REVSTAT – Statistical Journal
Volume 0, Number 0, Month 0000, 000-000
https://doi.org/00.00000/revstat.v00i0.000

The Wald-type confidence interval on the mean response function of the Poisson Inverse Gaussian Ridge regression

Authors: Muhammad Qasim
– Department of Economics, Finance and Statistics,
Jönköping International Business School,
Jönköping University,
Sweden
muhammad.qasim@ju.se

Y. Murat BULUT ✉
– Department of Statistics,
Science Faculty,
Eskişehir Osmangazi University,
Türkiye
ymbulut@ogu.edu.tr

Kristofer Månsson
– Department of Economics, Finance and Statistics,
Jönköping International Business School,
Jönköping University,
Sweden
kristofer.mansson@ju.se

Received: Month 0000 Revised: Month 0000 Accepted: Month 0000

Abstract:

- The negative binomial (NB) regression model is commonly used to model overdispersed count data. However, the NB regression model is not suitable for highly overdispersed data, for which the Poisson-inverse Gaussian (PIG) regression model is often used instead. The maximum likelihood (ML) estimator is typically used to estimate the coefficients of the PIG regression model. However, when multicollinearity exists among the explanatory variables, the ML estimator's variance can become inflated. To address this issue, we propose PIG ridge regression (PIGRR) and quantile-based ridge regression estimators for the PIG regression model. We also suggest using a Wald-type method to calculate the confidence interval on the mean response function of the PIGRR. To evaluate the performance of these proposed methods, we conducted a Monte Carlo simulation study, considering mean squared error and average confidence lengths as performance criteria. Additionally, we analyzed the traffic fatalities dataset

✉ Corresponding author

to demonstrate the benefits of the proposed estimators for practitioners dealing with multicollinearity issues in real datasets.

Keywords:

- *Multicollinearity; Poisson-inverse Gaussian distribution; Ridge Regression; Confidence Interval; Traffic fatalities.*

AMS Subject Classification:

- 62J07, 62J12.

1. INTRODUCTION

Count data is commonly analyzed in many real-world applications. This includes examples such as the number of days lost due to work accidents, the number of occurrences of thunderstorms in a calendar year, or the number of insurance claims. The Poisson distribution is the primary choice for modeling such data due to its simplicity, but it assumes that the mean and variance of the distribution are equal, known as the equidispersion assumption. This makes the Poisson distribution unsuitable for modeling overdispersed data, where the variance is larger than the mean. Various factors can cause overdispersion, such as an insufficient number of interaction terms in the regression model or the omission of important explanatory variables.

Various statistical distributions have been proposed to model overdispersed data, with the negative binomial (NB) distribution being the most popular. The NB distribution allows for the variance to exceed the mean of the data and can be obtained as a mixture of Poisson and gamma distributions. In regression models, when the dependent variable, y_i, takes the form of non-negative integers or counts, the NB regression model, which is a particular case of generalized linear models (GLMs), can be applied. The NB regression model is widely used in applied areas such as social, economic, and health sciences [30]. The coefficients of the NB regression model can be estimated using the maximum likelihood (ML) estimation method, which involves solving a non-linear equation through iterative algorithms such as the iterative weighted least squares (IWLS) algorithm.

The PIG distribution is another distribution used to model overdispersed data that is discussed in the literature. It is a special case of the Sichel distribution with two parameters [45]. Like the NB distribution, the PIG distribution can also be expressed as a mixture of distributions [48]. However, the PIG distribution uses a mixture of the Poisson and inverse Gaussian distributions, while the NB distribution uses a gamma distribution as the mixing distribution. The PIG distribution is particularly useful for count data with longer tails and larger kurtosis, making it an alternative to the NB distribution [14]. The PIG distribution is better suited for data with a high initial peak that may be skewed to the right, giving it a modeling capacity superior to that of the NB distribution [17]. Furthermore, the PIG distribution has an easily obtainable likelihood function that has a closed-form representation, making parameter estimation relatively simple [49]. These properties make the PIG distribution more effective than the NB distribution in dealing with highly overdispersed data. The PIG distribution has been applied in a variety of contexts, including modeling vehicle crash data in Zha et al. [49] and comparing NB and PIG regression models for horseshoe crabs data in Putri et al. [39].

In regression models, another issue that can arise is the problem of multicollinearity. This occurs when the explanatory variables are highly correlated with each other, as defined by Frisch [15]. Multicollinearity has a negative impact on the ML estimator, resulting in estimated coefficients with larger variances and unreliable inference. Confidence intervals for unknown parameters may also have a wider range than usual, and it becomes difficult to estimate the unique effects of the explanatory variables. To address this issue, Hoerl and Kennard introduced the Ridge estimator in the linear regression model (LRM). Since then, many different biased estimators have been proposed to combat multicollinearity, including the Liu estimator [23], Liu-type estimator [24], and two-parameter estimator [38] and [42], among others.

In GLMs, the problem of multicollinearity was first addressed by Schaefer et al. [43], who proposed the Ridge estimator for logistic regression, a particular case of GLMs [43]. Since then, the Ridge estimator has been adapted for other special cases of GLMs. For instance, Månsson and Shukur [29] introduced the Ridge estimator in the Poisson regression model and Locking et al. [25] defined it for the Probit regression model. Algamal [5] proposed the Ridge estimator for gamma and inverse Gaussian regression models, while Månsson [30] developed it for the negative binomial model. Khalaf et al. [21] introduced the Ridge estimator in the Tobit regression model, and Qasim et al. [41] proposed it for the Beta regression model. Recently, Awwad et al. [8] proposed two-parameter estimators for the logistic regression model. Other biased estimators have also been proposed by different authors to address multicollinearity in GLMs (see, e.g., [1]- [4], [6], [7], [9], [10], [11], [13], [22], [26], [27], [28], [35], [36], [37], [40], [42], [46], [47]). To the best of our knowledge, no Ridge regression or quantile-based ridge regression estimators have been defined for the PIG regression model. Moreover, we suggest using a Wald-type method to calculate the confidence interval for the mean response function of the PIG ridge regression (PIGRR). Gómez-Déniz and Calderín-Ojeda [16] investigated the mixture of Poisson-reciprocal inverse Gaussian distributions, which, under specific conditions, outperforms the PIG model. This paper introduces the PIGRR and a confidence interval that can be extended to a mixed Poisson regression model utilizing the reciprocal inverse Gaussian as the mixing distribution . The implementation of these enhancements holds substantial potential for advancing future research.

This paper is organized as follows. In Section 2, we introduce the PIG regression model, and an estimation of the parameters is given. In Section 3, we define the Ridge estimator for the PIG regression model and propose some estimators of the Ridge parameter, also, the Wald confidence interval is given. In Section 4, the Monte Carlo simulation study is conducted. In Section 5, we give a real data example. Finally, the paper ended with conclusions.

2. METHODOLOGY

In this section, we present the PIG regression model and introduce the ridge regression estimator, which was defined by Hoerl and Kennard [18, 19] in the PIG regression model. Furthermore, we introduce the properties of matrix mean squared error (MMSE) and scalar mean squared error (MSE). We will also propose certain biasing parameters to improve the performance of the proposed estimator.

2.1. The Poisson-inverse Gaussian regression model

The probability mass function (pmf) of the PIG distribution is given as

$$P\left(y_i;\mu_i,\tau\right)=\left(\frac{2\alpha_i}{\pi}\right)^{\frac{1}{2}}\frac{\mu_i^{y_i}\exp\left(\frac{1}{\tau}\right)K_{y_i-\frac{1}{2}}\left(\alpha_i\right)}{\left(\alpha_i\tau\right)^{y_i}y_i!},y_i=0,1,2,\ldots \tag{2.1}$$

where $\alpha_i=\frac{\sqrt{1+2\mu_i\tau}}{\tau}$ and $K_s\left(\cdot\right)$ is the modified Bessel function of the second kind [45]. The mean and variance are given as respectively

$$E\left(Y_i\right)=\mu_i, \tag{2.2}$$

$$Var\left(Y_i\right)=\mu_i\left(1+\tau\mu_i\right). \tag{2.3}$$

In the regression framework, generally, the mean of the response variable is modeled. Because $Var(Y_i)>E(Y_i)$, the PIG model is suitable to model overdispersed data.

