mathematics

MDPI

Article

K-L Estimator: Dealing with Multicollinearity in the Logistic Regression Model

Adewale F. Lukman [1,*], B. M. Golam Kibria [2], Cosmas K. Nziku [3], Muhammad Amin [4], Emmanuel T. Adewuyi [5] and Rasha Farghali [6]

[1] Department of Epidemiology and Biostatistics, University of Medical Sciences, Ondo 220282, Nigeria
[2] Department of Mathematics and Statistics, Florida International University, Miami, FL 33199, USA
[3] Department of Statistics, University of Dar es Salaam, Dar es Salaam 65015, Tanzania
[4] Department of Statistics, University of Sargodha, Sargodha 40100, Pakistan
[5] Department of Statistics, Ladoke Akintola University of Technology, Ogbomoso 210214, Nigeria
[6] Department of Mathematics, Insurance and Applied Statistics, Helwan University, Cairo 11732, Egypt
* Correspondence: fadewale@unimed.edu.ng

Abstract: Multicollinearity negatively affects the efficiency of the maximum likelihood estimator (MLE) in both the linear and generalized linear models. The Kibria and Lukman estimator (KLE) was developed as an alternative to the MLE to handle multicollinearity for the linear regression model. In this study, we proposed the Logistic Kibria-Lukman estimator (LKLE) to handle multicollinearity for the logistic regression model. We theoretically established the superiority condition of this new estimator over the MLE, the logistic ridge estimator (LRE), the logistic Liu estimator (LLE), the logistic Liu-type estimator (LLTE) and the logistic two-parameter estimator (LTPE) using the mean squared error criteria. The theoretical conditions were validated using a real-life dataset, and the results showed that the conditions were satisfied. Finally, a simulation and the real-life results showed that the new estimator outperformed the other considered estimators. However, the performance of the estimators was contingent on the adopted shrinkage parameter estimators.

Keywords: Kibria-Lukman estimator; logistic regression model; Liu estimator; multicollinearity; ridge regression estimator

MSC: 62JO5; 62J07; 62J12

check for updates

Citation: Lukman, A.F.; Kibria, B.M.G.; Nziku, C.K.; Amin, M.; Adewuyi, E.T.; Farghali, R. K-L Estimator: Dealing with Multicollinearity in the Logistic Regression Model. *Mathematics* **2023**, *11*, 340. https://doi.org/10.3390/math11020340

Academic Editor: Elvira Di Nardo

Received: 11 November 2022
Revised: 23 December 2022
Accepted: 29 December 2022
Published: 9 January 2023

Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Frisch [1] coined the term "multicollinearity" to describe the problem that occurs when the explanatory variables in a model are linearly related. This problem posed a severe threat to different regression models, e.g., the linear regression model (LRM), logistic regression model, Poisson regression model and gamma regression model. The parameters in the linear and logistic regression models are popularly estimated using the ordinary least squares (OLS) estimator and the maximum likelihood estimator (MLE), respectively. However, both estimators with multicollinearity possess high standard error, and occasionally, the estimated regression coefficients exhibit the wrong coefficient signs, making the conclusion doubtful [2,3]. The ridge regression estimator (RRE) and the logistic ridge estimator are notable alternatives to the OLS estimator in the LRM and the logistic regression model [4,5]. The Liu estimator is an alternative to the ridge estimator which accounts for multicollinearity in the LRM and the logistic regression model [6,7]. The modified ridge-type estimator is a two-parameter estimator that competes favorably with the ridge and Liu estimators [8,9]. Recently the K-L estimator emerged as another estimator in the ridge class and the Liu estimator with a single biasing parameter class [10]. The K-L estimator is a form of the Liu-type estimator with one parameter that minimizes the residual sum of squares with respect to the L2 norm with a prior information. The K-L

estimator outperforms the RRE and Liu estimators based on the theoretical conditions. In this study, we developed the K-L estimator for parameter estimation with the logistic regression model, derived its statistical properties, performed a theoretical comparison with other estimators, and validated its performance by a simulation and a real-life application.

The organization of this paper is as follows. The proposed estimator is discussed in Section 2. A theoretical comparison of various estimators is presented in Section 3. A simulation study is conducted in Section 4. Real-life data are analyzed in Section 5. Finally, some concluding remarks are given in Section 6.

2. Proposed Estimator

Given that y_i is a binary response variable, then the logistic regression model is defined as a Bernoulli distribution: $y_i \sim Be(\pi_i)$.

$$p(y_i) = \pi_i{}^{y_i}(1-\pi_i)^{(1-y_i)} \tag{1}$$

where $\pi_i = \frac{e^{x_i^T\beta}}{1+e^{x_i^T\beta}} = \frac{1}{1+e^{-(x_i^T\beta)}}, i = 1,2,\ldots,n$ and x_i is the ith row of X, which is an $n \times (p+1)$ matrix of explanatory variables, β is a $(p+1) \times 1$ vector of regression coefficients and $y_i \sim Be(\pi_i)$. The parameters in the logistic regression model are estimated by the method of MLE. The MLE of β is

$$\hat{\beta}_{MLE} = \left(X^T\hat{G}_nX\right)^{-1}X^T\hat{G}_n\hat{z} \tag{2}$$

where $\hat{G}_n = diag(\hat{\pi}_i(1-\hat{\pi}_i))$ and $\hat{z}_i = \log(\hat{\pi}_i) + \frac{y_i-\hat{\pi}_i}{\hat{\pi}_i(1-\hat{\pi}_i)}$.

Multicollinearity among the explanatory variables affects the MLE. The variance of the regression parameter is often influenced by the presence of multicollinearity [11,12]. The RRE is an alternative to the MLE in linear and logistic regression models [4,5]. The logistic ridge estimator (LRE) is defined as:

$$\hat{\beta}_{LRE} = \left(X^T\hat{G}_nX + kI_p\right)^{-1}X^T\hat{G}_nX\hat{\beta}_{MLE}, \tag{3}$$

where I is an identity matrix, k ($k > 0$) is the ridge parameter, and $\hat{G}_n$ is the estimate of G using $\hat{\beta}_{MLE}$. The ridge parameter [13] is defined as

$$k = \frac{(p+1)\sigma^2}{\sum_{j=1}^{p+1}\alpha_j^2} \tag{4}$$

while the logistic version [14] is as follows:

$$k = \frac{(p+1)}{\sum_{j=1}^{p}\alpha_j^2} \tag{5}$$

The Liu estimator [6] is an alternative to the ridge estimator in the linear regression model, while the logistic Liu estimator (LLE) [7] is expressed as follows:

$$\hat{\beta}_{LLE} = \left(X^T\hat{G}_nX + I_p\right)^{-1}\left(X^T\hat{G}_nX + dI_p\right)\hat{\beta}_{MLE}, \tag{6}$$

where d $(0 < d < 1)$ is the Liu parameter. Further, we adopted the following method to compute Liu parameter d [15]:

$$d = min\left(\frac{\alpha_j^2}{\frac{1}{\lambda_j} + \alpha_j^2}\right) \tag{7}$$

where max and min represent the maximum and minimum operators, respectively. Further, λ_j represents the jth eigenvalue of the $X^T\hat{G}_nX$ and $\alpha = Q^T\hat{\beta}_{MLE}$, where Q is the eigenvector of $X^T\hat{G}_nX$.

Liu [16] proposed a two-parameter estimator called the Liu-type estimator. Inan and Erdogan [17] extended this work to the logistic regression model. The logistic Liu-type estimator (LLTE) is as follows:

$$\hat{\beta}_{LLTE} = \left(X^T\hat{G}_nX + kI_p\right)^{-1}\left(X^T\hat{G}_nX - dI_p\right)\hat{\beta}_{MLE} \tag{8}$$

where k ($k > 0$) and d ($-\infty < d < \infty$) are the biasing parameters of the LLTE.

