# A New Ridge Type Estimator in the Logistic Regression Model with Correlated Regressors

OLADAPO O. J[1], IDOWU J. I.[1,*], OWOLABI A. T.[1], AYINDE K.[2], ADEJUMO T. J.[1]

[1]Department of Statistics,
Ladoke Akintola University of Technology,
Ogbomoso, Oyo state,
NIGERIA

[2]Department of Mathematics and Statistics,
Northwest Missouri State University,
Maryville, Missouri,
USA

*Corresponding Author

*Abstract:* - The maximum likelihood (ML) technique is always one of the most widely employed to estimate model parameters in logistic regression models. However, due to the problem of multicollinearity, unstable parameter estimates, and inaccurate variance which affects confidence intervals and hypothesis tests can be achieved. A new two-parameter biased estimator is proposed in this paper to handle multicollinearity in binary logistic regression models. The proposed estimator's properties were determined, and five (5) different types of biasing parameter k (generalized, maximum, median, mid-range, and arithmetic mean) were applied in this work. The necessary and sufficient criteria for the new two-parameter biased estimators to outperform the existing estimators is considered. In addition, Monte Carlo simulation studies are carried out to compare the performance of the proposed biased estimator. Finally, a numerical example is provided to support the theoretical and simulations findings.

*Key-Words:* - Multicollinearity, Logistics, Estimators, Simulations, Parameter, Biasing, Variance.

Received: July 5, 2024. Revised: November 13, 2024. Accepted: December 2, 2024. Published: December 20, 2024.

## 1 Introduction

Logistic regression models are widely used in the social sciences, in engineering, in economic research, and in the medical fields. The general form of the logistic regression model is defined as:

$$y_i = \pi_i + \varepsilon_i \qquad i = 1,...,n \tag{1}$$

Which is based on a Bernoulli distribution with parameter $\pi_i$ *as*:

$$\pi_i = \frac{\exp\left(x_i^{'}\beta\right)}{1+\exp\left(x_i^{'}\beta\right)} \tag{2}$$

Where $x_i^{'}$ is the ith row of X, which is a $n \times (p+1)$ data matrix with p explanatory variables and $\beta$ is a $(p+1) \times 1$ vector of coefficients, $\varepsilon_i$ is independent with mean zero and variance $\pi_i(1-\pi_i)$.

The most well-known estimation method for estimating the parameter $\beta$ is the maximum likelihood method. The maximum likelihood estimator (MLE) of $\beta$ can be obtained using the iterative weighted least square approach as follows.

$$\hat{\beta}_{MLE} = \left(X'\hat{W}X\right)^{-1} X'\hat{W}z \tag{3}$$

Where z is the ith element of the column vector, which equals $z_i = \log(\hat{\pi}_i) + \frac{y_i - \hat{\pi}_i}{\hat{\pi}_i(1-\hat{\pi}_i)}$ and $\hat{W}_i = diag\left[\hat{\pi}_i(1-\hat{\pi}_i)\right]$, is the asymptotically unbiased estimate of $\beta$. The asymptotic covariance matrix of $\hat{\beta}_{MLE}$ is $\text{cov}\left(\hat{\beta}_{MLE}\right) = \left(X'\hat{W}X\right)^{-1}$.

The most popular known method of parameter estimation in logistic regression is the concept of maximum likelihood (ML) method. However, when the explanatory variables are highly correlated, it is termed a multicollinearity problem. When there is multicollinearity, some component of the ML estimator may be imprecise and may not show the true effects of the corresponding covariates, and also, the asymptotic covariance of the ML estimator inflates. Therefore, estimators' alternative to the ML estimator were proposed to handle the problem of multicollinearity, [1], [2], [3], [4], [5], [6], [7], [8].

