

# Modified Jackknife Ridge Estimator for Beta Regression Model with Application to Chemical Data

Zakariya Yahya Algamal<sup>1,\*</sup>, Mohamed R. Abonazel<sup>2</sup>, Adewale F. Lukman<sup>3</sup>

<sup>1</sup>Department of Statistics and Informatics, University of Mosul, Mosul, Iraq; zakariya.algamal@uomosul.edu.iq

<sup>2</sup>Department of Applied Statistics and Econometrics, Faculty of Graduate Studies for Statistical Research, Cairo University, Egypt; mabonazel@cu.edu.eg

<sup>3</sup>Department of Epidemiology and Biostatistics, University of Medical Sciences, Ondo, Nigeria; adewale.folaranmi@lmu.edu.ng

Received 04.01.2023, Revised 05.02.2023, Accepted 15.02.2023, Published 20.02.2023

**ABSTRACT:** The linear regression model is not applicable when the response variable's value comes in the form of percentages, proportions, and rates, which are restricted to the interval (0, 1). In this situation, we applied the beta regression model (BRM) which is popularly used to model chemical, environmental and biological data. The parameters in the model are often estimated using the conventional method of maximum likelihood. However, this estimator is unreliable and inefficient when the explanatory variables are linearly correlated- a condition known as multicollinearity. Thus, we developed the Jackknife Beta ridge and the modified Jackknife Beta ridge estimator for efficient estimation of the regression coefficient when there is multicollinearity. The properties of the new estimators were derived. We compared the performance of the estimator with the existing estimators theoretically using the mean squared error criterion. Furthermore, we conducted a simulation study and a chemical data to evaluate the new estimators' performance. The theoretical comparison, simulation and real-life application results established the dominance of the proposed methods.

**Keywords:** Beta regression model, linear regression model, multicollinearity, ridge, modified Jackknife, Jackknife ridge.

## 1. INTRODUCTION

The beta regression model has been common in many areas, primarily economic and medical research, such as income share, unemployment rates in certain nations, the Gini index for each region, graduation rates in major universities, or the percentage of body fat in medical subjects. Beta regression model, like any regression model in the context of generalized linear models (GLMs), is used to examine the effect of certain explanatory variables on a non-normal response variable. However, in beta regression, the response component is restricted to an interval (0,1) such as proportions, percentages, and fractions.

Multicollinearity is a popular issue in econometric modelling introduced by Frisch [14]. It indicates that there is a strong association between the explanatory variables. It is well established that the covariance matrix of the maximum likelihood (ML) estimator is ill-conditioned in the case of severing multicollinearity. One of the negative consequences of this issue is that the variance of the regression coefficients gets inflated. Therefore, the significance and the magnitude of the coefficients are affected. Many of the conventional approaches used to address this issue include: gathering additional data, re-specifying the model, or removing the correlated variable/s.

Throughout recent years, shrinkage methods have become a commonly recognized and more effective methodology for solving this issue throughout regression models. To solve this issue, Hoerl and Kennard [17, 18] proposed the ridge estimator. The concept behind the ridge estimator is to apply a small definite amount ( $k$ ) to the diagonal entries of the covariance matrix to increase the conditioning of this matrix, reduce the MSE and achieve consistent coefficients. For a review of this method in both linear and GLMs, see, e.g., Kibria and Lukman [20], Algamal [4], Abonazel [1], Rady et al. [29], Abonazel and Farghali [2], Farghali et al. [12], and Lukman et al. [23].

One of the drawbacks of the ridge estimator is that estimated parameters are nonlinear functions of the ridge parameter and that the small selected might not be high enough to solve multicollinearity.

## 2. BETA RIDGE REGRESSION

The Beta regression model was first introduced by Ferrari and Cribari-Neto [13] by relating the mean function of its response variable to a set of linear predictors through a link function. This model includes a precision parameter whose reciprocal is considered as a dispersion measure.

Suppose that  $y$  is a continuous random variable that follows a beta distribution, the probability density function has the following form:

$$f(y; \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1}(1-y)^{(1-\mu)\phi-1}, \quad (1)$$

for  $0 < y < 1$ ;  $0 < \mu < 1$ ;  $\phi > 0$ , and  $\Gamma(\cdot)$  is the gamma function, and  $\phi$  is the precision parameter that can be written as in [6]:

$$\phi = \frac{1-\sigma^2}{\sigma^2}.$$

The mean and variance of the beta probability distribution are:  $E(y) = \mu$ ,  $\text{var}(y) = \mu(1-\mu)\sigma^2$ .

The model allows  $\mu_i$ , depending on covariates as follows:

$$g(\mu_i) = \log\left(\frac{\mu_i}{1-\mu_i}\right) = \mathbf{x}_i^T \boldsymbol{\beta} = \eta_i, \quad (2)$$

where  $g(\cdot)$  be a monotonic differentiable link function used to relate the systematic component with the random component,  $\boldsymbol{\beta} = (\beta_1, \dots, \beta_p)^T$  is a  $p \times 1$  vector of unknown parameters,  $\mathbf{x}_i = (x_{i1}, \dots, x_{ip})^T$  is the vector of  $p$  regressors, and  $\eta_i$  is a linear predictor.

Estimation of the beta regression parameters is done by using the ML method [10]. The log-likelihood function of the beta regression model is given by:

$$\mathcal{L}(\mu_i, \sigma^2; \mathbf{y}_i) = \sum_{i=1}^n \left\{ \log \Gamma\left(\frac{1-\sigma_i^2}{\sigma^2}\right) - \log \Gamma\left(\mu_i \left(\frac{1-\sigma_i^2}{\sigma^2}\right)\right) - \log \Gamma\left((1-\mu_i) \left(\frac{1-\sigma_i^2}{\sigma^2}\right)\right) + \left(\mu_i \left(\frac{1-\sigma_i^2}{\sigma^2}\right) - 1\right) \log(y_i) + \left((1-\mu_i) \left(\frac{1-\sigma_i^2}{\sigma^2}\right) - 1\right) \log(1-y_i) \right\}. \quad (3)$$