Let $Y_1,Y_2,\ldots,Y_n$ be n independent random variables from the PIG distribution with the parameters μ and τ ($Y_i\sim PIG(\mu,\tau)$). Then, we assume that the mean of Y_i satisfies the following functional relation:

$$g\left(\mu_i\right)=\eta_i=x_i^T\beta,\quad i=1,2,\ldots,n \tag{2.4}$$

where η_i is the linear predictor, $x_i^T=(x_{i1},x_{i2},\ldots,x_{in})$ denotes the vector of covariates, and $\beta=(\beta_1,\beta_2,\ldots,\beta_p)$ is the regression coefficient's vector with size p. It is assumed that the mean link function $g:(0,\infty)\to\mathbb{R}$ is strictly monotonic and twice differentiable. In literature, there are different link functions such as identity link function ($g(\mu)=\mu$), logarithmic link function ($g(\mu)=\log(\mu)$), and square root link function ($g(\mu)=\sqrt{\mu}$) for more details about the link function, please see [31]. In this study, we use the log link function.

Generally, the estimate of β is usually found using the ML estimation method. The likelihood function of the PIG is given as

$$\begin{aligned} L\left(\mu,\tau;y\right) &= \prod_{i=1}^{n}P\left(y_i;\mu,\tau\right) \\ &= \prod_{i=1}^{n}\left(\frac{2\alpha_i}{\pi}\right)^{\frac{1}{2}}\frac{\mu_{y_i}\exp\left(\frac{1}{\tau}\right)K_{y_i-\frac{1}{2}}\left(\alpha_i\right)}{\left(\alpha_i\tau\right)^{y_i}y_i!}. \end{aligned} \tag{2.5}$$

The log-likelihood function is obtained by taking the natural logarithm of both sides of the Eq. (2.5) as follows

$$\begin{aligned}\ell(\mu,\tau;y) = & \sum_{i=1}^{n}\left\{\frac{1}{2}\log(2\alpha_i)-\frac{1}{2}\log(\pi)+y_i\log(\mu_i)+\frac{1}{\tau}+\log\left(K_{y_i-\frac{1}{2}}(\alpha_i)\right)\right.\\ & \left. - y_i\log(\alpha_i)-y_i\log(\tau)-\log(y_i!)\right\}.\end{aligned} \tag{2.6}$$

When we take $\mu_i=\exp\left(x_i^T\beta\right)$, the log-likelihood function is rewritten as follows

$$\begin{aligned}\ell(\beta,\tau;y) = & \sum_{i=1}^{n}\left\{ \frac{1}{2}\log\left(2\frac{\sqrt{1+2\exp\left(x_i^T\beta\right)\tau}}{\tau}\right)-\frac{1}{2}\log(\pi)\right.\\ & + y_i\log\left(\exp\left(x_i^T\beta\right)\right)+\frac{1}{\tau}+\log\left(K_{y_i-\frac{1}{2}}\left(\frac{\sqrt{1+2\exp\left(x_i^T\beta\right)\tau}}{\tau}\right)\right)\\ & \left. - y_i\log\left(\frac{\sqrt{1+2\exp\left(x_i^T\beta\right)}}{\tau}\right)-y_i\log(\tau)-\log(y_i!)\right\}.\end{aligned} \tag{2.7}$$

The vector of coefficients using the ML estimation method is estimated by solving the following equation

$$\begin{aligned}S(\beta) = & \frac{\partial\ell(\beta,\tau;y)}{\partial\beta}\\ = & \sum_{i=1}^{n}\left\{x_i\left[y_i-\frac{R_{y_i-\frac{1}{2}}\left(\frac{\sqrt{1+2\exp\left(x_i^T\beta\right)\tau}}{\tau}\right)}{\left(1+2\exp\left(x_i^T\beta\right)\tau\right)^{\frac{1}{2}}}\exp\left(x_i^T\beta\right)\right]\right\}=0,\end{aligned} \tag{2.8}$$

where $R_\nu(z)=\frac{K_{\nu+1}(z)}{K_\nu(z)}$ is calculated recursively as $R_{-\frac{1}{2}}(z)=1$; $R_\nu(z)=\frac{2\nu}{z}+\frac{1}{R_{\nu-1}(z)}$.

Since the Eq. (2.8) is a nonlinear, $S(\beta)$ can not be solved explicitly. Iterative methods, such as Newton-Raphson or Iteratively Reweighted Least Squares (IRLS), are necessary to solve this equation. When estimating unknown coefficients using the IRLS algorithm, the unknown parameters are obtained in each iteration as follows:

$$\beta^{(m+1)}=\beta^{(m)}+I^{-1}\left(\beta^{(m)}\right)S\left(\beta^{(m)}\right), \tag{2.9}$$

where $S\left(\beta^{(m)}\right)$ is the score vector evaluated at $\beta^{(m)}$ and $I^{-1}\left(\beta^{(m)}\right)$ is the Fisher information matrix evaluated at $\beta^{(m)}$. When convergence holds, the vector of coefficients can be obtained as follows

$$\widehat{\beta}_{ML}=(\mathcal{D})^{-1}X^T\widehat{W}\widehat{u}, \tag{2.10}$$

where $\widehat{u}_i = \log(\widehat{\mu}_i) + \frac{(y_i - \widehat{\mu}_i)}{\widehat{\mu}_i}$, $W = diag\left(\frac{\widehat{\mu}_i}{1+\widehat{\tau}\widehat{\mu}_i}\right)$ and $\mathcal{D} = X^T\widehat{W}X$. Also, to estimate dispersion parameter τ, partial derivative of the log-likelihood function is obtained as

$$\begin{aligned} S(\tau) &= \frac{\partial \ell(\beta, \tau; y)}{\partial \beta} \\ &= \sum_{i=1}^{n} \left\{ -\frac{1}{\tau^2} - \frac{y_i}{\tau} \right. \\ &+ \left. R_{y_i - \frac{1}{2}} \left(\frac{\sqrt{1 + 2\exp\left(x_i^T \beta\right)\tau}}{\tau} \right) \frac{1 + \tau \exp\left(x_i^T\beta\right)}{\tau^2 \left(1 + 2\exp\left(x_i^T\beta\right)\right)^{\frac{1}{2}}} \right\} = 0. \end{aligned}$$

The covariance matrix of the ML estimator is obtained as follows,

$$Cov\left(\widehat{\beta}_{ML}\right) \cong (\mathcal{D})^{-1}. \tag{2.11}$$

The asymptotic mean squared error (MSE) of ML estimator can be defined as

$$MSE\left(\widehat{\beta}_{ML}\right) \cong tr\left[(\mathcal{D})^{-1}\right] = \sum_{j=1}^{p} \frac{1}{\lambda_j}, \tag{2.12}$$

where λ_j is the j^{th} eigenvalue of the $X^T\widehat{W}X$ matrix. When the explanatory variables in a PIG regression model are highly intercorrelated, it results in a multicollinearity problem. This often causes the matrix of weighted cross products to become ill-conditioned, with some eigenvalues being small. As a consequence, the estimated MSE of the ML estimator is inflated. To overcome this issue, we propose a biased estimator, the PIGRR estimator.

