Received: 12 February 2023 | Revised: 14 June 2023 | Accepted: 20 June 2023

DOI: 10.1111/stan.12313

ORIGINAL ARTICLE

WILEY

Poisson average maximum likelihood-centered penalized estimator: A new estimator to better address multicollinearity in Poisson regression

Sheng Li[1] | Wei Wang[1] | Menghan Yao[1] | Junyu Wang[1] | Qianqian Du[1] | Xuelin Li[1] | Xinyue Tian[1] | Jing Zeng[2] | Ying Deng[2] | Tao Zhang[1,3] | Fei Yin[1] | Yue Ma[1,3]

[1]Department of Epidemiology and Biostatistics, West China School of Public Health and West China Fourth Hospital, Sichuan University, Chengdu, China

[2]Department of Chronic Disease Surveillance, Sichuan Center for Disease Control and Prevention, Chengdu, China

[3]Institute of Systems Epidemiology, West China School of Public Health and West China Fourth Hospital, Sichuan University, Chengdu, China

Correspondence
Yue Ma, Department of Epidemiology and Biostatistics, West China School of Public Health and West China Fourth Hospital, Sichuan University, Chengdu, China.
Email: gordonrozen@scu.edu.cn

Funding information
National Natural Science Foundation of China, Grant/Award Numbers: 81803332, 81872713; Sichuan Province Science and Technology Support Program, Grant/Award Number: 2021YFS0181

Abstract

The Poisson ridge estimator (PRE) is a commonly used parameter estimation method to address multicollinearity in Poisson regression (PR). However, PRE shrinks the parameters toward zero, contradicting the real association. In such cases, PRE tends to become an insufficient solution for multicollinearity. In this work, we proposed a new estimator called the Poisson average maximum likelihood-centered penalized estimator (PAMLPE), which shrinks the parameters toward the weighted average of the maximum likelihood estimators. We conducted a simulation study and case study to compare PAMLPE with existing estimators in terms of mean squared error (MSE) and predictive mean squared error (PMSE). These results suggest that PAMLPE can obtain smaller MSE and PMSE (i.e., more accurate estimates) than the Poisson ridge estimator, Poisson Liu estimator, and Poisson K-L estimator when the true βs have the same sign and small variation. Therefore, we recommend using PAMLPE to address multicollinearity in PR when the signs of the true βs are known to be identical in advance.

KEYWORDS

multicollinearity, Poisson penalized estimator, Poisson regression, shrinkage center

1 | INTRODUCTION

Poisson regression (PR) is widely employed to analyze the relationship between a count response variable and regressors in epidemiology, sociology, economics, and engineering. The Poisson maximum likelihood estimator (PMLE) is the commonly used parameter estimation method and is asymptotically unbiased. However, the Poisson maximum likelihood estimate becomes unstable with a high variance when multicollinearity exists, which leads to low power, insignificant t-ratios and wide confidence intervals. When multicollinearity is strong, the sign of the maximum likelihood estimate can even be inconsistent with the sign of the true parameter value.

To date, several estimation methods have been proposed to address multicollinearity in PR. The Poisson ridge estimator (PRE) (Månsson & Shukur, 2011), the most commonly used method among them, penalizes the cost function by shrinking the coefficients toward zero. In the Bayesian view, PRE chooses $\beta \sim N(0, 1/k)$ as the penalization prior (Hsiang, 1975) where k is the penalization coefficient and determines the shrinkage strength. The zero-shrinkage center may be appropriate when there is no prior information about the parameter. However, since the variable included in a model is often of interest, it is a more common truth that the true value of the parameter is not equal to zero. In such cases, the zero-shrinkage center contradicts the prior information about the real association, and PRE tends to choose a small, near-zero penalization coefficient, resulting in an insufficient limitation for multicollinearity.

Swindel (1976) proved that the nearer the shrinkage center is to the true value of the parameter, the closer the estimate is to the true value of the parameter (i.e., the nonzero shrinkage center leads to a more accurate estimate than the zero-shrinkage center in some cases). However, he did not give the method to determine the nonzero shrinkage centers in real datasets. Although the near isotonic regularization estimator (Matsuda & Miyatake, 2021) and the fused LASSO estimator (Choi, Song, Hwang, & Lee, 2018; Tibshirani & Taylor, 2011), which penalize the cost functions by shrinking the differences between parameters, give a remedy for the nonzero shrinkage centers and better solve multicollinearity, their requirement of a known natural order of parameters in advance limits their application. The Poisson Liu estimator (PLE) (Mansson, Kibria, Sjolander, & Shukur, 2012) fixes the penalization coefficient $k = 1$ and then chooses the multiple of the Poisson maximum likelihood estimator as the shrinkage center. Some studies (Amin, Akram, & Kibria, 2021; Kibria & Lukman, 2020; Qasim, Kibria, Månsson, & Sjölander, 2020) have shown that PLE tends to choose the multiple equal to zero. In this case, PLE equals PRE with $k = 1$ and tends to be an insufficient solution for multicollinearity because the fixed penalization coefficient could be too small. The Poisson K-L estimator (PKLE) (Akay & Ertan, 2022; Aladeitan, Adebimpe, Lukman, Oludoun, & Abiodun, 2021; Lukman, Adewuyi, Månsson, & Kibria, 2021) uses the opposite number of the Poisson maximum likelihood estimator as the shrinkage center, which could be too far from the true value of the parameter. In such cases, PKLE tends to choose a small, near-zero penalization coefficient and therefore leads to difficulty restricting multicollinearity sufficiently. Additionally, some two-parameter ridge estimators (Akay & Ertan, 2022; Alkhateeb & Algamal, 2020; Amin et al., 2021; Asar & Genç, 2018; Kandemir Çetinkaya & Kaçıranlar, 2019; Lukman, Aladeitan, Ayinde, & Abonazel, 2022) have been proposed and can address multicollinearity better, but it is difficult to simultaneously select the optimal value of two parameters, which limits their application.

Other penalized estimators, such as the LASSO estimator (Hossain & Ahmed, 2012), adaptive LASSO estimator (Algamal & Lee, 2015; Hossain & Ahmed, 2012; Ivanoff, Picard, & Rivoirard, 2016), elastic net estimator (Noori Asl, Bevrani, & Arabi Belaghi, 2022), and bridge estimator (Chowdhury, Chatterjee, Mallick, Banerjee, & Garai, 2019), have been proposed, but they also use zero as the shrinkage center and mainly focus on variable selection rather than limiting multicollinearity. The Jackknifed version and modified Jackknifed version of PRE, PLE, and PKLE (Oranye & Ugwuowo, 2022; Rasheed, Sadik, & Algamal, 2022; Turkan & Ozel, 2016) were also proposed to obtain an almost unbiased estimator by applying the Jackknifed procedure of bias reduction. However, the Jackknifed procedure of bias reduction cannot improve the selection of the shrinkage center.

Therefore, in common cases in which the true parameters of regressors of interest are not equal to zero, an appropriate penalization method is still scarce in PR.

Recently, Wang et al. (2022) proposed the average ordinary least squares-centered penalized estimator (AOPE) based on the idea of nonzero shrinkage center, which uses the weighted average of the ordinary least squares estimators as the shrinkage center. In their study, AOPE obtained a more accurate estimate than the ridge estimator (RE) when the signs of the true parameters were identical. Additionally, AOPE only needs to determine one hyperparameter, the penalization coefficient k, which is easy to implement. This suggests that AOPE could be a potential method to better limit multicollinearity for cases in which the signs of the true parameters are identical. Although there are some similarities between ordinary linear regression and Poisson regression, the parameter estimation of Poisson regression involves an iterative process, and its estimator is related to a weight matrix. Therefore, a similar estimation method needs to be proposed and verified in PR.

In this study, we extended AOPE to PR by using the weighted average of the maximum likelihood estimators as the shrinkage center, and proposed the Poisson average maximum likelihood-centered penalized estimator (PAMLPE). The performance of this new estimator was compared with the existing estimators in terms of mean squared error (MSE) and predictive mean squared error (PMSE) in a simulation study and case study.

The rest of the paper is organized as follows: the statistical methodology is described in Section 2. The simulation study and case study are represented in Sections 3 and 4, respectively. Finally, a conclusion is given.

2 | METHODOLOGY

To better present the PAMLPE, the PMLE will be described first, and then the PRE, PLE, PKLE, and PAMLPE are derived from the perspective of the shrinkage center.

2.1 | Poisson maximum likelihood estimator

According to iterative weighed least squares (IWLS), the PMLE can be written as

$$\hat{\boldsymbol{\beta}}^{\text{PMLE}} = \left(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}\right)^{-1}\mathbf{X}'\widehat{\mathbf{W}}\widehat{\boldsymbol{Z}}, \tag{1}$$

where $\mathbf{X}$ is an $n \times (p+1)$ observed matrix composed of the regressors, $\widehat{\mathbf{W}} = \text{diag}[\hat{\boldsymbol{\mu}}]$ where $\hat{\boldsymbol{\mu}} = e^{\mathbf{X}\hat{\boldsymbol{\beta}}}$, and $\widehat{\boldsymbol{Z}}$ is a vector in which $\widehat{Z}_i = \ln\left(\hat{\mu}_i\right) + \frac{y_i - \hat{\mu}_i}{\hat{\mu}_i}$. The covariance matrix and MSE are obtained as follows:

$$\text{Cov}\left(\hat{\boldsymbol{\beta}}^{\text{PMLE}}\right) = \left(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}\right)^{-1}, \tag{2}$$

$$\text{MSE}\left(\hat{\boldsymbol{\beta}}^{\text{PMLE}}\right) = tr\left[\left(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}\right)^{-1}\right] = \sum_{j=1}^{p}\frac{1}{\lambda_j}, \tag{3}$$

where λ_j is the ith eigenvalue of $\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}$. When a strong correlation between regressors exists, the ill-conditioned $\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}$ leads to high variance and instability of the maximum likelihood estimate, and a large MSE.

