose an extension of this two-parameter approach to simultaneously address multicollinearity and outlier influence. The resulting estimator, denoted as RProp1, is formally defined as follows:

$$\hat{\beta}_{R\text{Pr}op1} = \left(\not{J} + I\right)^{-1}\left(\not{J} + dI\right)\left(\not{J} + kI\right)^{-1}\left(\not{J} - kI\right)\hat{\beta}_{RPMLE} \tag{14}$$

And the MSE of the RProp1 is:

$$MSE\left(\hat{\beta}_{R\text{Pr}op1}\right) = \sum_{i=1}^{p}\left[\frac{\left(\lambda_i^* - k\right)^2\left(\lambda_i^* + d\right)^2}{\lambda_i^*\left(\lambda_i^* + I\right)^2\left(\lambda_{i\,i}^* + k\right)^2}\right] + \sum_{i=1}^{p}\left[\frac{\left(\left(2k+1-d\right)\lambda_i^* + k\left(d+1\right)\right)^2}{\left(\lambda_i^* + I\right)^2\left(\lambda_i^* + k\right)^2}\right]\hat{\alpha}_i^{*2} \tag{15}$$

Superiority of Proposed Estimators

Lemma: Suppose that $\alpha_i = M_i y, i = 1,2$ be the two competing estimators of α. Assume that $I = Cov(\hat{\alpha}_1) - Cov(\hat{\alpha}_2) > 0$, then, $MSEM\left(\hat{\alpha}_1\right) - MSEM\left(\hat{\alpha}_2\right) > 0$ if and only if $v_2'\left(I + v_1 v_1'\right) \leq 1$, where v_i denotes the bias of $\hat{\alpha}_i$, according to Trenkler and Toutenburg (1990).

The mean square error matrix (MSEM) of an estimator $\hat{\beta}$ is defined as

MSEM $\left(\hat{\beta}\right) = Cov\left(\hat{\beta}\right) + bias\left(\hat{\beta}\right) bias\left(\hat{\beta}\right)'$

MSEM $\left(\hat{\beta}\right) = Cov\left(\hat{\beta}\right) + bias\left(\hat{\beta}\right) bias\left(\hat{\beta}\right)'$

Where $Cov\left(\hat{\beta}\right)$ is the dispersion matrix and $bias\left(\hat{\beta}\right) = E\left(\hat{\beta}\right) - \beta$

The bias and dispersion matrix of $\hat{\alpha}_{prop1}$ can be computed as follows:

ISSN NUMBER: 1116-249X

$$Bias\left(\hat{\beta}_{Rprop1}\right)=\left[\left(\Lambda^{*}+1\right)^{-1}\left(\Lambda^{*}+dI\right)\ \left(\Lambda^{*}+kI\right)^{-1}\left(\Lambda^{*}-kI\right)-I\right]\alpha^{*} \tag{16}$$

$$Cov\left(\hat{\beta}_{Rprop1}\right)=\left(\Lambda^{*}+1\right)^{-1}\left(\Lambda^{*}+dI\right)\ \left(\Lambda^{*}+kI\right)^{-1}\left(\Lambda^{*}-kI\right)\Lambda^{*-1}\left(\Lambda^{*}+1\right)^{-1}\left(\Lambda^{*}+dI\right)\ \left(\Lambda^{*}+kI\right)^{-1}\left(\Lambda^{*}-kI\right) \tag{17}$$

The MSEM and MSE in terms of eigenvalues are defined respectively as

$$MSEM\left(\hat{\beta}_{Rprop1}\right)=Cov\left(\hat{\beta}_{Rprop1}\right)+Bias\left(\hat{\beta}_{Rprop1}\right)Bias\left(\hat{\beta}_{Rprop1}\right)'$$

$$=\begin{array}{l}\left(\Lambda^{*}+1\right)^{-1}\left(\Lambda^{*}+dI\right)\left(\Lambda^{*}+kI\right)^{-1}\left(\Lambda^{*}-kI\right)\Lambda^{*-1}\left(\Lambda^{*}+1\right)^{-1}\left(\Lambda^{*}+dI\right)\left(\Lambda^{*}+kI\right)^{-1}\left(\Lambda^{*}-kI\right)+\\ \left[\left(\Lambda^{*}+1\right)^{-1}\left(\Lambda^{*}+dI\right)\left(\Lambda^{*}+kI\right)^{-1}\left(\Lambda^{*}-kI\right)-I\right]\alpha^{*}\alpha^{*'}\left[\left(\Lambda^{*}+1\right)^{-1}\left(\Lambda^{*}+dI\right)\left(\Lambda^{*}+kI\right)^{-1}\left(\Lambda^{*}-kI\right)-I\right]\end{array} \tag{18}$$

$$MSE\left(\hat{\beta}_{Rprop1}\right)=tr\left(MSEM\left(\hat{\beta}_{Rprop1}\right)\right)$$

$$MSE\left(\hat{\beta}_{Rprop1}\right)=\sum_{i=1}^{p+1}\left[\frac{\left(\lambda_i^{*}-k\right)^2\left(\lambda_i^{*}+d\right)^2}{\lambda_i^{*}\left(\lambda_i^{*}+I\right)^2\left(\lambda_i^{*}+k\right)^2}\right]+\sum_{i=1}^{p}\left[\frac{\left(\left(2k+1-d\right)\lambda_i^{*}+k\left(d+1\right)\right)^2}{\left(\lambda_i^{*}+I\right)^2\left(\lambda_i^{*}+k\right)^2}\right]\alpha_i^{*2} \tag{19}$$

