ISSN: 2094-0343
2326-9865

Jackknifed K-L estimator in Bell regression model

Luay Adil Abduljabbar [1]

Ministry of Higher Education and Scientific Research, Baghdad, Iraq

E-mail: Luay.yahya1989@gmail.com

Zakariya Yahya Algamal [2]

Department of Statistics and Informatics, University of Mosul, Mosul, Iraq

E-mail: zakariya.algamal@uomosul.edu.iq

ORCID: 0000-0002-0229-7958

Article Info
Page Number: 267-278
Publication Issue:
Vol 71 No. 3s2 (2022)

Abstract
When there is collinearity between the response variable and numerous explanatory variables, modeling the link between the response variable and several explanatory variables is difficult. Several shrinkage estimators have traditionally been presented to avoid this problem. The Kibria and Lukman estimator is one of them (K-L). In this paper, a jackknifed version of the K-L estimator in the Bell regression model is proposed, which combines the Jackknife process with the K-L estimator to reduce biasedness. In terms of absolute bias and mean squared error, our Monte Carlo simulation findings and real-world application of the Bell regression model imply that the suggested estimate can provide significant improvements over current competing estimators.

Keywords: Collinearity; K-L estimator; Bell regression model; Jackknife estimator; Monte Carlo simulation.

Article History
Article Received: 28 April 2022
Revised: 15 May 2022
Accepted: 20 June 2022
Publication: 21 July 2022

1. Introduction

Because it describes the relationship between the response variable of interest and a variety of explanatory variables, statistical modeling is significant in many scientific research domains. The response variable in a linear regression model is assumed to have a normal distribution. This assumption, however, may not hold true in many real-world applications. The response variable in medical sciences, for example, can be favorably count. As a result, employing a linear regression model may not be appropriate. The generalized linear model (GLM) is a type of regression model that is becoming increasingly popular as a statistical modeling tool for both continuous and discrete response variables. (Algamal, 2018).

In real applications, the design data matrix $\mathbf{X}$ has multicollinearity between explanatory variables, and, therefore, $\mathbf{X}^T\mathbf{X}$ is singular or can be inflating the variance of the maximum likelihood

ISSN: 2094-0343
2326-9865

estimator (MLE). Therefore, the traditional estimation methods, such as MLE, tend to perform poorly. The ridge, Liu, Liu-type, and others estimator that given by several authors is an alternative to MLE to overcome the multicollinearity in linear regression model (Hoerl & Kennard, 1970; K. Liu, 1993). These estimators have been extended to the GLMs (Akram, Amin, & Amanullah, 2020; B. M. G. Kibria, 2003; G. Kibria, Månsson, & Shukur, 2012; Kurtoğlu & Özkale, 2016; Mackinnon & Puterman, 1989; Månsson & Shukur, 2011; Nyquist, 1991; Segerstedt, 1992; Shamany, Alobaidi, & Algamal, 2019).

Although the powerful of these shrinkage estimators, but they have a smaller bias. It is possible to reduce bias by applying a jackknife procedure to these estimators. This procedure enables processing of experimental data to get statistical estimator for unknown parameters. The advantage of the jackknife procedure is that it presents an estimator that has a small bias while still providing beneficial properties of large samples (Alkhateeb & Algamal, 2020; Mansi Khurana, Yogendra P. Chaubey, & Shalini Chandra, 2014; Özkale & Arıcan, 2018).

The main objective given in this paper is to use Jackknife approach with the new ridge-type estimator (K-L estimator) of Kibria and Lukman (2020). Our proposed estimator will efficiently help to decrease the biasness of K-L estimator in Bell regression model. The superiority of our proposed estimator in different simulated examples and a real data application is proved.

2. K-L estimator in Bell regression model

Assume that $(y_i, \mathbf{x}_i)$, $i = 1, 2, ..., n$ is independent observed data with the predictor vector $\mathbf{x}_i \in R^{p+1}$ and the response variable $y_i \in R$ which follows a distribution that belongs to the Bell distribution. Then, the density function of y_i can be expressed as

$$P(Y = y) = \frac{\theta^y e^{-e^\theta + 1} B_y}{y!}, \; y = 0, 1, 2, ..., \tag{1}$$

where $\theta > 0$ and $B_y = (1/e)\sum_{d=0}^{\infty}(d^y / d!)$ is the Bell numbers (Eric T Bell, 1934; Eric Temple Bell, 1934; Castellares, Ferrari, & Lemonte, 2018). The mean and variance of the Bell distribution are respectively defined by

$$E(y) = \theta e^\theta, \tag{2}$$

$$Var(y) = \theta(1+\theta)e^\theta. \tag{3}$$

Assuming $\psi = \theta e^\theta$ and $\theta = W_\circ(\psi)$ where $W_\circ(.)$ is the Lambert function. Then Eq. (1) can be written in the new parameterization as

ISSN: 2094-0343
2326-9865

$$P\left(Y = y\right) = \exp\left(1 - e^{W_{\circ}(\psi)}\right)\frac{W_{\circ}\left(\psi\right)^{y}\ B_{y}}{y\,!},\ \ y = 0,1,2,..., \tag{4}$$