3. Poisson Inverse Gaussian Ridge Regression

We propose PIGRR method as a solution to the problem caused by multicollinearity applied to count data on overdispersion case. Let $\widehat{\mathcal{B}}$ be any estimator of the parameter vector β. The difference between the log-likelihood functions assessed at ML estimate and $\widehat{\mathcal{B}}$

$$diff(\ell) = \ell_{max} - \ell\left(\widehat{\mathcal{B}}\right).$$

By following Segerstedt [44], we define the PIGRR as the solution of the optimization problem

$$\min \widehat{\mathcal{B}}^T\widehat{\mathcal{B}} \qquad \text{subject to} \qquad diff(\ell) = c_0$$

where c_0 is the fixed number and by developing the Lagrange function, we define

$$\begin{aligned} \mathcal{Q}\left(\widehat{\mathcal{B}}, k\right) &= \widehat{\mathcal{B}}^T\widehat{\mathcal{B}} - \left(\frac{2}{k}\right)\left[\sum_{i=1}^{n}\left\{\frac{1}{2}\log\left(\frac{2\alpha_i}{\pi}\right) + y_i \log\left(\exp\left(x_i^T\beta\right)\right) + \frac{1}{\tau}\right.\right. \\ &+ \left.\left. \log\left(K_{y_i - \frac{1}{2}}(\alpha_i)\right) - y_i \log\left(\frac{\alpha_i}{\tau}\right) - \log(y_i!)\right\} + (c_0 - \ell_{max})\right], \end{aligned}$$

where $\alpha_i = \frac{\sqrt{1+2\exp\left(x_i^T\beta\right)\tau}}{\tau}$, $\frac{2}{k}$ is the Lagrange multiplier and $\mathcal{Q}\left(\widehat{\mathcal{B}},k\right)$ is a $p\times 1$ vector with elements $\mathfrak{q}_j\left(\widehat{\mathcal{B}},k\right)$

$$\begin{aligned} \mathfrak{q}_j\left(\widehat{\mathcal{B}},k\right) &= \frac{\partial \mathcal{Q}\left(\widehat{\mathcal{B}},k\right)}{\partial \widehat{\mathcal{B}}_j} \\ &= 2\widehat{\mathcal{B}}_j - \left(\frac{2}{k}\right)\frac{\partial \ell\left(\widehat{\mathcal{B}}\right)}{\partial \widehat{\mathcal{B}}_j} \\ &= 2\widehat{\mathcal{B}}_j - \left(\frac{2}{k}\right)\sum_{i=1}^{n}\left[\left(\frac{y_i - \exp\left(x_i^T\beta\right)}{\exp\left(x_i^T\beta\right)\left[1+\tau\exp\left(x_i^T\beta\right)\right]}\right)\right. \\ &\quad \left.\times\ \frac{\partial \exp\left(x_i^T\beta\right)}{\partial \eta_i}x_{ij}\right], \quad j=1,2,\ldots,p. \end{aligned} \tag{3.1}$$

Now define $\mathcal{H}\left(\widehat{\mathcal{B}},k\right)$ is $p\times p$ matrix with elements $\mathfrak{h}_{jq}\left(\widehat{\mathcal{B}},k\right)$ and taking the second order derivates of $\mathfrak{q}_j\left(\widehat{\mathcal{B}},k\right)$ as

$$\mathfrak{h}_{jq}\left(\widehat{\mathcal{B}},k\right) = \frac{\partial^2 \mathcal{Q}\left(\widehat{\mathcal{B}},k\right)}{\partial \widehat{\mathcal{B}}_j\partial \widehat{\mathcal{B}}_q} = 2\delta_{jq} - \left(\frac{2}{k}\right)\frac{\partial^2 \ell\left(\widehat{\mathcal{B}}\right)}{\partial \widehat{\mathcal{B}}_j\partial \widehat{\mathcal{B}}_q}$$

Taking the expectation of both sides of Eq. (3.1) as

$$\begin{aligned} E\left[\frac{\partial^2 \mathcal{Q}\left(\widehat{\mathcal{B}},k\right)}{\partial \widehat{\mathcal{B}}_j\partial \widehat{\mathcal{B}}_q}\right] &= 2\left[\delta_{jq} - \left(\frac{1}{k}\right)E\left\{-\frac{\partial \ell\left(\widehat{\mathcal{B}}\right)}{\partial \widehat{\mathcal{B}}_j}\frac{\partial \ell\left(\widehat{\mathcal{B}}\right)}{\partial \widehat{\mathcal{B}}_q}\right\}\right], \\ &= 2\left[\delta_{jq} + \left(\frac{1}{k}\right)\sum_{i=1}^{n}\frac{x_{ij}x_{iq}}{\exp\left(x_i^T\beta\right)\left[1+\tau\exp\left(x_i^T\beta\right)\right]}\left\{\frac{\partial \exp\left(x_i^T\beta\right)}{\partial \eta_i}\right\}^2\right], \end{aligned} \tag{3.2}$$

where $\delta_{jq} = \begin{cases} 1 & \text{if } j=q \\ 0 & \text{otherwise} \end{cases}$. By means of the Fisher scoring method in this case yields

$$\mathcal{H}\left(\widehat{\mathcal{B}}^{(m)},k\right)\widehat{\beta}\left(k\right)^{(m+1)} = \mathcal{H}\left(\widehat{\mathcal{B}}^{(m)},k\right)\widehat{\beta}\left(k\right)^{(m)} + \mathcal{Q}\left(\widehat{\mathcal{B}}^{(m)},k\right), \tag{3.3}$$

where $\mathfrak{q}_j\left(\widehat{\mathcal{B}},k\right)$ and $\mathfrak{h}_{jq}\left(\widehat{\mathcal{B}},k\right)$ are the elements of the vector $\mathcal{H}\left(\widehat{\mathcal{B}}^{(m)},k\right)$ and matrix $\mathcal{Q}\left(\widehat{\mathcal{B}}^{(m)},k\right)$, respectively, and both can be assessed at the preliminary estimate $\widehat{\beta}\left(k\right)^{(m)}$. From Eqs (3.1-3.2), identifying that

$$\mathcal{Q}\left(\widehat{\mathcal{B}}^{(m)},k\right) = 2\widehat{\beta}\left(k\right)^{(m)} - \left(\frac{2}{k}\right)\left[X^T\widehat{W}\widehat{u} - \mathcal{D}\widehat{\beta}\left(k\right)^{(m)}\right].$$

$$\mathcal{H}\left(\widehat{\mathcal{B}}^{(m)}, k\right) = \left(-\frac{2}{k}\right)[\mathcal{D} + kI].$$

Now, the Eq.(3.3) is equivalent to

$$\begin{aligned}(\mathcal{D} + kI)\,\widehat{\beta}(k)^{(m+1)} &= \left(2\widehat{\beta}(k)^{(m)} - \left(\frac{2}{k}\right)\left[X^T W u - \mathcal{D}\widehat{\beta}(k)^{(m)}\right]\right.\\ &- \left.\left(\frac{2}{k}\right)(\mathcal{D} + kI)\,\widehat{\beta}(k)^{(m)}\right)\left(-\frac{k}{2}\right)(\mathcal{D} + kI)^{-1}\end{aligned}$$

$$(\mathcal{D} + kI)\,\widehat{\beta}(k)^{(m+1)} = X^T W u.$$

When the successive estimates $\widehat{\beta}(k)^{(m)}$ converges to $\widehat{\beta}(k)$ as $m \to \infty$, then we find the following PIGRR estimator

$$\widehat{\beta}_{PIGRR} = \widehat{\beta}(k) = (\mathcal{D}(k))^{-1} X^T W u, \tag{3.4}$$

where k is the ridge parameter and $\mathcal{D}(k) = (\mathcal{D} + kI)$.

The mean squared error (MSE) of $\widehat{\beta}_{PIGRR}$ equals

$$\begin{aligned}MSE\left(\widehat{\beta}_{PIGRR}\right) &\cong E\left[\left(\widehat{\beta}_{PIGRR} - \beta\right)^T\left(\widehat{\beta}_{PIGRR} - \beta\right)\right]\\ &\cong tr\left[(\mathcal{D}(k))^{-1}\mathcal{D}(\mathcal{D}(k))^{-1}\right] + k^2\beta^T(\mathcal{D}(k))^{-2}\beta\\ &\cong \sum_{j=1}^{p}\frac{\lambda_j}{(\lambda_j + k)^2} + \sum_{j=1}^{p}\frac{k^2\alpha_j^2}{(\lambda_j + k)^2}, \end{aligned}\tag{3.5}$$

where $\lambda_j(\lambda_1, \lambda_2, \ldots, \lambda_p) > 0$ is the eigenvalues of the matrix $\mathcal{D}$, α_j^2 is the jth element of $\mathcal{T}^T\widehat{\beta}_{ML}$ and $\mathcal{T}$ is the matrix whose columns are the eigenvectors of the matrix $\mathcal{D}$ such that $\mathcal{D} = \xi\Lambda\xi^T$, where $\Lambda = diag(\lambda_j)$. Månsson and Shukur [29] investigated the MSE properties of the Poisson ridge regression, and these properties are same for the PIGRR. Månsson and Shukur [29] showed that Poisson ridge regression is superior to the ML estimator in terms of scalar MSE. Similarly, the $\widehat{\beta}_{PIGRR}$ is also superior to the $\widehat{\beta}_{ML}$ when $k > 0$.