2.2 | Poisson penalized estimator

PRE, PLE, PKLE, and PAMLPE are penalized estimators with different shrinkage centers. The general formula of cost functions can be written as

$$\ln l(\boldsymbol{\beta}) = -\sum_{i=1}^{n}\left[y_i \times \ln \mu_i - \mu_i - \ln y_i!\right] + k\|\boldsymbol{\beta} - \mathbf{c}\|_2^2, k \geq 0, \tag{4}$$

where k is the penalty coefficient and $\boldsymbol{c}$ is the shrinkage center. Then, the general Poisson penalized estimator is

$$\begin{aligned}\hat{\boldsymbol{\beta}}^{\text{GPPE}} &= \left(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X} + k\mathbf{I}\right)^{-1}\left(\mathbf{X}'\widehat{\mathbf{W}}\widehat{\boldsymbol{Z}} + k\boldsymbol{c}\right) \\ &= \left(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X} + k\mathbf{I}\right)^{-1}\left(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}\hat{\boldsymbol{\beta}}^{\text{PMLE}} + k\boldsymbol{c}\right) = \mathbf{H}\hat{\boldsymbol{\beta}}^{\text{PMLE}} + k\mathbf{M}\boldsymbol{c},\end{aligned} \tag{5}$$

where $\mathbf{M} = \left(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X} + \mathbf{kI}\right)^{\mathbf{-1}}$, $\mathbf{H} = \left(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X} + k\mathbf{I}\right)^{-1}\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X} = \mathbf{I} - k\mathbf{M}$, and $\mathbf{I}$ is a $(p+1) \times (p+1)$ identity matrix.

Let $\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X} = \mathbf{P}'\boldsymbol{\Lambda}\mathbf{P}$, where $\boldsymbol{\Lambda}$ is a diagonal matrix composed of the eigenvalues of $\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}$ and $\mathbf{P}'$ is an orthogonal matrix composed of the eigenvectors of $\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}$. Thus,

$$\mathbf{P}'\mathbf{P} = \mathbf{P}\mathbf{P}' = \mathbf{I}, \boldsymbol{\Lambda} = \text{diag}\left(\lambda_1, \lambda_2, \ldots, \lambda_p\right),$$

$$\xi_i(\mathbf{M}) = 1/\left(\lambda_i + k\right), \xi_i(\mathbf{H}) = \lambda_i/\left(\lambda_i + k\right),$$

where $\xi_i(\mathbf{M})$ and $\xi_i(\mathbf{H})$ are the ith eigenvalues of $\mathbf{M}$ and $\mathbf{H}$, respectively.

Let $\boldsymbol{S} = \boldsymbol{\beta} - \boldsymbol{c}$, $\boldsymbol{\alpha} = \mathbf{P}\boldsymbol{S}$, $\boldsymbol{\Lambda}_{\mathbf{HH}}$ and $\boldsymbol{\Lambda}_{\mathbf{MM}}$ be diagonal matrices composed of the eigenvalues of $\mathbf{H}'\mathbf{H}$ and $\mathbf{M}'\mathbf{M}$, respectively. Obtain the MSE:

$$\mathrm{MSE}\left(\hat{\boldsymbol{\beta}}^{\mathrm{GPPE}}\right)=\sum_{i=1}^{p}\frac{\lambda_i}{(\lambda_i+k)^2}+k^2\sum_{i=1}^{p}\frac{\alpha_i^2}{(\lambda_i+k)^2}. \tag{6}$$

The detailed calculation is shown in the Data S1.

Equation (6) suggests that the nearer the shrinkage center is to the true value of the parameter, the smaller the MSE is, that is, the more accurate the estimate is. Therefore, a more accurate estimate would be obtained by introducing a more suitable shrinkage center, and the accuracy of the estimate from different estimation methods can be roughly ranked by comparing the distance between the true parameters and the shrinkage center.

2.2.1 | Poisson ridge estimator

PRE uses zero as the shrinkage center, that is, $\boldsymbol{c}=\mathbf{0}$, and can be written as (Månsson & Shukur, 2011):

$$\hat{\boldsymbol{\beta}}^{\mathrm{PRE}}=\left(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}+k\mathbf{I}\right)^{-1}\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}\hat{\boldsymbol{\beta}}^{\mathrm{PMLE}}=\mathbf{H}\hat{\boldsymbol{\beta}}^{\mathrm{PMLE}}. \tag{7}$$

Let $\boldsymbol{b}=\mathbf{P}\boldsymbol{\beta}$, obtain the covariance matrix and MSE of the PRE:

$$\mathrm{Cov}\left(\hat{\boldsymbol{\beta}}^{\mathrm{PRE}}\right)=\left(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}+k\mathbf{I}\right)^{-1}\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}\left(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}+k\mathbf{I}\right)^{-1}, \tag{8}$$

$$\mathrm{MSE}\left(\hat{\boldsymbol{\beta}}^{\mathrm{PRE}}\right)=\sum_{i=1}^{p}\frac{\lambda_i}{(\lambda_i+k)^2}+k^2\sum_{i=1}^{p}\frac{b_i^2}{(\lambda_i+k)^2}. \tag{9}$$

When $k=0$, $\mathrm{MSE}\left(\hat{\boldsymbol{\beta}}^{\mathrm{PRE}}\right)=\mathrm{MSE}\left(\hat{\boldsymbol{\beta}}^{\mathrm{PMLE}}\right)$. Then,

$$\frac{\partial MSE\left(\hat{\boldsymbol{\beta}}^{\mathrm{PRE}}\right)}{\partial k}=\sum_{i=1}^{p}\frac{-1+b_i^2k}{\frac{1}{2\lambda_i}(\lambda_i+k)^3},$$

when $0\leq k\leq 1/b_{\max}^2$, which is a conservative range, $\partial MSE\left(\hat{\boldsymbol{\beta}}^{\mathrm{PRE}}\right)/\partial k\leq 0$ and $\mathrm{MSE}\left(\hat{\boldsymbol{\beta}}^{\mathrm{PRE}}\right)\leq \mathrm{MSE}\left(\hat{\boldsymbol{\beta}}^{\mathrm{PMLE}}\right)$. Therefore, $k=1/b_{\max}^2$ could be regarded as the optimal selection.

2.2.2 | Poisson Liu estimator

PLE was introduced by Mansson et al. (2012), which fixes the penalization coefficient $k=1$ and chooses the multiple of the Poisson maximum likelihood estimator as the shrinkage center. The PLE can be written as

$$\hat{\boldsymbol{\beta}}^{\mathrm{PLE}}=\left(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}+\mathbf{I}\right)^{-1}\left(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}+d\mathbf{I}\right)\hat{\boldsymbol{\beta}}^{\mathrm{PMLE}}=\mathbf{A}\hat{\boldsymbol{\beta}}^{\mathrm{PMLE}},0\leq d\leq 1, \tag{10}$$

where $\mathbf{A}=\left(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}+\mathbf{I}\right)^{-1}\left(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}+d\mathbf{I}\right)=\mathbf{I}-(1-d)\mathbf{M}$.

$$\xi_i(\mathbf{A}) = (\lambda_i + d) / (\lambda_i + 1).$$

Let $\mathbf{\Lambda_{AA}}$ be diagonal matrices composed of the eigenvalues of $\mathbf{A'A}$. Then,

$$\mathrm{Cov}\left(\hat{\boldsymbol{\beta}}^{\mathrm{PLE}}\right) = \mathbf{A}\left(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}\right)^{-1}\mathbf{A}'. \tag{11}$$

$$\mathrm{MSE}\left(\hat{\boldsymbol{\beta}}^{\mathrm{PLE}}\right) = \sum_{i=1}^{p}\frac{(\lambda_i + d)^2}{\lambda_i(\lambda_i + 1)^2} + \sum_{i=1}^{p}\frac{b_i^2(1-d)^2}{(\lambda_i + 1)^2}. \tag{12}$$

When $d = 1$, $\mathrm{MSE}\left(\hat{\boldsymbol{\beta}}^{\mathrm{PLE}}\right) = \mathrm{MSE}\left(\hat{\boldsymbol{\beta}}^{\mathrm{PMLE}}\right)$. Then,

$$\frac{\partial MSE\left(\hat{\boldsymbol{\beta}}^{\mathrm{PLE}}\right)}{\partial d} = \sum_{i=1}^{p}\frac{2d\left(1 + \lambda_i b_i^2\right) - 2\lambda_i\left(b_i^2 - 1\right)}{\lambda_i(\lambda_i + 1)^2},$$

when $\max\left(0, \max\left(\lambda_i\left(b_i^2 - 1\right) / \left(1 + \lambda_i b_i^2\right)\right)\right) < d < 1$, which is a conservative range, $\partial MSE\left(\hat{\boldsymbol{\beta}}^{\mathrm{PLE}}\right) / \partial d \geq 0$ and $\mathrm{MSE}\left(\hat{\boldsymbol{\beta}}^{\mathrm{PLE}}\right) \leq \mathrm{MSE}\left(\hat{\boldsymbol{\beta}}^{\mathrm{PMLE}}\right)$. Therefore, $d = \max\left(0, \max\left(\lambda_i\left(b_i^2 - 1\right) / \left(1 + \lambda_i b_i^2\right)\right)\right)$ could be regarded as the optimal selection.