To mitigate the negative impacts of multicollinearity, the biased estimators that are alternatives to the MLE are generalized in the same way that the linear regression model was. To begin, [9] presented the Logistic Ridge Estimator (LRE) shown below:

$$\hat{\beta}_{LRE} = \left(X'\hat{W}X + kI\right)^{-1} X'\hat{W}X\hat{\beta}_{MLE} \quad (4)$$

k is a shrinkage parameter and $\hat{W}$ is the estimate of W using.

According to them, this estimator was appealing because it did not rely too heavily on logistic regression software. They demonstrated that, under certain situations, the ridge type estimator has a lower mean square error (MSE) than the ML estimator. The influence of multicollinearity and sample size was investigated empirically by [10], because the major goal of ridge type estimators is to determine a value for the ridge parameter, [11] extended the approach of [12] to logistic regression in determining k.

Recently, the K-L estimator (KLE) introduced by [13] has demonstrated higher performance than the ordinary least squares, the RRE and the LE for parameter estimation in the linear regression model (LRM). [14], extended the KLE to the Logistic regression model and defined it as:

$$\hat{\beta}_{LKLE} = \left(X'\ddot{W}X + kI\right)^{-1}\left(X'\hat{W}X - kI\right)\hat{\beta}_{MLE} \quad (5)$$

where k (k > 0) is the LKLE biasing parameter. The LKLE was shown to have a smaller MSE than the ML, and other estimators used in comparison under some certain conditions used in their study.

[15], also proposed a modified ridge-type (MRT) estimator in a linear regression model. This research is also being applied to the logistic regression model. The Logistic modified ridge- type (LMRT) [16] estimator is defined as:

$$\hat{\beta}_{LMRT} = \left(X'\hat{W}X + k(1+d)I\right)^{-1} X'\hat{W}X\hat{\beta}_{MLE} \quad (6)$$

where (k > 0) and $0 < d < 1$ are the LMRT biasing parameters. The LMRT was shown to have a smaller MSE than the ML, and other estimators used in comparison under some certain conditions used in their study. It has been noted that the biasing parameter has an effect on the value of the MSE.

The use of biased estimators with two biasing parameters has become increasingly widespread as an alternative to the LRE and LLE. However, since the performances of these biased estimators do depend on the two biasing parameters, it is always kind of difficult to determine the optimum performances, [7]. The goal is to select appropriate values for k and d in order to achieve a reduction in MSE. There are numerous techniques available for estimating the biasing parameters k and d. In these circumstances, k can be estimated based on the biasing parameter d, or d can be computed based on the biasing parameter k, [3], [8], [10].

Therefore, the primary objective of this study is to introduce a new general biased estimator and propose some biasing parameters k, which would be an alternative method for overcoming multicollinearity in binary logistic regression models. The organization of the article is as follows: In Section 2, the methods used are explained following; a new biased estimator named the logistic modified Liu Ridge type estimator is defined and some of its properties are given; demonstrating the superiority of this estimator over the other biased estimators using the matrix mean square error (MMSE) criteria and Monte Carlo simulation studies are executed in Section results and discussion of the simulation results and a real data application is provided to illustrate the performances of the proposed biased estimators. Finally, the conclusions of the study are given in Section 4.

## 2 Methods

### 2.1 Proposed Estimator

Recently, [17] proposed the modified Liu Ridge type estimator in the linear regression model, this estimator is defined as:

$$\hat{\beta}_{MLRT} = \left(X'X + I\right)^{-1}\left(X'X + dI\right)\left(X'X + k(1+d)I\right)^{-1} X'y \quad (7)$$

Therefore, a logistic version of this estimator is developed and is called the logistic modified Liu Ridge type estimator which is defined as:

$$\hat{\beta}_{LMLRT} = (X'\hat{W}X + I_p)^{-1}(X'\hat{W}X + dI_p)(X'\hat{W}X + k(1+d)I_p)^{-1} X'\hat{W}X\hat{\beta}_{MLE} \quad (8)$$

k > 0 and 0 < d < 1, our proposed estimator $\hat{\beta}_{LMLRT}(k,d)$, is a general class of biased estimator, for different values of k and d it becomes the ML, LRE, LLE, and LMRTE, i.e.