Comparison of $\hat{\alpha}_{Rprop1}$ and $\hat{\alpha}_{RPMLE}$

$\hat{\beta}_{RPMLE}=\Lambda^{-1}Q^{T}X^{T}\hat{W}^{*}\hat{s}^{*}$ with $MSEM\left(\hat{\beta}_{RPMLE}\right)=\Lambda^{-1}$

Theorem 1: $\hat{\beta}_{Rprop1}$ is better than $\hat{\beta}_{RPMLE}$ if

$$\alpha'\left(\left(\Lambda^{*}+1\right)^{-1}\left(\Lambda^{*}+dI\right)\left(\Lambda^{*}+kI\right)^{-1}\left(\Lambda^{*}-kI\right)-I\right)'\left[\left(\Lambda^{*-1}-\left(\Lambda^{*}+1\right)^{-2}\left(\Lambda^{*}+dI\right)^{2}\left(\Lambda^{*}+kI\right)^{-2}\left(\Lambda^{*}-kI\right)^{2}\right)\right]^{-1}$$
$$\left(\left(\Lambda^{*}+1\right)^{-1}\left(\Lambda^{*}+dI\right)\left(\Lambda^{*}+kI\right)^{-1}\left(\Lambda^{*}-kI\right)-I\right)\alpha^{*}<1$$

Proof

The difference of the dispersion is

$$Cov\left(\hat{\beta}_{RPMLE}\right)-Cov\left(\hat{\beta}_{Rprop1}\right)=Q\left(\Lambda^{*-1}-\Lambda^{*-1}\left(\Lambda^{*}+1\right)^{-2}\left(\Lambda^{*}+dI\right)^{2}\left(\Lambda^{*}+kI\right)^{-2}\left(\Lambda^{*}-kI\right)^{2}\right)Q^{T}$$

ISSN NUMBER: 1116-249X

$$=Qdiag\left[\frac{1}{\lambda_i^*}-\frac{\left(\lambda_i^*-k\right)^2\left(\lambda_i^*+d\right)^2}{\lambda_i^*\left(\lambda_i^*+1\right)^2\left(\lambda_i^*+k\right)^2}\right]_{i=1}^{p}Q^T \tag{20}$$

It is observed that $\left(\Lambda^{*-1}-\Lambda^{*-1}\left(\Lambda^*+1\right)^{-2}\left(\Lambda^*+dI\right)^2\left(\Lambda^*+kI\right)^{-2}\left(\Lambda^*-kI\right)^2\right)$ is positive definite since for 0 < d < 1 and k > 0, $\lambda_i^*\left(\lambda_i+1\right)^2\left(\lambda_i^*+k\right)^2-\lambda_i^*\left(\lambda_i^*-k\right)^2\left(\lambda_i^*+d\right)^2>0$. Hence, by lemma the proof is completed.

Theorem 2: $\hat{\beta}_{Rprop1}$ is better than $\hat{\beta}_{RPRRE}$ if

$$\alpha^{*'}\left(\left(\Lambda^*+1\right)^{-1}\left(\Lambda^*+dI\right)\left(\Lambda^*+kI\right)^{-1}\left(\Lambda^*-kI\right)-I\right)^{'}\left[J_1+\left(\left(\Lambda^*+k\right)^{-1}-I\right)\alpha^*\alpha^{*'}\left(\left(\Lambda^*+k\right)^{-1}-I\right)\right]^{-1}$$
$$\left(\left(\Lambda^*+1\right)^{-1}\left(\Lambda^*+dI\right)\left(\Lambda^*+kI\right)^{-1}\left(\Lambda^*-kI\right)-I\right)\alpha^*<1$$

$$J_1=Q\left(\Lambda^*\left(\Lambda^*+k\right)^{-2}-\Lambda^{*-1}\left(\Lambda^*+1\right)^{-2}\left(\Lambda^*+dI\right)^2\left(\Lambda^*+kI\right)^{-2}\left(\Lambda^*-kI\right)^2\right)Q^T$$