2.2.3 | Poisson K-L estimator

PKLE was introduced by Lukman et al. (2021) and uses the opposite number of the Poisson maximum likelihood estimator as the shrinkage center. The estimator can be written as

$$\hat{\boldsymbol{\beta}}^{\mathrm{PKLE}} = \left(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X} + k\mathbf{I}\right)^{-1}\left(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X} - k\mathbf{I}\right)\hat{\boldsymbol{\beta}}^{\mathrm{PMLE}} = \mathbf{D}\hat{\boldsymbol{\beta}}^{\mathrm{PMLE}}, \tag{13}$$

where $\mathbf{D} = \left(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X} + k\mathbf{I}\right)^{-1}\left(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X} - k\mathbf{I}\right) = \mathbf{I} - 2k\mathbf{M}$,

$$\xi_i(\mathbf{D}) = (\lambda_i - k) / (\lambda_i + k).$$

Let $\mathbf{\Lambda_{DD}}$ be diagonal matrices composed of the eigenvalues of $\mathbf{D'D}$. Then,

$$\mathrm{Cov}\left(\hat{\boldsymbol{\beta}}^{\mathrm{PKLE}}\right) = \mathbf{D}\left(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}\right)^{-1}\mathbf{D}', \tag{14}$$

$$\mathrm{MSE}\left(\hat{\boldsymbol{\beta}}^{\mathrm{PKLE}}\right) = \sum_{i=1}^{p}\frac{(\lambda_i - k)^2}{\lambda_i(\lambda_i + k)^2} + 4k^2\sum_{i=1}^{p}\frac{b_i^2}{(\lambda_i + k)^2}. \tag{15}$$

When $k = 0$, $\mathrm{MSE}\left(\hat{\boldsymbol{\beta}}^{\mathrm{PKLE}}\right) = \mathrm{MSE}\left(\hat{\boldsymbol{\beta}}^{\mathrm{PMLE}}\right)$. Then,

$$\frac{\partial MSE\left(\hat{\boldsymbol{\beta}}^{\mathrm{PKLE}}\right)}{\partial k} = \sum_{i=1}^{p}\frac{4k\left(1 + 2\lambda_i b_i^2\right) - 4\lambda_i}{(\lambda_i + k)^3},$$

when $0 \leq k \leq \min\left(\lambda_i / \left(1 + 2\lambda_i b_i^2\right)\right)$, which is a conservative range, $\partial MSE\left(\widehat{\boldsymbol{\beta}}^{\text{PKLE}}\right)/\partial k \leq 0$ and $\text{MSE}\left(\widehat{\boldsymbol{\beta}}^{\text{PKLE}}\right) \leq \text{MSE}\left(\widehat{\boldsymbol{\beta}}^{\text{PMLE}}\right)$. Therefore, $k = \min\left(\lambda_i / \left(1 + 2\lambda_i b_i^2\right)\right)$ could be regarded as the optimal selection.

2.2.4 | Poisson average maximum likelihood-centered penalized estimator

PAMLPE chooses the weighted average of the maximum likelihood estimators as the shrinkage center, which is obtained by maximum likelihood estimation with restrictions on all coefficients being equal, that is,

$$\begin{aligned}\widehat{\boldsymbol{c}} &= \arg\max\left\{\sum_{i=1}^{n}\left[y_i \times \boldsymbol{x}_i'\boldsymbol{c} - e^{\boldsymbol{x}_i'\boldsymbol{c}} - \ln y_i!\right]\right\} \\ &= \boldsymbol{\vartheta}\left(\boldsymbol{\vartheta}'\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}\boldsymbol{\vartheta}\right)^{-1}\boldsymbol{\vartheta}'\mathbf{X}'\widehat{\mathbf{W}}\widehat{\boldsymbol{Z}} = \boldsymbol{\vartheta}\left(\boldsymbol{\vartheta}'\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}\boldsymbol{\vartheta}\right)^{-1}\boldsymbol{\vartheta}'\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}\widehat{\boldsymbol{\beta}}^{\text{PMLE}},\end{aligned} \tag{16}$$

where $\boldsymbol{\vartheta}$ is a $p \times 1$ vector in which all the elements are one. The shrinkage center of PAMLPE, the weighted average of the maximum likelihood estimators, is expected to be closer to the true parameter than $\mathbf{0}$ and the opposite number of the maximum likelihood estimator.

Then, obtain the PAMLPE

$$\widehat{\boldsymbol{\beta}}^{\text{PAMLPE}} = \left(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X} + k\mathbf{I}\right)^{-1}\left[\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X} + k\boldsymbol{\vartheta}\left(\boldsymbol{\vartheta}'\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}\boldsymbol{\vartheta}\right)^{-1}\boldsymbol{\vartheta}'\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}\right]\widehat{\boldsymbol{\beta}}^{\text{PMLE}} = \mathbf{F}\widehat{\boldsymbol{\beta}}^{\text{PMLE}}, \tag{17}$$

where $\mathbf{F} = \left(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X} + k\mathbf{I}\right)^{-1}\left[\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X} + k\boldsymbol{\vartheta}\left(\boldsymbol{\vartheta}'\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}\boldsymbol{\vartheta}\right)^{-1}\boldsymbol{\vartheta}'\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}\right]$.

The covariance of $\widehat{\boldsymbol{\beta}}^{\text{PAMLPE}}$ is

$$\text{Cov}\left(\widehat{\boldsymbol{\beta}}^{\text{PAMLPE}}\right) = \mathbf{F}\left(\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}\right)^{-1}\mathbf{F}'. \tag{18}$$

Let $\widehat{\boldsymbol{S}} = \boldsymbol{\beta} - \widehat{\boldsymbol{c}}$, $\widehat{\boldsymbol{\alpha}} = \mathbf{P}\widehat{\boldsymbol{S}}$, $\widehat{\boldsymbol{b}} = \mathbf{P}\widehat{\boldsymbol{\beta}}^{\textbf{PMLE}}$. Then, the MSE of $\widehat{\boldsymbol{\beta}}^{\text{PAMLPE}}$ can be calculated as follows:

$$\text{MSE}\left(\widehat{\boldsymbol{\beta}}^{\text{PAMLPE}}\right) = \sum_{i=1}^{p}\frac{\lambda_i}{(\lambda_i + k)^2} + E\left[\sum_{i=1}^{p}\frac{k^2\widehat{\alpha}_i^2 - 2k\lambda_i\left(\widehat{b}_i - b_i\right)\widehat{\alpha}_i}{(\lambda_i + k)^2}\right].$$

Let

$$\mathbf{h} = \mathbf{P}\boldsymbol{\vartheta}\left(\boldsymbol{\vartheta}'\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}\boldsymbol{\vartheta}\right)^{-1}\boldsymbol{\vartheta}'\mathbf{X}'\widehat{\mathbf{W}}\mathbf{X}\mathbf{P}',$$

$$\tau_i = b_i^2 + \sum_{j=1}^{p}\left(\frac{h_{ij}^2}{\lambda_j} + h_{ij}^2 b_j^2\right) - 2b_i\boldsymbol{h}_i'\boldsymbol{b},$$

where $\boldsymbol{h}_i$ is the ith row vector of $\mathbf{h}$. Then, the MSE of $\widehat{\boldsymbol{\beta}}^{\text{PAMLPE}}$ can be calculated as follows:

$$\mathrm{MSE}\left(\hat{\boldsymbol{\beta}}^{\mathrm{PAMLPE}}\right)=\sum_{i=1}^{p}\frac{k^2\tau_i+2kh_{ii}+\lambda_i}{(\lambda_i+k)^2}. \tag{19}$$

The detailed calculation is shown in the Data S1. When $k=0$, $\mathrm{MSE}\left(\hat{\boldsymbol{\beta}}^{\mathrm{PAMLPE}}\right)=\mathrm{MSE}\left(\hat{\boldsymbol{\beta}}^{\mathrm{PMLE}}\right)$. Then,

$$\frac{\partial MSE\left(\hat{\boldsymbol{\beta}}^{\mathrm{PAMLPE}}\right)}{\partial k}=\sum_{i=1}^{p}\frac{(\lambda_i\tau_i-h_{ii})\,k+\lambda_ih_{ii}-\lambda_i}{\frac{1}{2}(\lambda_i+k)^3}.$$

Let $(\lambda_i\tau_i-h_{ii})\,k+\lambda_ih_{ii}-\lambda_i\leq\mathbf{0}$, which reduces to

$$\begin{cases}k\leq\frac{\lambda_i-\lambda_ih_{ii}}{\lambda_i\tau_i-h_{ii}},\ \lambda_i\tau_i-h_{ii}>0\\ k\geq\frac{\lambda_i-\lambda_ih_{ii}}{\lambda_i\tau_i-h_{ii}},\ \lambda_i\tau_i-h_{ii}<0\end{cases}.$$

Set

$$f_i=\begin{cases}0, & \lambda_i\tau_i-h_{ii}\leq 0 \text{ and } h_{ii}>1\\ \max\left(0\frac{\lambda_i-\lambda_ih_{ii}}{\lambda_i\tau_i-h_{ii}}\right), & \text{others}\end{cases},$$

when $0\leq k\leq f_{\min}$, which is also a conservative range, $\mathrm{MSE}\left(\hat{\boldsymbol{\beta}}^{\mathrm{PAMLPE}}\right)\leq\mathrm{MSE}\left(\hat{\boldsymbol{\beta}}^{\mathrm{PMLE}}\right)$. Therefore, $k=f_{\min}$ could be the optimal selection for the penalization coefficient in PAMLPE. Additionally, $k=\mathrm{median}(f)$ or $k=\left(\prod_{i=1}^{p}f_i\right)^{\frac{1}{p}}$ can be regarded as an appropriate selection.

