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

Zakariya Yahya Algamal[1], Mohamed R. Abonazel[2], Adewale F. Lukman[3]

[1]Department of Statistics and Informatics, University of Mosul, Mosul, Iraq; zakariya.algamal@uomosul.edu.iq
[2]Department of Applied Statistics and Econometrics, Faculty of Graduate Studies for Statistical Research, Cairo University, Egypt; mabonazel@cu.edu.eg
[3]Department of Epidemiology and Biostatistics, University of Medical Sciences, Ondo, Nigeria; adewale.folaranmi@lmu.edu.ng

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:

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

for $0 < y < 1$; $0 < \mu < 1$; $\phi > 0$, and $\Gamma(.)$ is the gamma function, and ϕ 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 μ_i, depending on covariates as follows:

$$g(\mu_i) = \log\left(\frac{\mu_i}{1-\mu_i}\right) = x_i^T\beta = \eta_i, \tag{2}$$

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

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

Differentiating the log-likelihood in Eq. (3) with respect to β gives us the score function for β which is given by:

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

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

$$\beta^{(r+1)} = \beta^{(r)} + \left(I_{\beta\beta}^{(r)}\right)^{-1} S_\beta^{(r)}(\beta),$$

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

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

where $\hat{\mu}$ 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 β is obtained as:

$$\hat{\beta}_{BR} = \left(X^T\hat{W}X\right)^{-1}X^T\hat{W}\hat{z}, \tag{6}$$

where X is an $n \times p$ matrix of regressors, $\hat{z} = \hat{\eta} + \hat{W}^{-1}\hat{A}(y^* - \mu^*)$, and $\hat{W} = \text{diag}(\hat{w}_1, \dots, \hat{w}_n)$;
$$\hat{w}_i = \frac{(1-\hat{\sigma}_i^2)}{\hat{\sigma}_i^2}\left\{\psi'\left(\frac{\hat{\mu}_i(1-\hat{\sigma}_i^2)}{\hat{\sigma}_i^2}\right) + \psi'\left(\frac{(1-\hat{\mu}_i)(1-\hat{\sigma}_i^2)}{\hat{\sigma}_i^2}\right)\right\}\frac{1}{\{g'(\hat{\mu}_i)\}^2}.$$

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

$$\text{Cov}(\hat{\beta}_{BR}) = \frac{1}{\phi}\left(X^T\hat{W}X\right)^{-1}. \tag{7}$$

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

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

where λ_j is the eigenvalue of the $X^T\hat{W}X$ matrix.

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

$$\hat{\beta}_{\text{BRR}} = \left(X^T \widehat{W} X + k\mathrm{I}\right)^{-1} X^T \widehat{W} \hat{z}, \tag{9}$$

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 $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_p) = Q^T X^T \widehat{W} X Q$, 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 $X^T \widehat{W} X$. Then the MSE of the BRR estimator is:

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

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 $\mathbf{\Lambda}$ and $\mathbf{Q}$ matrices, the beta regression estimator of Eq. (6) can be re-written as

$$\hat{\boldsymbol{\beta}}_{\text{BR}} = \mathbf{Q}\hat{\alpha}_{\text{BR}}, \tag{11}$$

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

$$\hat{\alpha}_{\text{BRR}} = (\mathbf{\Lambda} + k\,\mathbf{I})^{-1}\mathbf{M}^T \hat{\mathbf{W}}\hat{z}. \tag{12}$$

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

$$\hat{\alpha}_{\text{BRR}(-i)} = (\mathbf{M}_{(-i)}^T \hat{\mathbf{W}}_{(-i)}\mathbf{M}_{(-i)} + k\,\mathbf{I})^{-1}\mathbf{M}_{(-i)}^T \hat{\mathbf{W}}_{(-i)}\hat{z}_{(-i)}, \tag{13}$$

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

$$(\mathbf{M}_{(-i)}^T \hat{\mathbf{W}}_{(-i)}\mathbf{M}_{(-i)} + k\,\mathbf{I})^{-1} = (\mathbf{M}^T \hat{\mathbf{W}}\mathbf{M} + k\,\mathbf{I})^{-1}$$
$$+ \frac{(\mathbf{M}^T \hat{\mathbf{W}}\mathbf{M} + k\,\mathbf{I})^{-1}\mathbf{m}_i \mathbf{m}_i^T (\mathbf{M}^T \hat{\mathbf{W}}\mathbf{M} + k\,\mathbf{I})^{-1}}{1 - \mathbf{m}_i^T (\mathbf{M}^T \hat{\mathbf{W}}\mathbf{M} + k\,\mathbf{I})^{-1}\mathbf{m}_i},$$