In GLM, the mean of the response variable, $\mu_i = E\,(y_i)$, is conditionally related to a linear function of predictors through a link function. The linear function is stated as $\eta_i = \beta_0 + \sum_{j=1}^{p} x_{ij}\beta_j = \mathbf{x}_i^T\boldsymbol{\beta}$ with $\mathbf{x}_i^T = (1, x_{i2}, x_{i3}, ..., x_{ip})$ and $\boldsymbol{\beta} = (\beta_0, \beta_1, ..., \beta_p)^T$. The link function is providing the relation of the mean and the natural parameter as $\mu_i = g^{-1}(\eta_i) = g^{-1}(\mathbf{x}_i^T\boldsymbol{\beta})$. The Bell regression model (BRM) can be modeled by assuming $\psi_i = \exp\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)\exp\left(\exp\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)\right)$ and $\log\psi_i = \mathbf{x}_i^T\boldsymbol{\beta}\,\exp\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)$ as

$y_i \sim \mathrm{Bell}\left(W_{\circ}\left(\psi_i\right)\right)$. The parameter estimation in the BRM is achieved through using the MLE based on the iteratively reweighted least-squares algorithm. The log-likelihood is defined

$$\begin{aligned}\ell(\boldsymbol{\beta},\psi) = {} & \sum_{i=1}^{n} y_i\,\log\left(\exp\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)\exp\left(e^{\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)}\right)\right) + \sum_{i=1}^{n}\left(1 - e^{e^{\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)e^{e^{\left(\mathbf{x}_i^T\boldsymbol{\beta}\right)}}}}\right) \\ & + \log B_y - \log\left(\prod_{i=1}^{n} y_i\,!\right).\end{aligned} \tag{5}$$

Then, the MLE is derived by equaling the first derivative of Eq. (5) to zero. This derivative cannot be solved analytically because it is nonlinear in $\boldsymbol{\beta}$. Fisher-scoring algorithm can be used to obtain the MLE where in each iteration, the parameter is updated by

$$\boldsymbol{\beta}^{(r+1)} = \boldsymbol{\beta}^{(r)} + I^{-1}(\boldsymbol{\beta}^{(r)})S\,(\boldsymbol{\beta}^{(r)}), \tag{6}$$

where $I^{-1}(\boldsymbol{\beta}) = \left(-E\,\left(\partial^2\ell(\boldsymbol{\beta},\phi)/\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^T\right)\right)^{-1}$. After that, the estimated coefficients are defined as

$$\hat{\boldsymbol{\beta}}_{\mathrm{MLE}} = (\mathbf{X}^T\,\hat{\mathbf{W}}\mathbf{X})^{-1}\mathbf{X}^T\,\hat{\mathbf{W}}\hat{\mathbf{u}}, \tag{7}$$

where $\hat{\mathbf{W}} = \mathrm{diag}\left[(\partial\mu_i/\partial\eta_i)^2/V\,(y_i)\right]$ and $\hat{\mathbf{u}}$ is a vector where i^th^ element equals to $\hat{u}_i = \log\hat{\psi}_i + [(y_i - \hat{\mu}_i)/\sqrt{\mathrm{var}(\hat{\psi}_i)}]$. The MLE is distributed asymptotically normal with a covariance matrix as

$$\mathrm{cov}(\hat{\boldsymbol{\beta}}_{\mathrm{MLE}}) = \left[-E\left(\frac{\partial^2\ell(\boldsymbol{\beta},\phi)}{\partial\boldsymbol{\beta}\,\partial\boldsymbol{\beta}^T}\right)\right]^{-1} = (\mathbf{X}^T\,\hat{\mathbf{W}}\mathbf{X})^{-1}. \tag{8}$$

In the presence of multicollinearity, the $rank\,(\mathbf{X}^T\,\hat{\mathbf{W}}\mathbf{X}) \le rank\,(\mathbf{X})$, and, therefore, the near singularity of $\mathbf{X}^T\,\hat{\mathbf{W}}\mathbf{X}$ makes the estimation unstable and enlarges the variance (G. Liu & Piantadosi, 2016). The ridge estimator (RE) (Hoerl & Kennard, 1970), Liu estimator (K. Liu, 1993) have been consistently demonstrated to be an attractive and alternative to the MLE, when the multicollinearity exists. In in Bell regression model, the ridge estimator and Liu estimator have been proposed by, respectively, Amin, Akram, and Majid (2021) and Majid, Amin, and Akram (2021).

3. Jackknifing K-L estimator in Bell regression model

In 2020, Kibria and Lukman proposed a new ridge-type estimator for the linear regression model. This proposed estimator is called as Kibria-Lukman (KL) estimator, which is defined as (B. M. Golam Kibria & Lukman, 2020):

$$\hat{\boldsymbol{\beta}}_{\text{KL}} = \left(\mathrm{I} + k\,(\mathbf{X}^T\mathbf{X})^{-1}\right)^{-1}\left(\mathrm{I} - k\,(\mathbf{X}^T\mathbf{X})^{-1}\right)\left(\mathbf{X}^T\mathbf{X}\right)^{-1}\mathbf{X}^T\,\mathbf{y}, \qquad (9)$$

where $k > 0$ is the shrinkage parameter. The estimator $\hat{\boldsymbol{\beta}}_{\text{KL}}$ is biased but more stable and has less mean square error than the ordinary least square estimator. For the BRM, Eq. (9), $\hat{\boldsymbol{\beta}}_{\text{KL-BRM}}$, can be defined as (Adewale F. Lukman, Algamal, Kibria, & Ayinde, 2021; A. F. Lukman, Dawoud, Kibria, Algamal, & Aladeitan, 2021)

$$\hat{\boldsymbol{\beta}}_{\text{KL-BRM}} = \left(\mathbf{I} + k\,(\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X})^{-1}\right)^{-1}\left(\mathbf{I} - k\,(\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X})^{-1}\right)(\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X})^{-1}\mathbf{X}^T\hat{\mathbf{W}}\hat{\mathbf{u}}. \qquad (10)$$