Because the equations for calculating k or d involve the actual parameter values which are unknown during the analysis, the maximum likelihood estimate $\hat{\boldsymbol{\beta}}^{\mathrm{PMLE}}$ is used in place of the unknown parameter $\boldsymbol{\beta}$. Additionally, the iterative estimation (Hoerl & Kennard, 1970), cross-validation (Noori Asl et al., 2022), generalized cross-validation (Algamal, 2020) and ridge trace (Hoerl & Kennard, 1970) methods can be used to select an appropriate k or d for the four Poisson penalized estimators.

Equation (6) suggests that the nearer the shrinkage center is to the true value of the parameter, the more accurate the estimate is. The shrinkage centers of PRE, PKLE and PAMLPE are 0, the opposite number of the maximum likelihood estimator and the weighted average value of the maximum likelihood estimators, respectively. The weighted average of the maximum likelihood estimators is expected to be closer to the true parameter than 0 and the opposite number of the maximum likelihood estimator. Therefore, PAMLPE is expected to obtain a more accurate estimate than PRE and PKLE.

It is also worth noting that before estimating the parameters, the explanatory variables should be standardized to avoid differences in units.

3 | SIMULATION STUDY

In this section, a simulation study was conducted to compare the performance of the PMLE, PRE, PLE, PKLE, and PAMLPE. The study was implemented using R software (version 4.1.2).

3.1 | Simulation datasets

The number of explanatory variables p, the correlation between the explanatory variables r, the sample size n and the intercept a may affect the accuracy of estimators (Lukman et al., 2021; Mansson et al., 2012; Månsson & Shukur, 2011). Since the demand of sample size is associated with the number of explanatory variables in practice, the ratio, g, of the sample size to the number of explanatory variables instead of the sample size was set as a given parameter. Additionally, the performance of PAMLPE will be affected by both the average value of the elements and the variability among the elements in $\boldsymbol{\beta}$ (Wang et al., 2022). In addition, $\boldsymbol{\beta}$ is set so that $\sum_{i=1}^{p} \beta_i^2 = 1$, which is a common restriction in simulation studies (Akay & Ertan, 2022; Aladeitan et al., 2021; Amin et al., 2021; Mansson et al., 2012; Månsson & Shukur, 2011; Qasim et al., 2020). Therefore, 756 simulation scenarios were generated with different p, r, g, a, and $\boldsymbol{\beta}$ values (in Tables 1 and 2). For brevity, each simulation scenario is mentioned in short by r, g, p, a, and $\boldsymbol{\beta}$. For example, $p4$-$beta2$-$g10$-$r0.9$-$a0$ denotes the simulation scenario where the value of p, the code for $\boldsymbol{\beta}$, and the values of g, r, and a are 4, 2, 10, 0.9, and 0, respectively.

TABLE 1 The available values for p, g, r, and a in the simulation datasets.

Simulation parameter	The available values
p	4, 8
g	10, 20, 40
r	0, 0.3, 0.6, 0.9, 0.95, 0.99
a	−1, 0, 1

TABLE 2 The available values for $\boldsymbol{\beta}$ in the simulation datasets.

p	β code	The available values for β_i^2
4	1	(0.25, 0.25, 0.25, 0.25)
	2	(0.1, 0.2, 0.3, 0.4)
	3	(0.01, 0.01, 0.49, 0.49)
	4	(0, 0, 0.5, 0.5)
	5	(−0.01, −0.01, 0.49, 0.49)*
	6	(−0.1, −0.2, 0.3, 0.4)
	7	(−0.25, −0.25, 0.25, 0.25)
8	1	(0.25, 0.25, 0.25, 0.25, 0.25, 0.25, 0.25, 0.25)/2**
	2	(0.1, 0.2, 0.3, 0.4, 0.1, 0.2, 0.3, 0.4)/2
	3	(0.01, 0.01, 0.49, 0.49, 0.01, 0.01, 0.49, 0.49)/2
	4	(0, 0, 0.5, 0.5, 0, 0, 0.5, 0.5)/2
	5	(−0.01, −0.01, 0.49, 0.49, −0.01, −0.01, 0.49, 0.49)/2
	6	(−0.1, −0.2, 0.3, 0.4, −0.1, −0.2, 0.3, 0.4)/2
	7	(−0.25, −0.25, 0.25, 0.25, −0.25, −0.25, 0.25, 0.25)/2

Given p, g, r, a, and $\boldsymbol{\beta}$, the 1,000 datasets of each simulation scenario were generated according to the following process:

Step 1. Set $x_{i0}^{*} \sim N(0,1)$ and $e_{ij}^{*} \sim N(0,1)$, where $i = 1, 2, 3, \cdots, n$ and $j = 1, 2, 3, \cdots, p$.

Step 2. Obtain $x_{ij} = r x_{i0}^{*} + \sqrt{(1-r^2)} e_{ij}^{*}$.

Step 3. Obtain the expected response variable μ_i without the observed error:

$$\mu_i = e^{a+\beta_1 x_{i1}+\beta_2 x_{i2}+\cdots+\beta_p x_{ip}}.$$

Step 4. Obtain the response variable y_i with the observed error:

$$y_i \sim Po(\mu_i).$$

Step 5. Repeat step 41,000 times to generate 1,000 simulation datasets.

The R code for generating the simulation dataset is shown in the Data S2.

3.2 | Method for estimating $\boldsymbol{\beta}$

PMLE, PRE, PLE, PKLE, and PAMLPE were employed to estimate $\boldsymbol{\beta}$ in Model 1 for each simulation dataset.

$$\boldsymbol{Y} \sim Po\left(e^{a+\mathbf{X}\boldsymbol{\beta}}\right) \tag{20}$$

For PRE, PLE, PKLE, and PAMLPRE, the penalization coefficient k or d was selected by 10-fold cross-validation. The evaluation criterion of cross-validation is the sum of the log likelihood (LLH) values of the test set, obtained through the following equation:

$$\hat{\mu}_i = e^{\hat{a}+\boldsymbol{x}_i'\hat{\boldsymbol{\beta}}},$$

$$\mathrm{LLH} = \sum_{i=1}^{n} \left[y_i \times \ln \hat{\mu}_i - \hat{\mu}_i - \ln y_i!\right],$$

where $\hat{\mu}_i$, $\hat{a}$, and $\hat{\boldsymbol{\beta}}$ are the estimates of μ_i, a, and $\boldsymbol{\beta}$. The greater LLH is, the higher the predictive accuracy of the estimate is. Given k, $\boldsymbol{\beta}$ is calculated by a quasi-Newton method. The maximum number of iterations and the relative convergence tolerance are set to 1,000 and 10^{-30}, respectively.

3.3 | Performance evaluation

The MSE is a comprehensive index to evaluate the accuracy of the parameter estimate, which simultaneously considers bias and variance. Therefore, the MSE was employed to evaluate the performance of the PMLE, PRE, PLE, PKLE, and PAMLPE in each simulation scenario.

$$\text{MSE} = \frac{\sum_{i=1}^{1,000}\left(\widehat{\boldsymbol{\beta}}_i^* - \boldsymbol{\beta}\right)'\left(\widehat{\boldsymbol{\beta}}_i^* - \boldsymbol{\beta}\right)}{1,000},$$

where $\widehat{\boldsymbol{\beta}}_i^*$ is the estimate from PMLE, PRE, PLE, PKLE, or PAMLPE in the ith simulation dataset, and $\boldsymbol{\beta}$ is the true value of the parameter.

In addition, the PMSE was employed to evaluate the predictive accuracy of the PMLE, PRE, PLE, PKLE, or PAMLPE.

$$\text{PMSE} = \frac{\sum_{i}^{1,000}\left(\widehat{\boldsymbol{\mu}}_i^* - \boldsymbol{\mu}^*\right)'\left(\widehat{\boldsymbol{\mu}}_i^* - \boldsymbol{\mu}^*\right)/n}{1,000},$$

where $\boldsymbol{\mu}^*$ is the expected response variable without the observed error, and $\widehat{\boldsymbol{\mu}}_i^* = e^{\mathbf{X}\widehat{\boldsymbol{\beta}}_i^* + \widehat{\beta}_{0i}}$.

To better evaluate the performance of the PAMLPE, the following relative indicators are calculated:

$$\text{MSE}^{\text{pen}}\% = \frac{\text{MSE}^{\text{pen}} - \text{MSE}^{\text{PAMLPE}}}{\text{MSE}^{\text{pen}}} \times 100, \text{PMSE}^{\text{pen}}\% = \frac{\text{PMSE}^{\text{pen}} - \text{PMSE}^{\text{PAMLPE}}}{\text{PMSE}^{\text{pen}}} \times 100,$$

where $\text{MSE}^{\text{pen}}\%$ and $\text{PMSE}^{\text{pen}}\%$ evaluate the relative improvement of PAMLPE over PMLE/PRE/PLE/PKLE. Larger relative indicators indicate more improvement.