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

$$\hat{\alpha}_{\text{BRR}(-i)} = \hat{\alpha}_{\text{BRR}} - \frac{(\mathbf{M}^T \hat{\mathbf{W}}\mathbf{M} + k\,\mathbf{I})^{-1}\mathbf{m}_i^T (\hat{z}_i - \mathbf{m}_i^T \hat{\alpha}_{\text{BRR}})}{1 - \mathbf{m}_i^T (\mathbf{M}^T \hat{\mathbf{W}}\mathbf{M} + k\,\mathbf{I})^{-1}\mathbf{m}_i}.$$

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

$$T_i = \hat{\alpha}_{\text{BRR}} + n(1 - \mathbf{m}_i^T (\mathbf{M}^T \hat{\mathbf{W}}\mathbf{M} + k\,\mathbf{I})^{-1}\mathbf{m}_i)(\hat{\alpha}_{\text{BRR}} - \hat{\alpha}_{\text{BRR}(-i)}).$$

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

$$\hat{\alpha}_{\text{JBRR}} = \hat{\alpha}_{\text{BRR}} + (\mathbf{M}^T \hat{\mathbf{W}}\mathbf{M} + k\,\mathbf{I})^{-1}\sum_{i=1}^{n}\mathbf{m}_i^T (\hat{z}_i - \mathbf{m}_i^T \hat{\alpha}_{\text{BRR}}).$$

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

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

(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 $\hat{\boldsymbol{\tau}}_{\text{JBRR}}$. Following Algamal [4], the proposed estimator, which is denoted by modified Jackknifed beta ridge estimator (MJBRR), is designated as in the case of the $\hat{\alpha}_{\text{JBRR}}$ by replacing $\hat{\alpha}_{\text{BR}}$ with $\hat{\alpha}_{\text{BRR}}$. The MJBRR estimator is defined as:

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

(15)

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

$$\text{Bias}\left(\hat{\alpha}_{\text{MJBRR}}\right) = E\left(\hat{\alpha}_{\text{MJBRR}}\right) - \alpha$$

$$= -k\left(\mathbf{I} + k\,B^{-1} - k^2 B^{-2}\right)B^{-1}\alpha,$$

(16)

$$\text{var}\left(\hat{\alpha}_{\text{MJBRR}}\right) = E\left[\hat{\alpha}_{\text{MJBRR}} - E(\hat{\alpha}_{\text{MJBRR}})\right]\left[\hat{\alpha}_{\text{MJBRR}} - E(\hat{\alpha}_{\text{MJBRR}})\right]^T$$

$$= \phi^{-1}\mathbf{H}\boldsymbol{\Lambda}^{-1}\mathbf{H}^T,$$

(17)

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

$$\text{MSE}\left(\hat{\alpha}_{\text{MJBRR}}\right) = \text{var}\left(\hat{\alpha}_{\text{MJBRR}}\right) + \text{Bias}\left(\hat{\alpha}_{\text{MJBRR}}\right)\text{Bias}\left(\hat{\alpha}_{\text{MJBRR}}\right)^T$$

$$= \phi^{-1}\mathbf{H}\boldsymbol{\Lambda}^{-1}\mathbf{H}^T + k\left(\mathbf{I} + k\,B^{-1} - k^2 B^{-2}\right)B^{-1}\alpha\alpha^T B^{-1}$$

$$\times k\left(\mathbf{I} + k\,B^{-1} - k^2 B^{-2}\right)^T.$$

(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 $\hat{\boldsymbol{\beta}}_1$ and $\hat{\boldsymbol{\beta}}_2$ of $\boldsymbol{\beta}$, the estimator $\hat{\boldsymbol{\beta}}_2$ is said to be superior to $\hat{\boldsymbol{\beta}}_1$ under the MSE criterion if and only if (iff) $\Delta = \text{MSE}(\hat{\boldsymbol{\beta}}_1) - \text{MSE}(\hat{\boldsymbol{\beta}}_2) \geq 0$.