The bias and variance of Eq. (10) are defined as, respectively,

$$\text{Bias}(\hat{\boldsymbol{\beta}}_{\text{KL-BRM}}) = -2k\,\mathbf{Q}(\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X} + k\,\mathbf{I})^{-1}\alpha \qquad (11)$$

$$\begin{aligned}\text{Variance}(\hat{\boldsymbol{\beta}}_{\text{KL-BRM}}) = {} & \mathbf{Q}(\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X} + k\,\mathbf{I})^{-1}(\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X} - k\,\mathbf{I})^{-1}(\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X})^{-1} \\ & (\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X} + k\,\mathbf{I})^{-1}(\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X} - k\,\mathbf{I})^{-1}\mathbf{Q}^T,\end{aligned} \qquad (12)$$

where $\mathbf{Q} = (q_1, q_2, ..., q_p)$ represents the matrix of eigenvectors of the $\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X}$ matrix, and $\alpha = \mathbf{Q}^T\boldsymbol{\beta}$. In simple way, the mean square error (MSE) of Eq. (10) can be written as

$$\text{MSE}(\hat{\boldsymbol{\beta}}_{\text{KL-BRM}}) = \sum_{j=1}^{p}\frac{(\lambda_j - k)^2}{\lambda_j(\lambda_j + k)^2} + 4k^2\sum_{j=1}^{p}\frac{\alpha_j^2}{(\lambda_j + k)^2}. \qquad (13)$$

Shrinkage estimators are biased estimators. In linear regression model, Singh, Chaubey, and Dwivedi (1986) proposed the Jackknife procedure to alleviate the problem of bias in generalized ridge estimator. The theoretical and application of the jackknife estimator have been studied by several authors (Akdeniz Duran & Akdeniz, 2012; Alkhateeb & Algamal, 2020; Gruber, 1991; Mansi Khurana, Yogendra P Chaubey, & Shalini Chandra, 2014; Nyquist, 1988; Özkale & Arıcan, 2018; Türkan & Özel, 2015; Yıldız, 2017).

The proposed estimator, Jackknifed K-L estimator (JKL-BRM), in BRM can be expressed and derived. Let $\mathbf{\Lambda} = \text{diag}(\lambda_1, \lambda_2, ..., \lambda_p)$ is the matrix of eigenvalues of the $\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X}$ matrix, such that

ISSN: 2094-0343
2326-9865

$\mathbf{Q}^T\mathbf{X}^T\hat{\mathbf{W}}\mathbf{X}\mathbf{Q}=\mathbf{M}^T\hat{\mathbf{W}}\mathbf{M}=\boldsymbol{\Lambda}$, where $\mathbf{M}=\mathbf{X}\mathbf{Q}$. Consequently, the MLE estimator of Eq. (7) can be re-written as

$$\hat{\boldsymbol{\beta}}_{\text{MLE}}=\mathbf{Q}\,\hat{\boldsymbol{\upsilon}}_{\text{MLE}}, \tag{14}$$

where $\hat{\boldsymbol{\upsilon}}_{\text{MLE}}=\boldsymbol{\Lambda}^{-1}\mathbf{M}^T\hat{\mathbf{W}}\hat{\mathbf{u}}$. As a result, the KL-BRM estimator of Eq. (10) is re-written as

$$\hat{\boldsymbol{\upsilon}}_{\text{KL-GLM}}=(\boldsymbol{\Lambda}+k\,\mathbf{I})^{-1}(\boldsymbol{\Lambda}-k\,\mathbf{I})^{-1}\mathbf{M}^T\hat{\mathbf{W}}\hat{\mathbf{u}}. \tag{15}$$

Following the idea of Jackknife approach (Hinkley, 1977), let $\mathbf{u}_{(-i)}$, $\mathbf{m}_{(-i)}$, and $\mathbf{W}_{(-i)}$, respectively, are the i$^{\text{th}}$ row deleted from the vector $\mathbf{u}$, the i$^{\text{th}}$ row deleted from the matrix $\mathbf{M}$, and the i$^{\text{th}}$ row and column deleted from the matrix $\mathbf{W}$. Let $\hat{\boldsymbol{\theta}}_{\text{KL-GLM}(-i)}$ be given by Eq. (12) with replacing $\mathbf{M}$, $\mathbf{W}$, and $\mathbf{u}$ by $\mathbf{M}_{(-i)}$, $\mathbf{W}_{(-i)}$, and $\mathbf{u}_{(-i)}$, thus ,