Ozkale and Kaciranlar [15] developed the two-parameter estimator (TPE) to mitigate multicollinearity in the LRM. Huang [18] developed the logistic TPE estimator (LTPE), defined as follows:

$$\hat{\beta}_{LTPE} = \left(X^T\hat{G}_nX + kI_p\right)^{-1}\left(X^T\hat{G}_nX + kdI_p\right)\hat{\beta}_{MLE} \tag{9}$$

where k ($k > 0$) and d ($-\infty < d < \infty$) are the biasing parameters. The biasing parameters are defined in Equations (5) and (7), respectively.

Recently, the K-L estimator (KLE) [10] has shown better performance than the ordinary least squares, the RRE and the LE for parameter estimation in the LRM. The KLE is defined as

$$\hat{\beta}_{KLE} = \left(X^TX + kI_p\right)^{-1}\left(X^TX - kI_p\right)\hat{\beta}_{MLE} \tag{10}$$

where k ($k > 0$) is the KLE biasing parameter, which, as will be discussed in Section 3.6, was obtained by minimizing the mean squared error (MSE). However, in this study, we propose the logistic K-L estimator (LKLE) as

$$\hat{\beta}_{LKLE} = \left(X^T\hat{G}_nX + kI_p\right)^{-1}\left(X^T\hat{G}_nX - kI_p\right)\hat{\beta}_{MLE} \tag{11}$$

The bias and the matrix mean squared error (MMSE) of the LKLE is obtained as follows:
The bias of the LKLE is as follows:

$$B\left(\hat{\beta}_{LKLE}\right) = -2kQ\Lambda^k\alpha \tag{12}$$

where $\Lambda^k = \left(\Lambda + kI_p\right)^{-1}$.

The variance of the LKLE is defined as follows:

$$Cov\left(\hat{\beta}_{LKLE}\right) = Q\left(\Lambda - kI_p\right)\Lambda^k\Lambda^{-1}\Lambda^k\left(\Lambda - kI_p\right)Q^T \tag{13}$$

where $\Lambda^k = \left(\Lambda + kI_p\right)^{-1}$.

Therefore, the MMSE and the scalar mean squared error (MSE) are, respectively, defined by

$$MMSE\left(\hat{\beta}_{LKLE}\right) = Q\left(\Lambda - kI_p\right)\Lambda^k\Lambda^{-1}\Lambda^k\left(\Lambda - kI_p\right)Q^T + 4k^2Q\Lambda^k\alpha\alpha^T\Lambda^kQ^T \tag{14}$$

and

$$MSE\left(\hat{\beta}_{LKLE}\right) = \sum_{j=1}^{p}\left(\frac{\left(\lambda_j - k\right)^2}{\lambda_j\left(\lambda_j + k\right)^2}\right) + 4k^2\sum_{j=1}^{p}\left(\frac{\alpha_j^2}{\left(\lambda_j + k\right)^2}\right) \tag{15}$$

The MMSE and MSE of the MLE, LRE, LLE, LLTE and LTPE are given, respectively, as follows:

$$MMSE\left(\hat{\beta}_{MLE}\right) = Q\Lambda^{-1}Q^T \tag{16}$$

$$MSE\left(\hat{\beta}_{MLE}\right) = \sum_{i=1}^{p}\frac{1}{\lambda_j} \tag{17}$$

$$MMSE(\hat{\beta}_{LRE}) = Q\Lambda^k\Lambda\Lambda^kQ^T + k^2Q\Lambda^k\alpha\alpha^T\Lambda^kQ^T \tag{18}$$

$$MSE(\hat{\beta}_{LRE}) = \sum_{j=1}^{p}\left(\frac{\lambda_j}{(\lambda_j+k)^2}\right) + k^2\sum_{j=1}^{p}\left(\frac{\alpha_j^2}{(\lambda_j+k)^2}\right) \tag{19}$$

$$MMSE(\hat{\beta}_{LLE}) = Q\Lambda_d\Lambda^{-1}\Lambda_d^TQ^T + Q(\Lambda_d - I_p)\alpha\alpha^T(\Lambda_d - I_p)^TQ^T, \tag{20}$$

where $\Lambda_d = (\Lambda + I_p)^{-1}(\Lambda + dI_p)$.

$$MSE(\hat{\beta}_{LLE}) = \sum_{j=1}^{p}\left(\frac{(\lambda_j+d)^2}{\lambda_j(\lambda_j+1)^2} + \frac{(d-1)^2\alpha_j^2}{(\lambda_j+1)^2}\right) \tag{21}$$

$$MMSE(\hat{\beta}_{LLTE}) = Q\Lambda_{kd}\Lambda^{-1}\Lambda_{kd}^TQ^T + Q(d+k)^2\Lambda^k\alpha\alpha^T\Lambda^kQ^T \tag{22}$$

where $\Lambda_{kd} = (\Lambda + kI_p)^{-1}(\Lambda - dI_p)$.

$$MSE(\hat{\beta}_{LLTE}) = \sum_{j=1}^{p}\left(\frac{(\lambda_j-d)^2}{\lambda_j(\lambda_j+k)^2} + \frac{(d+k)^2\alpha_j^2}{(\lambda_j+k)^2}\right) \tag{23}$$

$$MMSE(\hat{\beta}_{LTPE}) = Q\Lambda_{k-d}\Lambda^{-1}\Lambda_{k-d}^TQ^T + Qk^2(1-d)^2\Lambda^k\alpha\alpha^T\Lambda^kQ^T \tag{24}$$

where $\Lambda_{k-d} = (\Lambda + kI_p)^{-1}(\Lambda + kdI_p)$.

$$MSE(\hat{\beta}_{LTPE}) = \sum_{j=1}^{p}\left(\frac{(\lambda_j+kd)^2}{\lambda_j(\lambda_j+k)^2} + \frac{k^2(1-d)^2\alpha_j^2}{(\lambda_j+k)^2}\right) \tag{25}$$

The following lemmas are needful to prove the statistical properties of the proposed estimator.

Lemma 1. *Let M be a positive definite matrix, that is, $M > 0$, and α be some vector. Then $M - \alpha\alpha^T \geq 0$ if and only if $\alpha^TM^{-1}\alpha \leq 1$ [19].*

Lemma 2. *Let $\hat{\beta}_j = A_jy$, $j = 1, 2$ be two linear estimators of β [20]. Suppose that $D = Cov(\hat{\beta}_1) - Cov(\hat{\beta}_2) > 0$, where $Cov(\hat{\beta}_j)$, $\rightleftarrows$ $j = 1, 2$ denotes the covariance matrix of $\hat{\beta}_j$ and $b_j = Bias(\hat{\beta}_j) = (A_jX - I)\beta$, $j = 1, 2$. Consequently,*

$$\Delta(\hat{\beta}_1 - \hat{\beta}_2) = MMSE(\hat{\beta}_1) - MMSE(\hat{\beta}_2) = \sigma^2D + b_1b_1^T - b_2b_2^T > 0 \tag{26}$$

if and only if $b_2^T[\sigma^2D + b_1b_1^T]^{-1}b_2 < 1$, where $MSE(\hat{\beta}_j) = Cov(\hat{\beta}_j) + b_j^Tb_j$.

3. Comparison among the Estimators

In this section, we will perform a theoretical comparison of the proposed estimator with the available estimators in terms of MMSEs.