Differentiating the log-likelihood in Eq. (3) with respect to  $\boldsymbol{\beta}$  gives us the score function for  $\boldsymbol{\beta}$  which is given by:

$$S(\boldsymbol{\beta}) = \boldsymbol{\phi} X^T A (\mathbf{y}^* - \boldsymbol{\mu}^*), \quad (4)$$

where  $A = \text{diag}\left(\frac{1}{g'(\mu_1)}, \dots, \frac{1}{g'(\mu_n)}\right)$ ,  $\mathbf{y}^* = (y_1^*, \dots, y_n^*)^T$ ,  $\boldsymbol{\mu}^* = (\mu_1^*, \dots, \mu_n^*)^T$ ,  $y_i^* = \log\left(\frac{y_i}{1-y_i}\right)$ , and  $\mu_i^* = \psi\left(\mu_i \left(\frac{1-\sigma_i^2}{\sigma^2}\right)\right) - \psi\left((1-\mu_i) \left(\frac{1-\sigma_i^2}{\sigma^2}\right)\right)$ , such that  $\psi(\cdot)$  denoting the digamma function. The iterative reweighted least-squares (IWLS) algorithm or Fisher scoring algorithm used for estimating  $\boldsymbol{\beta}$  [8, 9]. The form of this algorithm can be written as:

$$\boldsymbol{\beta}^{(r+1)} = \boldsymbol{\beta}^{(r)} + (\mathbf{I}_{\boldsymbol{\beta}\boldsymbol{\beta}}^{(r)})^{-1} S_{\boldsymbol{\beta}}^{(r)}(\boldsymbol{\beta}),$$

where  $S_{\boldsymbol{\beta}}^{(r)}$  is the score function defined in Eq. (4), and  $\mathbf{I}_{\boldsymbol{\beta}\boldsymbol{\beta}}^{(r)}$  is the information matrix for  $\boldsymbol{\beta}$ , see Espinheira et al. (2019) for more details. The initial value of  $\boldsymbol{\beta}$  can be obtained by the least-squares estimation, while the initial value for each precision parameter is:

$$\hat{\phi}_i = \frac{\hat{\mu}_i(1-\hat{\mu}_i)}{\hat{\sigma}_i^2}, \quad (5)$$

where  $\hat{\mu}_i$  and  $\hat{\sigma}_i^2$  values are obtained from linear regression. Given  $r = 0, 1, 2, \dots$  is the number of iterations that are performed, convergence occurs when the difference between successive estimates becomes smaller than a given small constant. At the final step, the ML estimator of  $\boldsymbol{\beta}$  is obtained as:

$$\hat{\boldsymbol{\beta}}_{\text{BR}} = (X^T \widehat{W} X)^{-1} X^T \widehat{W} \hat{\mathbf{z}}, \quad (6)$$

where  $X$  is an  $n \times p$  matrix of regressors,  $\hat{\mathbf{z}} = \hat{\boldsymbol{\eta}} + \widehat{W}^{-1} \hat{A} (\mathbf{y}^* - \boldsymbol{\mu}^*)$ , and  $\widehat{W} = \text{diag}(\widehat{w}_1, \dots, \widehat{w}_n)$ ;

$$\widehat{w}_i = \frac{(1-\hat{\sigma}_i^2)}{\hat{\sigma}_i^2} \left\{ \psi' \left( \frac{\hat{\mu}_i(1-\hat{\sigma}_i^2)}{\hat{\sigma}_i^2} \right) + \psi' \left( \frac{(1-\hat{\mu}_i)(1-\hat{\sigma}_i^2)}{\hat{\sigma}_i^2} \right) \right\} \frac{1}{\{g'(\hat{\mu}_i)\}^2}.$$

Here,  $\widehat{W}$  and  $\hat{A}$  are the matrices  $W$  and  $A$ , respectively, evaluated at the ML estimator. The ML estimator of  $\boldsymbol{\beta}$  is normally distributed with asymptotic mean vectors  $E(\hat{\boldsymbol{\beta}}_{\text{BR}}) = \boldsymbol{\beta}$  and asymptotic covariance matrix:

$$\text{Cov}(\hat{\boldsymbol{\beta}}_{\text{BR}}) = \frac{1}{\phi} (X^T \widehat{W} X)^{-1}. \quad (7)$$

Hence the asymptotic mean squared error (MSE) [11] of the ML estimator based on the asymptotic covariance matrices is:

$$\text{MSE}(\hat{\boldsymbol{\beta}}_{\text{BR}}) = \text{tr} \left[ \frac{1}{\phi} (X^T \widehat{W} X)^{-1} \right] = \frac{1}{\phi} \sum_{j=1}^p \frac{1}{\lambda_j} \quad (8)$$

where  $\lambda_j$  is the eigenvalue of the  $X^T \widehat{W} X$  matrix.

Recently, Abonazel and Taha [3] and Qasim et al. [27] introduced the beta ridge regression (BRR) estimator as follows:

$$\hat{\beta}_{\text{BRR}} = (X^T \hat{W} X + kI)^{-1} X^T \hat{W} \hat{z}, \quad (9)$$

where  $k > 0$ , and  $\|\hat{\beta}_{\text{BRR}}\| < \|\hat{\beta}_{\text{BR}}\|$ . It can note that if  $k = 0$ , then  $\hat{\beta}_{\text{BRR}} = \hat{\beta}_{\text{BR}}$ . Now, let  $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_p) = Q^T X^T \hat{W} X Q$ , and  $\alpha = Q^T \beta$ , where  $\lambda_1 \geq \dots \geq \lambda_p \geq 0$ , and  $Q$  is the matrix whose columns are the eigenvectors of the matrix  $X^T \hat{W} X$ . Then the MSE of the BRR estimator is:

$$\begin{aligned} \text{MSE}(\hat{\beta}_{\text{BRR}}) &= E(\hat{\beta}_{\text{BRR}} - \beta)^T E(\hat{\beta}_{\text{BRR}} - \beta), \\ &= \frac{1}{\phi} \sum_{j=1}^p \frac{\lambda_j}{(\lambda_j + k)^2} + k^2 \sum_{j=1}^p \frac{\alpha_j^2}{(\lambda_j + k)^2}. \end{aligned} \quad (10)$$

The first term in Eq. (10) is the asymptotic variance, and the second term is the squared bias.