$$\hat{\beta}_{LMLRT}(0,1)=\hat{\beta}_{MLE}$$

$$\hat{\beta}_{LMLRT}\left(\frac{k}{2},1\right)=\hat{\beta}_{LRE}$$

$$\hat{\beta}_{LMLRT}(0,d)=\hat{\beta}_{LLE}=\left(X'\hat{W}X+I\right)^{-1}\left(X'\hat{W}X+dI\right)\hat{\beta}_{MLE}$$

$$\hat{\beta}_{LMLRT}(k(1+d),1)=\hat{\beta}_{LMRT}$$

## 2.2 MSE and MMSE Properties of the Estimators

The MMSE, which contains all important information about the estimators, can be used as a comparison criteria. The MMSE and MSE, which is the trace of the MMSE of an estimator $\hat{\beta}$, are defined by:

$$\text{MSEM}\left(\hat{\beta}\right)=Cov\left(\hat{\beta}\right)+bias\left(\hat{\beta}\right)\left(bias\left(\hat{\beta}\right)\right)^{T} \quad (9)$$

$$\text{MSE}\left(\hat{\beta}\right)=\text{trace}\left(Cov\left(\hat{\beta}\right)\right)+\left(bias\left(\hat{\beta}\right)\right)^{T}bias\left(\hat{\beta}\right) \quad (10)$$

If $\hat{\beta}_1$ *and* $\hat{\beta}_2$ are two estimators of the coefficient vector, then $\hat{\beta}_2$ is superior to $\hat{\beta}_1$ if and only if $MMSE\left(\hat{\beta}_1\right)-MMSE\left(\hat{\beta}_2\right)\geq 0$. It's proved that $MMSE\left(\hat{\beta}_1\right)-MMSE\left(\hat{\beta}_2\right)\geq 0$ implies $MSE\left(\hat{\beta}_1\right)-MSE\left(\hat{\beta}_2\right)\geq 0$. But the converse is not always true.

By the use of spectral decomposition of the matrix, $X'\hat{W}X=T\Lambda T'$, where T is said to be the matrix whose columns are always the eigenvectors of $X'\hat{W}X$ and $\Lambda$ is the matrix of eigenvalues of $X'\hat{W}X$ .the MSEM of the estimators is:

$$MSEM\left(\hat{\alpha}_{MLE}\right)=T\Lambda^{-1}T^{T} \quad (11)$$

$$MSEM\left(\hat{\alpha}_{LRR}\right)=T\Lambda_k\Lambda\Lambda_k^{T}T^{T}+k^{2}\Lambda_k\alpha\alpha^{T}\Lambda_k^{T} \quad (12)$$

$$\Lambda_k=\left(\Lambda+kI\right)^{-1}$$

$$MSEM\left(\hat{\alpha}_{LKL}\right)=T\Lambda_k^{*}\Lambda_k\Lambda^{-1}\Lambda^{T}{}_{k}\Lambda_k^{*T}T^{T}+4k^{2}T\Lambda_k\alpha\alpha^{T}\Lambda^{T}{}_{k}T^{T} \quad (13)$$

$$\Lambda_k^{*}=\left(\Lambda-kI_p\right)$$

$$MSEM\left(\hat{\alpha}_{LMRT}\right)=T\tilde{\Lambda}_k\Lambda^{-1}\tilde{\Lambda}_kT^{T}+\left(\tilde{\Lambda}_k-1\right)\alpha\alpha^{T}\left(\tilde{\Lambda}_k-1\right)^{T} \quad (14)$$