3.4 | Results and discussion

As shown in Table 3, all four penalized estimators, that is, PRE, PLE, PKLE, and PAMLPE, have smaller average MSE and PMSE than PMLE for almost all given simulation scenarios. In addition, in the simulation scenarios with $\boldsymbol{\beta}$ codes of 1 and 2, where the elements in $\boldsymbol{\beta}$ have the same sign and small variations, PAMLPE has the smallest average MSE and PMSE. In the simulation scenarios with $\boldsymbol{\beta}$ codes of 3–7, where the elements in $\boldsymbol{\beta}$ have greater variations, even different

TABLE 3 The average MSE and PMSE of five estimators for given $\boldsymbol{\beta}$ code.

	Average MSE					Average PMSE				
$\boldsymbol{\beta}$ code	PMLE	PRE	PLE	PKLE	PAMLPE	PMLE	PRE	PLE	PKLE	PAMLPE
1	0.1294	0.0551	0.0841	0.0643	**0.0246**	9.5885	8.4889	9.4521	8.4705	**6.0036**
2	0.1396	0.0623	0.0886	0.0727	**0.0516**	8.2565	7.7323	8.1326	7.7319	**7.5167**
3	0.2033	**0.1032**	0.1200	0.1224	0.1176	3.0255	2.9794	**2.9465**	2.9695	3.0455
4	0.2505	**0.1284**	0.1417	0.1529	0.1456	1.8516	1.8229	**1.7880**	1.8145	1.8361
5	0.3189	**0.1630**	0.1709	0.1948	0.1798	1.1213	1.1004	**1.0740**	1.0979	1.0982
6	0.6339	0.2817	**0.2708**	0.3495	0.2887	0.3334	0.3314	**0.3180**	0.3311	0.3294
7	0.7144	0.3303	**0.2948**	0.4500	0.3209	0.2540	0.2507	**0.2411**	0.2520	0.2504

Note: The numbers in bold indicate the minimum average MSE or PMSE given $\boldsymbol{\beta}$ code.
Abbreviations: MSE, mean squared error; PAMLPE, Poisson average maximum likelihood-centered penalized estimator; PKLE, Poisson K-L estimator; PLE, Poisson Liu estimator; PMLE, Poisson maximum likelihood estimator; PMSE, predictive mean squared error; PRE, Poisson ridge estimator.

signs, PAMLPE does not achieve the smallest average MSE and PMSE, but its disadvantage is slight. Furthermore, a same result is presented in Tables 4–7 for given the number of explanatory variables p, the ratio of the sample size to the number of explanatory variables g, the correlation between the explanatory variables r and the intercept a, showing that the result is robust. These results indicate that PRE, PLE, PKLE, and PAMLPE are all efficient in addressing multicollinearity in PR, but PAMLPE is more efficient when the true βs have the same sign and small variations. The detailed results are presented in Table S2 (see Data S1).

Specifically, the frequency of positive MSE and PMSE improvements of PAMLPE over PMLE/PRE/PLE/PKLE and the average relative improvement in the simulation scenarios with $\boldsymbol{\beta}$ codes of 1 and 2 were calculated and are presented in Tables 8 and 9. In the simulation scenarios with $\boldsymbol{\beta}$ codes of 1 and 2, where the elements in $\boldsymbol{\beta}$ have the same sign and small variation, PAMLPE achieves positive MSE and PMSE improvements over PMLE/PRE/PLE/PKLE in at least 74.1% (160/216) of the simulation scenarios, and the average relative MSE and PMSE improvements are greater than 35.9% and 19.9%, respectively. In addition, as the ratio of the sample size to the number of explanatory variables g and the intercept a decrease and the number of explanatory variables p increases, the average relative MSE and PMSE improvement of PAMLPE over PMLE/PRE/PLE/PKLE increase. As the correlation between the explanatory variables r increases, the average relative MSE and PMSE improvement increases for PAMLPE versus PMLE/PLE but decreases for PAMLPE versus PRE/PKLE. These results show that PAMLPE should be used more when the ratio of the sample size to the number of explanatory variables

TABLE 4 The average MSE and PMSE of five estimators for given p and $\boldsymbol{\beta}$ code.

		Average MSE					Average PMSE				
p	**Beta**	**PMLE**	**PRE**	**PLE**	**PKLE**	**PAMLPE**	**PMLE**	**PRE**	**PLE**	**PKLE**	**PAMLPE**
4	1	0.1376	0.0677	0.0878	0.0754	**0.0368**	3.3571	3.1459	3.2682	3.1170	**2.6039**
	2	0.1454	0.0740	0.0906	0.0824	**0.0646**	3.3563	3.1964	3.2749	3.1892	**3.1787**
	3	0.2012	**0.1119**	0.1183	0.1289	0.1317	1.9099	1.8831	**1.8538**	1.8770	1.9017
	4	0.2385	**0.1348**	0.1364	0.1575	0.1585	1.3934	1.3908	**1.3489**	1.3822	1.3918
	5	0.2896	0.1649	**0.1589**	0.1929	0.1885	0.9978	1.0023	**0.9639**	0.9966	0.9980
	6	0.5033	0.2713	**0.2396**	0.3307	0.2808	0.3844	0.3919	**0.3699**	0.3903	0.3889
	7	0.5744	0.3201	**0.2615**	0.4309	0.3091	0.2902	0.2916	**0.2768**	0.2929	0.2919
8	1	0.1213	0.0426	0.0805	0.0531	**0.0124**	15.8200	13.8320	15.6359	13.8241	**9.4033**
	2	0.1337	0.0507	0.0867	0.0631	**0.0387**	13.1568	12.2682	12.9903	12.2747	**11.8546**
	3	0.2055	**0.0946**	0.1216	0.1158	0.1035	4.1410	4.0756	**4.0392**	4.0621	4.1894
	4	0.2625	**0.1219**	0.1470	0.1483	0.1326	2.3097	2.2551	**2.2271**	2.2468	2.2803
	5	0.3482	**0.1610**	0.1829	0.1966	0.1710	1.2448	1.1986	**1.1842**	1.1993	1.1985
	6	0.7645	**0.2922**	0.3020	0.3683	0.2966	0.2824	0.2709	**0.2661**	0.2719	0.2698
	7	0.8544	0.3406	**0.3281**	0.4691	0.3327	0.2179	0.2097	**0.2054**	0.2111	0.2089

Note: The numbers in bold indicate the minimum average MSE or PMSE given p and $\boldsymbol{\beta}$ code.
Abbreviations: MSE, mean squared error; PAMLPE, Poisson average maximum likelihood-centered penalized estimator; PKLE, Poisson K-L estimator; PLE, Poisson Liu estimator; PMLE, Poisson maximum likelihood estimator; PMSE, predictive mean squared error; PRE, Poisson ridge estimator.

TABLE 5 The average MSE and PMSE of five estimators for given g and $\boldsymbol{\beta}$ code.

		Average MSE					**Average PMSE**				
g	**Beta**	**PMLE**	**PRE**	**PLE**	**PKLE**	**PAMLPE**	**PMLE**	**PRE**	**PLE**	**PKLE**	**PAMLPE**
10	1	0.2539	0.0964	0.1488	0.1191	**0.0416**	13.0644	12.0321	12.8348	12.0787	**9.8570**
	2	0.2717	0.1026	0.1541	0.1271	**0.0792**	11.6925	10.8529	11.4726	10.9169	**9.7629**
	3	0.3804	**0.1679**	0.1984	0.2162	0.2056	4.6246	**4.3616**	4.4698	4.3722	4.3787
	4	0.4631	**0.2082**	0.2307	0.2700	0.2521	2.9305	**2.7652**	2.8008	2.7688	2.7674
	5	0.5780	**0.2623**	0.2715	0.3404	0.3034	1.7899	**1.6841**	1.6923	1.6885	1.6775
	6	1.1223	0.4283	**0.4080**	0.5791	0.4446	0.5021	0.4868	**0.4697**	0.4861	0.4864
	7	1.2701	0.4891	**0.4406**	0.7420	0.4885	0.3783	0.3650	**0.3507**	0.3685	0.3669
20	1	0.0992	0.0482	0.0721	0.0533	**0.0240**	8.4809	7.3937	8.3501	7.3852	**4.8268**
	2	0.1086	0.0578	0.0776	0.0648	**0.0509**	6.9917	6.3745	6.8757	**6.3576**	6.6534
	3	0.1679	**0.0943**	0.1092	0.1029	0.0972	2.5865	2.6676	**2.5223**	2.6400	2.8347
	4	0.2083	**0.1155**	0.1289	0.1261	0.1202	1.6149	1.6633	**1.5674**	1.6404	1.6985
	5	0.2698	**0.1453**	0.1572	0.1600	0.1509	1.0266	1.0462	**0.9920**	1.0371	1.0485
	6	0.5373	0.2564	**0.2516**	0.2939	0.2591	0.3560	0.3622	**0.3453**	0.3619	0.3579
	7	0.6034	0.3106	**0.2786**	0.3901	0.2953	0.2614	0.2637	**0.2527**	0.2641	0.2614
40	1	0.0352	0.0208	0.0315	0.0204	**0.0082**	7.2202	6.0411	7.1713	5.9478	**3.3270**
	2	0.0384	0.0266	0.0343	0.0263	**0.0247**	6.0854	5.9694	6.0495	**5.9213**	6.1338
	3	0.0617	**0.0475**	0.0523	0.0480	0.0500	1.8653	1.9090	**1.8474**	1.8964	1.9232
	4	0.0802	**0.0613**	0.0654	0.0625	0.0645	1.0093	1.0403	**0.9958**	1.0343	1.0424
	5	0.1088	**0.0813**	0.0840	0.0838	0.0849	0.5473	0.5711	**0.5379**	0.5682	0.5687
	6	0.2419	0.1605	**0.1528**	0.1756	0.1624	0.1421	0.1453	**0.1389**	0.1454	0.1438
	7	0.2697	0.1912	**0.1653**	0.2178	0.1788	0.1225	0.1233	**0.1198**	0.1233	0.1230

Note: The numbers in bold indicate the minimum average MSE or PMSE given g and $\boldsymbol{\beta}$ code.
Abbreviations: MSE, mean squared error; PAMLPE, Poisson average maximum likelihood-centered penalized estimator; PKLE, Poisson K-L estimator; PLE, Poisson Liu estimator; PMLE, Poisson maximum likelihood estimator; PMSE, predictive mean squared error; PRE, Poisson ridge estimator.

g and the intercept a are small but the number of explanatory variables p and the correlation between the explanatory variables r are large, in which the multicollinearity problem is usually severe.