### 3. PROPOSED ESTIMATORS

#### Jackknife Beta Ridge Regression

In the context of the linear regression model, the Jackknife procedure has been proposed by Singh et al. [31] to alleviate the problem of bias in generalized ridge estimator. The theoretical and application of the jackknife estimator have been studied by several authors [24, 15, 7, 32, 33, 19, 34, 35].

Now, we introduce the Jackknifed ridge estimator of the beta regression. Based on  $\Lambda$  and  $Q$  matrices, the beta regression estimator of Eq. (6) can be re-written as

$$\hat{\beta}_{\text{BR}} = Q \hat{\alpha}_{\text{BR}}, \quad (11)$$

where  $\hat{\alpha}_{\text{BR}} = \Lambda^{-1} M^T \hat{W} \hat{z}$ ; with  $M = XQ$ . As a result, the BRR estimator of Eq. (9) is re-written as:

$$\hat{\alpha}_{\text{BRR}} = (\Lambda + kI)^{-1} M^T \hat{W} \hat{z}. \quad (12)$$

Following the idea of Jackknife approach [16], let  $z^{(-i)}$ ,  $m^{(-i)}$ , and  $W^{(-i)}$ , respectively, are the  $i^{\text{th}}$  row deleted from the vector  $z$ , the  $i^{\text{th}}$  row deleted from the matrix  $M$ , and the  $i^{\text{th}}$  row and column deleted from the matrix  $W$ . Let  $\hat{\alpha}_{\text{BRR}(-i)}$  be given by Eq. (12) with replacing  $M$ ,  $W$ , and  $z$  by  $M^{(-i)}$ ,  $W^{(-i)}$ , and  $z^{(-i)}$ , thus,

$$\hat{\alpha}_{\text{BRR}(-i)} = (M_{(-i)}^T \hat{W}_{(-i)} M_{(-i)} + kI)^{-1} M_{(-i)}^T \hat{W}_{(-i)} \hat{z}_{(-i)}, \quad (13)$$

where  $(M_{(-i)}^T \hat{W}_{(-i)} M_{(-i)} + kI)^{-1}$  is calculated according to Sherman-Morrison Woodbury theorem [30] as

$$\begin{aligned} (M_{(-i)}^T \hat{W}_{(-i)} M_{(-i)} + kI)^{-1} &= (M^T \hat{W} M + kI)^{-1} \\ &\quad + \frac{(M^T \hat{W} M + kI)^{-1} m_i m_i^T (M^T \hat{W} M + kI)^{-1}}{1 - m_i^T (M^T \hat{W} M + kI)^{-1} m_i}, \end{aligned}$$

and  $M_{(-i)}^T \hat{W}_{(-i)} \hat{z}_{(-i)} = M^T \hat{W} \hat{z} - m_i \hat{z}_i$ . Then, Eq. (13) can be expressed as

$$\hat{\alpha}_{\text{BRR}(-i)} = \hat{\alpha}_{\text{BRR}} - \frac{(M^T \hat{W} M + kI)^{-1} m_i^T (\hat{z}_i - m_i^T \hat{\alpha}_{\text{BRR}})}{1 - m_i^T (M^T \hat{W} M + kI)^{-1} m_i}.$$

Using the weighted pseudo values [16], which are calculated as

$$T_i = \hat{\alpha}_{\text{BRR}} + n(1 - m_i^T (M^T \hat{W} M + kI)^{-1} m_i)(\hat{\alpha}_{\text{BRR}} - \hat{\alpha}_{\text{BRR}(-i)}).$$

Then, the Jackknifed beta ridge regression estimator (JBRR) is defined as

$$\hat{\alpha}_{\text{JBRR}} = \hat{\alpha}_{\text{BRR}} + (M^T \hat{W} M + kI)^{-1} \sum_{i=1}^n m_i^T (\hat{z}_i - m_i^T \hat{\alpha}_{\text{BRR}}).$$

Let  $B = (M^T \hat{W} M + kI)$ , then  $\hat{\alpha}_{\text{JBRR}}$  can be further simplified as

$$\hat{\alpha}_{\text{JBRR}} = (I - k B^{-1}) \hat{\alpha}_{\text{BRR}} = (I - k^2 B^{-2}) \hat{\alpha}_{\text{BR}}, \quad (14)$$

### Modified Jackknife Beta Ridge Regression

In the linear regression model, the modified Jackknife ridge estimator has been proposed by Batah, et al. [9] by combining the ideas of the generalized ridge estimator and Jackknifed ridge estimator, which is given by Singh et al. [31]. In a similar way, we propose a new estimator of  $\hat{\tau}_{JBRR}$ . Following Algamal [4], the proposed estimator, which is denoted by modified Jackknifed beta ridge estimator (MJBRR), is designated as in the case of the  $\hat{\alpha}_{JBRR}$  by replacing  $\hat{\alpha}_{BR}$  with  $\hat{\alpha}_{BRR}$ . The MJBRR estimator is defined as:

$$\hat{\alpha}_{MJBRR} = (\mathbf{I} - k^2 \mathbf{B}^{-2})(\mathbf{I} - k \mathbf{B}^{-1})\hat{\alpha}_{BR}. \quad (15)$$

The asymptotic properties of  $\hat{\alpha}_{MJBRR}$  are defined as, respectively

$$\begin{aligned} \text{Bias}(\hat{\alpha}_{MJBRR}) &= E(\hat{\alpha}_{MJBRR}) - \alpha \\ &= -k(\mathbf{I} + k \mathbf{B}^{-1} - k^2 \mathbf{B}^{-2})\mathbf{B}^{-1}\alpha, \end{aligned} \quad (16)$$

$$\begin{aligned} \text{var}(\hat{\alpha}_{MJBRR}) &= E[\hat{\alpha}_{MJBRR} - E(\hat{\alpha}_{MJBRR})][\hat{\alpha}_{MJBRR} - E(\hat{\alpha}_{MJBRR})]^T \\ &= \phi^{-1} \mathbf{H} \mathbf{\Lambda}^{-1} \mathbf{H}^T, \end{aligned} \quad (17)$$

where  $\mathbf{H} = (\mathbf{I} - k^2 \mathbf{B}^{-2})(\mathbf{I} - k \mathbf{B}^{-1})$ , and

$$\begin{aligned} \text{MSE}(\hat{\alpha}_{MJBRR}) &= \text{var}(\hat{\alpha}_{MJBRR}) + \text{Bias}(\hat{\alpha}_{MJBRR})\text{Bias}(\hat{\alpha}_{MJBRR})^T \\ &= \phi^{-1} \mathbf{H} \mathbf{\Lambda}^{-1} \mathbf{H}^T + k(\mathbf{I} + k \mathbf{B}^{-1} - k^2 \mathbf{B}^{-2})\mathbf{B}^{-1}\alpha\alpha^T \mathbf{B}^{-1} \\ &\quad \times k(\mathbf{I} + k \mathbf{B}^{-1} - k^2 \mathbf{B}^{-2})^T. \end{aligned} \quad (18).$$