$$\tilde{\Lambda}_k=\Lambda\left(\Lambda+k(1+d)I\right)^{-1}$$

$$MSEM\left(\hat{\alpha}_{LMLRT}\right)=T\Lambda_c\Lambda_d\Lambda\Lambda_c^{T}T^{T}+T\left(\Lambda_c-1\right)\alpha\alpha^{T}\left(\Lambda_c-1\right)^{T}T^{T} \quad (15)$$

$$\Lambda_c=\left(\Lambda+I\right)^{-1}\left(\Lambda+dI\right)\left(\Lambda+k(1+d)I\right)^{-1}$$

$$\Lambda_d=\left(\Lambda+dI\right)^{-1}$$

Where $\alpha=T^{T}\beta$, then the MSE of the estimators are:

$$MSE\left(\hat{\alpha}_{MLE}\right)=\sum_{j=1}^{p}\frac{1}{\lambda_j} \quad (16)$$

$$MSE\left(\hat{\alpha}_{LRE}\right)=\sum_{j=1}^{p}\frac{\lambda_j}{\left(\lambda_j+k\right)^2}+k^2\sum_{j=1}^{p}\frac{\alpha_j^2}{\left(\lambda_j+k\right)^2} \quad (17)$$

$$MSE\left(\hat{\alpha}_{LKL}\right)=\sum_{j=1}^{p}\frac{\left(\lambda_j-k\right)^2}{\lambda_j\left(\lambda_j+k\right)^2}+4k^2\sum_{j=1}^{p}\frac{\alpha_j^2}{\left(\lambda_j+k\right)^2} \quad (18)$$

$$MSE\left(\alpha_{LMRT}\right)=\sum_{j=1}^{p}\frac{\lambda_j\left[k(1+d)\right]^2\hat{\alpha}_j^2}{\left(\lambda_j+k(1+d)\right)^2} \quad (19)$$

$$MSE\left(\hat{\alpha}_{LMLRT}\right)=\sum_{j=1}^{p}\frac{\lambda_j\left(\lambda_j+d\right)}{\left(\lambda_j+1\right)^2\left(\lambda_j+k\left(1+d\right)\right)^2}+\sum_{j=1}^{p}\frac{\left[\left(1-d\right)\lambda_j+k\left(1+d\right)\left(\lambda_j+1\right)\right]^2\alpha_j^2}{\left(\lambda_j+1\right)^2\left(\lambda_j+k\left(1+d\right)\right)^2} \quad (20)$$

## 2.3 The Superiority of the Modified Liu Ridge Type Estimator in Logistic Regression

Using the following lemma, the following theorems are presented to compare the MSE and MMSEs of the estimators:

**Lemma 2.3.1:** Let M be a positive definite (pd) matrix, that is, M>0, and α be some vector, then $M-\alpha\alpha^{T}\geq 0$ if and only if $\alpha^{T}M^{-1}\alpha\leq 1$, [18].

**Lemma 2.3.2:** Let $\hat{\beta}_j = A_{jy}, j = 1,2$ be two linear estimators of $\beta$. Suppose that $D = Cov(\hat{\beta}_1) - Cov(\hat{\beta}_2) > 0,\ where\ Cov(\hat{\beta}_j), j = 1,2$ denotes the covariance matrix of $\hat{\beta}_j$ and $b = Bias(\hat{\beta}_j) = (A_j X - 1)\beta,\ j = 1,2.$

consequently,

$$\Delta(\hat{\beta}_1 - \hat{\beta}_2) = MSEM(\hat{\beta}_1) - MSEM(\hat{\beta}_2) = \sigma^2 D + b_1 b_1^t - b_2 b_2^T > 0 \qquad (21)$$

If and only if $b_2^T[\sigma^2 D + b_1 b_1^T]^{-1} b_2 < 1$,
where $MSE(\hat{\beta}_j) = Cov(\hat{\beta}_j) + b_j^T b_j$ [19].

### 2.3.1 The Comparison of MLE and LMLRT

The following theorem provides a criterion for comparing MLE to LMLRT in the MMSE sense.