3.1. Proposed estimators of ridge parameter k

The performance of the PIGRR method depends on the choice of ridge parameter (k). Therefore, we suggest some new methods for estimating k. First, we derive the optimal value of k by taking the derivative of $MSE\left(\widehat{\beta}_{PIGRR}\right)$ with respect to k

$$\frac{\partial MSE\left(\widehat{\beta}_{PIGRR}\right)}{\partial k} = -2\sum_{j=1}^{p}\frac{\lambda_j\left(1 - k\alpha_j^2\right)}{(\lambda_j + k)^3}. \tag{3.6}$$

By equating Eq. (3.6) to zero, we find the value of k that minimizes $MSE\left(\widehat{\beta}_{PIGRR}\right)$ as

$$k_j = \frac{1}{\alpha_j^2}. \tag{3.7}$$

Different estimators of k recommended by Månsson and Shukur [29] for the Poisson regression and some of these estimators are generalized for the PIGRR:

$$\widehat{k}_1 = \frac{1}{\alpha_{max}^2}; \quad \widehat{k}_2 = \left(\prod_{j=1}^{p} \frac{1}{m_j}\right)^{\frac{1}{p}}; \quad \widehat{k}_3 = max\left(\frac{1}{m_j}\right),$$

where α_{max}^2 is the maximum value of α_j^2 and $m_j = (k)^{\frac{1}{2}}$. By following Qasim et al. [41], we suggest $\widehat{k}_4$ estimator for the PIGRR:

$$\widehat{k}_4 = \frac{p}{\sum_{j=1}^{p} \alpha_j^2}.$$

Now, we propose a quantile-based ridge estimation. Let $\mathcal{X}$ be a continuous random variable with a cumulative distribution function (cdf) $\mathcal{F}(x)$. Then the q^{th} quantile of a population is denoted as $\mathbb{Q}(\mathsf{q})$ and characterized the functional inverse of the cdf taken at q, such as

$$\mathbb{Q}(\mathsf{q}) = \mathcal{F}^{-1}(\mathsf{q}) = inf\left[\mathsf{x} : \mathcal{F}(\mathsf{x}) \geq \mathsf{q}\right] \tag{3.8}$$

for predetermined $0 < q < 1$. The Eq.(3.8) can also be defined as $P\left[\mathsf{x} \leq \mathbb{Q}(\mathsf{q})\right] = \mathsf{q}$, which indicates that $100q\%$ of the observations less than or equal to the population quantile $\mathbb{Q}(\mathsf{q})$.

By means of Eq. (3.7), we begin by setting $\mathcal{S}_1 = \left(L, k_{(1)}\right)$, $\mathcal{S}_2 = \left[\ k_{(1)}, k_{(2)}\right)$, $\ldots$, $\mathcal{S}_{p-1} = \left[\ k_{(p),U}\right)$, where $k_{(1)}, k_{(2)}, \ldots, k_{(p)}$ are order statistics of $k_j (j = 1, 2, \ldots, p)$, and L and U are lower and upper bounds of k (which can be > 0 and ∞, respectively). In addition, the q^{th} quantile of $\mathbb{Q}_\mathsf{q}$ lies in one of these intervals. As a result, the quantile-based ridge estimator can be computed as

$$P\left[k \leq \mathbb{Q}(\mathsf{q})\right] = \mathsf{q}, \qquad 0 < \mathsf{q} < 1.$$

The q^{th} quantile can also be expressed as $P\left[k > \mathbb{Q}(\mathsf{q})\right] = 1 - \mathsf{q}$. We set four levels ($\mathsf{q} = 0.25, 0.50, 0.75, 0.95$) of the quantiles and therefore, we propose four different quantile-based ridge estimators for the PIGRR, namely, $\mathbb{Q}_{0.25}, \mathbb{Q}_{0.50}, \mathbb{Q}_{0.75}$ and $\mathbb{Q}_{0.95}$. These ridge estimators are proposed in order to get the lowest MSE of the PIGRR.

3.2. Wald confidence interval

We first examine the coverage properties of Wald-type confidence intervals on the mean response function, $\mu_i = \exp\left(x_i^T \beta\right)$. Recall that in the LRM, a confidence interval on $E\left(y|x = x_0\right)$ with p parameter is

$$\widehat{y}(x_0) \pm t_{\frac{\alpha}{2}, n-p} \widehat{\sigma} \sqrt{x_0^T \left(X^T X\right)^{-1} x_0}.$$

Myers and Montgomery [33] provide an analogous illustration for the GLMs. The asymptotic $100\,(1-\alpha)\,\%$ confidence interval on the mean response function at the point x_0 as follows

$$\widehat{\mu}\,(x_0) \pm z_{\frac{\alpha}{2}} \sqrt{c_0^T \left(\mathcal{C}^T \widehat{W}^{-1} \mathcal{C}\right)^{-1} c_0}, \tag{3.9}$$

where $\widehat{\mu}\,(x_0)$ is the estimated mean response function at the point of x_0, $\mathcal{C}$ is the matrix of derivatives stimulated by the Taylor series expansion applied in the GLMs, $W = diag\,\{Var\,(y_i)\}$, where $Var\,(y_i) = \ddot{b}\,(\theta)\,a\,(\phi)$ from the exponential family of distribution, and c_0 is the vector of these derivatives. By following McCullagh and Nelder [31] and Myers et al. [34], the development of Eq. (3.9) is relatively simple. The PIG regression is the member of the exponential family of distribution. For a link, say, $g\,(\mu) = x^T\beta$, we have $\mu = g^{-1}\left(x^T\beta\right)$ and its estimated mean function $\widehat{\mu} = g^{-1}\left(x^T\widehat{\beta}\right)$. We know that μ is a nonlinear function of β. We make use of the delta method to approximate $Var\,[\widehat{\mu}\,(x_0)]$ for the confidence interval on $\mu\,(x_0)$. The delta method allows for approximation of the variance of a quantity that is a nonlinear function. Since $\widehat{\mu}\,(x_0)$ is a nonlinear function of parameter in estimates $\widehat{\beta}$ and by means of the delta method

$$Var\,(\widehat{\mu}\,(x_0)) = c_0^T \left\{Cov\left(\widehat{\beta}\right)\right\} c_0,$$

where $c_0 = \frac{\partial\widehat{\mu}(x_0)}{\partial\widehat{\beta}_{ML}}$ is a vector of derivatives and $Cov\left(\widehat{\beta}\right)$ is the asymptotic variance-covariance matrix of the ML estimator for the GLMs, such as $I\left(\widehat{\beta}\right)^{-1} = \left(\mathcal{C}^T W^{-1}\mathcal{C}\right)^{-1}$ and therefore, $Var\,(\widehat{\mu}\,(x_0)) = c_0^T$ is approximated by $c_0^T\left(\mathcal{C}^T W^{-1}\mathcal{C}\right)^{-1} c_0$. We define the pivotal quantity

$$z = \frac{\widehat{\mu}\,(x_0) - \mu\,(x_0)}{\left[c_0^T\left(\mathcal{C}^T W^{-1}\mathcal{C}\right)^{-1} c_0\right]^{\frac{1}{2}}}$$

has a normal distribution with mean 0 and variance 1. Hence the useful general form of the approximate $100(1-\alpha)\%$ confidence interval on $\mu\,(x_0)$ for the GLMs is $\widehat{\mu}\,(x_0) \pm z_{\frac{\alpha}{2}}\sqrt{c_0^T\left(\mathcal{C}^T\widehat{W}^{-1}\mathcal{C}\right)^{-1} c_0}$. Then, we consider the special case of the GLMs is PIG regression. By using the canonical link function $\mathcal{C} = \left(\frac{WX}{a(\phi)}\right)\left(\mathcal{C}^T W^{-1}\mathcal{C}\right)^{-1} = \left(X^T W^{-1} X\right)^{-1}\{a\,(\phi)\}^2$, $c_0 = \frac{\{Var(y_0)\}x_0}{a(\phi)}$, and finally the required confidence interval is defined as

$$\widehat{\mu}\,(x_0) \pm z_{\frac{\alpha}{2}}\widehat{Var}\,(y_0)\sqrt{x_0^T\left(X^T\widehat{W}^{-1}X\right)^{-1} x_0}.$$