3.1. Comparison between $\hat{\beta}_{MLE}$ and $\hat{\beta}_{LKLE}$

Theorem 1. *If $k > 0$, the estimator $\hat{\beta}_{LKLE}$ is preferable to the estimator $\hat{\beta}_{MLE}$ in the MMSE sense, if and only if, $b^T\left[\left(\Lambda^{-1} - \Lambda^k(\Lambda - kI_p)\Lambda^{-1}\Lambda^k(\Lambda - kI_p)\right)\right]^{-1}b < 1$, where $b = -2k\Lambda^k\alpha$.*

Proof.

$$MMSE(\hat{\beta}_{MLE}) - MMSE(\hat{\beta}_{LKLE}) = \left[\Lambda^{-1} - \Lambda^k(\Lambda - kI_p)\Lambda^{-1}\Lambda^k(\Lambda - kI_p)\right] - 4k^2\Lambda^k\Lambda^k\alpha\alpha^T$$

$$Cov(\hat{\beta}_{MLE}) - Cov(\hat{\beta}_{LKLE}) = \Lambda^{-1} - \Lambda^k(\Lambda - kI_p)\Lambda^{-1}\Lambda^k(\Lambda - kI_p)$$

can be written in scalar form as follows:

$$Cov(\hat{\beta}_{MLE}) - Cov(\hat{\beta}_{LKLE}) = Q\ diag\left\{\frac{1}{\lambda_j} - \frac{(\lambda_j - k)^2}{\lambda_j(\lambda_j + k)^2}\right\}_{j=1}^{p} Q^T \tag{27}$$

$$= Q\ diag\left\{\frac{[(\lambda_j + k) - (\lambda_j - k)][(\lambda_j + k) + (\lambda_j - k)]}{\lambda_j(\lambda_j + k)^2}\right\}_{j=1}^{p} Q^T$$

$Cov(\hat{\beta}_{MLE}) - Cov(\hat{\beta}_{LKLE})$ is positive definite since $4k\lambda_j > 0$. Hence, using Lemma 2, $MMSE(\hat{\beta}_{MLE}) - MMSE(\hat{\beta}_{LKLE}) > 0$ if and only if

$$4k^2\alpha^T\Lambda^k\left[\left(\Lambda^{-1} - \Lambda^k(\Lambda - kI_p)\Lambda^{-1}\Lambda^k(\Lambda - kI_p)\right)\right]^{-1}\Lambda^k\alpha < 1 \tag{28}$$

Simplifying Equation (27), we have $4k^2\alpha^T\Lambda[(\Lambda + k)^2 - (\Lambda - k)^2]^{-1}\alpha < 1$. This was practically illustrated in Section 3.5 (Proof completed). □

3.2. Comparison between $\hat{\beta}_{LRE}$ and $\hat{\beta}_{LKLE}$

Theorem 2. *If $k > 0$, the estimator $\hat{\beta}_{LKLE}$ is preferable to the estimator $\hat{\beta}_{LRE}$ in the MMSE sense, if and only if, $b^T\left[\left(\Lambda^k\Lambda\Lambda^k - \Lambda^k(\Lambda - kI_p)\Lambda^{-1}\Lambda^k(\Lambda - kI_p)\right) + b_1^T b_1\right]^{-1} b < 1$, where $b = -2k\Lambda^k\alpha$ and $b_1 = -k\Lambda^k\alpha$.*

Proof.

$$MMSE(\hat{\beta}_{LRE}) - MMSE(\hat{\beta}_{LKLE}) = \left[\Lambda^k\Lambda\Lambda^k - \Lambda^k(\Lambda - kI_p)\Lambda^{-1}\Lambda^k(\Lambda - kI_p)\right] + b_1b_1^T - bb^T$$

$$Cov(\hat{\beta}_{LRE}) - Cov(\hat{\beta}_{LKLE}) = \Lambda^k\Lambda\Lambda^k - \Lambda^k(\Lambda - kI_p)\Lambda^{-1}\Lambda^k(\Lambda - kI_p)$$

can be written in scalar form as follows:

$$Cov(\hat{\beta}_{LRE}) - Cov(\hat{\beta}_{LKLE}) = Q\ diag\left\{\frac{\lambda_j}{(\lambda_j + k)^2} - \frac{(\lambda_j - k)^2}{\lambda_j(\lambda_j + k)^2}\right\}_{j=1}^{p} Q^T \tag{29}$$

$$= Q\ diag\left\{\frac{\lambda_j^2(\lambda_j + k)^2 - (\lambda_j - k)^2(\lambda_j + k)^2}{\lambda_j(\lambda_j + k)^2}\right\}_{j=1}^{p} Q^T$$

$Cov(\hat{\beta}_{LRE}) - Cov(\hat{\beta}_{LKLE})$ is positive definite since $\lambda_j^2(\lambda_j + k)^2 > (\lambda_j - k)^2(\lambda_j + k)^2$ for $k > 0$.

$$B(\hat{\beta}_{LRE}) - B(\hat{\beta}_{LKLE}) = -k(\lambda_j + k)^{-1}\alpha + 2k(\lambda_j + k)^{-1}\alpha \tag{30}$$

$$= k(\lambda_j + k)^{-1}\alpha$$

$$B(\hat{\beta}_{LRE}) - B(\hat{\beta}_{LKLE}) > 0$$

Hence, using Lemma 2, $MMSE(\hat{\beta}_{LRE}) - MMSE(\hat{\beta}_{LKLE}) > 0$ if and only if

$$b^T\left[\left(\Lambda^k\Lambda\Lambda^k - \Lambda^k(\Lambda - kI_p)\Lambda^{-1}\Lambda^k(\Lambda - kI_p)\right) + b_1^T b_1\right]^{-1} b < 1 \tag{31}$$

Simplifying Equation (29), we have $4k^2\alpha^T\left[\Lambda+k^2\alpha\alpha^T-\Lambda^{-1}(\Lambda-kI_p)^2\right]^{-1}\alpha<1$. This was practically illustrated in Section 5 (Proof completed). □

3.3. Comparison between $\hat{\beta}_{LLE}$ and $\hat{\beta}_{LKLE}$

Theorem 3. *If $k>0$ and $0<d<1$, the estimator $\hat{\beta}_{LKLE}$ is preferable to the estimator $\hat{\beta}_{LLE}$ in MMSE sense if and only if, $b^T\left[\left(\Lambda_d\Lambda^{-1}\Lambda_d^T-\Lambda^k(\Lambda-kI_p)\Lambda^{-1}\Lambda^k(\Lambda-kI_p)\right)+b_2^Tb_2\right]^{-1}b<1$, where $b=-2k\Lambda^k\alpha$ and $b_2=-(1-d)(\Lambda+I)^{-1}\alpha$.*

Proof.

$$MMSE(\hat{\beta}_{LLE})-MMSE(\hat{\beta}_{LKLE})=\left[\Lambda_d\Lambda^{-1}\Lambda_d^T-\Lambda^k(\Lambda-kI_p)\Lambda^{-1}\Lambda^k(\Lambda-kI_p)\right]+b_2b_2^T-\mathrm{b}b^T$$

$$Cov(\hat{\beta}_{LLE})-Cov(\hat{\beta}_{LKLE})=\Lambda_d\Lambda^{-1}\Lambda_d^T-\Lambda^k(\Lambda-kI_p)\Lambda^{-1}\Lambda^k(\Lambda-kI_p)$$

can be written in scalar form as follows:

$$\begin{aligned}Cov(\hat{\beta}_{LLE})-Cov(\hat{\beta}_{LKLE})&=Q\,diag\left\{\frac{(\lambda_j+d)^2}{\lambda_j(\lambda_j+1)^2}-\frac{(\lambda_j-k)^2}{\lambda_j(\lambda_j+k)^2}\right\}_{j=1}^{p}Q^T\\&=Q\,diag\left\{\frac{(\lambda_j+d)^2(\lambda_j+K)^2-(\lambda_j-K)^2(\lambda_j+1)^2}{\lambda_j(\lambda_j+K)^2(\lambda_j+1)^2}\right\}_{j=1}^{p}Q^T\end{aligned}\tag{32}$$