**Theorem 2.3.1:**
$MMSE(MLE) - MMSE(LMLRT) > 0$
$iff\ b_{LMLRT}^T T(\Lambda^{-1} - \Lambda_c \Lambda_d \Lambda \Lambda_c^T) T^T b_{LMLRT} < 1$

**Proof:** The following equations (11) and (15) can be used to calculate the difference between the MMSEs of the estimators MLE and LMLRT:

$$MMSE(MLE) - MMSE(LMLRT) = T(\Lambda^{-1} - \Lambda_c \Lambda_d \Lambda \Lambda_c^T) T^T$$

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

$\Lambda^{-1} - \Lambda_c \Lambda_d \Lambda \Lambda_c^T$ is said to be p.d iff $(\lambda_j + 1)^2 (\lambda_j + k(1+d))^2 - \lambda_j^2 (\lambda_j + d) > 0$. The result is valid for k>0 and 0<d<1 therefore Lemma 2.3.2 completes the proof.

### 2.3.2 The Comparison of LRE and LMLRT

The following theorem provides a criterion for comparing LRE to LMLRT in the MMSE sense.

**Theorem 2.3.2:**
$MMSE(LRE) - MMSE(LMLRT) > 0$
$iff\ b_{LMLRT}^T T(\Lambda_k \Lambda \Lambda_k^T - \Lambda_c \Lambda_d \Lambda \Lambda_c^T) T^T b_{LMLRT} < 1$

**Proof:** The difference between the MMSEs of the estimators LRE and LMLRT can be calculated using the following equations (12) and (15):

$$MMSE(LRE) - MMSE(LMLRT) = T(\Lambda_k \Lambda \Lambda_k^T - \Lambda_c \Lambda_d \Lambda \Lambda_c^T) T^T$$

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

$\Lambda_k \Lambda \Lambda_k^T - \Lambda_c \Lambda_d \Lambda \Lambda_c^T$ is said to be p.d iff $(\lambda_j + 1)^2 (\lambda_j + k(1+d))^2 \lambda_j - \lambda_j(\lambda_j + d)(\lambda_j + k)^2 > 0$. The result is valid for k>0 and 0<d<1 therefore Lemma 2.3.2 completes the proof.

### 2.3.3 The Comparison of LKL and LMLRT

The following theorem provides a criterion for comparing LKL to LMLRT in the MMSE sense.

**Theorem 2.3.3:**
$MMSE(LKL) - MMSE(LMLRT) > 0$
$iff\ b_{LMLRT}^T T(\Lambda_k^* \Lambda_k \Lambda^{-1} \Lambda^T{}_k \Lambda_k^{*T} - \Lambda_c \Lambda_d \Lambda \Lambda_c^T) T^T b_{LMLRT} < 1$

**Proof:** The difference between the MMSEs of the estimators LKL and LMLRT can be calculated using the following equations (13) and (15):

$$MMSE(LKL) - MMSE(LMLRT) = T(\Lambda_k^* \Lambda_k \Lambda^{-1} \Lambda^T{}_k \Lambda_k^{*T} - \Lambda_c \Lambda_d \Lambda \Lambda_c^T) T^T$$

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

$\Lambda_k^* \Lambda_k \Lambda^{-1} \Lambda^T{}_k \Lambda_k^{*T} - \Lambda_c \Lambda_d \Lambda \Lambda_c^T$ is said to be p.d iff $(\lambda_j + 1)^2 (\lambda_j + k(1+d))^2 (\lambda_j - k)^2 - \lambda_j^2 (\lambda_j + d)(\lambda_j + k)^2 > 0$. The result is valid for k>0 and 0<d<1 therefore Lemma 2.3.2 completes the proof.

### 2.3.4 The Comparison of LMRT and LMLRT

The following theorem provides a criterion for comparing LMRT to LMLRT in the MMSE sense.