Proof

The difference of the dispersion is

$$Cov\left(\hat{\beta}_{RPRRE}\right)-Cov\left(\hat{\beta}_{Rprop1}\right)=Q\left(\Lambda^*\left(\Lambda^*+k\right)^{-2}-\Lambda^{*-1}\left(\Lambda^*+1\right)^{-2}\left(\Lambda^*+dI\right)^2\left(\Lambda^*+kI\right)^{-2}\left(\Lambda^*-kI\right)^2\right)Q^T$$

$$=Qdiag\left[\frac{\lambda_i^*}{\left(\lambda_i^*+k\right)^2}-\frac{\left(\lambda_i^*-k\right)^2\left(\lambda_i^*+d\right)^2}{\lambda_i^*\left(\lambda_i^*+1\right)^2\left(\lambda_i^*+k\right)^2}\right]_{i=1}^{p}Q^T \tag{21}$$

It is observed that $\left(\Lambda^*\left(\Lambda^*+k\right)^{-2}-\Lambda^{*-1}\left(\Lambda^*+1\right)^{-2}\left(\Lambda^*+dI\right)^2\left(\Lambda^*+kI\right)^{-2}\left(\Lambda^*-kI\right)^2\right)$ will be positive definite if and only if $\lambda_i^{*2}\left(\lambda_i^*+1\right)^2-\left(\lambda_i^*-k\right)^2\left(\lambda_i^*+d\right)^2>0$. For 0 < d < 1 and k > 0, by lemma 3 the proof is completed.

Theorem 3: $\hat{\alpha}_{Rprop1}$ is better than $\hat{\alpha}_{RPLE}$ if

ISSN NUMBER: 1116-249X

$$\alpha^{*'}\left(\left(\Lambda^*+1\right)^{-1}\left(\Lambda^*+dI\right)\left(\Lambda^*+kI\right)^{-1}\left(\Lambda^*-kI\right)-I\right)'\left[J_2+\left(\left(\Lambda^*+1\right)^{-1}\left(\Lambda^*+d\right)-I\right)\alpha\alpha'\left(\left(\Lambda^*+1\right)^{-1}\left(\Lambda^*+d\right)-I\right)\right]^{-1}$$
$$\left(\left(\Lambda^*+1\right)^{-1}\left(\Lambda^*+dI\right)\left(\Lambda^*+kI\right)^{-1}\left(\Lambda^*-kI\right)-I\right)\alpha^*<1$$

$$J_2=Q\left(\Lambda^{*-1}\left(\Lambda^*+d\right)^2\left(\Lambda^*+1\right)^{-2}-\Lambda^{*-1}\left(\Lambda^*+1\right)^{-2}\left(\Lambda^*+dI\right)^2\left(\Lambda^*+kI\right)^{-2}\left(\Lambda^*-kI\right)^2\right)Q^T$$

Proof

The difference of the dispersion is

$$Cov\left(\hat{\beta}_{RPLE}\right)-Cov\left(\hat{\beta}_{Rprop1}\right)=Q\left(\Lambda^{*-1}\left(\Lambda^*+d\right)^2\left(\Lambda^*+1\right)^{-2}-\Lambda^{*-1}\left(\Lambda^*+1\right)^{-2}\left(\Lambda^*+dI\right)^2\left(\Lambda^*+kI\right)^{-2}\left(\Lambda^*-kI\right)^2\right)Q^T$$

$$=Qdiag\left[\frac{\left(\lambda_i^*+d\right)^2}{\lambda_i^*\left(\lambda_i^*+1\right)^2}-\frac{\left(\lambda_i^*-k\right)^2\left(\lambda_i^*+d\right)^2}{\lambda_i^*\left(\lambda_i^*+1\right)^2\left(\lambda_i^*+k\right)^2}\right]_{i=1}^p Q^T \tag{22}$$

$\Lambda^{*-1}\left(\Lambda^*+d\right)^2\left(\Lambda^*+1\right)^{-2}-\Lambda^{*-1}\left(\Lambda^*+1\right)^{-2}\left(\Lambda^*+dI\right)^2\left(\Lambda^*+kI\right)^{-2}\left(\Lambda^*-kI\right)^2$ will be positive definite if and only if $\left(\lambda_i^*+k\right)^2-\left(\lambda_i^*-k\right)^2>0$ for k > 0 and 0 < d < 1. Hence, by lemma the proof is completed.

Theorem 4: $\hat{\alpha}_{Rprop1}$ is better than $\hat{\alpha}_{PKLE}$ if

$$\alpha^{*'}\left(\left(\Lambda^*+1\right)^{-1}\left(\Lambda^*+dI\right)\left(\Lambda^*+kI\right)^{-1}\left(\Lambda^*-kI\right)-I\right)'\left[J_3+\left(\left(\Lambda^*+k\right)^{-1}\left(\Lambda^*-k\right)-I\right)\alpha\alpha'\left(\left(\Lambda^*+k\right)^{-1}\left(\Lambda^*-k\right)-I\right)\right]^{-1}$$
$$\left(\left(\Lambda^*+1\right)^{-1}\left(\Lambda^*+dI\right)\left(\Lambda^*+kI\right)^{-1}\left(\Lambda^*-kI\right)-I\right)\alpha<1$$

$$J_3=Q\left(\Lambda^{*-1}\left(\Lambda^*-k\right)^2\left(\Lambda^*+k\right)^{-2}-\Lambda^{*-1}\left(\Lambda^*+1\right)^{-2}\left(\Lambda^*+dI\right)^2\left(\Lambda^*+kI\right)^{-2}\left(\Lambda^*-kI\right)^2\right)Q^T$$