$$\hat{\boldsymbol{\upsilon}}_{\text{KL-GLM(-i)}}=(\mathbf{M}^T_{(-i)}\hat{\mathbf{W}}_{(-i)}\mathbf{M}_{(-i)}+k\,\mathbf{I})^{-1}(\mathbf{M}^T_{(-i)}\hat{\mathbf{W}}_{(-i)}\mathbf{M}_{(-i)}-k\,\mathbf{I})^{-1}\mathbf{M}^T_{(-i)}\hat{\mathbf{W}}_{(-i)}\hat{\mathbf{u}}_{(-i)}, \tag{16}$$

where $(\mathbf{M}^T_{(-i)}\hat{\mathbf{W}}_{(-i)}\mathbf{M}_{(-i)}\mp k\,\mathbf{I})^{-1}$ is calculated according to Sherman-Morrison Woodbury theorem. Consequently, Eq. (13) can be expressed as

$$\hat{\boldsymbol{\upsilon}}_{\text{KL-GLM(-i)}}=\hat{\boldsymbol{\upsilon}}_{\text{KL-GLM}}-\frac{(\mathbf{M}^T\hat{\mathbf{W}}\mathbf{M}+k\,\mathbf{I})^{-1}(\mathbf{M}^T\hat{\mathbf{W}}\mathbf{M}-k\,\mathbf{I})^{-1}\mathbf{m}_i^T\,(\hat{\mathbf{u}}_i-\mathbf{m}_i^T\,\hat{\boldsymbol{\upsilon}}_{\text{KL-GLM}})}{1-\mathbf{m}_i^T\,(\mathbf{M}^T\hat{\mathbf{W}}\mathbf{M}+k\,\mathbf{I})^{-1}(\mathbf{M}^T\hat{\mathbf{W}}\mathbf{M}-k\,\mathbf{I})^{-1}\mathbf{m}_i}. \tag{17}$$

Using the weighted pseudo values (Hinkley, 1977), which are calculated as

$$T_i=\hat{\boldsymbol{\upsilon}}_{\text{KL-GLM}}+n(1-\mathbf{m}_i^T\,(\mathbf{M}^T\hat{\mathbf{W}}\mathbf{M}+k\,\mathbf{I})^{-1}(\mathbf{M}^T\hat{\mathbf{W}}\mathbf{M}-k\,\mathbf{I})^{-1}\mathbf{m}_i)(\hat{\boldsymbol{\upsilon}}_{\text{KL-GLM}}-\hat{\boldsymbol{\upsilon}}_{\text{KL-GLM(-i)}}). \tag{18}$$

Then, our proposed estimator, JKL-BRM, is defined as

$$\hat{\boldsymbol{\upsilon}}_{\text{JKL-BRM}}=\hat{\boldsymbol{\upsilon}}_{\text{KL-BRM}}+(\mathbf{M}^T\hat{\mathbf{W}}\mathbf{M}+k\,\mathbf{I})^{-1}(\mathbf{M}^T\hat{\mathbf{W}}\mathbf{M}-k\,\mathbf{I})^{-1}\sum_{i=1}^{n}\mathbf{m}_i^T\,(\hat{\mathbf{u}}_i-\mathbf{m}_i^T\,\hat{\boldsymbol{\upsilon}}_{\text{KL-BRM}}). \tag{19}$$

The bias, variance and MSE of $\hat{\boldsymbol{\upsilon}}_{\text{JKL-BRM}}$ is respectively defined as

$$\text{Bias}(\hat{\boldsymbol{\upsilon}}_{\text{JKL-BRM}})=\begin{bmatrix}\left[\mathbf{I}-2k\,(\mathbf{M}^T\hat{\mathbf{W}}\mathbf{M}+k\,\mathbf{I})^{-1}\right]^2\\ \left[\mathbf{I}+2k\,(\mathbf{M}^T\hat{\mathbf{W}}\mathbf{M}+k\,\mathbf{I})^{-1-1}\right]-\mathbf{I}\end{bmatrix}\boldsymbol{\upsilon}, \tag{20}$$

$$\text{Variance}(\hat{\boldsymbol{\upsilon}}_{\text{JKL-BRM}}) = \left[\mathbf{I} - \left(2k\,(\mathbf{M}^T\,\hat{\mathbf{W}}\mathbf{M} + k\,\mathbf{I})^{-1}\right)^2\right]^2 (\mathbf{M}^T\,\hat{\mathbf{W}}\mathbf{M} + k\,\mathbf{I})^{-1} \left[\mathbf{I} - \left(2k\,(\mathbf{M}^T\,\hat{\mathbf{W}}\mathbf{M} + k\,\mathbf{I})^{-1}\right)\right]^2, \tag{21}$$

$$\text{MSE}(\hat{\boldsymbol{\upsilon}}_{\text{JKL-BRM}}) = \sum_{j=1}^{p} \frac{\left(\left(\lambda_j + k\right)^2 - 4k^2\right)^2 \left(\lambda_j - k\right)^2}{\lambda_j\,(\lambda_j + k)^6} + \sum_{j=1}^{p} \frac{\left(\left(\lambda_j - k\right)^2 \left(\lambda_j + 3k\right) - \left(\lambda_j + k\right)^3\right)^2 \alpha_j^2}{(\lambda_j + k)^6}. \tag{22}$$

4. Theoretical comparison between $\hat{\boldsymbol{\upsilon}}_{\text{JKL-BRM}}$ and $\hat{\boldsymbol{\upsilon}}_{\text{KL-BRM}}$

With availability of different estimators for a parameter in the regression model, it is of interest to compare their performances in terms of MSE. For two given estimators $\hat{\boldsymbol{\beta}}_A$ and $\hat{\boldsymbol{\beta}}_B$ of $\boldsymbol{\beta}$, the estimator $\hat{\boldsymbol{\beta}}_B$ is said to be superior to $\hat{\boldsymbol{\beta}}_A$ under the MSE criterion if and only if $\Delta = \text{MSE}(\hat{\boldsymbol{\beta}}_A) - \text{MSE}(\hat{\boldsymbol{\beta}}_B) \geq 0$.