$Cov(\hat{\beta}_{LLE})-Cov(\hat{\beta}_{LKLE})$ since $(\lambda_j+k)^2(\lambda_j+d)^2-(\lambda_j-k)^2(\lambda_j+1)^2>0$ for $k>\frac{\lambda_j(1-d)}{d+\lambda_j+1}$ and $d>\frac{\lambda_j(1-k)-k}{\lambda_j+k}$.

$$\begin{aligned}B(\hat{\beta}_{LLE})-B(\hat{\beta}_{LKLE})&=-(1-d)(\lambda_j+1)^{-1}\alpha+2k(\lambda_j+k)^{-1}\alpha\\&=-\alpha diag\left\{\frac{(1-d)}{(\lambda_j+1)^2}-\frac{2k}{\lambda_j+k}\right\}_{j=1}^{p}\end{aligned}\tag{33}$$

$$B(\hat{\beta}_{LLE})-B(\hat{\beta}_{LKLE})=(k+d\lambda_j+kd+2k\lambda_j-\lambda_j)\alpha>0$$

Hence, using Lemma 2.2, $MMSE(\hat{\beta}_{LLE})-MMSE(\hat{\beta}_{LKLE})>0$ if and only if

$$b^T\left[\left(\Lambda_d\Lambda^{-1}\Lambda_d^T-\Lambda^k(\Lambda-kI_p)\Lambda^{-1}\Lambda^k(\Lambda-kI_p)\right)+b_2^Tb_2\right]^{-1}b<1\tag{34}$$

This was practically illustrated in Section 5 (Proof completed). □

3.4. Comparison between $\hat{\beta}_{LLTE}$ and $\hat{\beta}_{LKLE}$

Theorem 4. *If $k>0$ and $-\infty<d<\infty$, the estimator $\hat{\beta}_{LKLE}$ is preferable to the estimator $\hat{\beta}_{LLTE}$ in MMSE sense if and only if, $b^T\left[\left(\Lambda_{kd}\Lambda^{-1}\Lambda_{kd}^T-\Lambda^k(\Lambda-kI_p)\Lambda^{-1}\Lambda^k(\Lambda-kI_p)\right)+b_2^Tb_2\right]^{-1}b<1$, where $\Lambda_{kd}=(\Lambda+k)^{-1}(\Lambda-dI)$, $b=-2k\Lambda^k\alpha$ and $b_2=-(d+k)(\Lambda+k)^{-1}\alpha$.*

Proof.

$$MMSE(\hat{\beta}_{LLTE})-MMSE(\hat{\beta}_{LKLE})=\left[\Lambda_{kd}\Lambda^{-1}\Lambda_{kd}^T-\Lambda^k(\Lambda-kI_p)\Lambda^{-1}\Lambda^k(\Lambda-kI_p)\right]+b_2b_2^T-\mathrm{b}b^T$$

$$Cov(\hat{\beta}_{LLTE})-Cov(\hat{\beta}_{LKLE})=\Lambda_{kd}\Lambda^{-1}\Lambda_{kd}^T-\Lambda^k(\Lambda-kI_p)\Lambda^{-1}\Lambda^k(\Lambda-kI_p)$$

can be written in scalar form as follows:

$$\begin{aligned} Cov(\hat{\beta}_{LLTE}) - Cov(\hat{\beta}_{LKLE}) &= Q\, diag\left\{ \frac{(\lambda_j-d)^2}{\lambda_j(\lambda_j+k)^2} - \frac{(\lambda_j-k)^2}{\lambda_j(\lambda_j+k)^2} \right\}_{j=1}^{p} Q^T \\ &= Q\, diag\left\{ \frac{(\lambda_j-d)^2-(\lambda_j-k)^2}{\lambda_j(\lambda_j+k)^2} \right\}_{j=1}^{p} Q^T + b_2^T b_2 - b^T b. \end{aligned} \tag{35}$$

$Cov(\hat{\beta}_{LLTE}) - Cov(\hat{\beta}_{LKLE})$ is non-negative (nn), since $(\lambda_j - d)^2 > (\lambda_j - k)^2$.

$$\begin{aligned} B(\hat{\beta}_{LLTE}) - B(\hat{\beta}_{LKLE}) &= -(\mathrm{d}+k)(\lambda_j+k)^{-1}\alpha + 2k(\lambda_j+k)^{-1}\alpha \\ &= (k-d)(\lambda_j+k)^{-1}\alpha \\ B(\hat{\beta}_{LLTE}) - B(\hat{\beta}_{LKLE}) &> 0 \text{fork} > \mathrm{d}. \end{aligned} \tag{36}$$

Hence, using Lemma 2.2, $MMSE(\hat{\beta}_{LLTE}) - MMSE(\hat{\beta}_{LKLE}) > 0$ if and only if

$$b^T\left[\left(\Lambda_{kd}\Lambda^{-1}\Lambda_{kd}^T - \Lambda^k(\Lambda - kI_p)\Lambda^{-1}\Lambda^k(\Lambda - kI_p)\right) + b_2^T b_2\right]^{-1} b < 1 \tag{37}$$

This was practically illustrated in Section 5 (Proof completed). □

3.5. Comparison between $\hat{\beta}_{LTPE}$ and $\hat{\beta}_{LKLE}$

Theorem 5. *If $k > 0$ and $-\infty < d < \infty$, the estimator $\hat{\beta}_{LKLE}$ is preferable to the estimator $\hat{\beta}_{LTPE}$ in MMSE sense if and only if, $b^T\left[\left(\Lambda_{k-d}\Lambda^{-1}\Lambda_{k-d}^T - \Lambda^k(\Lambda - kI_p)\Lambda^{-1}\Lambda^k(\Lambda - kI_p)\right) + b_2^T b_2\right]^{-1} b < 1$, where $\Lambda_{k-d} = (\Lambda + k)^{-1}(\Lambda + kdI)$, $b = -2k\Lambda^k\alpha$ and $b_2 = -k(1-d)(\Lambda+k)^{-1}\alpha$.*

Proof.

$$MMSE(\hat{\beta}_{LTPE}) - MMSE(\hat{\beta}_{LKLE}) = \left[\Lambda_{k-d}\Lambda^{-1}\Lambda_{k-d}^T - \Lambda^k(\Lambda - kI_p)\Lambda^{-1}\Lambda^k(\Lambda - kI_p)\right] + b_2 b_2^T - \mathrm{b}b^T$$

$$Cov(\hat{\beta}_{LTPE}) - Cov(\hat{\beta}_{LKLE}) = \Lambda_{k-d}\Lambda^{-1}\Lambda_{k-d}^T - \Lambda^k(\Lambda - kI_p)\Lambda^{-1}\Lambda^k(\Lambda - kI_p)$$

can be written in scalar form as follows:

$$\begin{aligned} Cov(\hat{\beta}_{LTPE}) - Cov(\hat{\beta}_{LKLE}) &= Q\, diag\left\{ \frac{(\lambda_j+kd)^2}{\lambda_j(\lambda_j+k)^2} - \frac{(\lambda_j-k)^2}{\lambda_j(\lambda_j+k)^2} \right\}_{j=1}^{p} Q^T \\ &= Q\, diag\left\{ \frac{(\lambda_j+kd)^2-(\lambda_j-k)^2}{\lambda_j(\lambda_j+k)^2} \right\}_{j=1}^{p} Q^T. \end{aligned} \tag{38}$$

$Cov(\hat{\beta}_{LTPE}) - Cov(\hat{\beta}_{LKLE})$ is non-negative (nn) since $(\lambda_j + kd)^2 > (\lambda_j - k)^2$.

$$\begin{aligned} B(\hat{\beta}_{LTPE}) - B(\hat{\beta}_{LKLE}) &= -k(1-\mathrm{d})(\lambda_j+k)^{-1}\alpha + 2k(\lambda_j+k)^{-1}\alpha \\ &= k(1+d)(\lambda_j+k)^{-1}\alpha \\ B(\hat{\beta}_{LTPE}) - B(\hat{\beta}_{LKLE}) &> 0 \end{aligned} \tag{39}$$

Hence, using Lemma 2, $MMSE(\hat{\beta}_{LTPE}) - MMSE(\hat{\beta}_{LKLE}) > 0$ if and only if

$$b^T\left[\left(\Lambda_{k-d}\Lambda^{-1}\Lambda_{k-d}^T - \Lambda^k(\Lambda - kI_p)\Lambda^{-1}\Lambda^k(\Lambda - kI_p)\right) + b_2^T b_2\right]^{-1} b < 1 \tag{40}$$

This was practically illustrated in Section 5 (Proof completed). □

3.6. Selection of k

Since the shrinkage parameter plays a significant role in estimating biased estimators such as the LRE, LLE and LKLE, several researchers have introduced various shrinkage parameter estimation methods for the different regression models [21–29]. Based on these studies, we propose some shrinkage estimators of the parameter k for the LKLE.