Moreover, as shown in Tables 3–7, the accuracy of the estimate from the five estimation methods is affected by the number of explanatory variables p, the ratio of the sample size to the number of explanatory variables g, the correlation between the explanatory variables r, the intercept a and the variation of the real βs value. As the $\boldsymbol{\beta}$ code increases, that is, the variation of the real βs value increases and the distance from the average of the real βs to 0 decreases, the average MSE increases, but the average PMSE decreases. Moreover, as the ratio of the sample size to the number of explanatory variables g and the intercept a increase and the correlation between the explanatory variables r decreases, the average MSE decreases. As the correlation between the explanatory variables r and the intercept a increase and the ratio of the sample size to

TABLE 6 The average MSE and PMSE of five estimators for given r and $\boldsymbol{\beta}$ code.

		Average MSE					Average PMSE				
r	Beta	PMLE	PRE	PLE	PKLE	PAMLPE	PMLE	PRE	PLE	PKLE	PAMLPE
0	1	0.0752	0.0836	0.0680	0.0825	**0.0255**	0.2614	0.2607	0.2560	0.2601	**0.1210**
	2	0.0741	0.0824	0.0671	0.0809	**0.0489**	0.2881	0.2878	0.2825	0.2873	**0.2726**
	3	0.0717	0.0699	**0.0638**	0.0698	0.0656	0.4169	0.4263	**0.4109**	0.4252	0.4214
	4	0.0707	0.0670	**0.0628**	0.0670	0.0668	0.4630	0.4752	**0.4568**	0.4743	0.4690
	5	0.0697	0.0646	**0.0617**	0.0647	0.0669	0.5068	0.5224	**0.5002**	0.5214	0.5146
	6	0.0664	0.0601	**0.0585**	0.0601	0.0643	0.4812	0.4918	**0.4736**	0.4913	0.4864
	7	0.0660	0.0617	**0.0584**	0.0620	0.0653	0.4456	0.4509	**0.4377**	0.4507	0.4505
0.3	1	0.0588	0.0548	0.0532	0.0543	**0.0168**	0.4279	0.4308	0.4215	0.4303	**0.2135**
	2	0.0584	0.0520	0.0527	0.0513	**0.0387**	0.4869	0.4935	0.4807	0.4923	**0.4712**
	3	0.0597	**0.0507**	0.0535	0.0507	0.0595	0.6642	0.6857	**0.6581**	0.6830	0.6741
	4	0.0608	**0.0527**	0.0546	0.0530	0.0629	0.6892	0.7128	**0.6832**	0.7096	0.7024
	5	0.0613	**0.0549**	0.0552	0.0553	0.0640	0.6792	0.7024	**0.6736**	0.6999	0.6918
	6	0.0644	0.0608	**0.0577**	0.0611	0.0650	0.5129	0.5261	**0.5064**	0.5245	0.5235
	7	0.0662	0.0650	**0.0593**	0.0652	0.0668	0.4136	0.4185	**0.4065**	0.4181	0.4202
0.6	1	0.0435	0.0358	0.0397	0.0354	**0.0109**	1.7513	1.7527	1.7403	1.7472	**1.1168**
	2	0.0441	0.0363	0.0406	0.0360	**0.0304**	1.7994	1.8168	**1.7892**	1.8096	1.7962
	3	0.0547	**0.0444**	0.0488	0.0446	0.0604	1.4172	1.4410	**1.4074**	1.4352	1.4551
	4	0.0608	**0.0509**	0.0537	0.0518	0.0667	1.1761	1.2004	**1.1664**	1.1944	1.2046
	5	0.0670	**0.0588**	0.0591	0.0603	0.0727	0.9534	0.9725	**0.9451**	0.9690	0.9763
	6	0.0866	0.0812	**0.0755**	0.0823	0.0846	0.4226	0.4310	**0.4157**	0.4289	0.4307
	7	0.0903	0.0895	**0.0785**	0.0890	0.0888	0.3036	0.3063	**0.2963**	0.3059	0.3075
0.9	1	0.0621	0.0342	0.0526	0.0335	**0.0120**	12.6252	12.2344	12.5637	12.1534	**8.1843**
	2	0.0661	0.0380	0.0556	0.0376	**0.0354**	11.0997	10.9213	11.0449	10.8634	**10.6451**
	3	0.1025	**0.0719**	0.0829	0.0770	0.0950	4.2425	4.3124	**4.2013**	4.2850	4.4664
	4	0.1251	**0.0929**	0.0994	0.1000	0.1171	2.5668	2.5870	**2.5299**	2.5698	2.6236
	5	0.1567	**0.1203**	0.1215	0.1302	0.1426	1.4365	1.4496	**1.4073**	1.4394	1.4508
	6	0.2890	0.2221	**0.2063**	0.2329	0.2248	0.2314	0.2280	**0.2185**	0.2264	0.2268
	7	0.3212	0.2710	**0.2263**	0.2803	0.2548	0.1436	0.1403	**0.1324**	0.1394	0.1389
0.95	1	0.1040	0.0404	0.0795	0.0418	**0.0174**	18.2699	16.9991	18.1343	16.8158	**11.6704**
	2	0.1106	0.0491	0.0839	0.0522	**0.0453**	15.6149	15.0632	15.4959	14.9735	**14.4715**
	3	0.1722	**0.1059**	0.1250	0.1194	0.1291	5.2619	5.2572	**5.1792**	5.2280	5.4379
	4	0.2199	**0.1421**	0.1558	0.1600	0.1681	2.9412	2.9514	**2.8731**	2.9264	2.9991
	5	0.2848	**0.1867**	0.1960	0.2089	0.2108	1.5463	1.5310	**1.4946**	1.5209	1.5354
	6	0.5669	0.3567	**0.3373**	0.3956	0.3617	0.1923	0.1834	**0.1740**	0.1827	0.1822
	7	0.6322	0.4398	**0.3711**	0.4968	0.4096	0.1184	0.1131	**0.1037**	0.1129	0.1110

(Continues)

TABLE 6 (Continued)

r	Beta	Average MSE PMLE	PRE	PLE	PKLE	PAMLPE	Average PMSE PMLE	PRE	PLE	PKLE	PAMLPE
0.99	1	0.4330	0.0820	0.2118	0.1381	**0.0648**	24.1954	19.2560	23.5966	19.4165	**14.7155**
	2	0.4840	0.1162	0.2320	0.1784	**0.1112**	20.2502	17.8111	19.7024	17.9655	**17.4435**
	3	0.7589	**0.2766**	0.3457	0.3729	0.2960	6.1502	**5.7536**	5.8222	5.7609	5.8184
	4	0.9657	**0.3646**	0.4238	0.4855	0.3918	3.2732	**3.0108**	3.0184	3.0126	3.0179
	5	1.2738	**0.4926**	0.5319	0.6491	0.5215	1.6055	1.4248	**1.4235**	1.4369	1.4205
	6	2.7298	0.9097	**0.8894**	1.2649	0.9318	0.1600	0.1283	**0.1197**	0.1329	0.1266
	7	3.1104	1.0550	**0.9752**	1.7066	1.0400	0.0994	0.0747	**0.0698**	0.0849	0.0744

Note: The numbers in bold indicate the minimum average MSE or PMSE given r and β code.
Abbreviations: MSE, mean squared error; PAMLPE, Poisson average maximum likelihood-centered penalized estimator; PKLE, Poisson K-L estimator; PLE, Poisson Liu estimator; PMLE, Poisson maximum likelihood estimator; PMSE, predictive mean squared error; PRE, Poisson ridge estimator.