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 $\hat{\alpha}_{\text{MJBRR}}$ with $\hat{\alpha}_{\text{BRR}}$

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

$$\text{Bias}\left(\hat{\alpha}_{\text{BRR}}\right) = E\left(\hat{\alpha}_{\text{BRR}}\right) - \alpha$$

$$= k\,B^{-1}\alpha,$$

$$\text{var}\left(\hat{\alpha}_{\text{BRR}}\right) = E\left[\hat{\alpha}_{\text{BRR}} - E(\hat{\alpha}_{\text{BRR}})\right]\left[\hat{\alpha}_{\text{BRR}} - E(\hat{\alpha}_{\text{BRR}})\right]^T$$

$$= \phi^{-1}(\mathbf{I} - k\,B^{-1})\boldsymbol{\Lambda}^{-1}(\mathbf{I} - k\,B^{-1})^T,$$

and

$$\mathrm{MSE}(\hat{\alpha}_{\mathrm{BRR}}) = \mathrm{var}(\hat{\alpha}_{\mathrm{BRR}}) + \mathrm{Bias}(\hat{\alpha}_{\mathrm{BRR}})\mathrm{Bias}(\hat{\alpha}_{\mathrm{BRR}})^T$$

$$= \phi^{-1}(\mathbf{I} - k\,B^{-1})\mathbf{\Lambda}^{-1}(\mathbf{I} - k\,B^{-1})^T + k^2 B^{-1}\alpha\,\alpha^T B^{-1}. \quad (19)$$

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

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

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

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

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

$$\Delta 1 = \mathrm{MSE}(\hat{\alpha}_{\mathrm{BRR}}) - \mathrm{MSE}(\hat{\alpha}_{\mathrm{MJBRR}})$$

$$= \phi^{-1}\mathbf{J} + k^2 B^{-1}\alpha\,\alpha^T B^{-1} - k^2 \mathbf{\Omega}B^{-1}\alpha\,\alpha^T B^{-1}\mathbf{\Omega}^T,$$

$\mathbf{\Omega} = (\mathbf{I} + k\,B^{-1} - k^2 B^{-2})$ 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,

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

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

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

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

$$\mathrm{Bias}(\hat{\alpha}_{\mathrm{JBRR}}) = E(\hat{\alpha}_{\mathrm{JBRR}}) - \alpha$$

$$= k^2 B^{-2}\alpha,$$

$$\mathrm{var}(\hat{\alpha}_{\mathrm{JBRR}}) = E\left[\hat{\alpha}_{\mathrm{JBRR}} - E(\hat{\alpha}_{\mathrm{JBRR}}) \right]\left[\hat{\alpha}_{\mathrm{JBRR}} - E(\hat{\alpha}_{\mathrm{JBRR}}) \right]^T$$

$$= \phi^{-1}(\mathbf{I} - k^2 B^{-2})\mathbf{\Lambda}^{-1}(\mathbf{I} - k^2 B^{-2})^T,$$

and

$$\mathrm{MSE}(\hat{\alpha}_{\mathrm{JBRR}}) = \mathrm{var}(\hat{\alpha}_{\mathrm{JBRR}}) + \mathrm{Bias}(\hat{\alpha}_{\mathrm{JBRR}})\mathrm{Bias}(\hat{\alpha}_{\mathrm{JBRR}})^T$$

$$= \phi^{-1}(\mathbf{I} - k^2 B^{-2})\mathbf{\Lambda}^{-1}(\mathbf{I} - k^2 B^{-2})^T + k^4 B^{-2}\alpha\,\alpha^T B^{-2}. \quad (20)$$

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

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

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

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

$$\Delta 2 = \mathrm{MSE}(\hat{\alpha}_{\mathrm{JBRR}}) - \mathrm{MSE}(\hat{\alpha}_{\mathrm{MJBRR}})$$

$$= \phi^{-1}\mathbf{L} + k^4 B^{-2}\alpha\alpha^T B^{-2} - k^2 \mathbf{\Omega}B^{-1}\alpha\alpha^T B^{-1}\mathbf{\Omega}^T,$$

$\mathbf{\Omega} = (\mathbf{I} + k\,B^{-1} - k^2 B^{-2})$ 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,

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

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

5. SIMULATION STUDY

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

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

where ϖ_{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:

$$y_i \sim B(\mu_i,\phi) \tag{22}$$

where $\mu_i = \exp(x_i'\beta)/(1+\exp(x_i'\beta))$ for $i = 1,2,...,n.$ The regression parameter β is chosen such that $\sum_{i=1}^{p}\hat{\beta}_i^2 = 1$ [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:

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

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

where j is the number of replications which is set to be 2000 and β_j^* is the estimated vector of β in the jth replication.