To estimate parameter k, following [4], we will consider the generalized version of KL-estimator, which is given as follows:

$$\hat{\beta}_{LKLE} = \left(X^T\hat{G}_nX + KI_p\right)^{-1}\left(X^T\hat{G}_nX - KI_p\right)\hat{\beta}_{MLE}, \tag{41}$$

where K = diag(k_1, k_2, … , k_p).

The MSE of KL estimator in (44) would be

$$MSE(\hat{\beta}_{LKLE}) = \sum_{j=1}^{p}\left(\frac{(\lambda_j - k_j)^2}{\lambda_j(\lambda_j + k_j)^2}\right) + 4\sum_{j=1}^{p}\left(\frac{\alpha_j^2k_j^2}{(\lambda_j + k_j)^2}\right) \tag{42}$$

Differentiating Equation (42) with respect to k_j, (all terms except k_j will be 0) and equating to 0, we have

$$-2\left(\frac{\lambda_j - k_j}{\lambda_j(\lambda_j + k_j)^2}\right) - 2\left(\frac{(\lambda_j - k_j)^2}{\lambda_j(\lambda_j + k_j)^3}\right) + 8\left(\frac{k_j\alpha_j^2}{(\lambda_j + k_j)^2}\right) - 8\left(\frac{\alpha_j^2k_j^2}{(\lambda_j + k_j)^3}\right) = 0$$

$$\frac{-2(\lambda_j - k_j)(\lambda_j + k_j) - 2(\lambda_j - k_j)^2 + 8\lambda_jk_j\alpha_j^2(\lambda_j + k_j) - 8\lambda_j\alpha_j^2k_j^2}{\lambda_j(\lambda_j + k_j)^3} = 0$$

$$-2(\lambda_j - k_j)(\lambda_j + k_j) - 2(\lambda_j - k_j)^2 + 8\lambda_jk_j\alpha_j^2(\lambda_j + k_j) - 8\lambda_j\alpha_j^2k_j^2 = 0$$

$$(\lambda_j - k_j)(\lambda_j + k_j) + (\lambda_j - k_j)^2 - 4\lambda_jk_j\alpha_j^2(\lambda_j + k_j) + 4\lambda_j\alpha_j^2k_j^2 = 0 \tag{43}$$

Simplifying further Equation (43), we have

$$2\lambda_j^2 - 2\lambda_jk_j - 4k_j\lambda_j^2\alpha_j^2 = 0 \tag{44}$$

Dividing both sides of Equation (44) by $2\lambda_j$, we obtain

$$k_j = \frac{\lambda_j}{1 + 2\lambda_j\alpha_j^2} \tag{45}$$

By replacing α and λ with its unbiased estimates, Equation (45) becomes

$$\hat{k}_j = \frac{\hat{\lambda_j}}{1 + 2\hat{\lambda}_j\hat{\alpha}_j^2} \ (\mathrm{j} = 1, 2, \ldots, \mathrm{p}) \tag{46}$$

Following Hoerl et al. [13], and based on the study of Mansson et al. [7], Lukman and Ayinde [3] and Qasim et al. [22,30], we suggest the following biasing parameter estimators for the logistic regression model:

LKLE 1: $k = min\left(\frac{1}{\hat{\alpha}_j^2}\right)$

LKLE 2: $k = min\left(\hat{k}_j\right)$

LRE 1: $k = min\left(\frac{1}{\hat{\alpha}_j^2}\right)$

LRE 2: $k = \frac{p}{\sum_{j=1}^{p}\hat{\alpha}_j^2}$

$$\text{LLE: } d = \min\left(\frac{\alpha_j^2}{\frac{1}{\lambda_j}+\alpha_j^2}\right)$$

$$\text{LLTE: } k = \frac{p}{\sum_{j=1}^{p}\hat{\alpha}_j^2},\ d = \min\left(\frac{\alpha_j^2}{\frac{1}{\lambda_j}+\alpha_j^2}\right)$$

$$\text{LTPE: } k = \frac{p}{\sum_{j=1}^{p}\hat{\alpha}_j^2},\ d = \min\left(\frac{\alpha_j^2}{\frac{1}{\lambda_j}+\alpha_j^2}\right)$$

4. Monte Carlo Simulation

In this section, we compare the performance of the logistic regression estimators using a simulation study. A significant number of simulation studies have been conducted to compare the performance of estimators for both linear and logistic regression models [24–35]. The MSE is a function of β,σ^2, p and is minimized subject to constraint $\beta'\beta = 1$ [36,37]. Schaefer [14] showed that the logistic regression model can be designed employing a similar approach to that of the linear regression model. The correlated explanatory variables can be obtained using the simulation procedure given in [38,39].

$$x_{ij} = \left(1-\rho^2\right)^{1/2} w_{ij} + \rho w_{i(j+1)},\ i = 1,2,\ldots,n;\ j = 1,2,\ldots,p, \tag{47}$$

where w_{ij} are independent standard normal pseudo-random numbers and ρ is the correlation between the explanatory variables. The values of ρ are chosen to be 0.9, 0.95, 0.99 and 0.999. The response variable is generated from the Bernoulli distribution, i.e., $y_i \sim Be(\pi_i)$, where $\pi_i = \frac{e^{x_i^T\beta}}{1+e^{x_i^T\beta}}$. Sample size, n, is varied, i.e., 50, 100, 250 or 300. The estimated MSE is calculated as

$$MSE(\hat{\beta}) = \frac{1}{2000}\sum_{i=1}^{2000}\left(\hat{\beta}_i-\beta\right)^T\left(\hat{\beta}_i-\beta\right) \tag{48}$$

where $\hat{\beta}_i$ denotes the vector of the estimated regression coefficient in ith replication and β is the vector of the true parameter values, chosen such that $\beta^{'}\beta = 1$. The experiment was replicated 2000 times. We present the estimated MSEs and the bias of each of the estimators for p = 3 in Tables 1 and 2, respectively. For p = 7, the results are provided in Tables 3 and 4, respectively. We observed that increasing the sample size resulted in a decrease in the MSE values for each case. The following observations were obtained from the simulation result. The MSE values of the estimators increased as the degree of correlation and the number of explanatory variables increased. The simulation results show that the LKLE performed best at most levels of multicollinearity, sample sizes and the number of explanatory variables with few exceptions. The LTPE competed favorably in most cases, except on a few occasions. Upon comparing the performance of the shrinkage parameters in the LKLE, we found that LKLE 1 performed well except in a few cases. The MLE performed least well when there was multicollinearity in the data. Of the two-parameter estimators (LTPE and LLTE), LTPE performed better. Additionally, it is obvious that the bias of the proposed estimator was the lowest in most cases. Generally, the LKLE estimator is preferred over the two-parameter estimator.