Lemma: (Farebrother, 1976) Let $\mathbf{G}$ is a $p \times p$ positive definite matrix, $\mathbf{b}$ is a $p \times 1$ vector, and c is a positive constant. Then $c\,\mathbf{G} - \mathbf{b}\mathbf{b}^T$ is a nonnegative definite if and only if $\mathbf{b}^T\,\mathbf{G}^{-1}\mathbf{b} \leq c$ is hold.

Theorem. The proposed estimator $\hat{\boldsymbol{\upsilon}}_{\text{JKL-BRM}}$ is superior to estimator $\hat{\boldsymbol{\upsilon}}_{\text{KL-BRM}}$ if and only if

$$\begin{aligned}
&\boldsymbol{\upsilon}^T \left[\mathbf{I} - 2k\,(\mathbf{M}^T\,\hat{\mathbf{W}}\mathbf{M} + k\,\mathbf{I})^{-1}\right]^2 \left[\mathbf{I} + 2k\,(\mathbf{M}^T\,\hat{\mathbf{W}}\mathbf{M} + k\,\mathbf{I})^{-1}\right] - \mathbf{I}\Big] \\
&\Big[\ \left(\mathbf{M}^T\,\hat{\mathbf{W}}\mathbf{M} - k\mathbf{I}\right)^2 (\mathbf{M}^T\,\hat{\mathbf{W}}\mathbf{M})^{-1} \left(\mathbf{M}^T\,\hat{\mathbf{W}}\mathbf{M} + k\mathbf{I}\right)^{-2} - \left[\mathbf{I} - \left(2k\,(\mathbf{M}^T\,\hat{\mathbf{W}}\mathbf{M} + k\,\mathbf{I})^{-1}\right)^2\right]^2 \\
&(\mathbf{M}^T\,\hat{\mathbf{W}}\mathbf{M}) \left[\mathbf{I} - \left(2k\,(\mathbf{M}^T\,\hat{\mathbf{W}}\mathbf{M} + k\,\mathbf{I})^{-1}\right)\right]^2 + 4k^2 \left(\mathbf{M}^T\,\hat{\mathbf{W}}\mathbf{M} + k\mathbf{I}\right)^{-2} \boldsymbol{\upsilon}^T\,\boldsymbol{\upsilon}\Big] \\
&\left[\mathbf{I} - 2k\,(\mathbf{M}^T\,\hat{\mathbf{W}}\mathbf{M} + k\,\mathbf{I})^{-1}\right]^2 \left[\mathbf{I} + 2k\,(\mathbf{M}^T\,\hat{\mathbf{W}}\mathbf{M} + k\,\mathbf{I})^{-1}\right] - \mathbf{I}\Big]\boldsymbol{\upsilon} < 1
\end{aligned} \tag{23}$$

Proof. The difference between $\text{MSE}(\hat{\boldsymbol{\upsilon}}_{\text{JKL-BRM}})$ and $\text{MSE}(\hat{\boldsymbol{\upsilon}}_{\text{KL-BRM}})$ is

$$= diag \left\{ \frac{1}{\lambda_j} \left(\frac{\lambda_j - k}{\lambda_i + k}\right)^2 - \frac{\left(\left(\lambda_j + k\right)^2 - 4k^2\right)^2 \left(\lambda_j - k\right)^2}{\lambda_j \left(\lambda_j + k\right)^6} \right\}_{j=1}^{p} \tag{24}$$

Consequently,

ISSN: 2094-0343
2326-9865

$$\left[\left(\mathbf{M}^T \hat{\mathbf{W}}\mathbf{M} + k\,\mathbf{I}\right)^2 (\mathbf{M}^T \hat{\mathbf{W}}\mathbf{M})^{-1} \left(\mathbf{M}^T \hat{\mathbf{W}}\mathbf{M} + k\,\mathbf{I}\right)^{-2} - \left[\mathbf{I} - \left(2k\,(\mathbf{M}^T \hat{\mathbf{W}}\mathbf{M} + k\,\mathbf{I})^{-1}\right)^2\right]^2 \right. \\ \left. \mathbf{M}^T \hat{\mathbf{W}}\mathbf{M}\left[\mathbf{I} - \left(2k\,(\mathbf{M}^T \hat{\mathbf{W}}\mathbf{M} + k\,\mathbf{I})^{-1}\right)\right]^2 \quad (25)$$

is positive definite provided

$$\left(\lambda_j - k\right)^2 \left(\lambda_j + k\right)^4 > \left(\left(\lambda_j + k\right)^2 - 4k^2\right)^2 \left(\lambda_j - k\right)^2. \quad (26)$$

the proof is completed.