## 4. SUPERIORITY OF PROPOSED ESTIMATORS IN TERMS OF MSE

In general, to compare the two different estimators, one should be concerned with their performance in terms of MSE. For two given estimators  $\hat{\beta}_1$  and  $\hat{\beta}_2$  of  $\beta$ , the estimator  $\hat{\beta}_2$  is said to be superior to  $\hat{\beta}_1$  under the MSE criterion if and only if (iff)  $\Delta = \text{MSE}(\hat{\beta}_1) - \text{MSE}(\hat{\beta}_2) \geq 0$ .

**Lemma 1:** [21] Let  $\mathbf{G}$  is a  $P \times P$  positive definite matrix,  $a$  is a  $P \times 1$  vector, and  $c$  is a positive constant. Then  $c\mathbf{G} - aa^T$  is a nonnegative definite if and only if  $a^T \mathbf{G}^{-1} a \leq c$  is hold (Sherman and Morrison, 1950).

**Comparison of  $\hat{\alpha}_{MJBRR}$  with  $\hat{\alpha}_{BRR}$**

The bias, variance, and the MSE of  $\hat{\alpha}_{BRR}$  are respectively defined as

$$\begin{aligned} \text{Bias}(\hat{\alpha}_{BRR}) &= E(\hat{\alpha}_{BRR}) - \alpha \\ &= k \mathbf{B}^{-1}\alpha, \\ \text{var}(\hat{\alpha}_{BRR}) &= E[\hat{\alpha}_{BRR} - E(\hat{\alpha}_{BRR})][\hat{\alpha}_{BRR} - E(\hat{\alpha}_{BRR})]^T \\ &= \phi^{-1}(\mathbf{I} - k \mathbf{B}^{-1})\mathbf{\Lambda}^{-1}(\mathbf{I} - k \mathbf{B}^{-1})^T, \end{aligned}$$

and

$$\begin{aligned} \text{MSE}(\hat{\alpha}_{\text{BRR}}) &= \text{var}(\hat{\alpha}_{\text{BRR}}) + \text{Bias}(\hat{\alpha}_{\text{BRR}})\text{Bias}(\hat{\alpha}_{\text{BRR}})^T \\ &= \phi^{-1}(\mathbf{I} - k\mathbf{B}^{-1})\mathbf{\Lambda}^{-1}(\mathbf{I} - k\mathbf{B}^{-1})^T + k^2\mathbf{B}^{-1}\alpha\alpha^T\mathbf{B}^{-1}. \end{aligned} \quad (19)$$

**Theorem 1:** The proposed estimator  $\hat{\alpha}_{\text{MJBRR}}$  is superior to the  $\hat{\alpha}_{\text{BRR}}$  estimator if and only if (iff)

$$a^T \left[ F(\phi^{-1}\mathbf{J} + k^2\mathbf{B}^{-1}\alpha\alpha^T\mathbf{B}^{-1})F^{-1} \right]^{-1} a \leq 1,$$

where  $\mathbf{J} = (\mathbf{I} - k\mathbf{B}^{-1})\mathbf{\Lambda}^{-1}(\mathbf{I} - k\mathbf{B}^{-1})^T + k^2\mathbf{B}^{-2}(\mathbf{I} + (\mathbf{I} - k^2\mathbf{B}^{-2}))$  and

$$F = \left[ k(\mathbf{I} + k\mathbf{B}^{-1} - k^2\mathbf{B}^{-2})\mathbf{B}^{-1} \right]^{-1}.$$

*proof* From Eq. (18) and Eq. (19) the difference in MSE is

$$\begin{aligned} \Delta 1 &= \text{MSE}(\hat{\alpha}_{\text{BRR}}) - \text{MSE}(\hat{\alpha}_{\text{MJBRR}}) \\ &= \phi^{-1}\mathbf{J} + k^2\mathbf{B}^{-1}\alpha\alpha^T\mathbf{B}^{-1} - k^2\mathbf{\Omega}\mathbf{B}^{-1}\alpha\alpha^T\mathbf{B}^{-1}\mathbf{\Omega}^T, \end{aligned}$$

$\mathbf{\Omega} = (\mathbf{I} + k\mathbf{B}^{-1} - k^2\mathbf{B}^{-2})$  is a positive definite matrix. It is clear that  $\mathbf{J}$  is a positive definite matrix. As a result, the  $\Delta 1$  is a nonnegative definite iff  $F^T \Delta F$  is a nonnegative definite matrix. Thus,

$$F^T \Delta F = F^T (\phi^{-1}\mathbf{J} + k^2\mathbf{B}^{-1}\alpha\alpha^T\mathbf{B}^{-1})F - \alpha\alpha^T.$$

Because  $(\phi^{-1}\mathbf{J} + k^2\mathbf{B}^{-1}\alpha\alpha^T\mathbf{B}^{-1})$  is a symmetric positive definite matrix, using Lemma 1, it is concluded that  $F^T \Delta F$  is nonnegative definite iff  $a^T \left[ F(\phi^{-1}\mathbf{J} + k^2\mathbf{B}^{-1}\alpha\alpha^T\mathbf{B}^{-1})F^{-1} \right]^{-1} a \leq 1$  is satisfied. Consequently, the proof is completed.