Proof

The difference of the dispersion is

ISSN NUMBER: 1116-249X

$$Cov\left(\hat{\beta}_{RPKLE}\right)-Cov\left(\hat{\beta}_{Rprop1}\right)=Q\left(\Lambda^{*-1}\left(\Lambda^{*}-k\right)^{2}\left(\Lambda^{*}+k\right)^{-2}-\Lambda^{*-1}\left(\Lambda^{*}+1\right)^{-2}\left(\Lambda^{*}+dI\right)^{2}\left(\Lambda^{*}+kI\right)^{-2}\left(\Lambda^{*}-kI\right)^{2}\right)Q^{T}$$

$$=Qdiag\left[\frac{\left(\lambda_i^{*}-k\right)^{2}}{\lambda_i^{*}\left(\lambda_i^{*}+k\right)^{2}}-\frac{\left(\lambda_i^{*}-k\right)^{2}\left(\lambda_i^{*}+d\right)^{2}}{\lambda_i^{*}\left(\lambda_i^{*}+1\right)^{2}\left(\lambda_i^{*}+k\right)^{2}}\right]_{i=1}^{p}Q^{T} \tag{23}$$

$\Lambda^{*-1}\left(\Lambda^{*}-k\right)^{2}\left(\Lambda^{*}+k\right)^{-2}-\Lambda^{*-1}\left(\Lambda^{*}+1\right)^{-2}\left(\Lambda^{*}+dI\right)^{2}\left(\Lambda^{*}+kI\right)^{-2}\left(\Lambda^{*}-kI\right)^{2}$ will be positive definite if and only if $\left(\lambda_i^{*}+1\right)^{2}-\left(\lambda_i^{*}+d\right)^{2}>0$ for k > 0 and 0 < d < 1. Hence, by lemma the proof is completed.

Selection of biasing parameters k and d for $\hat{\beta}_{Rprop1}$

MSE $\left(\alpha(k,d)\right)=E\left[\left(\hat{\alpha}(k,d)-\alpha\right)'\left(\alpha(k,d)-\alpha\right)\right]$

g(k,d)= $MSE\left(\hat{\alpha}(k,d)\right)=tr\left[MSEM\left(\hat{\alpha}(k,d)\right)\right]$

$$MSE\left(\hat{\beta}_{Rprop1}\right)=\sum_{i=1}^{p}\left[\frac{\left(\lambda_i^{*}-k\right)^{2}\left(\lambda_i^{*}+d\right)^{2}}{\lambda_i^{*}\left(\lambda_i^{*}+I\right)^{2}\left(\lambda_i^{*}+k\right)^{2}}\right]+\sum_{i=1}^{p}\left[\frac{\left(\left(2k+1-d\right)\lambda_i^{*}+k\left(d+1\right)\right)^{2}}{\left(\lambda_i^{*}+I\right)^{2}\left(\lambda_i^{*}+k\right)^{2}}\right]\alpha_i^{*2}$$

Considering d to be fixed, an optimal value of k is the value that minimizes $MSE\left(\hat{\beta}_{Rprop1}\right)$.

Then, by differentiating g(k, d) w.r.t. k and equating to 0, we have

$$k=\frac{\lambda_i^{*}\left(\lambda_i^{*}+d\right)+\left(d-1\right)\lambda_i^{*2}\alpha_i^{*2}}{\left(\lambda_i^{*}+d\right)+\lambda_i^{*}\left(2\lambda_i^{*}+d+1\right)\alpha_i^{*2}} \tag{24}$$

However, k depends on the unknown α_i. For practical purposes, it will be replaced by its unbiased estimator $\hat{\alpha}_i$. Hence, this will be obtained

$$\hat{k}=\frac{\lambda_i^{*}\left(\lambda_i^{*}+d\right)+\left(d-1\right)\lambda_i^{*2}\hat{\alpha}_i^{*2}}{\left(\lambda_i^{*}+d\right)+\lambda_i^{*}\left(2\lambda_i^{*}+d+1\right)\hat{\alpha}_i^{*2}}$$