**Theorem 2.3.4:**
$MMSE(LMRT) - MMSE(LMLRT) > 0$
$iff\ b_{LMLRT}^T T(\tilde{\Lambda}_k \Lambda^{-1} \tilde{\Lambda}_k - \Lambda_c \Lambda_d \Lambda \Lambda_c^T) T^T b_{LMLRT} < 1$

**Proof:** The difference between the MMSEs of the estimators LMRT and LMLRT can be calculated using the following equations (14) and (15):

$$MMSE\left(LMRT\right)-MMSE\left(LMLRT\right)=$$
$$T\left(\tilde{\Lambda}_k\Lambda^{-1}\tilde{\Lambda}_k-\Lambda_c\Lambda_d\Lambda\Lambda_c^T\right)T^T$$
$$=T\,diag\left(\frac{\lambda_j}{\left(\lambda_j+k\left(1+d\right)\right)^2}-\frac{\lambda_j\left(\lambda_j+d\right)}{\left(\lambda_j+1\right)^2\left(\lambda_j+k\left(1+d\right)\right)^2}\right)_{j=1}^{p}T^T$$

$\Lambda_k^*\Lambda_k\Lambda^{-1}\Lambda^T{}_k\Lambda_k^{*T}-\Lambda_c\Lambda_d\Lambda\Lambda_c^T$ is said to be p.d iff $\left(\lambda_j+1\right)^2\left(\lambda_j+k(1+d)\right)^2\lambda_j$ - $\lambda_j\left(\lambda_j+d\right)\left(\lambda_j+1\right)^2$ > 0. The result is valid for k>0 and 0<d<1 therefore Lemma 2.3.2 completes the proof.

### 2.3.5 Selection of Parameters k and d for LMLRT

There was development of some iterative methods to estimate k and d such that the mean squared error is minimized. Following, [17] their k and d in logistic form is gotten following the works of [20], differentiating the equation (20) with respect to the biasing parameters as shown below:

g(k,d)= $MSE\left(\hat{\alpha}(k,d)\right)=tr\left[MSEM\left(\hat{\alpha}(k,d)\right)\right]$

$$MSE\left(\hat{\alpha}_{LMLRT}\right)=\sum_{j=1}^{p}\frac{\lambda_j\left(\lambda_j+d\right)}{\left(\lambda_j+1\right)^2\left(\lambda_j+k\left(1+d\right)\right)^2}+\sum_{j=1}^{p}\frac{\left[\left(1-d\right)\lambda_j+k\left(1+d\right)\left(\lambda_j+1\right)\right]^2\alpha_j^2}{\left(\lambda_j+1\right)^2\left(\lambda_j+k\left(1+d\right)\right)^2}$$

With d fixed, an optimal value of k is the one that minimizes $MSE\left(\hat{\alpha}_{LMLRT}\right)$.

Then, by differentiating g(k, d) w.r.t. k and equating to 0, then:

$$k=\frac{\left(\lambda_i+d\right)-\left(1-d\right)\lambda_i\alpha_i^2}{\left(1+d\right)\left(\lambda_i+1\right)\alpha_i^2} \tag{22}$$

However, k depends on the unknown $\alpha_i$. For practical purposes, it will be replaced by its unbiased estimator $\hat{\alpha}_i$. Hence, this will be obtained

$$\hat{k}_{opt}=\frac{\left(\lambda_i+d\right)-\left(1-d\right)\lambda_i\hat{\alpha}_i^2}{\left(\lambda_i+1\right)\left(1+d\right)\hat{\alpha}_i^2} \tag{23}$$