Now we define the Wald confidence interval for $\mu_i = \exp\left(x_i^T\beta\right)$ by using the mean and variance functions of the PIG regression

$$\exp\left(x_i^T\widehat{\beta}_{ML}\right) \pm z_{\frac{\alpha}{2}}\left[\exp\left(x_i^T\widehat{\beta}_{ML}\right)\left\{1 + \widehat{\tau}\exp\left(x_i^T\widehat{\beta}_{ML}\right)\right\}\right]\sqrt{x_i^T\mathcal{D}^{-1}x_i}, \tag{3.10}$$

where $\mathcal{D} = X^T \widehat{W}^{-1} X$. By using the asymptotic variance-covariance matrix of the PIGRR, we define the Wald confidence interval for $\mu_i = \exp\left(x_i^T \beta\right)$

(3.11)

$$\exp\left(x_i^T \widehat{\beta}_{PIGRR}\right) \pm z_{\frac{\alpha}{2}} \left[\exp\left(x_i^T \widehat{\beta}_{PIGRR}\right)\left\{1 + \widehat{\tau}\exp\left(x_i^T \widehat{\beta}_{PIGRR}\right)\right\}\right] \sqrt{x_i^T \mathcal{D}(k)^{-1} \mathcal{D}\mathcal{D}(k)^{-1} x_i},$$

where $\mathcal{D}(k) = X^T \widehat{W}^{-1} X + kI$.

4. Monte Carlo simulation

In this section, we conduct a Monte Carlo simulation to demonstrate the performance of the proposed estimator compared to the ML estimator. In the following subsection, we present the simulation design.

4.1. Design of the simulation

The factors considered in the simulation to evaluate the performance of the proposed estimator include sample size (n), degree of correlation (ρ), and the number of explanatory variables (p). To investigate the impact of these factors, we set up a simulation design that comprises four different sample sizes: $50, 100, 150,$ and 200, with three distinct correlation coefficient values (ρ) of $0.90, 0.95,$ and 0.95. Moreover, we use three different values for the number of explanatory variables, namely $3, 5,$ and 7.

We generate the dependent variable of the PIG regression model as follows

$$y_i \sim PIG(\mu, \tau) \tag{4.1}$$

where $\mu = \exp\left(X^T \beta\right)$, $\beta = (\beta_1, \beta_2, \ldots, \beta_p)$ with $\sum_{i=1}^{p} \beta_j^2 = 1$. We use the following method, which was given by McDonald and Galarneau [32], to generate explanatory variables:

$$x_{ij} = \left(1 - \rho^2\right)^{\frac{1}{2}} z_{ij} + \rho z_{ip} \tag{4.2}$$

where z_{ij} are independent pseudo-random numbers following the standard normal distribution, ρ represents the correlation between the explanatory variables and $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, p$. To compare the performance of the estimators, we estimate the MSE, which is calculated based on 1000 simulation replications for various combinations of n, ρ, and p.

$$MSE\left(\widehat{\beta}\right) = \frac{\sum_{r=1}^{1000}\left(\widehat{\beta}_r - \beta\right)^T\left(\widehat{\beta}_r - \beta\right)}{1000} \tag{4.3}$$

where $\widehat{\beta}_r$ is the estimated value of coefficients in the r^{th} replication. Also, we obtain the average confidence lengths on the means response functions of ML estimator and PIGRR methods.

4.2. Results of the simulation study

Tables 1-2 present the results of simulation study, which aimed to investigate the impact of several factors on the accuracy of estimation. Specifically, we provide MSE values for different combinations of p, n, and ρ values. These factors include the number of explanatory variables (p), sample sizes (n), and correlation degrees (ρ), all of which can significantly affect the reliability of statistical analysis. In addition to the MSE values, we also report the average confidence lengths for the means response functions of ML and PIGRR methods. These measures provide insight into the precision and stability of the estimated response functions under varying conditions.

Table 1 presents several important findings from our simulation study. Firstly, the results show that PIGRR methods with all ridge parameters outperform the ML method in terms of MSE values, which confirms our initial expectation. Additionally, the MSE values of the PIGRR estimator with the ridge parameter $\mathbb{Q}_{0.95}$ are consistently smaller than the MSE values of other PIGRR estimators and ML across all conditions. Furthermore, our simulation results reveal that correlation degrees have a negative impact on the MSE values for fixed values of p and n. Specifically, as the degree of correlation increases, the MSE values of all estimators increase, indicating reduced accuracy and reliability of statistical inference. Conversely, increasing the sample size has a positive effect on the MSE values, as larger samples provide more reliable estimates and reduce sampling variability. It is also important to note that the MSE values of ML are approximately twice as large as those of the PIGRR method with $\mathbb{Q}_{0.95}$ across all simulation scenarios. This suggests that the PIGRR method is a more accurate and reliable estimator than the ML method, particularly in high-dimensional settings. Finally, we observe that increasing the number of explanatory variables also leads to an increase in MSE values when other factors are held constant. This highlights the importance of careful variable selection and regularization techniques in statistical modeling.

Table 2 presents the average confidence lengths for the ML and PIGRR methods with $\mathbb{Q}_{0.95}$. Similar to the MSE results, we observe several important findings related to the average confidence lengths. Firstly, the results show that the average confidence lengths of the PIGRR method with $\mathbb{Q}_{0.95}$ are generally shorter than those of the ML method, indicating a higher precision and more accurate inference. Moreover, as the degree of correlation increases, the average confidence lengths of both methods also increase, highlighting the impact of correlation on statistical inference. We also find that the number of explanatory variables has an adverse effect on the average confidence lengths. Specifically, as the number of explanatory variables increases from 3 to 7, the average confidence length of both methods also increases, suggesting that careful variable selection and regularization are important in minimizing the impact of overfitting. Additionally, the simulation results show that increasing the sample size generally leads to a decrease in the average confidence lengths for both methods, as larger samples provide more reliable estimates and reduce sampling variability. However, we note that there are some exceptions to this trend, as the average confidence lengths do not always decrease when the sample size changes from 150 to 200.

Overall, these findings provide important insights into the impact of various factors on the precision and accuracy of statistical inference and have important implications for researchers and practitioners in selecting appropriate methods and interpreting their results.

5. Application: Traffic fatalities

In this application, we model the traffic fatalities for 48 US states (excluding Alaska and Hawaii) during 1988. The data is taken from the AER package in R and it is denoted Fatalities. We fit PIG regression for the number of night-time vehicle fatalities of the age group 21-24 years old (y_i). We consider the following regressors: per capita personal income in 1987 dollars (X_1), spirits that measures the spirits consumption (X_2), population which is the population in the respective age group such as, population age group 21-24 years old (X_3), total population (X_4) and miles-total measured as the total vehicle miles in millions (X_5). We see the impact of per capita personal income, spirits consumption, population age group 21-24 years old, miles-total, and total population on the number of night-time vehicle fatalities. The square root transformation is used to find the positive predictive confidence interval values of the number of night-time vehicle fatalities. Our analysis aimed to understand how these variables are related to the number of night-time vehicle fatalities. The impact of per capita personal income, spirits consumption, population age group 21-24 years old, miles-total, and total population on the number of night-time vehicle fatalities was explored using PIG regression. Our computations are carried out by using R programing language. The problem of multicollinearity is tested by the condition index,which is calculated by taking the square root of the maximum eigenvalue divided by the minimum eigenvalue. CI $= \sqrt{\frac{\lambda_{max}}{\lambda_{min}}} = 95.12$, which indicate a strong multicollinearity in the dataset (for more details, please see [1], [10] and [12]). The eigenvalues are 2605.7, 92.5, 40.3, 11.6, 0.57 and 0.29.

The MSE of the conventional ML estimator is inflated in the presence of multicollinearity. Therefore, we used biased estimation methods to overcome the problem of multicollinearity. The benefit of our proposed estimator is shown by using this dataset in the PIG regression model. Table 3 presents the estimates of the parameters and scalar MSE of the ML and PIGRR estimators. Application results reveal that a substantial decrease of the MSE when applying PIGRR. This is in line with the simulated results where PIGRR always have a better performance than the ML in almost all considered conditions. Previous empirical results according to OECD (2015) is mixed but in general robustness as shown by our new estimator is desirable. The estimated coefficient using ML and PIGRR of per capita personal income is positive which shows the higher number of night-time vehicle fatalities. Furthermore, spirits are negative which is counterintuitive sine a higher consumption of alcohol should lead to more accidents on average. However, when using proposed PIGRR with quantile based ridge estimation the value of the coefficients is shrunken towards zero. Finally, total population is positive in age group 21-24 which according to OECD (2015) is in line with previous research. The values using PIGRR is shrunken towards zero but not as much as for income and spirits variables. Miles-total is actually positive, and it increases the number of night-time vehicle fatalities.