Table 1. Estimated MSEs and Bias for p = 3.

n		50	50	50	50	100	100	100	100
ρ		0.9	0.95	0.99	0.999	0.9	0.95	0.99	0.999
MLE	MSE	1.4837	2.5668	11.6012	112.132	0.8648	1.3979	5.6404	53.8
LLE	MSE	1.0597	1.4893	5.0029	47.674	0.7492	1.0409	2.6449	21.6
	BIAS	−0.7965	−0.8000	−0.8029	−0.8034	−0.8014	−0.8030	−0.8080	−0.8126

Table 1. *Cont.*

LRE 1	MSE	0.9867	1.4492	5.4291	49.507	0.6677	0.9090	2.7611	23.9
	BIAS	−0.8349	−0.8190	−0.7934	−0.7802	−0.8323	−0.8234	−0.8098	−0.8039
LRE 2	MSE	0.8332	1.1187	3.7215	32.719	0.6040	0.7538	1.9624	15.9
	BIAS	−0.8967	−0.8651	−0.8150	−0.7906	−0.8695	−0.8517	−0.8225	−0.8089
LLTE	MSE	0.6927	0.8063	7.3947	50.125	0.5983	0.6394	2.6679	48
	BIAS	−0.9207	−0.9011	−0.9448	−1.9698	−0.8751	−0.8595	−0.8447	−1.0067
LTPE	MSE	0.7238	0.8732	2.3002	18.585	0.5911	0.6384	1.3563	9.82
	BIAS	−0.9322	−0.8974	−0.8384	−0.8075	−0.8897	−0.8697	−0.8342	−0.8164
LKLE 1	MSE	0.7458	0.8043	1.6510	11.563	0.5918	0.6303	1.0533	6.21
	BIAS	**−1.0436**	**−0.9752**	**−0.8734**	**−0.8256**	**−0.9502**	**−0.9131**	**−0.8536**	**−0.8251**
LKLE 2	MSE	0.9598	1.4017	5.2142	47.268	0.6582	0.8916	2.6610	22.7
	BIAS	−0.8186	−0.8032	−0.7848	−0.7790	−0.8224	−0.8136	−0.8038	−0.8029

Table 2. Estimated MSEs and Bias for $p = 3$.

n		250	250	250	250	300	300	300	300
ρ		0.9	0.95	0.99	0.999	0.9	0.95	0.99	0.999
MLE	MSE	0.5112	0.7083	2.2958	20.4	0.4834	0.6532	1.9918	16.931
LLE	MSE	0.4997	0.6530	1.4819	8.36	0.4757	0.6119	1.3344	6.708
	BIAS	−0.8273	−0.8256	−0.8239	−0.8271	−0.8189	−0.8180	−0.8187	−0.8198
LRE 1	MSE	0.4847	0.5844	1.2803	9.07	0.4659	0.5550	1.1523	7.558
	BIAS	−0.8497	−0.8417	−0.8290	−0.8244	−0.8389	−0.8329	−0.8247	−0.8182
LRE 2	MSE	0.4825	0.5421	0.9858	6.14	0.4656	0.5192	0.9011	5.118
	BIAS	−0.8746	−0.8612	−0.8381	−0.8275	−0.8603	−0.8501	−0.8331	−0.8207
LLTE	MSE	0.4794	0.5239	0.6785	10.23	0.4634	0.5059	0.6580	7.865
	BIAS	−0.8766	−0.8637	−0.8438	−0.8609	−0.8618	−0.8519	−0.8370	−0.8414
LTPE	MSE	0.4785	0.5083	0.7690	3.98	0.4631	0.4894	0.7201	3.398
	BIAS	−0.8873	−0.8723	−0.8456	−0.8314	−0.8726	−0.8608	−0.8404	−0.8241
LKLE 1	MSE	0.2344	0.3576	0.6250	2.81	0.3087	0.3285	0.5771	2.364
	BIAS	**−0.9275**	**−0.9024**	**−0.8594**	−0.8367	**−0.9053**	**−0.8856**	**−0.8515**	−0.8281
LKLE 2	MSE	0.4344	0.5027	1.2531	8.66	0.4625	0.5533	1.1303	7.200
	BIAS	−0.8469	−0.8385	−0.8261	−0.8237	−0.8358	−0.8291	−0.8211	−0.8170

Table 3. Estimated MSEs and Bias for $p = 7$.

n		50	50	50	50	100	100	100	100
ρ		0.9	0.95	0.99	0.999	0.9	0.95	0.99	0.999
MLE	MSE	6.2754	11.5251	61.5	585	2.5865	4.9251	23.922	236
LLE	MSE	1.9336	2.3581	10.2	117	1.5505	2.0802	4.720	40.9
	BIAS	−1.4187	−1.4699	−1.5091	−1.5071	−1.3487	−1.3548	−1.4039	−1.4231
LRE 1	MSE	3.5013	6.0056	31.8	297	1.6169	2.8887	13.218	128
	BIAS	−1.3122	−1.3329	−1.3042	−1.3101	−1.3431	−1.3225	−1.3280	−1.3272

Table 3. *Cont.*

LRE 2	MSE	1.6420	2.6788	13.2	122	0.9322	1.5026	6.273	59.3
	BIAS	−1.4628	−1.4578	−1.4198	−1.4181	−1.4314	−1.3944	−1.3825	−1.3761
LLTE	MSE	1.2682	1.8409	1200	477	0.8472	1.2163	15.024	188
	BIAS	−1.5043	−1.5429	−2.2742	−10.5093	−1.4396	−1.4161	−1.5587	−3.0291
LTPE	MSE	1.3488	2.0536	8.96	78	0.8076	1.2358	4.747	43.2
	BIAS	−1.4979	−1.4971	−1.4701	−1.4802	−1.4483	−1.4162	−1.4102	−1.4045
LKLE 1	MSE	0.7149	0.9988	3.43	31	0.6939	0.7074	1.710	14.5
	BIAS	**−1.7501**	**−1.7118**	**−1.6656**	**−1.6547**	**−1.5857**	**−1.5257**	−1.4894	−1.4755
LKLE 2	MSE	3.3646	5.7296	30.1	281	1.5856	2.8236	12.878	125
	BIAS	−1.3135	−1.3373	−1.3132	−1.3202	−1.3416	−1.3222	−1.3293	−1.3294

Table 4. Estimated MSEs and Bias for $p = 7$.

n		250	250	250	250	300	300	300	300
ρ		0.9	0.95	0.99	0.999	0.9	0.95	0.99	0.999
MLE	MSE	1.0284	1.7662	7.58	72.9	0.8894	1.4841	6.2274	60.1
LLE	MSE	0.8894	1.3001	2.64	13.5	0.7944	1.1611	2.4346	9.43
	BIAS	−1.4056	−1.4033	−1.4287	−1.4415	−1.3846	−1.3865	−1.3990	−1.4195
LRE 1	MSE	0.7632	1.1834	4.40	40.5	0.6815	1.0195	3.6523	33.6
	BIAS	−1.4242	−1.4107	−1.4121	−1.4083	−1.4047	−1.3981	−1.3904	−1.3917
LRE 2	MSE	0.5474	0.7325	2.23	19.2	0.5088	0.6572	1.8915	16.2
	BIAS	−1.4755	−1.4483	−1.4333	−1.4245	−1.4518	−1.4331	−1.4089	−1.4063
LLTE	MSE	0.5392	0.7005	1.62	115	0.5033	0.6341	1.4637	47
	BIAS	−1.4768	−1.4502	−1.4422	−1.5435	−1.4529	−1.4353	−1.4163	−1.4908
LTPE	MSE	0.5117	0.6530	1.78	14.2	0.4808	0.5892	1.5361	12.6
	BIAS	−1.4849	−1.4559	−1.4403	−1.4332	−1.4619	−1.4433	−1.4169	−1.4139
LKLE 1	MSE	0.5344	0.5576	0.7250	2.81	0.5087	0.5285	0.6771	2.364
	BIAS	**−1.5577**	**−1.5105**	**−1.4735**	−1.4582	**−1.5285**	**−1.4925**	−1.4471	−1.4375
LKLE 2	MSE	0.4844	0.5827	1.2531	8.66	0.4655	0.5533	1.1303	7.200
	BIAS	−1.4217	−1.4089	−1.4117	−1.4085	−1.4026	−1.3968	−1.3900	−1.3919