**Comparison of  $\hat{\alpha}_{\text{MJBRR}}$  with  $\hat{\alpha}_{\text{JBRR}}$**

The bias, variance, and the MSE of  $\hat{\alpha}_{\text{JBRR}}$  are respectively defined as

$$\begin{aligned} \text{Bias}(\hat{\alpha}_{\text{JBRR}}) &= E(\hat{\alpha}_{\text{JBRR}}) - \alpha \\ &= k^2\mathbf{B}^{-2}\alpha, \end{aligned}$$

$$\begin{aligned} \text{var}(\hat{\alpha}_{\text{JBRR}}) &= E \left[ \hat{\alpha}_{\text{JBRR}} - E(\hat{\alpha}_{\text{JBRR}}) \right] \left[ \hat{\alpha}_{\text{JBRR}} - E(\hat{\alpha}_{\text{JBRR}}) \right]^T \\ &= \phi^{-1}(\mathbf{I} - k^2\mathbf{B}^{-2})\mathbf{\Lambda}^{-1}(\mathbf{I} - k^2\mathbf{B}^{-2})^T, \end{aligned}$$

and

$$\begin{aligned} \text{MSE}(\hat{\alpha}_{\text{JBRR}}) &= \text{var}(\hat{\alpha}_{\text{JBRR}}) + \text{Bias}(\hat{\alpha}_{\text{JBRR}})\text{Bias}(\hat{\alpha}_{\text{JBRR}})^T \\ &= \phi^{-1}(\mathbf{I} - k^2\mathbf{B}^{-2})\mathbf{\Lambda}^{-1}(\mathbf{I} - k^2\mathbf{B}^{-2})^T + k^4\mathbf{B}^{-2}\alpha\alpha^T\mathbf{B}^{-2}. \end{aligned} \quad (20)$$

**Theorem 2:** The proposed estimator  $\hat{\alpha}_{\text{MJBRR}}$  is superior to the  $\hat{\alpha}_{\text{JBRR}}$  estimator if and only if (iff)

$$a^T \left[ F(\phi^{-1}\mathbf{L} + k^4\mathbf{B}^{-2}\alpha\alpha^T\mathbf{B}^{-2})F^{-1} \right]^{-1} a \leq 1,$$

$$\mathbf{L} = (\mathbf{I} - k^2\mathbf{B}^{-2})\mathbf{\Lambda}^{-1}(\mathbf{I} - k^2\mathbf{B}^{-2})^T \left[ k\mathbf{B}^{-1}(\mathbf{I} + (\mathbf{I} - k\mathbf{B}^{-1})) \right].$$

where

*proof* From Eq. (18) and Eq. (20) the difference in MSE is

$$\begin{aligned} \Delta 2 &= \text{MSE}(\hat{\alpha}_{\text{JBRR}}) - \text{MSE}(\hat{\alpha}_{\text{MJBRR}}) \\ &= \phi^{-1}\mathbf{L} + k^4\mathbf{B}^{-2}\alpha\alpha^T\mathbf{B}^{-2} - k^2\mathbf{\Omega}\mathbf{B}^{-1}\alpha\alpha^T\mathbf{B}^{-1}\mathbf{\Omega}^T, \end{aligned}$$

$\mathbf{\Omega} = (\mathbf{I} + k\mathbf{B}^{-1} - k^2\mathbf{B}^{-2})$  is a positive definite matrix. It is clear that  $\mathbf{L}$  is a positive definite matrix. As a result, the  $\Delta 2$  is a nonnegative definite iff  $F^T \Delta F$  is a nonnegative definite matrix. Thus,

$$F^T \Delta F = F^T (\phi^{-1}\mathbf{L} + k^4\mathbf{B}^{-2}\alpha\alpha^T\mathbf{B}^{-2})F - \alpha\alpha^T.$$

Because  $(\phi^{-1}\mathbf{L} + k^4 \mathbf{B}^{-2}\alpha\alpha^T \mathbf{B}^{-2})$  is a symmetric positive definite matrix, using Lemma 1, it is concluded that  $F^T \Delta F$  is nonnegative definite iff  $a^T \left[ F(\phi^{-1}\mathbf{L} + k^4 \mathbf{B}^{-2}\alpha\alpha^T \mathbf{B}^{-2})F^{-1} \right]^{-1} a \leq 1$  is satisfied. Consequently, the proof is completed.

## 5. SIMULATION STUDY

In this section, the estimators' performance is investigated through a simulation study. Correlated explanatory variables are generated in line with the work of Kibria and Lukman [20]:

$$x_{ij} = (1 - \rho^2)^{1/2} \varpi_{ij} + \rho \varpi_{i,p+1}, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, p \quad (21)$$

where  $\varpi_{ij}$  are independent standard normal pseudo-random numbers and  $\rho = 0.90, 0.99$  and  $0.999$  are the degree of correlation between any two explanatory variables. The  $n$  observations for the response variable  $y$  are determined by the following equation:

$$y_i \sim B(\mu_i, \phi) \quad (22)$$

where  $\mu_i = \exp(x_i' \beta) / (1 + \exp(x_i' \beta))$  for  $i = 1, 2, \dots, n$ . The regression parameter  $\beta$  is

$$\sum_{i=1}^p \hat{\beta}_i^2 = 1$$

chosen such that [21, 12]. The sample sizes  $n$  are 30, 50 and 100, while the number of explanatory variables is taken to be 4 and 8. The simulation study is conducted by adopting the R programming language with the help of `betareg` package [28]. Following Qasim et al. [27], the MSE and the median squared error (MdSE) were employed to evaluate the estimators' performance:

$$MSE(\beta^*) = \frac{1}{2000} \sum_{j=1}^{2000} (\beta_j^* - \beta)' (\beta_j^* - \beta) \quad (23)$$

$$MdSE(\beta^*) = Median \left[ (\beta_j^* - \beta)' (\beta_j^* - \beta) \right]_j^{2000} \quad (24)$$

where  $j$  is the number of replications which is set to be 2000 and  $\beta_j^*$  is the estimated vector of  $\beta$  in the  $j$ th replication.

The initial values for parameter vector  $\beta$  and  $\phi$  are estimated by following Ferrari and Cribari-Neto [13], Abonazel and Taha [3], and Qasim et al. [27], in the `betareg` package.