The results presented in Table 4 (appendix) demonstrate the Wald-type 95% confidence interval on the mean response function for each observation of the number of night-time vehicle fatalities using the ML and PIGRR with a shrinkage parameter k_2. It can be seen that in the presence of multicollinearity, the confidence intervals for the ML estimator are wider compared to those of PIGRR. This indicates that the PIGRR method is better at shrinking the coefficients towards zero and producing narrower confidence intervals. Moreover, the average 95% confidence length on the mean response function using the ML and PIGRR are 6.56 and 5.44, respectively, which further confirms the superiority of the PIGRR method in terms of producing narrower confidence intervals and coefficient shrinkage. These findings provide additional evidence that the proposed PIGRR method is more effective in reducing the impact of multicollinearity on the estimation of the mean response function. Furthermore, it is worth noting that the estimated MSE of the proposed PIGRR estimator is smaller than that of the traditional ML estimator, which highlights the advantage of using the PIGRR method in this context.

6. Conclusion

The Poisson inverse Gaussian ridge (PIGRR) regression and Wald-type confidence interval for the predictive response function are proposed and is compared to alternative ML estimator. Also, quantile-based ridge estimation ($\mathbb{Q}_{0.25}$, $\mathbb{Q}_{0.50}$, $\mathbb{Q}_{0.75}$, $\mathbb{Q}_{0.95}$) and shrinkage estimators(k_1, k_2, k_3, k_4) are suggested to obtain the optimal value of the PIGRR for the GLM. Results show that the PIGRR method outperforms the traditional ML estimator, providing narrower confidence intervals for the predictive response function. Monte Carlo simulations and an application to a traffic fatalities dataset further demonstrated the effectiveness of the proposed PIGRR estimator. In conclusion, the results of this study indicate that the PIGRR method can be used as an alternative estimator for the PIG regression and GLMs, especially in the presence of multicollinearity.

ACKNOWLEDGMENTS

The authors would like to thank the referees and the associate editor for their constructive comments that led to substantial improvement of the paper.

REFERENCES

[1] Abbasi, A. and Özkale, M. R. (2021). The r-k class estimator in generalized linear models applicable with simulation and empirical study using a Poisson and Gamma responses, *Hacettepe Journal of Mathematics and Statistics*, **50(2)**, 594–611.

[2] Abonazel, M. R., Dawoud, I., Awwad, F. A. and Eldin, E. T. (2023). New Estimators for the Probit Regression Model with Multicollinearity, *Scientific African*, **e01565**.

[3] Ahmed, E., Belaghi, R. A., Asar, Y. and Hussein, A. (2021). Ridge-Type Shrinkage Estimators in Low and High Dimensional Beta Regression Model with Application in Econometrics and Medicine, *arXiv*, **preprint arXiv:2111.13887**.

[4] Akram, M. N., Amin, M., Kibria, G. B. M., Arashi, M., Lukman, A. F. and Afzal, N. (2022). A new improved Liu estimator for the QSAR model with inverse Gaussian response, *Communications in Statistics-Simulation and Computation*, 1-16.

[5] Algamal, Z. Y. (2019). Performance of ridge estimator in inverse Gaussian regression model, *Communications in Statistics-Theory and Methods*, **48(15)**, 3836–3849.

[6] Algamal, Z. Y., Lukman, A. F., Abonazel, M. R. and Awwad, F. A. (2022). Performance of the Ridge and Liu Estimators in the zero-inflated Bell Regression Model, *Journal of Mathematics*.

[7] Alheety, M. I., Qasim, M., Månsson, K. and Kibria, G. B. M. (2021). Modified almost unbiased two-parameter estimator for the Poisson regression model with an application to accident data, *SORT-Statistics and Operations Research Transactions*, **45(2)**, 121–142.

[8] Awwad, F. A., Odeniyi, K. A., Dawoud, I., Algamal, Z., Abonazel, M. R., Kibria, G. B. M. and Eldin, E. T. (2022). New two-parameter estimators for the logistic regression model with multicollinearity, *WSEAS Transactions on Mathematics*, **21**, 403–414.

[9] Amin, M., Qasim, M. and Amanullah, M. (2019). Performance of Asar and Genç and Huang and Yang's two-parameter estimation methods for the gamma regression model, *Iranian Journal of Science and Technology, Transactions A: Science*, **43(6)**, 2951–2963.

[10] Amin, M., Qasim, M., Amanullah, M. and Afzal, S. (2020). Performance of some ridge estimators for the gamma regression model, *Statistical Papers*, **61(3)**, 997–1026.

[11] Asar, Y. and Genç, A. (2018). A new two-parameter estimator for the Poisson regression model, *Iranian Journal of Science and Technology, Transactions A: Science*, **42(2)**, 793–803.

[12] Bulut, Y. M. and Işilar, M. (2021). Two parameter ridge estimator in the inverse Gaussian regression model, *Hacettepe Journal of Mathematics and Statistics*, **50(3)**, 895–910.

[13] Bulut, Y. M. (2021). Inverse Gaussian liu-type estimator, *Communications in Statistics-Simulation and Computation*, 1–16.

[14] Dean, C., Lawless, J. F. and Willmot, G. E. (1989). A Mixed Poisson-Inverse-Gaussian Regression Model, *The Canadian Journal of Statistics*, **17(2)**, 171–181.

[15] Frisch, R. (1934). *Statistical Confluence Analysis by means of Complete Regression Systems*, University Institute of Economics, Oslo.

[16] Gómez-Deniz, E. and Calderín-Ojeda, E. (2018). Properties and applications of the Poisson-reciprocal inverse Gaussian Distribution, *Journal of Statistical Computation and Simulation*, **88(2)**, 269–289.

[17] Hilbe, J. M. (2014). *Modeling Count Data*, New York: Cambridge University Press.

[18] Hoerl, A.E. and Kennard, R.W. (1970a). Ridge regression: Biased estimation for nonorthogonal problems, *Technometrics*, **12(1)**, 55–67.

[19] Hoerl, A.E. and Kennard, R.W. (1970b). Ridge regression: applications to nonorthogonal problems, *Technometrics*, **12(1)**, 69–82.

[20] Kaçiranlar, S. and Dawoud, I. (2018). On the performance of the Poisson and the Negative Binomial ridge predictors, *Communications in Statistics-Simulation and Computation*, **47(6)**, 1751–1770.

[21] Khalaf, G., Månsson, K., Sjölander, P. and Shukur, G. (2014). A Tobit Ridge Regression Estimator, *Communications in Statistics-Theory and Methods*, **43**, 131–140.

[22] Kibria, G. B. M. (2022). More than hundred (100) estimators for estimating the shrinkage parameter in a linear and generalized linear ridge regression models, *Journal of Econometrics and Statistics*, **2(2)**, 233–252.

[23] Liu, K. (1993). A new class of biased estimate in linear regression, *Communications in Statistics-Theory and Methods*, **22(2)**, 393–402.

[24] Liu, K. (2003). Using Liu-type estimator to combat collinearity, *Communications in Statistics-Theory and Methods*, **32(5)**, 1009–1020.

[25] Locking, H., Månsson, K. and Shukur, G. (2013). Performance of some Ridge parameters for Probit regression: with application to Swedish job search data, *Communications in Statistics-Simulation and Computation*, **42**, 698–710.

[26] Lukman, A. F., Aladeitan, B. Ayinde, K. and Abonazel, M. R. (2022). Modified ridge-type for the Poisson regression model: simulation and application, *Communications in Statistics-Simulation and Computation*, **49(8)**, 2124–2136.

[27] Lukman, A. F., Ayinde, K., Kibria, G. B. M. and Adewuyi, E. T. (2022). Modified ridge-type estimator for the gamma regression model, *Communications in Statistics-Simulation and Computation*, **51(9)**, 5009–5023.