TABLE 7 The average MSE and PMSE of five estimators for given a and β code.

a	Beta	Average MSE PMLE	PRE	PLE	PKLE	PAMLPE	Average PMSE PMLE	PRE	PLE	PKLE	PAMLPE
−1	1	0.2596	0.1051	0.1520	0.1257	**0.0493**	2.6134	2.1479	2.4972	2.1808	**1.6419**
	2	0.2834	0.1166	0.1607	0.1400	**0.0878**	2.2301	1.9509	2.1233	1.9629	**1.8975**
	3	0.4126	**0.1807**	0.2145	0.2212	0.1986	0.8104	**0.7432**	0.7465	0.7435	0.7500
	4	0.5054	**0.2191**	0.2492	0.2717	0.2452	0.4972	0.4521	**0.4487**	0.4527	0.4521
	5	0.6454	**0.2777**	0.2985	0.3481	0.3060	0.3026	0.2732	**0.2681**	0.2750	0.2720
	6	1.2939	0.4783	**0.4613**	0.6363	0.4921	0.0906	0.0841	**0.0791**	0.0839	0.0825
	7	1.4611	0.5462	**0.4998**	0.8279	0.5441	0.0700	0.0622	**0.0600**	0.0637	0.0631
0	1	0.0942	0.0425	0.0703	0.0481	**0.0177**	7.0344	6.0488	6.8932	6.0673	**4.4031**
	2	0.0990	0.0492	0.0736	0.0552	**0.0441**	6.1091	5.6112	5.9803	5.6225	**5.4388**
	3	0.1450	**0.0868**	0.1010	0.0994	0.1022	2.2556	2.1739	2.1745	**2.1648**	2.2001
	4	0.1806	**0.1116**	0.1216	0.1275	0.1282	1.3678	1.3169	**1.3013**	1.3120	1.3245
	5	0.2286	**0.1422**	0.1479	0.1614	0.1574	0.8161	0.7816	**0.7681**	0.7808	0.7818
	6	0.4471	0.2459	**0.2369**	0.2834	0.2510	0.2439	0.2385	**0.2285**	0.2383	0.2367
	7	0.5015	0.2954	**0.2585**	0.3622	0.2790	0.1870	0.1799	**0.1738**	0.1813	0.1804
1	1	0.0344	0.0178	0.0302	0.0189	**0.0067**	19.1177	17.2701	18.9658	17.1635	**11.9658**
	2	0.0363	**0.0212**	0.0316	0.0230	0.0229	16.4304	15.6349	16.2942	15.6104	**15.2137**
	3	0.0523	**0.0422**	0.0443	0.0466	0.0520	6.0105	6.0210	**5.9185**	6.0004	6.1866
	4	0.0655	0.0543	**0.0542**	0.0594	0.0633	3.6897	3.6998	**3.6139**	3.6789	3.7317
	5	0.0826	0.0690	**0.0663**	0.0748	0.0759	2.2452	2.2465	**2.1859**	2.2380	2.2408
	6	0.1606	0.1210	**0.1142**	0.1289	0.1230	0.6657	0.6717	**0.6463**	0.6711	0.6689
	7	0.1805	0.1494	**0.1262**	0.1597	0.1396	0.5051	0.5099	**0.4894**	0.5109	0.5079

Note: The numbers in bold indicate the minimum average MSE or PMSE given a and β code.
Abbreviations: MSE, mean squared error; PAMLPE, Poisson average maximum likelihood-centered penalized estimator; PKLE, Poisson K-L estimator; PLE, Poisson Liu estimator; PMLE, Poisson maximum likelihood estimator; PMSE, predictive mean squared error; PRE, Poisson ridge estimator.

TABLE 8 The relative MSE improvements of PAMLPE over other estimators for β codes of 1 and 2.

		$MSE^{PMLE}\%$		$MSE^{PRE}\%$		$MSE^{PLE}\%$		$MSE^{PKLE}\%$	
par0	**Par**	**Fre**[a]	**Mean**	**Fre**	**Mean**	**Fre**	**Mean**	**Fre**	**Mean**
β	1	100.0	76.2	99.1	64.4	100.0	73.3	100.0	66.6
	2	85.2	27.5	68.5	7.4	84.3	22.0	73.1	10.0
p	4	88.0	44.1	73.1	26.7	88.0	37.9	78.7	29.1
	8	97.2	59.6	94.4	45.0	96.3	57.5	94.4	47.4
g	10	93.1	58.0	79.2	34.5	91.7	51.2	83.3	40.1
	20	91.7	51.4	84.7	36.7	91.7	47.2	88.9	39.0
	40	93.1	46.2	87.5	36.4	93.1	44.7	87.5	35.8
r	0	88.9	40.3	88.9	42.5	86.1	38.2	88.9	42.3
	0.3	91.7	42.7	88.9	41.3	91.7	40.4	88.9	41.1
	0.6	86.1	44.1	80.6	39.1	86.1	42.2	80.6	38.9
	0.9	88.9	52.2	80.6	35.9	88.9	49.2	83.3	35.3
	0.95	100.0	59.1	80.6	34.4	100.0	54.9	83.3	34.8
	0.99	100.0	72.7	83.3	22.0	100.0	61.3	94.4	37.3
a	−1	100.0	61.6	97.2	40.3	100.0	54.0	100.0	44.5
	0	95.8	52.0	83.3	36.4	95.8	48.7	87.5	38.5
	1	81.9	41.9	70.8	30.9	80.6	40.4	72.2	31.9

Abbreviations: MSE, mean squared error; PAMLPE, Poisson average maximum likelihood-centered penalized estimator; PKLE, Poisson K-L estimator; PLE, Poisson Liu estimator; PMLE, Poisson maximum likelihood estimator; PMSE, predictive mean squared error; PRE, Poisson ridge estimator.

[a] Freq indicates the frequency of $MSE^{pen}\% \geq 0$.

the number of explanatory variables g decreases, the mean PMSE increases. The effect of the number of explanatory variables p on MSE and PMSE is related to the variation in the real βs value increases and the distance from the average of the real βs to 0 decreases. These results suggest that the above parameters should be considered when designing simulation studies to compare the performance of different estimation methods in PR, particularly the variation in the real βs value, which was not considered in previous studies.

Additionally, the results of analytical methods for calculating the penalization coefficient k or d are presented in the Data S1, in which similar results were found.

4 | CASE STUDY

To illustrate the performance of PAMLPE in practice, the aircraft damage dataset was adopted (see Data S2). The dataset was initially discussed by Myers, Montgomery, Vining, and Robinson (2012). Recently, the dataset was frequently used to assess the performance of different methods for the multicollinear PR model (Akay & Ertan, 2022; Aladeitan et al., 2021; Asar & Genç, 2018; Lukman et al., 2021). The dataset provides the information about 30 strike missions of two types

TABLE 9 The relative PMSE improvements of PAMLPE over other estimators for $\boldsymbol{\beta}$ codes of 1 and 2.

		$\text{PMSE}^{\text{PMLE}}\%$		$\text{PMSE}^{\text{PRE}}\%$		$\text{PMSE}^{\text{PLE}}\%$		$\text{PMSE}^{\text{PKLE}}\%$	
par0	**Par**	**Fre**[a]	**Mean**	**Fre**	**Mean**	**Fre**	**Mean**	**Fre**	**Mean**
β	1	100.0	42.0	100.0	37.5	100.0	40.2	100.0	37.2
	2	63.9	6.7	49.1	2.8	61.1	4.3	48.1	2.6
p	4	75.0	18.9	72.2	15.1	74.1	15.9	71.3	14.8
	8	88.9	29.9	76.9	25.1	87.0	28.7	76.9	25.1
g	10	93.1	23.1	93.1	19.8	91.7	20.6	91.7	19.6
	20	84.7	26.6	69.4	19.5	81.9	23.8	69.4	19.7
	40	68.1	23.5	61.1	21.0	68.1	22.4	61.1	20.5
r	0	80.6	32.4	80.6	31.7	80.6	30.2	80.6	31.5
	0.3	80.6	28.8	75.0	28.6	75.0	27.2	75.0	28.4
	0.6	72.2	20.5	75.0	20.9	72.2	19.5	75.0	20.6
	0.9	77.8	19.1	69.4	15.7	75.0	17.9	66.7	15.2
	0.95	83.3	20.6	69.4	14.6	83.3	18.9	69.4	13.7
	0.99	97.2	24.9	77.8	9.2	97.2	19.9	77.8	10.3
a	−1	91.7	28.2	84.7	21.2	91.7	24.1	84.7	21.1
	0	80.6	23.9	69.4	19.9	79.2	22.3	69.4	19.7
	1	73.6	21.0	69.4	19.2	70.8	20.4	68.1	19.0

Abbreviations: PAMLPE, Poisson average maximum likelihood-centered penalized estimator; PKLE, Poisson K-L estimator; PLE, Poisson Liu estimator; PMLE, Poisson maximum likelihood estimator; PMSE, predictive mean squared error; PRE, Poisson ridge estimator.

[a] Freq indicates the frequency of $\text{PMSE}^{\text{pen}}\% \geq 0$.

of aircraft (the McDonnell Douglas A-4 Skyhawk and the Grumman A-6 Intruder). The explanatory variables are as follows: x_1 is a binary variable representing the aircraft type (A-4 coded as 0 and A-6 coded as 1), and x_2 and x_3 denote bomb load in tons and total months of aircrew experience, respectively. The response variable, y which denotes the number of locations with damage to the aircraft follows a Poisson distribution. The condition index of $\mathbf{X}'\mathbf{W}\mathbf{X}$ is 260.98, showing that multicollinearity exists. Moreover, the correlation matrix between the variables obtained by the columns of $\mathbf{W}^{\frac{1}{2}}\mathbf{X}$ is given as

$$\begin{bmatrix} 1 & 0.8969 & 0.8788 \\ & 1 & 0.8453 \\ & & 1 \end{bmatrix}.$$

The common sense of mechanics hints that the coefficients of x_1 and x_2 are positive, but the coefficient of x_3 is negative, which is also supported by univariate analysis and some studies (Akay & Ertan, 2022; Aladeitan et al., 2021; Asar & Genç, 2018; Lukman et al., 2021). The results of the simulation study show that PAMLPE is more suitable for scenarios with βs having the same sign and the performance of the other four methods is not affected by the sign of

TABLE 10 The point estimates (*SD*s) for the PMLE, PRE, PLE, PKLE and PAMLPE for the aircraft damage dataset.