5. Application: Cancer Data

The performance of LKLE and the other estimators was evaluated using a cancer remission dataset [34,40]. In the dataset, the binary response variable y_i is 1 if the patient experiences complete cancer remission and 0 otherwise. There are five explanatory variables. These explanatory variables include cell index (x_1), smear index (x_2), infil index (x_3), blast index (x_4) and temperature (x_5). There were 27 patients, of which nine experienced complete remission. The eigenvalues of the $X^{'}\hat{G}_nX$ matrix were found to be $\lambda_1 = 9.2979$, $\lambda_2 = 3.8070$, $\lambda_3 = 3.0692$, $\lambda_4 = 2.2713$ and $\lambda_5 = 0.0314$. To test the multicollinearity among the explanatory variables, we use condition index (CI), computed as $CI = \sqrt{\frac{\max(\lambda_j)}{\min(\lambda_j)}} = 17.2$. There was moderate collinearity when CI was between 10 and 30 and severe multicollinearity when CI exceeded 30 [41]. Thus, the results provide evidence of moderate multicollinearity among the explanatory variables. Next, we compared the performance of the estimators using the previously described dataset. The estimated regression coefficients and the corresponding scalar MSE values are given in Table 5. The scalar MSEs of each of the estimators

under study were obtained using Equations (17), (19), (21) and (23)–(25), respectively. The proposed LKLE estimator surpassed the other estimators in this study in terms of MSE.

Table 5. Validation of the theoretical conditions for the cancer data.

Theorems	Conditions	Value
1	$b^T\left[\left(\Lambda^{-1}-\Lambda^k(\Lambda-kI_p)\Lambda^{-1}\Lambda^k(\Lambda-kI_p)\right)\right]^{-1}b<1$	0.2413
2	$b^T\left[\left(\Lambda^k\Lambda\Lambda^k-\Lambda^k(\Lambda-kI_p)\Lambda^{-1}\Lambda^k(\Lambda-kI_p)\right)+b_1^Tb_1\right]^{-1}b<1$	0.8866
3	$b^T\left[\left(\Lambda_d\Lambda^{-1}\Lambda_d^T-\Lambda^k(\Lambda-kI_p)\Lambda^{-1}\Lambda^k(\Lambda-kI_p)\right)+b_2^Tb_2\right]^{-1}b<1$	0.6443
4	$b^T\left[\left(\Lambda_{kd}\Lambda^{-1}\Lambda_{kd}^T-\Lambda^k(\Lambda-kI_p)\Lambda^{-1}\Lambda^k(\Lambda-kI_p)\right)+b_2^Tb_2\right]^{-1}b<1$	0.0958
5	$b^T\left[\left(\Lambda_{k-d}\Lambda^{-1}\Lambda_{k-d}^T-\Lambda^k(\Lambda-kI_p)\Lambda^{-1}\Lambda^k(\Lambda-kI_p)\right)+b_2^Tb_2\right]^{-1}b<1$	0.7540

Moreover, we also evaluated the theoretical conditions as stated in Theorems 1 to 5 for the actual dataset. The validation results of these conditions are given in Table 6. As shown, all the theorem conditions hold for the cancer data, because all the inequalities in the theorems were less than one, as expected.

Table 6. Regression coefficients and MSEs of the logistic regression estimators for the cancer dataset *.

Estimators	$\hat{\beta}_1$	$\hat{\beta}_2$	$\hat{\beta}_3$	$\hat{\beta}_4$	$\hat{\beta}_5$	Estimated MSEs
MLE	−0.1966 (0.328)	−1.5957 (0.513)	1.8139 (0.571)	1.3073 (0.664)	−0.4208 (5.639)	32.9393
LLE	0.1940 (0.305)	−0.4751 (0.498)	0.5808 (0.571)	1.0259 (0.592)	−0.3078 (1.645)	5.0989
LRE 1	0.3706 (0.318)	−0.2218 (0.489)	0.2266 (0.544)	1.1366 (0.605)	−0.3457 (0.652)	1.4278
LRE 2	0.3503 (0.303)	−0.0999 (0.504)	0.1472 (0.584)	0.9838 (0.596)	−0.2883 (0.461)	1.2544
LLTE	0.5373 (0.295)	0.4116 (0.536)	−0.4227 (0.643)	0.8732 (0.624)	−0.2430 (1.720)	4.1372
LTPE	0.4949 (0.310)	0.2955 (0.491)	−0.2933 (0.556)	0.8983 (0.592)	−0.2533 (1.679)	3.8161
LKLE 1	0.8972 (0.278)	1.3960 (0.400)	−1.5195 (0.446)	0.6604 (0.418)	−0.1558 (5.208)	28.9236
LKLE 2	0.4696 (0.326)	−0.1212 (0.505)	0.0576 (0.560)	1.2298 (0.646)	−0.3815 (0.084)	1.1350

* Standard error for each of the estimators is in parenthesis.

The logistic ridge estimator competed favorably in the simulation and the real-life application. The real-life application result agreed with the simulation study. However, the performance of the estimators in both the simulation and real life was a function of the biasing parameter. For instance, LKLE 1 performed best in the simulation study, while in the real-life analysis, LKLE 2 outperformed LKLE 1. Among the two-parameter estimators, the logistic two-parameter estimator (LTPE) performed best. Of the one-parameter estimators, LKLE outperformed the ridge and the Liu estimator. Generally, LKLE dominated among both the one- and two-parameter estimator. The performance of these estimators is a function of biasing parameters k and d. Additionally, as shown in Table 5. $\hat{\beta}_2$ and $\hat{\beta}_3$ did not fit well for the following estimators: MLE, LLE, LLTE, LTPE and LKLE 1.

6. Some Concluding Remarks

Kibria and Lukman (2020) developed the K-L estimator to circumvent the multicollinearity problem for the linear regression model. In this paper, we described the logistic Kibria-Lukman estimator (LKLE) to address the challenge of multicollinearity for the logistic regression model. We theoretically determined the superiority of LKLE over other existing estimators in terms of the MSE. The performance of the estimators was evaluated using the Monte Carlo simulation study. In the design of the experiment, factors such as the degree of correlation, the sample size and the number of explanatory variables were varied. The results showed that the performance of the estimators was highly dependent on these factors. Finally, to illustrate the efficiency of the proposed estimator, we applied a cancer dataset and observed that the results agreed with those of the simulation study to some extent. The findings of this study will be helpful for practitioners and applied researchers who use a logistic regression model with correlated explanatory variables.

Author Contributions: A.F.L.: Conceptualization, Methodology, Formal analysis, Software, Writing—original draft. B.M.G.K.: Conceptualization, Supervision, Review. R.F.: Writing—original draft, Resources, Review. M.A.: Methodology, Super vision, Writing—original draft. E.T.A.: Methodology, Formal analysis, Software, Writing—original draft. C.K.N.: Writing—original draft, Review. All authors have read and agreed to the published version of the manuscript.

Funding: Authors received no funding for this work.