ISSN NUMBER: 1116-249X

Equation (24) returns the biasing parameter for the PKL estimator when d =1, which is defined as follows:

$$\hat{k} = \frac{\lambda_i^*}{1 + 2\lambda_i^*{}_i\hat{\alpha}_i^{*2}}$$

Building upon the foundational contributions of Kibra (2003), Lukman and Ayinde (2017), as well as Oladapo *et al*. (2022, 2024), the conceptual frameworks presented in their studies have been adapted to inform the development of the proposed estimator. Consequently, the corresponding shrinkage parameters considered in this study are defined as follows:

$$\hat{k}_{HM} = p\sum_{i=1}^{p}\left(\frac{\lambda_i^*\left(\lambda_i^* + d\right) + \left(d-1\right)\lambda_i^{*2}\hat{\alpha}_i^{*2}}{\left(\lambda_i^* + d\right) + \lambda_i^*\left(2\lambda_i^* + d + 1\right)\hat{\alpha}_i^{*2}}\right) \tag{25}$$

$$\hat{k}_{MED} = Median\left(\frac{\lambda_i^*\left(\lambda_i^* + d\right) + \left(d-1\right)\lambda_i^{*2}\hat{\alpha}_i^{*2}}{\left(\lambda_i^* + d\right) + \lambda_i^*\left(2\lambda_i^* + d + 1\right)\hat{\alpha}_i^{*2}}\right) \tag{26}$$

Simulation Experiment

Simulation Design : Recognizing the limitations of relying solely on theoretical analysis to assess estimator performance, a detailed Monte Carlo simulation study was implemented. The response variable was drawn from a Poisson distribution defined as $Y_i \sim Poisson\left(\mu_i\right)$, where $\mu_i = \exp(X_i\beta),\ i = 1,2,...,n$, $\beta = \left(\beta_0, \beta_1, \beta_2, ..., \beta_P\right)$, and represent model parameters. Here, X_{ith} denotes the ith row of the design matrix X. In line with the approaches used by Kibra (2003), Oladapo *et al*. (2023), and Owolabi *et al*. (2022), the explanatory variables in X were generated using the following procedure:

$$x_{ij} = (1 - rho^2)^{\frac{1}{2}} z_{ij} + rhoz_{ip+1},\ i = 1,\ 2,\ ...,\ n,\ j = 1,\ 2, ...,\ p. \tag{27}$$

ISSN NUMBER: 1116-249X

where rho $\left(\rho^2\right)$ is the correlation between the explanatory variables. The values of ρ are chosen to be 0.8, 0.9, 0.95 and 0.99. The mean function is obtained for *p* =2, 4, 10 and 12 regressors, respectively. The slope coefficients chosen so that $\sum_j^p \beta_j^2 = 1$ and $\beta_1 = \beta_2 = \ldots = \beta_p$ for sample sizes 200,400,600,800 and 1000. Some outliers were generated with different percentages (τ = 0%, 10%, and 20%) in the model by randomly replacing some values due to the selected percentage τ% from the response variable. Point of outliers; $Y_{i(Outlier)} = f * Max\left(Y_i\right) + Y_i$, where magnitude of outliers in the y-direction (f) and percentage of outliers in the y-direction (τ). Simulation experiment conducted through R programming language .

The Mean Squared Error (MSE) for each estimator is computed as follows:

$$MSE\left(\beta^*\right) = \frac{1}{1000}\sum_{i=1}^{1000}\left(\beta_i^* - \beta\right)'\left(\beta_i^* - \beta\right) \tag{28}$$

Here β_i^*, denotes the vector of predicted values obtained from the ith simulation run using one of the evaluated estimators namely PMLE, PRRE, PLE, PKLE, and PMTPLE. Among these, the estimator yielding the smallest MSE is regarded as the most effective.

Simulation Results

Table 1 reports the Mean Squared Error (MSE) of several estimators under various simulation settings characterized by different intercept values, sample sizes (n = 200, 300, 400), and levels of multicollinearity (ρ = 0.8, 0.9, 0.95, 0.99), with 20% contamination (τ = 20%) and two predictors (P = 2). Across all scenarios, the Robust Poisson Maximum Likelihood Estimator (RPMLE) consistently produced the highest MSE values, confirming its sensitivity to both multicollinearity and outlier contamination. In contrast, RPR, RPL, and RPKL estimators significantly outperformed RPMLE, with notably lower MSEs across all configurations.

ISSN NUMBER: 1116-249X

Among all the estimators, the proposed methods particularly prop1HM demonstrated the most robust and stable performance. Prop1HM frequently yielded the lowest MSE values, especially at higher correlation levels ($\rho \geq 0.95$) and larger sample sizes. This suggests a high degree of robustness and efficiency in the presence of multicollinearity and contamination. While increasing the correlation among predictors generally led to higher MSEs for most estimators, the proposed methods remained resilient, showing minimal performance degradation. Additionally, increasing the sample size led to a general reduction in MSE for all methods, with the largest improvements observed in the robust and proposed estimators.

Overall, the findings indicate that the proposed estimators, particularly prop1HM, offer substantial improvements in estimation accuracy over traditional methods under challenging data conditions, including high multicollinearity and data contamination with outliers.