The biasing parameter for beta ridge and Jackknife beta ridge was determined using the following biasing parameters:

$$\hat{k}_1 = \frac{p\hat{\phi}}{\sum_{j=1}^p \hat{\beta}_j^2} \quad (25)$$

$$\hat{k}_2 = \sqrt{\frac{p\hat{\phi}}{\sum_{j=1}^p \hat{\beta}_j^2}} \quad (26)$$

The simulation results in terms of the MSE and MdSE are presented in Table 1 and 2. Table 1 displays the results when  $p = 4$ ,  $\phi = 1$  &  $5$ ,  $\rho = 0.90, 0.99$  &  $0.999$ , and  $n = 30, 50, \& 100$ . While the results when  $p = 8$ , with the same values of  $\phi, \rho$ , and  $n$ , are presented in Table 2.

The results from Table 1 and 2 show that the Jackknife Beta ridge regression with biasing parameter  $k_2$  has the smallest mean and median squared error for all the considered sample sizes. The maximum likelihood estimator possesses the least performance due to the presence of multicollinearity. The ridge regression estimator outperforms the maximum likelihood estimator. It is also evident from the result that the median squared error exhibit smaller values than the mean squared error. The result in this study agrees with

the conclusion made in the context of the linear regression model. The jackknife version of the ridge estimator produces an estimator with smaller MSE and MdSE. When the values of  $\rho$  and  $\phi$  increases the MSE and MdSE also increases but decreases as the sample sizes increases from 30 to 100. In general, the proposed estimator performs better than the other estimators for all simulated cases.

We observed that the MSE and MdSE values of all the estimators increased as  $p$  increases. The Jackknife beta ridge estimator outperform BR and BRR under all specifications Overall, the improvement in the proposed over the beta ridge regression is significant.

**Table 1: Simulated MSE and MdSE of the estimators when p=4**

Estimators	$\phi = 1$						$\phi = 5$					
	$\rho = 0.90$		$\rho = 0.99$		$\rho = 0.999$		$\rho = 0.90$		$\rho = 0.99$		$\rho = 0.999$	
	MSE	MdSE	MSE	MdSE	MSE	MdSE	MSE	MdSE	MSE	MdSE	MSE	MdSE
<b>n = 30</b>												
BR	30.58 6	23.41 7	289.99 6	225.03 9	2942.37 4	2282.07 8	40.11 2	30.49 5	396.44 6	305.27 6	4072.00 7	3202.01 8
BRR1	29.38 2	22.19 4	278.03 8	212.13 3	2821.53 3	2138.16 1	36.93 6	26.79 6	364.74 0	272.92 6	3757.09 0	2852.53 9
BRR2	29.40 1	22.29 6	256.88 0	193.16 3	2098.05 3	1449.93 3	33.68 3	24.01 2	245.55 9	161.49 4	1329.63 6	733.566 7
JBRR1	29.33 5	22.17 0	277.54 8	211.40 2	2816.60 2	2137.15 6	36.74 0	26.54 4	362.74 9	270.92 6	3737.74 7	2822.82 1
JBRR2	29.36 8	22.27 4	254.57 7	190.11 4	1989.67 3	1331.21 4	33.12 5	23.41 1	223.79 4	141.48 7	946.013	395.758
<b>n = 50</b>												
BR	15.99 8	12.82 8	150.22 7	122.69 9	1504.98 3	1160.33 7	22.59 1	17.71 1	221.17 1	170.77 8	2215.90 3	1696.67 2
BRR1	15.32 1	12.06 8	143.17 8	114.72 2	1433.36 0	1079.62 2	20.87 1	15.86 7	203.11 0	152.04 2	2030.64 9	1498.33 9
BRR2	15.47 8	12.26 4	134.69 7	105.37 2	1100.90 8	781.109	19.70 9	14.93 7	148.02 4	100.22 7	801.490	452.185
JBRR1	15.29 5	12.03 2	142.88 7	114.33 6	1430.32 9	1077.08 8	20.76 8	15.79 2	201.95 8	150.89 4	2018.86 1	1487.48 9
JBRR2	15.46 6	12.25 3	133.76 0	104.37 9	1051.64 6	721.596	19.50 3	14.68 3	138.10 7	89.916	592.430	267.100
<b>n = 100</b>												
BR	15.24 6	11.82 9	150.64 3	121.29 4	1516.27 6	1157.22 6	21.60 9	17.30 0	215.72 9	164.71 7	2176.91 2	1736.25 4
BRR1	14.54 1	11.13 5	143.19 9	112.91 6	1445.69 2	1079.24 8	19.78 5	15.29 6	196.79 5	144.60 9	1994.67 1	1536.78 0
BRR2	14.69 9	11.30 4	133.81 5	104.53 3	1099.87 9	757.181	18.51 8	14.20 9	137.71 7	91.772	740.585	424.652
JBRR1	14.51 2	11.11 3	142.88 4	112.72 8	1442.75 4	1076.49 3	19.67 2	15.16 4	195.57 9	143.36 4	1983.35 2	1521.91 2
JBRR2	14.68 7	11.29 2	132.78 7	103.25 9	1049.48 2	701.777	18.28 7	14.00 6	126.72 3	80.106	524.170	246.487

**Table 2: Simulated MSE and MdSE of the estimators when p=8**

Estimators	$\phi = 1$						$\phi = 5$					
	$\rho = 0.90$		$\rho = 0.99$		$\rho = 0.999$		$\rho = 0.90$		$\rho = 0.99$		$\rho = 0.999$	
	MSE	MdSE	MSE	MdSE	MSE	MdSE	MSE	MdSE	MSE	MdSE	MSE	MdSE
<b>n = 30</b>												
BR	81.30 5	66.29 1	822.32 8	678.71 2	8432.46 7	7146.82 6	103.40 3	86.73 4	1053.46 7	867.26 5	10509.4 20	8847.77 3
BRR1	77.70 1	63.19 5	785.94 6	642.88 9	8059.13 0	6757.85 6	94.143	77.81 3	963.408	779.32 0	9602.44 3	7914.73 0
BRR2	77.50 2	63.24 8	715.79 5	583.29 9	5860.64 3	4722.80 2	84.202	69.20 6	636.187	491.23 5	3502.55 0	2437.06 2
JBRR1	77.54 1	63.15 6	784.44 1	642.11 7	8043.53 9	6739.79 0	93.496	77.28 6	957.439	774.15 9	9542.86 7	7794.37 7
JBRR2	77.34 5	63.05 6	706.36 7	570.46 3	5506.20 3	4370.40 2	82.139	66.90 3	576.019	434.45 5	2661.35 3	1682.29 2
<b>n = 50</b>												
BR	49.80 9	43.52 7	494.12 9	434.88 5	4964.02 5	4483.55 9	66.895	57.74 2	651.306	559.32 3	6685.72 7	5758.25 9
BRR1	48.03 8	41.36 4	476.43 5	418.96 8	4786.66 1	4294.84 2	62.228	52.83 7	605.365	511.67 8	6220.53 2	5324.82 3