5. Simulation results

Table 1: Mean squared error of simulated data when p=3

n	Estimator	r=0.7	r=0.8	r=0.9	r=0.99
30	MLE	4.8489	5.0375	5.4917	13.4242
	RIDGE	1.4124	1.4424	1.5614	1.7544
	Liu	1.4548	1.4574	1.5767	9.2924
	KL	1.2532	1.4152	1.496	1.6553
	JKL	**1.2116**	**1.2501**	**1.3282**	**1.4879**
50	MLE	1.6083	1.8681	2.0006	7.1706
	RIDGE	1.1729	1.2498	1.271	1.2737
	Liu	1.3381	1.3737	1.4066	7.0068
	KL	1.1718	1.2457	1.2562	1.2654
	JKL	**1.1448**	**1.2058**	**1.1944**	**1.2125**
100	MLE	1.3103	1.3704	1.4676	3.1757
	RIDGE	1.0974	1.1547	1.1669	1.1753
	Liu	1.2814	1.2858	1.2895	1.3467
	KL	1.0748	1.1517	1.1665	1.1745
	JKL	**1.0543**	**1.1085**	**1.0834**	**1.0719**

Table 2: Mean squared error of simulated data when p=8

n	Estimator	r=0.7	r=0.8	r=0.9	r=0.99
30	MLE	4.87	5.0586	5.5128	13.4453
	RIDGE	1.4335	1.4635	1.5825	1.7755
	Liu	1.4759	1.4785	1.5978	9.3135
	KL	1.2743	1.4363	1.5171	1.6764
	JKL	**1.2327**	**1.2712**	**1.3493**	**1.509**
50	MLE	1.6294	1.8892	2.0217	7.1917
	RIDGE	1.194	1.2709	1.2921	1.2948

ISSN: 2094-0343
2326-9865

	Liu	1.3592	1.3948	1.4277	7.0279
	KL	1.1929	1.2668	1.2773	1.2865
	JKL	**1.1659**	**1.2269**	**1.2155**	**1.2336**
100	MLE	1.3314	1.3915	1.4887	3.1968
	RIDGE	1.1185	1.1758	1.188	1.1964
	Liu	1.3025	1.3069	1.3106	1.3678
	KL	1.0959	1.1728	1.1876	1.1956
	JKL	**1.0754**	**1.1296**	**1.1045**	**1.093**

Table 3: Mean squared error of simulated data when p=12

n	Estimator	r=0.7	r=0.8	r=0.9	r=0.99
30	MLE	4.9027	5.0913	5.5455	13.478
	RIDGE	1.4662	1.4962	1.6152	1.8082
	Liu	1.5086	1.5112	1.6305	9.3462
	KL	1.307	1.469	1.5498	1.7091
	JKL	**1.2654**	**1.3039**	**1.382**	**1.5417**
50	MLE	1.6621	1.9219	2.0544	7.2244
	RIDGE	1.2267	1.3036	1.3248	1.3275
	Liu	1.3919	1.4275	1.4604	7.0606
	KL	1.2256	1.2995	1.31	1.3192
	JKL	**1.1986**	**1.2596**	**1.2482**	**1.2663**
100	MLE	1.3641	1.4242	1.5214	3.2295
	RIDGE	1.1512	1.2085	1.2207	1.2291
	Liu	1.3352	1.3396	1.3433	1.4005
	KL	1.1286	1.2055	1.2203	1.2283
	JKL	**1.1081**	**1.1623**	**1.1372**	**1.1257**

Table 4: Mean squared error of simulated data when p=16

n	Estimator	r=0.7	r=0.8	r=0.9	r=0.99
30	MLE	4.9411	5.1297	5.5839	13.5164
	RIDGE	1.5046	1.5346	1.6536	1.8466
	Liu	1.547	1.5496	1.6689	9.3846
	KL	1.3454	1.5074	1.5882	1.7475
	JKL	**1.3038**	**1.3423**	**1.4204**	**1.5801**
50	MLE	1.7005	1.9603	2.0928	7.2628
	RIDGE	1.2651	1.342	1.3632	1.3659
	Liu	1.4303	1.4659	1.4988	7.099
	KL	1.264	1.3379	1.3484	1.3576

ISSN: 2094-0343
2326-9865

	JKL	**1.237**	**1.298**	**1.2866**	**1.3047**
100	MLE	1.4025	1.4626	1.5598	3.2679
	RIDGE	1.1896	1.2469	1.2591	1.2675
	Liu	1.3736	1.378	1.3817	1.4389
	KL	1.167	1.2439	1.2587	1.2667
	JKL	**1.1465**	**1.2007**	**1.1756**	**1.1641**

6. Real-life application

6.1. Aircraft Data

This dataset is originally assumed to follow the Poisson regression model (see Myers et al., 2012; Asar & Genc, 2017; Amin et al. 2020; Lukman et al., 2021a,b), among others. The response variable y represent the number of locations with damage on the aircraft and it follows a Poisson distribution (Myers et al., 2012; Asar & Genc, 2017; Amin et al. 2020; Lukman et al., 2021a,b). The explanatory variables are described as follows: x_1 denotes aircraft type (A-4 coded as 0 and A-6 coded as 1), x_2 and x_3 denote bomb load in tons and total months of aircrew experience, respectively. Lukman et al. (2021a,b) diagnosed the model and conclude that the model suffers from multicollinearity because the condition number is 219.3654. The output of the Poisson regression model using the maximum likelihood method is presented in Table 5.