Table 1: MSE of different estimators when τ=20% and P = 2

intercept	n	rho	RPML	RPR	RPL	RPK-L	prop1	prop1ME	prop1HM
		0.8	4.0264	0.6908	0.7091	0.7013	0.7091	0.7269	1.6191
		0.9	4.1941	0.7024	0.7124	0.7106	0.7124	0.7363	1.5745
		0.95	4.2060	**0.7305**	0.7308	0.7357	0.7308	0.7504	1.9863
	200	0.99	4.5897	0.9382	0.8871	0.9273	0.8871	**0.8704**	2.8476
		0.8	4.2628	0.6825	0.6856	0.6882	0.6856	0.7077	1.3076
-1		0.9	4.4382	0.6914	**0.6775**	0.6950	**0.6775**	0.7126	1.2702
		0.95	4.6848	0.7430	**0.6687**	0.7442	**0.6687**	0.7293	1.4628
	300	0.99	6.0073	1.2024	**0.7356**	1.1878	**0.7356**	0.7800	1.5890
		0.8	3.8199	0.6465	0.6518	0.6524	0.6518	0.6646	1.0581
		0.9	3.9811	0.6343	0.6363	0.6391	0.6363	0.6503	1.0115
		0.95	4.1237	0.6388	**0.6266**	0.6422	**0.6266**	0.6454	0.9906
	400	0.99	4.6541	0.7936	**0.6503**	0.7888	**0.6503**	0.6489	0.9998
		0.8	3.0391	0.2472	0.2468	0.2470	0.2468	0.2468	**0.2452**
		0.9	3.1428	0.2560	0.2546	0.2556	0.2546	0.2547	**0.2457**
0		0.95	3.1589	0.2698	0.2641	0.2689	0.2641	0.2649	**0.2419**
	200	0.99	3.2618	0.3836	0.3012	0.3769	0.3012	0.2749	**0.2626**
		0.8	3.0586	0.2440	0.2435	0.2438	0.2435	0.2434	**0.2386**
	300	0.9	3.0492	0.2495	0.2476	0.2491	0.2476	0.2477	**0.2343**

ISSN NUMBER: 1116-249X

		0.95	3.1515	0.2726	0.2622	0.2716	0.2622	0.2656	**0.2247**
		0.99	3.5738	0.4928	0.2533	0.4863	0.2533	0.2937	**0.2489**
		0.8	2.7329	0.2204	0.2202	0.2203	0.2202	0.2202	**0.2193**
		0.9	2.8863	0.2153	0.2149	0.2152	0.2149	0.2149	**0.2109**
		0.95	3.0574	0.2194	0.2178	0.2190	0.2178	0.2178	**0.2067**
	400	0.99	3.1996	0.2859	0.2360	0.2831	0.2360	0.2427	**0.2270**
		0.8	6.0140	0.0932	0.0913	0.0930	0.0913	0.0914	**0.0839**
		0.9	5.8138	0.0957	0.0937	0.0955	0.0937	0.0939	**0.0856**
		0.95	5.7914	0.1020	0.0996	0.1018	0.0996	0.1001	**0.0894**
	200	0.99	6.1245	0.1497	0.1276	0.1482	0.1276	0.1443	**0.0816**
		0.8	5.7674	0.08910	0.08799	0.08896	0.08799	0.08797	**0.0826**
1		0.9	6.0084	0.0908	0.0896	0.0906	0.0896	0.0897	**0.0836**
		0.95	6.4139	0.1022	0.1003	0.1020	0.1003	0.1009	**0.0916**
	300	0.99	6.2366	0.1939	0.1435	0.1920	0.1435	0.1886	**0.0865**
		0.8	6.0934	0.07853	0.07768	0.07843	0.07768	0.07771	**0.0740**
		0.9	6.2382	0.0767	0.0759	0.0766	0.0759	0.0759	**0.0722**
		0.95	5.9348	0.0778	0.0768	0.0777	0.0768	0.0770	**0.0727**
	400	0.99	5.8016	0.1039	0.0975	0.1035	0.0975	0.1020	**0.0801**

The MSE that is smallest is bold

Table 2 presents the mean squared error (MSE) performance of various logistic regression estimators under 20% contamination and four predictors. Across all configurations, the RPMLE consistently exhibits the highest MSE, highlighting its vulnerability to both multicollinearity and data contamination. Classical shrinkage-based alternatives such as RPR, RPL, and RPK-L estimators offer modest improvements over RPMLE, yet their MSEs increase markedly under stronger multicollinearity ($\rho \geq 0.95$) and larger sample sizes. In contrast, the proposed estimators (prop1, prop1ME, and especially prop1HM) demonstrate superior performance across all simulation settings. Notably, prop1HM achieves the lowest or near-lowest MSE values in nearly all scenarios, particularly excelling under severe multicollinearity ($\rho = 0.99$) and small to moderate sample sizes (n = 200–400). Even as the intercept value changes from −1 to 1, the ranking of estimator performance remains consistent, with the proposed estimators retaining their edge. The

ISSN NUMBER: 1116-249X

stability and robustness of prop1HM even in high-dimensional, noisy conditions underscore its practical advantage over traditional methods. These results affirm the effectiveness of the proposed methodology in mitigating the adverse effects of outlier contamination and multicollinearity in logistic regression.