	PMLE	PRE	PLE	PKLE	PAMLPE[a]	PAMLPE[b]
β_0	−0.406(0.877)	−0.296(0.721)	−0.386(0.845)	−0.288(0.694)	−0.714(0.792)	0.227(0.253)
β_1	0.569(0.504)	0.500(0.325)	0.557(0.461)	0.497(0.284)	0.580(0.325)	0.742(0.152)*
β_2	0.165(0.068)*	0.139(0.048)*	0.160(0.063)*	0.135(0.043)*	0.149(0.048)*	0.128(0.026)*
β_3	−0.014(0.008)	−0.011 (0.007)	−0.013(0.008)	−0.010(0.007)	−0.007(0.007)	−0.019(0.004)*
MSE	1.029	0.631	0.931	0.568	0.771	0.437
PMSE	1.833	1.806	1.801	1.750	1.854	1.477
LLH	−45.858	−43.985	−45.112	−43.267	−44.278	−42.116
k/d	—	6.295	0	3.293	6.945	138.128

Note: When calculating LLH and PMSE, the penalization coefficient is fixed and calculated using the complete dataset. k or d were chosen by 10-fold cross-validation.

Abbreviations: MSE, mean squared error; LLH, log likelihood; PAMLPE, Poisson average maximum likelihood-centered penalized estimator; PKLE, Poisson K-L estimator; PLE, Poisson Liu estimator; PMLE, Poisson maximum likelihood estimator; PMSE, predictive mean squared error; PRE, Poisson ridge estimator.

[a] PAMLPE is applied before transforming x_3 into its opposite number.

[b] PAMLPE is applied after transforming x_3 into its opposite number.

[*] The point estimate is statistically different from 0; that is, $p \le .05$.

the true coefficient. Therefore, x_3 was transformed into its opposite number to make the signs of βs identical before estimating the parameters. The penalization coefficient k or d was chosen by 10-fold cross-validation. As the true βs are unknown in real-world datasets, a qualitative comparison based on professional knowledge, and a quantitative comparison based on the predictive mean squared error (PMSE) and the sum of LLH value from leave-one-out cross-validation were employed to evaluate the performance of different methods (Lukman et al., 2022). When computing the MSE values by Equations (3), (9), (12), (15), and (19), the maximum likelihood estimate $\hat{\boldsymbol{\beta}}^{\text{PMLE}}$ is used in place of the unknown parameter $\boldsymbol{\beta}$ (Akay & Ertan, 2022). On the other hand, sometimes there is no prior information about βs, and the univariate analysis may also misjudge the signs of βs. Therefore, to evaluate the risk of misjudging signs, PAMLPE was also applied without transforming x_3 into its opposite number.

As shown in Table 10, all six methods have the same result on the signs of β_1, β_2, and β_3, which is consistent with professional knowledge. After transforming x_3 into its opposite number, PAMLPE achieves the smallest standard deviations and MSE, identifies the most statistically significant coefficients (i.e., $p \le .05$), and has the greatest LLH. When x_3 is not transformed into its opposite number, PAMLPE has a greater LLH than PMLE and PLE, but a smaller LLH than PRE and PKLE. However, the disadvantage of PAMLPE versus PRE and PKLE is slight. The result suggests that PAMLPE outperforms the other four methods when the signs of βs are identical, even when the signs are different the disadvantage of PAMLPE versus PRE and PKLE is slight. In addition, the result also indicates that PAMLPE can be better used by changing the sign of the explanatory variable to make the signs of βs identical according to prior knowledge. Even if there is no prior information about the signs and the univariate analysis misjudges the signs of βs, which may lead to the signs of βs being different, the cost of misjudging signs is small. Additionally, the results of analytical methods for calculating the penalization coefficient k or d are presented in the Data S1, in which similar results were found.

5 | CONCLUSION

In this study, a new estimator called the Poisson average maximum likelihood-centered penalized estimator (PAMLPE) was proposed. PAMLPE uses the weighted average of the maximum likelihood estimators as the shrinkage center, which is more suitable than the shrinkage center of PRE and PKLE. A simulation study and case study indicate that PAMLPE is more efficient than PRE, PLE and PKLE in addressing multicollinearity in PR when the true βs have the same sign and small variation. Additionally, PAMLPE should be used more when the ratio of the sample size to the number of explanatory variables g and the intercept a are small but the number of explanatory variables p and the correlation between the explanatory variables r are large, in which the multicollinearity problem is usually severe. Therefore, we recommend using PAMLPE to address the multicollinearity problem when the signs of the true βs are known to be identical in advance. Moreover, PAMLPE can be better used by changing the sign of the explanatory variable to make the signs of the βs identical according to prior knowledge.

ACKNOWLEDGEMENTS

We appreciate our colleagues, Siwei Zhai, Yi Zhang, Jingfei Huang, Caiying Luo, Jian Qian, Jianping Wang and Wenhui Zhu for their constructive comments, which improved the quality of this work.

CONFLICT OF INTEREST STATEMENT

The authors report that there are no competing interests to declare.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available in the Data S1 and S2 of this article.

REFERENCES

Akay, K. U., & Ertan, E. (2022). A new improved Liu-type estimator for Poisson regression models. *Hacettepe Journal of Mathematics and Statistics*, *51*, 1–20.

Aladeitan, B. B., Adebimpe, O., Lukman, A. F., Oludoun, O., & Abiodun, O. E. (2021). Modified Kibria-Lukman (MKL) estimator for the Poisson regression model: Application and simulation. *F1000Research*, *10*, 548.

Algamal, Z. Y. (2020). Shrinkage parameter selection via modified cross-validation approach for ridge regression model. *Communications in Statistics: Simulation and Computation*, *49*, 1922–1930.

Algamal, Z. Y., & Lee, M. H. (2015). Adjusted adaptive lasso in high-dimensional poisson regression model. *Modern Applied Science*, *9*, 170.

Alkhateeb, A., & Algamal, Z. (2020). Jackknifed Liu-type estimator in Poisson regression model. *Journal of the Iranian Statistical Society*, *19*, 21–37.

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*, e6340.

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*, 793–803.

Choi, H., Song, E., Hwang, S.-S., & Lee, W. (2018). A modified generalized lasso algorithm to detect local spatial clusters for count data. *AStA Advances in Statistical Analysis*, *102*, 537–563.

Chowdhury, S., Chatterjee, S., Mallick, H., Banerjee, P., & Garai, B. (2019). Group regularization for zero-inflated poisson regression models with an application to insurance ratemaking. *Journal of Applied Statistics*, *46*, 1567–1581.

Hoerl, A., & Kennard, R. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics*, *12*, 55–67.

Hossain, S., & Ahmed, E. (2012). Shrinkage and penalty estimators of a Poisson regression model. *Australian & New Zealand Journal of Statistics*, *54*, 359–373.

Hsiang, T. C. (1975). Bayesian view on ridge regression. *Statistician*, *24*, 267–268.

Ivanoff, S., Picard, F., & Rivoirard, V. (2016). Adaptive lasso and group-lasso for functional Poisson regression. *The Journal of Machine Learning Research*, *17*, 1903–1948.

Kandemir Çetinkaya, M., & Kaçıranlar, S. (2019). Improved two-parameter estimators for the negative binomial and Poisson regression models. *Journal of Statistical Computation and Simulation*, *89*, 2645–2660.

Kibria, B. M. G., & Lukman, A. F. (2020). A new ridge-type estimator for the linear regression model: Simulations and applications. *Scientifica*, *2020*, 9758378.

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*, 3732.

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*, 2124–2136.

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

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

Matsuda, T., & Miyatake, Y. (2021). Generalized nearly isotonic regression. *arXiv Preprint arXiv:2108.13010*.

Myers, R. H., Montgomery, D. C., Vining, G. G., & Robinson, T. J. (2012). *Generalized linear models: With applications in engineering and the sciences*. Hoboken, New Jersey: John Wiley & Sons.

Noori Asl, M., Bevrani, H., & Arabi Belaghi, R. (2022). Penalized and ridge-type shrinkage estimators in Poisson regression model. *Communications in Statistics: Simulation and Computation*, *51*, 4039–4056.

Oranye, H. E., & Ugwuowo, F. I. (2022). Modified jackknife Kibria-Lukman estimator for the Poisson regression model. *Concurrency and Computation: Practice and Experience*, *34*, e6757.

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*, 2258–2271.

Rasheed, H. A., Sadik, N. J., & Algamal, Z. Y. (2022). Jackknifed Liu-type estimator in the Conway-Maxwell Poisson regression model. *International Journal of Nonlinear Analysis and Applications*, *13*, 3153–3168.

Swindel, B. F. (1976). Good ridge estimators based on prior information. *Communications in Statistics-Theory and Methods*, *5*, 1065–1075.

Tibshirani, R. J., & Taylor, J. (2011). The solution path of the generalized lasso. *The Annals of Statistics*, *39*, 1335–1371.

Turkan, S., & Ozel, G. (2016). A new modified jackknifed estimator for the Poisson regression model. *Journal of Applied Statistics*, *43*, 1892–1905.

Wang, W., Li, L., Li, S., Yin, F., Liao, F., Zhang, T., … Ma, Y. (2022). Average ordinary least squares-centered penalized regression: A more efficient way to address multicollinearity than ridge regression. *Statistica Neerlandica*, *76*, 347–368.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Li, S., Wang, W., Yao, M., Wang, J., Du, Q., Li, X., Tian, X., Zeng, J., Deng, Y., Zhang, T., Yin, F., & Ma, Y. (2023). Poisson average maximum likelihood-centered penalized estimator: A new estimator to better address multicollinearity in Poisson regression. *Statistica Neerlandica*, 1–20. https://doi.org/10.1111/stan.12313