Data Availability Statement: The Data is available as this article.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Frisch, R. *Statistical Confluence Analysis by Means of Complete Regression Systems*; University Institute of Economics: Oslo, Norway, 1934.
2. Kibria, B.M.G.; Mansson, K.; Shukur, G. Performance of some logistic ridge regression estimators. *Comp. Econ.* **2012**, *40*, 401–414. [CrossRef]
3. Lukman, A.F.; Ayinde, K. Review and classifications of the ridge parameter estimation techniques. *Hacet. J. Math. Stat.* **2017**, *46*, 953–967. [CrossRef]
4. Hoerl, A.E.; Kennard, R.W. Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics* **1970**, *12*, 55–67. [CrossRef]
5. Schaeffer, R.L.; Roi, L.D.; Wolfe, R.A. A ridge logistic estimator. *Commun. Stat. Theory Methods* **1984**, *13*, 99–113. [CrossRef]
6. Liu, K. A new class of biased estimate in linear regression. *Commun. Stat.* **1993**, *22*, 393–402.
7. Mansson, K.; Kibria, B.M.G.; Shukur, G. On Liu estimators for the logit regression model. *Econ. Model.* **2012**, *29*, 1483–1488. [CrossRef]
8. Lukman, A.F.; Ayinde, K.; Binuomote, S.; Onate, A.C. Modified ridge-type estimator to combat multicollinearity: Application to chemical data. *J. Chemomet.* **2019**, *33*, e3125. [CrossRef]
9. Lukman, A.F.; Adewuyi, E.; Onate, A.C.; Ayinde, K. A Modified Ridge-Type Logistic Estimator. *Iran. J. Sci. Technol. Trans. A Sci.* **2020**, *44*, 437–443. [CrossRef]
10. Kibria, B.M.G.; Lukman, A.F. A new ridge type estimator for the linear regression model: Simulations and applications. *Scientifica* **2020**, *2020*, 9758378. [CrossRef]
11. Lukman, A.F.; Ayinde, K.; Aladeitan, B.; Bamidele, R. An unbiased estimator with prior information. *Arab. J. Basic Appl. Sci.* **2020**, *27*, 45–55. [CrossRef]
12. Dawoud, I.; Lukman, A.F.; Haadi, A. A new biased regression estimator: Theory, simulation and application. *Sci. Afr.* **2022**, *15*, e01100. [CrossRef]
13. Hoerl, A.E.; Kennard, R.W.; Baldwin, K.F. Ridge regression: Some simulation. *Commun. Stat. Theory Methods* **1975**, *4*, 105–123. [CrossRef]
14. Schaeffer, R.L. Alternative estimators in logistic regression when the data is collinear. *J. Stat. Comput. Simul.* **1986**, *25*, 75–91. [CrossRef]
15. Özkale, M.R.; Kaciranlar, S. The restricted and unrestricted two-parameter estimators. *Commun. Statist. Theor. Meth* **2007**, *36*, 2707–2725. [CrossRef]
16. Liu, K. Using Liu-type estimator to combat collinearity. *Commun. Stat.-Theory Methods* **2003**, *32*, 1009–2003. [CrossRef]
17. Inan, D.; Erdogan, B.E. Liu-Type logistic estimator. *Commun. Stat. Simul. Comput.* **2013**, *42*, 1578–1586. [CrossRef]

18. Huang, J. A Simulation Research on a Biased Estimator in Logistic Regression Model. In *Computational Intelligence and Intelligent Systems. ISICA 2012. Communications in Computer and Information Science*; Li, Z., Li, X., Liu, Y., Cai, Z., Eds.; Springer: Berlin/Heidelberg, Germany, 2012; Volume 316. [CrossRef]
19. Farebrother, R.W. Further results on the mean square error of ridge regression. *J. R. Stat. Soc. Ser. B* **1976**, *38*, 248–250. [CrossRef]
20. Trenkler, G.; Toutenburg, H. Mean squared error matrix comparisons between biased estimators—An overview of recent results. *Stat. Pap.* **1990**, *31*, 165–179. [CrossRef]
21. Kibria, B.M.G. Performance of some new ridge regression estimators. *Commun. Stat. Simul. Comput.* **2003**, *32*, 419–435. [CrossRef]
22. Qasim, M.; Amin, M.; Ullah, M.A. On the performance of some new liu parameters for the gamma regression model. *J. Stat. Comput Simul* **2018**, *88*, 3065–3080. [CrossRef]
23. Amin, M.; Akram, M.N.; Majid, A. On the estimation of Bell regression model using ridge estimator. *Commun. Stat. Simul. Comput.* **2021**. [CrossRef]
24. Lukman, A.F.; Zakariya, A.; Kibria, B.M.G.; Ayinde, K. The KL estimator for the inverse gaussian regression model. *Concurrency Computat. Pract. Exper.* **2021**, *33*, e6222. [CrossRef]
25. Lukman, A.F.; Aladeitan, B.; Ayinde, K.; Abonazel, M.R. Modified ridge-type for the Poisson Regression Model: Simulation and Application. *J. Appl. Stat.* **2021**, *49*, 2124–2136. [CrossRef] [PubMed]
26. Lukman, A.F.; Adewuyi, E.; Månsson, K.; Kibria, B.M.G. A new estimator for the multicollinear poisson regression model: Simulation and application. *Sci. Rep.* **2021**, *11*, 3732. [CrossRef]
27. Amin, M.; Qasim, M.; Amanullah, M.; Afzal, S. Performance of some ridge estimators for the gamma regression model. *Stat. Pap.* **2020**, *61*, 997–1026. [CrossRef]
28. Amin, M.; Qasim, M.; Afzal, S.; Naveed, M. New ridge estimators in the inverse Gaussian regression: Monte Carlo simulation and application to chemical data. *Commun. Stat. Simul. Comput.* **2020**, *51*, 6170–6187. [CrossRef]
29. Naveed, M.; Amin, M.; Afzal, S.; Qasim, M. New shrinkage parameters for the inverse Gaussian liu regression. *Commun. Stat. Theory Methods* **2020**, *51*, 3216–3236. [CrossRef]
30. Qasim, M.; Amin, M.; Omer, T. Performance of some new Liu parameters for the linear regression model. *Commun. Stat. Theory Methods* **2020**, *49*, 4178–4196. [CrossRef]
31. Ayinde, K.; Lukman, A.F.; Samuel, O.O.; Ajiboye, S.A. Some new adjusted ridge estimators of linear regression model. *Int. J. Civil Eng. Technol.* **2018**, *9*, 2838–2852.
32. Asar, Y.; Genç, A. Two-parameter ridge estimator in the binary logistic regression. *Commun. Stat.-Simul. Comput.* **2017**, *46*, 7088–7099. [CrossRef]
33. Kibria, B.M.G.; Banik, S. Some ridge regression estimators and their performances. *J. Mod. Appl. Stat. Methods* **2016**, *15*, 206–238. [CrossRef]
34. Özkale, M.R.; Arıcan, E. A new biased estimator in logistic regression model. *Statistics* **2016**, *50*, 233–253. [CrossRef]
35. Varathan, N.; Wijekoon, P. Optimal generalized logistic estimator. *Commun. Stat. Theory Methods* **2018**, *47*, 463–474. [CrossRef]
36. Saleh, A.K.; Md, E.; Arashi, M.; Kibria, B.M.G. *Theory of Ridge Regression Estimation with Applications*; John Wiley: Hoboken, NJ, USA, 2019.
37. Newhouse, J.P.; Oman, S.D. *An Evaluation of Ridge Estimators*; P-716-PR; Rand Corporation: Santa Monica, CA, USA, 1971; pp. 1–28.
38. Gibbons, D.G. A simulation study of some ridge estimators. *J. Am. Stat. Assoc.* **1981**, *76*, 131–139. [CrossRef]
39. McDonald, G.; Galarneau, D.I. A Monte Carlo evaluation of some ridge-type estimators. *J. Am. Stat. Assoc.* **1975**, *70*, 407–416. [CrossRef]
40. Lesaffre, E.; Marx, B.D. Collinearity in generalized linear regression. *Commun. Stat. Theory Methods* **1993**, *22*, 1933–1952. [CrossRef]
41. Gujarati, D.N. *Basic Econometrics*; McGraw-Hill: New York, NY, USA, 1995.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.