Table 2: MSE of different estimators when τ=20% and P = 4

intercept	n	rho	RPML	RPRR	RPL	RPK-L	prop1	prop1ME	prop1HM
-1	200	0.8	3.4667	**0.5469**	0.5811	0.5607	0.5811	0.5864	0.5745
		0.9	3.3149	**0.5388**	0.5596	0.5470	0.5596	0.5646	0.5587
		0.95	3.6570	0.5966	**0.5733**	0.5967	**0.5733**	0.5853	0.5833
		0.99	3.6014	1.3257	0.8113	1.2600	0.8113	**0.7195**	0.7845
	300	0.8	9.5276	9.8089	**9.7487**	9.8025	**9.7487**	9.7825	9.7924
		0.9	20.9870	28.8728	28.8300	28.8596	28.8300	**28.8120**	28.9870
		0.95	50.1574	63.9683	63.9141	63.9471	63.9141	**63.8415**	63.9365
		0.99	277.8851	295.8444	295.4740	295.7473	295.4740	**295.3904**	295.4786
	400	0.8	3.5941	**0.4090**	0.4266	0.4176	0.4266	0.4268	1.2698
		0.9	3.3816	**0.4105**	0.4412	0.4167	0.4412	0.4244	1.3496
		0.95	3.5974	**0.4998**	0.6276	0.5021	0.6276	0.5057	1.8273
		0.99	5.5655	4.0942	3.8063	4.0698	**3.8063**	3.8226	**1.9309**
0	200	0.8	2.9709	0.2399	0.2325	0.2387	0.2325	0.2363	**0.1551**
		0.9	9.2140	6.9268	6.5542	6.9244	6.5542	6.8652	**4.8557**
		0.95	31.3100	29.0709	28.2200	29.0635	28.2200	28.7831	**17.3894**
		0.99	181.3467	179.2644	158.2840	179.2187	158.2840	175.3685	178.3064
	300	0.8	13.9722	17.9229	17.9203	17.9219	17.9203	**17.9182**	18.1054
		0.9	34.7387	30.5778	30.5750	30.5767	30.5750	**30.5705**	31.3246
		0.95	59.4891	53.8385	53.8312	53.8366	53.8312	**53.8242**	56.9313
		0.99	245.9560	239.8896	239.8198	239.8780	239.8198	**239.8150**	240.1475
	400	0.8	2.8354	0.1448	0.1443	0.1445	0.1443	0.1443	**0.1248**
		0.9	3.1807	0.3408	0.3353	0.3404	0.3353	0.3386	**0.1881**
		0.95	3.9825	1.5793	1.5708	1.5782	1.5708	1.5569	**0.5647**

ISSN NUMBER: 1116-249X

		0.99	37.0352	37.5041	36.5771	37.4978	36.5771	37.2338	**20.0850**
		0.8	17.8673	9.3570	9.3140	9.3565	9.3140	9.3557	**8.7022**
		0.9	32.7416	20.8962	20.8671	20.8957	20.8671	20.8954	**19.2829**
		0.95	41.6098	33.6252	33.5766	33.6242	33.5766	33.6239	**28.8128**
	200	0.99	129.5441	117.5816	116.6704	117.5760	116.6704	117.5546	**89.5307**
		0.8	30.2117	18.7017	18.7016	18.7017	18.7016	18.7015	**18.6991**
		0.9	47.5881	26.7076	26.7075	26.7076	26.7075	26.7074	**26.7035**
1		0.95	59.9399	36.1348	36.1344	36.1348	36.1344	36.1344	**36.0830**
	300	0.99	133.3771	101.6136	101.6078	101.6122	101.6078	101.6109	**100.8533**
		0.8	6.9160	0.3041	0.2988	0.3040	0.2988	0.3034	**0.2491**
		0.9	18.6088	7.6656	7.6620	7.6654	7.6620	7.6628	**7.2214**
		0.95	29.9047	15.6965	15.6864	15.6964	15.6864	15.6949	**15.4356**
	400	0.99	55.2891	39.7585	39.7013	39.7577	39.7013	39.7554	**39.6931**

The MSE that is smallest is bold

Conclusion

The simulation results provide compelling evidence that the proposed estimators particularly prop1HM offer significant improvements over conventional methods in the presence of outlier contamination and multicollinearity in Poisson regression models. While RPMLE and classical shrinkage estimators (RPR, RPL, RPK-L) show performance degradation under these adverse conditions, the proposed methods demonstrate greater robustness, stability, and efficiency. These findings highlight the practical utility of the proposed estimators, especially prop1HM, as reliable alternatives for real-world applications where data contamination and predictor correlation are common challenges.