Table 5: Poisson regression estimates using MLE

Coef.	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-0.4060	0.8775	-0.463	0.6436
x1	0.5688	0.5044	1.128	0.2595
x2	0.1654	0.0675	2.449	0.0143
x3	-0.0135	0.0083	-1.633	0.1025

However, the variance of the number of locations with damage on the aircraft is more than twice the mean (2.0569). With this, it is evident that the data exhibit over-dispersion. Bell Regression models account for over-dispersion in count data (Castellares et al., 2018; Lemonte et al. 2020). Recently, Amin et al. (2021) employed the bell regression model to model the same dataset. Table 6 provides the regression estimates and the mean squared error of each of the adopted estimators in this study. The biasing parameter k proposed by Hoerl et al. (1975) was adopted as the biasing parameter for the Bell ridge and the Bell KL estimators.

$$\hat{k} = \frac{p}{\sum_{j=1}^{p} \hat{\upsilon}_j^2} \quad (25)$$

The Scalar mean squared error (MSE) for the other adopted method of estimation in this study are as follows:

$$\mathrm{MSE}(\hat{\boldsymbol{\beta}}_{\mathrm{MLE}}) = \sum_{j=1}^{p} \frac{1}{\lambda_j} \quad (26)$$

where λ_j is the eigenvalue of $\mathbf{X}^T\widehat{\mathbf{W}}\mathbf{X}$.

$$\mathrm{MSE}(\widehat{\boldsymbol{\beta}}_{k-BRM}) = \sum_{j=1}^{p} \frac{\lambda_j}{(\lambda_j+k)^2} + k^2 \sum_{j=1}^{p} \frac{\widehat{\upsilon}_j^2}{(\lambda_j+k)^2} \qquad (27)$$

where $\widehat{\alpha}_j^2$ is the jth squared of the maximum likelihood estimate.

$$\mathrm{MSE}(\widehat{\boldsymbol{\beta}}_{d-BRM}) = \sum_{j=1}^{p} \frac{(\lambda_j+d)^2}{\lambda_j(\lambda_j+1)^2} + (1-d)^2 \sum_{j=1}^{p} \frac{\widehat{\upsilon}_j^2}{(\lambda_j+1)^2} \qquad (28)$$

6.2. Table 6: Bell regression estimates for Aircraft Data

Coef.	$\widehat{\boldsymbol{\beta}}_{MLE}$	$\widehat{\boldsymbol{\beta}}_{k-BRM}$	$\widehat{\boldsymbol{\beta}}_{d-BRM}$	$\widehat{\boldsymbol{\beta}}_{KL-BRM}$	$\widehat{\boldsymbol{\beta}}_{JKL-BRM}$
Intercept	-0.5422	-0.1509	-0.3006	-0.0375	-0.0534
x1	0.5990	0.3433	0.0034	0.3285	0.0021
x2	0.1630	0.1665	0.0119	0.1605	0.0521
x3	-0.0117	-0.0146	-0.0023	-0.0161	-0.0362
MSE	1.7447	0.1609	0.5327	0.1493	**0.1125**

7. Conclusions

We have presented a new proposed estimator of K-L estimator for Bell regression model in the presence of collinearity. The proposed estimator combines Jackknife procedure with K-L estimator to reduce the biasedness. Our experimental results with both simulated and real application, which is related to the Bell regression model, demonstrated that the proposed estimator could successfully deal with collinearity. Moreover, compared with MLE, Ridge, and KL-BRM, the proposed estimator can efficiently reduce the MSE.

References

1. Akdeniz Duran, E., & Akdeniz, F. (2012). Efficiency of the modified jackknifed Liu-type estimator. Statistical Papers, 53(2), 265-280. doi:10.1007/s00362-010-0334-5
2. Akram, M. N., Amin, M., & Amanullah, M. (2020). Two-parameter estimator for the inverse Gaussian regression model. Communications in Statistics - Simulation and Computation, 1-19. doi:10.1080/03610918.2020.1797797
3. Algamal, Z. Y. (2018). A new method for choosing the biasing parameter in ridge estimator for generalized linear model. Chemometrics and Intelligent Laboratory Systems, 183, 96-101. doi:10.1016/j.chemolab.2018.10.014
4. Alkhateeb, A., & Algamal, Z. (2020). Jackknifed Liu-type Estimator in Poisson Regression Model. Journal of the Iranian Statistical Society, 19(1), 21-37. doi:10.29252/jirss.19.1.21
5. Amin, M., Akram, M. N., & Majid, A. (2021). On the estimation of Bell regression model using ridge estimator. Communications in Statistics - Simulation and Computation, 1-14. doi:10.1080/03610918.2020.1870694