<b>BRR2</b>	48.19 8	41.70 5		446.85 6	390.23 7		3719.03 6	3185.61 7		57.909	49.39 2		435.625	353.62 3		2479.65 8	1858.98 2
<b>JBRR1</b>	47.98 7	41.31 5		475.93 7	418.49 4		4781.60 5	4289.20 5		61.999	52.69 1		603.120	509.92 8		6198.03 3	5309.11 4
<b>JBRR2</b>	48.16 2	41.63 6		444.11 2	387.71 0		3570.72 4	2999.81 1		57.244	48.93 5		406.105	325.17 8		1856.71 1	1247.04 1
<b>n = 100</b>																	
<b>BR</b>	45.62 9	41.49 5		446.49 8	399.97 4		4451.45 2	3938.16 3		63.153	56.45 4		629.173	568.14 6		6365.43 2	5794.77 0
<b>BRR1</b>	43.75 4	39.71 1		427.68 8	381.14 1		4262.34 8	3765.36 6		58.315	51.62 8		580.080	516.48 4		5873.27 7	5302.99 3
<b>BRR2</b>	43.86 2	39.65 7		395.26 8	349.16 9		3129.50 7	2697.43 7		53.599	46.84 8		393.772	336.24 4		2026.01 3	1590.71 5
<b>JBRR1</b>	43.69 7	39.66 6		427.11 4	380.54 6		4256.68 4	3761.80 3		58.067	51.42 1		577.581	514.17 9		5848.56 5	5277.01 7
<b>JBRR2</b>	43.81 9	39.62 4		392.01 4	346.18 6		2959.99 0	2535.45 2		52.841	46.19 2		359.924	302.36 3		1397.91 4	1008.48 2

**6. GASOLINE YIELD DATA**

This chemical dataset was originally obtained by Prater [26], and later used by the following authors: Ospina et al. [25], Karlsson et al. [21]. The dataset contains 32 observations on the response and on the independent variables. The variables in the study are described as follows: The dependent variable *y* is the proportion of crude oil after distillation and fractionation while the explanatory variables are crude oil gravity (gravity), vapor pressure of crude oil (pressure), temperature at which 10% of the crude oil has vaporized (temp10) and temperature at which all petrol in the amount of crude oil vaporizes (temp). Atkinson [5] analyzed this dataset using the linear regression model and observed some anomaly in the distribution of the error. Recently, Karlsson et al. [21] shows that the beta regression model is more suitable to model the data. We used the condition index to diagnose if the model is multicollinearity free. Condition index (CI) is defined as follows:

$$CI = \sqrt{\frac{\text{max eigenvalue}}{\text{min eigenvalue}}}$$

(27)

According to literature if the CI exceeds 1000, there is severe multicollinearity. The CI for the dataset under study is 11281.4 which signal severe multicollinearity. The Beta regression estimates for each of the estimation techniques and their corresponding mean squared error are shown in Table 3.

Table 3: The estimated coefficients and MSE values

Coefficient	BR	BRR1	BRR2	JBRR1	JBRR2
Intercept	-2.6949	-0.0280	-0.3142	-0.0697	-0.0017
Gravity	0.0045	-0.0118	-0.0096	-0.0111	-0.0120
Pressure	0.0304	-0.0385	-0.0331	-0.0396	-0.0391
temp10	-0.0110	-0.0182	-0.0175	-0.0182	-0.0183
temp	0.0106	0.0106	0.0106	0.0106	0.0106
MSE	144.5621	0.13019	2.071377	0.193196	0.08801

The result show that the ridge and Jackknife ridge regression estimates possess the same regression coefficient sign except the maximum likelihood estimate. This is traceable to one of the consequences of multicollinearity on the maximum likelihood estimator as mentioned by Lukman and Ayinde [22]. The proposed method has the lowest mean squared error. This agree with the literature that the jackknife ridge estimator produced a reduced bias which in turn result to a lower mean squared error than the Beta ridge estimator. The estimator performance is also a function of the Biasing parameter. It is also evident that jackknifing the ridge estimator reduced the mean squared error of the Beat ridge estimator.

**7. CONCLUSIONS**

Multicollinearity is a threat to parameter estimation and inferences in the linear regression model and the generalized regression model such as the Beta regression model (BRM). The regression parameters are popularly estimated using the method of maximum likelihood. However, the performance of the maximum likelihood suffers setbacks when there is multicollinearity. The Beta ridge estimator produces a more reliable

estimate with minimum mean squared error than the earlier mentioned estimator. The limitation of the Beta ridge estimator is high bias. An effective means of reducing the bias is to employ the Jackknife procedure. In this study, we proposed the Jackknife Beta ridge and Modified Jackknife Beta ridge estimators. We compared the estimators' performance in the following ways: theoretically, simulation study and real-life application. In conclusion, the new methods generally produced a more efficient estimates when there is multicollinearity in the BRM.