The initial values for parameter vector β and ϕ 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:

$$\hat{k}_1 = \frac{p\hat{\phi}}{\sum_{j=1}^{p}\hat{\beta}_i^2}$$

$$\hat{k}_2 = \sqrt{\frac{p\hat{\phi}}{\sum_{j=1}^{p}\hat{\beta}_i^2}} \tag{26}$$

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 ρ and ϕ 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
							n = 30					
BR	30.586	23.417	289.999	225.039	2942.374	2282.078	40.112	30.495	396.446	305.276	4072.007	3202.018
BRR1	29.382	22.194	278.038	212.133	2821.533	2138.161	36.936	26.796	364.740	272.926	3757.090	2852.539
BRR2	29.401	22.296	256.880	193.163	2098.053	1449.933	33.683	24.012	245.559	161.494	1329.636	733.566
JBRR1	29.335	22.170	277.548	211.402	2816.602	2137.156	36.740	26.544	362.749	270.926	3737.747	2822.821
JBRR2	29.368	22.274	254.577	190.114	1989.673	1331.214	33.125	23.411	223.794	141.487	946.013	395.758
							n = 50					
BR	15.998	12.828	150.227	122.699	1504.983	1160.337	22.591	17.711	221.171	170.778	2215.903	1696.672
BRR1	15.321	12.068	143.178	114.722	1433.360	1079.622	20.871	15.867	203.110	152.042	2030.649	1498.339
BRR2	15.478	12.264	134.697	105.372	1100.908	781.109	19.709	14.937	148.024	100.227	801.490	452.185
JBRR1	15.295	12.032	142.887	114.336	1430.329	1077.088	20.768	15.792	201.958	150.894	2018.861	1487.489
JBRR2	15.466	12.253	133.760	104.379	1051.646	721.596	19.503	14.683	138.107	89.916	592.430	267.100
							n = 100					
BR	15.246	11.829	150.643	121.294	1516.276	1157.226	21.609	17.300	215.729	164.717	2176.912	1736.254
BRR1	14.541	11.135	143.199	112.916	1445.692	1079.248	19.785	15.296	196.795	144.609	1994.671	1536.780
BRR2	14.699	11.304	133.815	104.533	1099.879	757.181	18.518	14.208	137.717	91.772	740.585	424.652
JBRR1	14.512	11.113	142.884	112.728	1442.754	1076.493	19.672	15.164	195.579	143.364	1983.352	1521.912
JBRR2	14.687	11.292	132.787	103.259	1049.482	701.777	18.287	14.006	126.723	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
							n = 30					
BR	81.305	66.291	822.328	678.712	8432.467	7146.826	103.403	86.734	1053.467	867.265	10509.420	8847.773
BRR1	77.701	63.195	785.946	642.889	8059.130	6757.856	94.143	77.813	963.408	779.320	9602.443	7914.730
BRR2	77.502	63.248	715.795	583.299	5860.643	4722.802	84.202	69.206	636.187	491.235	3502.550	2437.062
JBRR1	77.541	63.156	784.441	642.117	8043.539	6739.790	93.496	77.286	957.439	774.159	9542.867	7794.377
JBRR2	77.345	63.056	706.367	570.463	5506.203	4370.402	82.139	66.903	576.019	434.455	2661.353	1682.292
							n = 50					
BR	49.809	43.527	494.129	434.885	4964.025	4483.559	66.895	57.742	651.306	559.323	6685.727	5758.259
BRR1	48.038	41.364	476.435	418.968	4786.661	4294.842	62.228	52.837	605.365	511.678	6220.532	5324.823

BRR2	48.198	41.705		446.856	390.237		3719.036	3185.617		57.909	49.392		435.625	353.623		2479.658	1858.982
JBRR1	47.987	41.315		475.937	418.494		4781.605	4289.205		61.999	52.691		603.120	509.928		6198.033	5309.114
JBRR2	48.162	41.636		444.112	387.710		3570.724	2999.811		57.244	48.935		406.105	325.178		1856.711	1247.041
	n = 100															
BR	45.629	41.495		446.498	399.974		4451.452	3938.163		63.153	56.454		629.173	568.146		6365.432	5794.770
BRR1	43.754	39.711		427.688	381.141		4262.348	3765.366		58.315	51.628		580.080	516.484		5873.277	5302.993
BRR2	43.862	39.657		395.268	349.169		3129.507	2697.437		53.599	46.848		393.772	336.244		2026.013	1590.715
JBRR1	43.697	39.666		427.114	380.546		4256.684	3761.803		58.067	51.421		577.581	514.179		5848.565	5277.017
JBRR2	43.819	39.624		392.014	346.186		2959.990	2535.452		52.841	46.192		359.924	302.363		1397.914	1008.482

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.

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-proje ct.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., Karaibrahimoğlu, 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.