[28] Mahmoudi, A., Arabi Belaghi, R. and Mandal, S. (2020). A comparison of preliminary test, Stein-type and penalty estimators in gamma regression model, *Journal of Statistical Computation and Simulation*, **90(17)**, 3051–3079.

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

[30] Månsson, K. (2012). On ridge estimators for the negative binomial regression model, *Economic Modelling*, **29(2)**, 178–184.

[31] McCullagh, P. and Nelder, J. (1989). *Generalized Liner Models*, London: Chapman and Hall.

[32] McDonald, G. C. and Galarneau, D. I. (1975). A Monte Carlo evaluation of some ridge-type estimators, *Journal of the American Statistical Association*, **70(350)**, 407–416.

[33] Myers, R. H. and Montgomery, D. C. (1997). A Tutorial on Generalized Linear Models, *Journal of Quality Technology*, **29(3)**, 274–291.

[34] Myers, R. H., Montgomery, D. C. and Vining, G. G. (2002). *Generalized linear models: with applications in engineering and the sciences*, New York: J. Wiley.

[35] Nooi Asl, M. Bevrani, H., Arabi Belaghi, R.and Månsson, K. (2021). Ridge-type shrinkage estimators in generalized linear models with an application to prostate cancer data, *Statistical papers*, **62**, 1043–1085.

[36] Omer, T., Sjölander, P., Månsson, K. and Kibria, G. B. M. (2021). Improved estimators for the zero-inflated Poisson regression model in the presence of multicollinearity: simulation and application of maternal death data, *Communications in Statistics: Case Studies, Data Analysis and Applications*, **7(3)**, 394–412.

[37] Özkale, M. R. (2019). The r-d class estimator in generalized linear models: applications on gamma, Poisson and binomial distributed responses, *Journal of Statistical Computation and Simulation*, **89(4)**, 615–640.

[38] Özkale, M. R. and Kaçiranlar, S. (2007). The restricted and unrestricted two-parameter estimators, *Communications in Statistics-Theory and Methods*, **36(15)**, 2707–2725.

[39] Putri, G. N., Nurrohmah, S. and Fithriani, I. (2020). Comparing Poisson-Inverse Gaussian Model and Negative Binomial Model on case study: Horseshoe crabs data, *Journal of Physics: Conference Series*, **1442**, 1–5.

[40] Qasim, M., Månsson, K., Amin, M., Kibria, G. B. M. and Sjölander, P. (2020a). Biased adjusted Poisson ridge estimators-method and application, *Iranian Journal of Science and Technology, Transactions A: Science*, **44(6)**, 1775–1789.

[41] Qasim, M., Månsson, K. and Kibria, G. B. M. (2021). On some beta ridge regression estimators: method, simulation and applications, *Journal of Statistical Computation and Simulation*, **91(9)**, 1699–1712.

[42] Qasim, M., Månsson, K., Sjölander, P. and Kibria, G. B. M. (2022). A new class of efficient and debiased two-step shrinkage estimators: method and application, *Journal of Applied Statistics*, **49(16)**, 4181–4205.

[43] Schaefer, R. L., Roi, L. D. and Wolfe, R. A. (1984). A ridge logistic estimator, *Communications in Statistics-Theory and Methods*, **13**, 99–113.

[44] Segerstedt, B. (1992). On ordinary ridge regression in generalized linear models, *Communications in Statistics-Theory and Methods*, **21(8)**, 2227–2246.

[45] Stein, G. Z., Zucchini, W. and Juritz, J. M. (1987). Parameter Estimation for the Sichel Distribution and its Multivariate Extension, *Journal of the American Statistical Association*, **82(399)**, 938–944.

[46] Toker, S., Şiray, G. Ü., and Qasim, M. (2019). Developing a first order two parameter estimator for generalized linear model, *In 11th International Statistics Congress.*

[47] Toker, S., Özbay, N., Şiray, G. Ü. and Yenilmez, ?. (2021). Tobit Liu estimation of censored regression model: an application to Mroz data and a Monte Carlo simulation study, *Journal of Statistical Computation and Simulation*, **91(6)**, 1061–1091.

[48] Willmot, G. E. (1987). The Poisson-inverse Gaussian distribution as an alternative to the negative binomial, *Scandinavian Actuarial Journal*, **3-4**, 113–127.

[49] Zha, L., Lord, D. and Zou, Y. (2016). The Poisson inverse Gaussian (PIG) generalized linear regression model for analyzing motor vehicle crash data, *Journal of Transportation Safety and Security*, **8(1)**, 18–35.

A. Appendix

Table 1: MSE values of ML and PIGRR estimators with different ridge parameters

	n	ρ	ML	PIGRR							
				k_1	k_2	k_3	k_4	$\mathbb{Q}_{0.25}$	$\mathbb{Q}_{0.50}$	$\mathbb{Q}_{0.75}$	$\mathbb{Q}_{0.95}$
		0.90	3.9629	3.6578	3.7991	3.5100	3.4228	3.3892	3.1003	2.5715	**2.3345**
	50	0.95	5.3212	4.8090	4.6749	3.9837	4.3833	4.2103	3.7195	3.0916	**2.8166**
		0.99	16.4580	14.2634	8.6817	7.3257	12.0915	10.3336	7.9850	6.8134	**6.2833**
		0.90	3.2531	3.0851	3.2157	3.1278	3.0607	3.0814	2.9287	2.4258	**2.2284**
	100	0.95	4.0130	3.7358	3.8521	3.6021	3.5256	3.5081	3.2874	2.8365	**2.6232**
p=3		0.99	9.8350	8.6379	6.4043	5.3192	7.5223	6.8228	5.6322	4.9925	**4.7043**
		0.90	3.1694	3.0541	3.1495	3.0964	2.9636	2.9847	2.8415	2.4136	**2.2074**
	150	0.95	3.6547	3.4477	3.5705	3.4123	3.2946	3.2951	3.1528	2.6992	**2.4918**
		0.99	7.3670	6.5738	5.5960	4.6865	5.8786	5.4267	4.7384	4.2610	**4.0352**
		0.90	3.1255	3.0402	3.1128	3.0654	2.9555	2.9785	2.8237	2.3700	**2.1659**
	200	0.95	3.5378	3.3804	3.4881	3.3811	3.2627	3.2783	3.1339	2.6926	**2.4826**
		0.99	6.3801	5.7209	5.1625	4.3956	5.1791	4.8949	4.4018	3.9482	**3.7147**
		0.90	5.9567	5.4809	5.6316	4.5465	4.8643	4.7143	4.0507	3.2758	**2.3346**
	50	0.95	9.7346	8.7749	8.0036	5.4935	7.4877	7.0213	5.7144	4.5515	**3.0213**
		0.99	42.3171	37.2487	14.5175	10.6302	29.5380	23.9425	15.8654	11.3122	**8.6180**
		0.90	4.2455	4.0043	4.2025	3.9079	3.7171	3.7109	3.4010	2.8741	**2.1255**
	100	0.95	5.9535	5.4818	5.6709	4.6344	4.8770	4.7407	4.1433	3.5055	**2.5924**
p=5		0.99	20.7169	18.2691	10.4827	7.1336	14.6964	12.5546	9.1581	7.2057	**5.8155**
		0.90	3.7633	3.6038	3.7489	3.5885	3.4036	3.4129	3.1439	2.7543	**2.0351**
	150	0.95	5.0467	4.7354	4.9369	4.3738	4.3428	4.2698	3.8872	3.3723	**2.5377**
		0.99	15.1806	13.5483	9.4951	6.5243	11.1794	9.8631	7.6467	6.3033	**5.1333**
		0.90	3.6360	3.5172	3.6278	3.5229	3.3640	3.3716	3.1369	2.6800	**1.9944**
	200	0.95	4.5471	4.3075	4.4927	4.1243	4.0071	3.9783	3.6564	3.1451	**2.3418**
		0.99	11.8073	10.5958	8.6331	5.8646	8.8742	7.9941	6.4428	5.3743	**4.2514**
		0.90	9.1405	8.4541	8.5601	6.0465	7.2886	7.0655	5.8112	4.1206	**2.3914**
	50	0.95	15.6450	14.2603	12.2415	6.9601	11.6641	10.6449	7.9813	5.6598	**3.4269**
		0.99	77.3166	69.2965	20.3857	14.8155	53.4698	44.5839	28.7129	17.5815	**11.8956**
		0.90	5.5511	5.2208	5.5038	4.7570	4.6679	4.6057	4.0392	3.0931	**1.9922**
	100	0.95	8.6180	7.9574	8.1922	5.8951	6.8185	6.5105	5.4209	4.1414	**2.7296**
p=7		0.99	35.9445	32.2017	15.3524	8.6824	24.9931	22.1400	14.8486	9.5913	**6.4953**
		0.90	4.6103	4.3901	4.5963	4.2505	4.0346	4.0119	3.6539	2.8875	**1.8452**
	150	0.95	6.5878	6.1419	6.4549	5.2634	5.4149	5.2693	4.5961	3.6723	**2.3812**
		0.99	22.3684	20.0214	13.8088	7.5938	15.7448	14.5071	10.2449	7.2399	**5.2942**
		0.90	4.4384	4.2867	4.4326	4.1912	4.0171	4.0027	3.6143	2.8580	**1.8242**
	200	0.95	6.1816	5.8494	6.1113	5.1619	5.2816	5.1938	4.5491	3.6021	**2.3417**
		0.99	18.6922	16.8889	12.5442	7.5197	13.6177	12.4779	9.3996	7.0053	**5.0791**