As an operational estimator for k. Furthermore, when d = 1, in the above equation $\hat{k}$ *is replaced by* $\frac{\hat{k}}{2}$ and the above equation reduces to the biasing parameter proposed by[9]and its stated below as:

$$\hat{k}=\frac{1}{\alpha_i^2} \tag{24}$$

Following the works of [12], and applying their concepts the below shrinkage parameters are examined:

$$\hat{k}_{MED}=Median\left(\frac{\left(\lambda_i+d\right)-\lambda_i\left(1-d\right)\hat{\alpha}_i^2}{\left(1+d\right)\left(\lambda_i+1\right)\hat{\alpha}_i^2}\right) = \text{LMRT1} \tag{25}$$

$$\hat{k}_{MAX}=Maximum\left(\frac{\left(\lambda_i+d\right)-\lambda_i\left(1-d\right)\hat{\alpha}_i^2}{\left(1+d\right)\left(\lambda_i+1\right)\hat{\alpha}_i^2}\right) = \text{LMRT2} \tag{26}$$

$$\hat{k}_{MR}=\frac{1}{2}\left(\hat{k}_{MAX}+\hat{k}_{MIN}\right) = \text{LMRT3} \tag{27}$$

$$\hat{k}_{AM}=\frac{1}{p}\sum_{i=1}^{p}\frac{\left(\lambda_i+d\right)-\lambda_i\left(1-d\right)\hat{\alpha}_i^2}{\left(1+d\right)\left(\lambda_i+1\right)\hat{\alpha}_i^2} = \text{LMRT4} \tag{28}$$

Furthermore, the optimal value for d can also be obtained by minimizing $MSE\left(\hat{\alpha}_{LMLRT}\right)$

Let

$$MSE\left(\hat{\alpha}_{LMLRT}\right)=\sum_{j=1}^{p}\frac{\lambda_j\left(\lambda_j+d\right)}{\left(\lambda_j+1\right)^2\left(\lambda_j+k\left(1+d\right)\right)^2}+\sum_{j=1}^{p}\frac{\left[\left(1-d\right)\lambda_j+k\left(1+d\right)\left(\lambda_j+1\right)\right]^2\alpha_j^2}{\left(\lambda_j+1\right)^2\left(\lambda_j+k\left(1+d\right)\right)^2}$$

Then, by differentiating $MSE\left(\hat{\alpha}_{LMLRT}\right)$ w.r.t. d and equating to 0, *d* is obtained as:

$$d=\sum_{i=1}^{p}\frac{\left(\lambda_i+k\left(\lambda_i+1\right)\right)\left(\lambda_i^2-k\lambda_i^2-k\lambda_i\right)\alpha_i^2-\lambda_i^3}{\lambda_i^2+\left(\lambda_i-k\left(\lambda_i+1\right)\right)\left(\lambda_i^2-k\lambda_i^2-k\lambda_i\right)\alpha_i^2} \tag{29}$$

However, d depends on the unknown $\alpha_i$. For practical purposes, it will be replaced by its unbiased estimator $\hat{\alpha}_i$. Hence, this will be obtained:

$$\hat{d}_{opt}=\sum_{i=1}^{p}\frac{\left(\lambda_i+k\left(\lambda_i+1\right)\right)\left(\lambda_i^2-k\lambda_i^2-k\lambda_i\right)\hat{\alpha}_i^2-\lambda_i^3}{\lambda_i^2+\left(\lambda_i-k\left(\lambda_i+1\right)\right)\left(\lambda_i^2-k\lambda_i^2-k\lambda_i\right)\hat{\alpha}_i^2} \tag{30}$$

Noting that when k =0 the above equation leads to the Liu biasing estimator, the above estimator would be:

$$\hat{d} = \frac{\sum_{i=1}^{p} \lambda_i \left( \hat{\alpha}_i^2 - 1 \right)}{\sum_{i=1}^{p} \left( 1 + \lambda_i \alpha_i^2 \right)} = \hat{d}_{opt} \qquad (31)$$

By adopting the same method as [20] to select k and d at the same time. Considering:

$$k = \frac{(\lambda_i + d) - (1-d)\lambda_i \alpha_i^2}{(1+d)(\lambda_i + 1)\alpha_i^2} > 0;$$