ISSN NUMBER: 1116-249X

REFERENCES

Abonazel, M. R., & Saber, O. (2020). A comparative study of robust estimators for Poisson regression model with outliers. *Journal of Statistics Applications & Probability, 9*(3), 279–286.

Abonazel, M. R., El-Sayed, S. M., & Saber, O. M. (2021). Performance of robust count regression estimators in the case of overdispersion, zero-inflation, and outliers. *Communications in Mathematical Biology and Neuroscience, 2021*(55).

Albert, S. O., Olayemi, O. S., Oguntade, E. S., & Onifade, O. C. (2025). A new class of Poisson biasing ridge type estimator: Simulation and theory approach. *Journal of the Royal Statistical Society – Nigeria Group, 2*(1), 165–185. ISSN 1116-249X.

Cantoni, E., & Ronchetti, E. (2001). Robust inference for generalized linear models. *Journal of the American Statistical Association, 96*(456), 1022–1030.

Chen, W., Qian, L., Shi, J., & Franklin, M. (2018). Comparing performance between log-binomial and robust Poisson regression models. *BMC Medical Research Methodology, 18*, 63.

Chen, W., Shi, J., Qian, L., & Azen, S. P. (2014). Comparison of robustness to outliers between robust Poisson models and log-binomial models. *BMC Medical Research Methodology, 14*, 82.

Dawoud, I., & Abonazel, M. R. (2022). Developing robust ridge estimators for Poisson regression model. *Concurrency and Computation: Practice and Experience, 34*, e6779.

Dawoud, I., Awwad, F. A., Tag Eldin, E., & Abonazel, M. R. (2022). New robust estimators for handling multicollinearity and outliers in the Poisson model: Methods, simulation and applications. *Axioms, 11*(11), 612. https://doi.org/10.3390/axioms11110612

Hall, D. B., & Shen, J. (2010). Robust estimation for zero-inflated Poisson regression. *Scandinavian Journal of Statistics, 37*(2), 237–252.

Hosseinian, S. (2009). *Robust inference for generalized linear models: Binary and Poisson regression* (Doctoral dissertation). EPFL, Lausanne, Switzerland.

Hosseinian, S., & Morgenthaler, S. (2011). Weighted maximum likelihood estimates in Poisson regression. In *Proceedings of the International Conference on Robust Statistics*.

Kibria, B. M. G. (2003). Performance of some new ridge regression estimators. *Communications in Statistics - Simulation and Computation, 32*(2), 419–435.

Lukman, A. F., Adewuyi, E., Månsson, K., & Kibria, B. G. (2021). A new estimator for the multicollinear Poisson regression model: Simulation and application. *Scientific Reports, 11*, 3732.

Lukman, A. F., & Ayinde, K. (2017). Review and classifications of the ridge parameter estimation techniques. *Hacettepe Journal of Mathematics and Statistics, 46*(5), 953–967.

Marazzi, A. (2021). Improving the efficiency of robust estimators for the generalized linear model. *Stats, 4*(1), 88–107.

Månsson, K., & Shukur, G. (2011). A Poisson ridge regression estimator. *Economic Modelling, 28*(4), 1475–1481.

Oladapo, O. J., Alabi, O. O., & Ayinde, K. (2024). Another new two parameter estimator in dealing with multicollinearity in the logistic regression model. *International Journal of Mathematical Sciences and Optimization, 10*(2), 22–35. https://doi.org/10.5281/zenodo.10937145

ISSN NUMBER: 1116-249X

Oladapo, O. J., Owolabi, A. T., Idowu, J. I., & Ayinde, K. (2022). A new modified Liu ridge-type estimator for the linear regression model: Simulation and application. *Clinical Biostatistics and Biometrics, 8*(2), 1–14. https://doi.org/10.23937/2469-5831/1510048

Oladapo, J., Idowu, J., Owolabi, A., Ayinde, K., & Adejumo, J. (2024). A new ridge type estimator in the logistic regression model with correlated regressors. *WSEAS Transactions on Business and Economics, 18*, 612–635.

Oladapo, J., Idowu, J., Owolabi, A., Ayinde, K., Adejumo, J., Oshuoporu, O., & Alao, A. (2023). Mitigating multicollinearity in linear regression model with two parameter Kibria-Lukman estimators. *WSEAS Transactions on Systems and Control, 18*, 612–635.

Owolabi, A. T., Ayinde, K., Idowu, J. I., Oladapo, O. J., & Lukman, A. F. (2022). A new two-parameter estimator in the linear regression model with correlated regressors. *Journal of Statistics Applications & Probability, 11*(2), 499–512.

Trenkler, G., & Toutenburg, H. (1990). Mean squared error matrix comparisons between biased estimators—An overview of recent results. *Statistical Papers, 31*(1), 165–179.

ISSN NUMBER: 1116-249X