**Funding:** This research received no external funding.

**Conflict of interest:** The authors declare no conflicts of interest.

## REFERENCES

- [1] Abonazel, M. R. (2019). New ridge estimators of SUR model when the errors are serially correlated. *International Journal of Mathematical Archive*, 10(7), 53-62.
- [2] Abonazel, M. R., & Farghali, R. A. (2019). Liu-type multinomial logistic estimator. *Sankhya B*, 81(2), 203-225.
- [3] Abonazel, M. R., & Taha, I. (2021). Beta Ridge Regression Estimators: Simulation and Application. *Communications in Statistics-Simulation and Computation*. 1-20. <https://doi.org/10.1080/03610918.2021.1960373>
- [4] Algamal, Z. Y. (2018). Developing a ridge estimator for the gamma regression model. *Journal of Chemometrics*, 32(10), e305
- [5] Atkinson, A.C. (1985). *Plots, Transformations and Regression: An Introduction to Graphical Methods of Diagnostic Regression Analysis*. New York: Oxford University Press.
- [6] Batah, F. S. M., Ramanathan, T. V., & Gore, S. D. (2008). The efficiency of modified jackknife and ridge type regression estimators: a comparison. *Surveys in Mathematics and its Applications*, 3, 111-122.
- [7] Duran, E. A., & Akdeniz, F. (2012). Efficiency of the modified jackknifed Liu-type estimator. *Statistical Papers*, 53(2), 265-280.
- [8] Espinheira, P. L., da Silva, L. C. M., & Silva, A. D. O. (2015). Prediction measures in beta regression models. arXiv preprint arXiv:1501.04830.
- [9] Espinheira, P. L., da Silva, L. C. M., Silva, A. D. O., & Ospina, R. (2019). Model selection criteria on beta regression for machine learning. *Machine Learning and Knowledge Extraction*, 1(1), 427-449.
- [10] Espinheira, P. L., Ferrari, S. L., & Cribari-Neto, F. (2008). On beta regression residuals. *Journal of Applied Statistics*, 35(4), 407-419.
- [11] Farebrother, R. W. (1976). Further results on the mean square error of ridge regression. *Journal of the Royal Statistical Society. Series B (Methodological)*, 248-250.
- [12] Farghali, R. A., Qasim, M., Kibria, B. G., & Abonazel, M. R. (2021). Generalized two-parameter estimators in the multinomial logit regression model: methods, simulation and application. *Communications in Statistics-Simulation and Computation*, 1-16. <https://doi.org/10.1080/03610918.2021.1934023>
- [13] Ferrari, S., & Cribari-Neto, F. (2004). Beta regression for modelling rates and proportions. *Journal of applied statistics*, 31(7), 799-815.
- [14] Frisch, R. (1934). *Statistical confluence analysis by means of complete regression systems*. University Institute of Economics. Cambridge University Press.
- [15] Gruber, M. H. (1991). The efficiency of jack-knifed and usual ridge type estimators: A comparison. *Statistics & probability letters*, 11(1), 49-51.
- [16] Hinkley, D. V. (1977). Jackknifing in unbalanced situations. *Technometrics*, 19(3), 285-292.
- [17] Hoerl, A. E., & Kennard, R. W. (1970a). Ridge regression: Biased estimation for non-orthogonal problems. *Technometrics*, 12(1), 55-67.
- [18] Hoerl, A. E., & Kennard, R. W. (1970b). Ridge regression: applications to non-orthogonal problems. *Technometrics*, 12(1), 69-82.
- [19] 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.
- [20] Kibria, B. M. G., & Lukman, A. F. (2020). A new ridge-type estimator for the linear regression model: Simulations and applications. *Scientifica*. <https://doi.org/10.1155/2020/9758378>
- [21] Karlson, P., Mansson, K., & Kibria, B. M. G. (2020). A Liu estimator for the beta regression model and its application to chemical data. *Journal of Chemometrics*, 34: e3300. <https://doi.org/10.1002/cem.3300>
- [22] Lukman, A. F., & Ayinde, K. (2017). Review and Classifications of the ridge parameter estimation techniques. *Hacettepe Journal of Mathematics and Statistics*, 46(5), 953-967.
- [23] Lukman, A. F., Aladeitan, B., Ayinde, K., & Abonazel, M. R. (2021). Modified ridge-type for the Poisson regression model: simulation and application. *Journal of Applied Statistics*, 1-13.
- [24] Nyquist, H. (1988). Applications of the jackknife procedure in ridge regression. *Computational Statistics & Data Analysis*, 6(2), 177-183.
- [25] Ospina R., Cribari-Neto, F., & Vasconcellos, K. (2006). Improved point and interval estimation for a beta regression model. *Comput Stat Data An.*, 51(2):960-981.
- [26] Prater, N. H. (1956). Estimate gasoline yields from crude. *PetroLium Refiner*. 35:236-238.
- [27] Qasim, M., Månsson, K., & Golam Kibria, B. M. (2021). On some beta ridge regression estimators: method, simulation and application. *Journal of Statistical Computation and Simulation*, 1-14.

- 
- [28] R Core Team (2020). R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <https://www.R-project.org>.
- [29] Rady, E. A., Abonazel, M. R., & Taha, I. M. (2019). A new biased estimator for zero-inflated count regression models. *Journal of Modern Applied Statistical Methods*, Accepted paper (July 2019).
- [30] Sherman, J., & Morrison, W. J. (1950). Adjustment of an inverse matrix corresponding to a change in one element of a given matrix. *The Annals of Mathematical Statistics*, 21(1), 124-127.
- [31] Singh, B., Chaubey, Y. P., & Dwivedi, T. D. (1986). An almost unbiased ridge estimator. *Sankhyā: The Indian Journal of Statistics, Series B*, 342-346.
- [32] Türkan, S., & Özel, G. (2016). A new modified Jackknifed estimator for the Poisson regression model. *Journal of Applied Statistics*, 43(10), 1892-1905.
- [33] Türkan, S., & Özel, G. (2018). A Jackknifed estimators for the negative binomial regression model. *Communications in Statistics-Simulation and Computation*, 47(6), 1845-1865.
- [34] Yasin, A. S. A. R., Karabrahimoğlu, A., & Aşır, G. E. N. Ç. (2013). Modified ridge regression parameters: A comparative Monte Carlo study. *Hacettepe Journal of Mathematics and Statistics*, 43(5), 827-841.
- [35] Yıldız, N. (2018). On the performance of the Jackknifed Liu-type estimator in linear regression model. *Communications in Statistics-Theory and Methods*, 47(9), 2278-2290.