ISSN: 2094-0343
2326-9865

6. Bell, E. T. (1934). Exponential numbers. The American Mathematical Monthly, 41(7), 411-419.
7. Bell, E. T. (1934). Exponential polynomials. Annals of Mathematics, 35(2), 258-277.
8. Castellares, F., Ferrari, S. L. P., & Lemonte, A. J. (2018). On the Bell distribution and its associated regression model for count data. Applied Mathematical Modelling, 56, 172-185. doi:10.1016/j.apm.2017.12.014
9. Farebrother, R. (1976). Further results on the mean square error of ridge regression. Journal of the Royal Statistical Society. Series B (Methodological), 38(3), 248-250.
10. Gruber, M. H. (1991). The efficiency of jack-knifed and usual ridge type estimators: A comparison. Statistics & Probability Letters, 11(1), 49-51.
11. Hinkley, D. V. (1977). Jackknifing in unbalanced situations. Technometrics, 19(3), 285-292.
12. Hoerl, A. E., & Kennard, R. W. (1970). Ridge regression: Biased estimation for nonorthogonal problems. Technometrics, 12(1), 55-67.
13. Khurana, M., Chaubey, Y. P., & Chandra, S. (2014). Jackknifing the Ridge Regression Estimator: A Revisit. Communications in Statistics - Theory and Methods, 43(24), 5249-5262. doi:10.1080/03610926.2012.729640
14. Khurana, M., Chaubey, Y. P., & Chandra, S. (2014). Jackknifing the ridge regression estimator: A revisit. Communications in Statistics-Theory and Methods, 43(24), 5249-5262.
15. Kibria, B. M. G. (2003). Performance of some new ridge regression estimators. Communications in Statistics - Simulation and Computation, 32(2), 419-435. doi:10.1081/SAC-120017499
16. Kibria, B. M. G., & Lukman, A. F. (2020). A New Ridge-Type Estimator for the Linear Regression Model: Simulations and Applications. Scientifica, 2020, 9758378. doi:10.1155/2020/9758378
17. Kibria, G., Månsson, K., & Shukur, G. (2012). Performance of some logistic ridge regression estimators. Computational Economics, 40(4), 401–414.
18. Kurtoğlu, F., & Özkale, M. R. (2016). Liu estimation in generalized linear models: application on gamma distributed response variable. Statistical Papers, 57(4), 911-928. doi:10.1007/s00362-016-0814-3
19. Liu, G., & Piantadosi, S. (2016). Ridge estimation in generalized linear models and proportional hazards regressions. Communications in Statistics - Theory and Methods, 46(23), 11466-11479. doi:10.1080/03610926.2016.1267767
20. Liu, K. (1993). A new class of biased estimate in linear regression. Communication in Statistics -Theory and Methods, 22, 393–402.
21. Lukman, A. F., Algamal, Z. Y., Kibria, B. M. G., & Ayinde, K. (2021). The KL estimator for the inverse Gaussian regression model. Concurrency and Computation: Practice and Experience, 33(13). doi:10.1002/cpe.6222
22. Lukman, A. F., Dawoud, I., Kibria, B. M. G., Algamal, Z. Y., & Aladeitan, B. (2021). A New Ridge-Type Estimator for the Gamma Regression Model. Scientifica (Cairo), 2021, 5545356. doi:10.1155/2021/5545356

ISSN: 2094-0343
2326-9865

23. Mackinnon, M. J., & Puterman, M. L. (1989). Collinearity in generalized linear models. Communications in Statistics - Theory and Methods, 18(9), 3463-3472. doi:10.1080/03610928908830102
24. Majid, A., Amin, M., & Akram, M. N. (2021). On the Liu estimation of Bell regression model in the presence of multicollinearity. Journal of Statistical Computation and Simulation, 1-21. doi:10.1080/00949655.2021.1955886
25. Månsson, K., & Shukur, G. (2011). A Poisson ridge regression estimator. Economic Modelling, 28(4), 1475-1481. doi:10.1016/j.econmod.2011.02.030
26. Nyquist, H. (1988). Applications of the jackknife procedure in ridge regression. Computational Statistics & Data Analysis, 6(2), 177-183.
27. Nyquist, H. (1991). Restricted estimation of generalized linear models. Journal of the Royal Statistical Society. Series C (Applied Statistics), 40(1), 133-141. doi:10.2307/2347912
28. Özkale, M. R., & Arıcan, E. (2018). A first-order approximated jackknifed ridge estimator in binary logistic regression. Computational Statistics, 34(2), 683-712. doi:10.1007/s00180-018-0851-6
29. Segerstedt, B. (1992). On ordinary ridge regression in generalized linear models. Communications in Statistics - Theory and Methods, 21(8), 2227-2246. doi:10.1080/03610929208830909
30. Shamany, R., Alobaidi, N. N., & Algamal, Z. Y. (2019). A new two-parameter estimator for the inverse Gaussian regression model with application in chemometrics. Electronic Journal of Applied Statistical Analysis, 12(2), 453-464.
31. Singh, B., Chaubey, Y., & Dwivedi, T. (1986). An almost unbiased ridge estimator. Sankhyā: The Indian Journal of Statistics, Series B, 13, 342-346.
32. Türkan, S., & Özel, G. (2015). A new modified Jackknifed estimator for the Poisson regression model. Journal of Applied Statistics, 43(10), 1892-1905. doi:10.1080/02664763.2015.1125861
33. Yıldız, N. (2017). On the performance of the Jackknifed Liu-type estimator in linear regression model. Communications in Statistics - Theory and Methods, 0-0. doi:10.1080/03610926.2017.1339087