Table 2: Comparison of 95% average confidence lengths on the means response functions of ML and PIGRR estimators

	ML				PIGRR[$\mathbb{Q}_{0.95}$]			
	n				n			
Degree of collinearity	50	100	150	200	50	100	150	200
				p=3				
0.90	139.84	48.42	40.63	58.54	26.68	14.84	13.01	11.00
0.95	239.95	65.23	46.57	63.21	40.92	21.44	14.54	12.19
0.99	273.31	66.00	59.94	68.10	194.36	44.19	21.76	18.22
				p=5				
0.90	1613.70	227.33	309.02	269.58	423.15	27.24	30.05	26.33
0.95	2677.17	468.77	317.27	330.86	955.23	67.50	31.49	29.73
0.99	14011.06	529.45	605.12	482.90	963.27	151.18	105.99	37.39
				p=7				
0.90	19381.12	1554.10	626.95	965.84	453.31	41.55	42.09	18.07
0.95	25894.30	1563.55	716.83	1143.08	718.47	122.27	37.51	50.53
0.99	29457.73	3702.62	10852.13	2389.50	4633.72	600.87	175.96	258.32

Table 3: Estimated coefficient and MSE of estimators

	Intercept	X_1	X_2	X_3	X_4	X_5	
Estimators	β_0	β_1	β_2	β_3	β_4	β_5	MSE
ML	1.644	0.117	−0.126	−0.628	0.777	0.130	5.349
PIGRR							
k_1	1.641	0.097	−0.094	−0.220	0.350	0.154	2.111
k_2	1.639	0.094	−0.092	−0.194	0.326	0.154	2.063
k_3	1.617	0.066	−0.075	−0.058	0.201	0.155	2.151
k_4	1.615	0.063	−0.074	−0.051	0.195	0.155	2.174
$\mathbb{Q}_{0.25}$	1.615	0.063	−0.074	−0.051	0.195	0.155	2.174
$\mathbb{Q}_{0.50}$	1.564	0.017	−0.058	0.015	0.148	0.160	2.437
$\mathbb{Q}_{0.75}$	1.504	−0.025	−0.049	0.044	0.136	0.166	2.577
$\mathbb{Q}_{0.95}$	1.145	−0.128	−0.048	0.109	0.148	0.193	2.864

Note: The explanatory variables are defined as: the per capita personal income in 1987 dollars (X_1), spirits that measures the spirits consumption (X_2), population which is the population in the respective age group such as, population age group 21-24 years old (X_3), total population (X_4) and miles-total measured as the total vehicle miles in millions (X_5)

Table 4: Comparison of 95%Confidence Interval (CI) on the means response function using ML and PIGRR estimators

Observation	MLE		PIGRR		Length of 95% CI	
	μ_i	95% CI	μ_i	95% CI	MLE	PIGRR
1	4.84	(2.86, 6.82)	4.84	(3.08, 66.60)	3.96	3.52
2	4.66	(3.20, 6.12)	4.67	(3.42, 5.92)	2.91	2.50
3	4.47	(2.57, 6.37)	4.36	(2.58, 6.15)	3.81	3.56
4	20.30	(14.10, 26.51)	20.45	(14.25, 26.64)	12.41	12.39
5	4.76	(2.81, 6.71)	4.73	(2.88, 6.57)	3.90	3.69
6	5.78	(0.00, 14.00)	5.54	(0.00, 13.13)	14.00	13.13
7	3.85	(1.12, 6.59)	3.91	(1.35, 6.47)	5.47	5.12
8	8.34	(4.12, 12.57)	7.77	(4.98, 10.55)	8.45	5.57
9	5.22	(2.82, 7.63)	5.47	(3.94, 7.00)	4.81	3.06
10	4.18	(2.63, 5.73)	4.07	(2.64, 5.51)	3.10	2.87
11	7.47	(0.32, 15.27)	7.34	(2.80, 11.89)	14.95	9.09
12	5.53	(3.68, 7.38)	5.51	(3.97, 7.04)	3.70	3.07
13	5.09	(2.68, 7.50)	4.87	(2.74, 7.00)	4.82	4.26
14	5.15	(2.24, 8.06)	4.91	(2.38, 7.44)	5.82	5.06
15	4.63	(3.00, 6.26)	4.65	(3.04, 6.25)	3.27	3.21
16	4.37	(1.94, 6.81)	4.51	(2.42, 6.60)	4.87	4.18
17	4.01	(2.63, 5.38)	4.02	(2.73, 5.31)	2.75	2.59
18	5.31	(0.83, 9.79)	5.35	(1.19, 9.52)	8.96	8.34
19	5.63	(0.50, 11.76)	5.72	(0.22, 11.23)	11.26	11.01
20	6.50	(4.28, 8.71)	6.62	(4.83, 8.40)	4.43	3.57
21	5.11	(3.05, 7.17)	5.05	(3.09, 7.01)	4.12	3.92
22	4.02	(1.68, 6.35)	4.08	(1.82, 6.34)	4.67	4.52
23	5.65	(3.59, 7.70)	5.47	(3.75, 7.19)	4.11	3.44
24	3.87	(2.60, 5.14)	3.85	(2.61, 5.09)	2.54	2.48
25	4.52	(2.91, 6.14)	4.40	(2.90, 5.90)	3.24	3.00
26	2.83	(0.00, 7.41)	3.12	(0.00, 7.47)	7.41	7.47
27	3.12	(0.00, 7.88)	3.39	(0.00, 7.90)	7.88	7.90
28	7.44	(0.00, 16.28)	7.05	(0.94, 15.04)	16.28	14.10
29	4.10	(2.60, 5.60)	4.08	(2.63, 5.52)	3.01	2.89
30	9.91	(0.00, 33.90)	9.75	(0.00, 22.63)	33.90	22.63
31	5.45	(3.85, 7.05)	5.56	(4.21, 6.91)	3.20	2.71
32	3.63	(2.10, 5.16)	3.67	(2.24, 5.10)	3.06	2.86
33	7.82	(3.51, 12.13)	7.48	(4.44, 10.51)	8.62	6.07
34	4.93	(2.67, 7.19)	4.80	(3.04, 6.57)	4.52	3.53
35	4.84	(3.20, 6.48)	4.67	(3.31, 6.04)	3.28	2.74
36	8.31	(0.26, 16.37)	7.81	(3.33, 12.29)	16.10	8.96
37	4.09	(1.77, 6.40)	4.07	(1.95, 6.19)	4.63	4.23
38	4.06	(1.94, 6.18)	4.25	(2.38, 6.11)	4.24	3.73
39	3.80	(2.49, 5.11)	3.80	(2.52, 5.07)	2.62	2.55
40	5.24	(3.58, 6.90)	5.19	(3.68, 6.70)	3.32	3.02
41	10.63	(2.53, 18.72)	10.77	(5.06, 16.47)	16.18	11.41
42	4.38	(2.37, 6.39)	4.28	(2.37, 6.20)	4.03	3.84
43	3.85	(2.38, 5.33)	3.87	(2.50, 5.25)	2.94	2.75
44	6.01	(1.60, 10.42)	6.11	(2.40, 9.82)	8.82	7.42