It implies that $d. > \frac{\lambda_i \left( \alpha_i^2 - 1 \right)}{1 + \lambda_i \alpha_i^2}$

Now considering

$$d_{\max} = \max\left( \frac{\lambda_i \left( \alpha_i^2 - 1 \right)}{1 + \lambda_i \alpha_i^2} \right)$$

Therefore, the following algorithm are used:

Step 1: Obtain $\hat{d}_{\max}$ after replacing $\alpha_i$ with their unbiased estimates as started earlier

Step 2: Compute $\hat{k}_{opt}$ using above computed $\hat{d}_{\max}$ in step 1

Step 3: Compute $\hat{d}_{opt}$ using $\hat{k}_{opt}$ of step 2

Step 4: Finally, if $\hat{d}_{opt}$<1 or $\hat{d}_{opt}$>1 consider $\hat{d}_{opt}$= $\hat{d}_{\max}$ .

# 3 A Monte Carlo Simulation Study

## 3.1 Design of the Simulation

The simulation's effective factors are the degree of correlation among the explanatory variables, the sample size n, and the number of explanatory variables p. Following the works of [21] and [22], the explanatory variable was generated as:

$$x_{ij} = \left(1 - \rho^2\right)^{1/2} z_{ij} + \rho z_{ip} \quad i=1, 2,\dots, n, \; j=1,2,\dots,p \qquad (32)$$

where $z_{ij}$ denotes an independent standard normal distribution with a mean zero and a variance of one unit, $\rho$ also denotes the correlation between any two explanatory factors or variables and p denotes the number of explanatory variables. $\rho$ will take the values of 0.8, 0.9, 0.95 and 0.99 respectively. Moreover, the values of degree of correlations between the variables will be the same. In addition, sample sizes equal to 150,200,250,300,350,400,450,500,600 and 700 are considered. Finally, the logit model consisting of *p* =4, 5, 6, and 8 are considered in the design of the experiment. The response variable is generated from the Bernoulli distribution such that:

$y_i \sim \; Ber(\nu_i)$ and $\pi_i = \frac{e^{x_i \beta}}{1 + e^{x_i \beta}}$. According to [23], the parameter values were set so that $\beta' \beta = 1$. The estimated MSE is calculated as

$$MSE\left(\hat{\beta}\right) = \frac{1}{1000} \sum_{i=1}^{1000} \sum_{j=1}^{p} \left( \hat{\beta}_{ij} - \beta_j \right) \qquad (33)$$

## 3.2 Results and Discussion

The experiment is repeated 1000 times. Table 1, Table 2, Table 3, Table 4 and Table 5 (Appendix) show the simulated SMSE values of the estimators and summarized the findings as follows:

- Based on the simulated SMSE values shown in Appendix in Table 1 and Table 2, the proposed estimator LMLRT outperforms MLE and thereby the proposed LMLRT1 outperforms all other estimators compared within this study when the degree of correlation is at 0.8 and 0.9. In addition, when rho is at 0.95 LMLRT2 dominates other estimators and when rho is at 0.99 LMLRT3 shows better performance than other estimators used in comparison.
- Based on the simulated SMSE values shown in Table 3, Table 4 and Table 5 (Appendix) the proposed estimator LMLRT outperforms MLE and thereby the proposed LMLRT1 outperforms all other estimators compared with in this study when the rho value is at 0.8 and 0.9. In addition, the estimator LMLRT1 when rho is at 0.95 is competing favorably with LMLRT2 and when rho is at 0.99 LMLRT3 still maintains its position as being the best.
- It is also worth noting that when the correlation between the explanatory variables is as high as 0.99, LMLRT3 outperforms all other estimators in this study.
- It is also worth noting that when the correlation between the explanatory variables is between 0.8 and 0.9, LMLRT1 outperforms all other estimators in this study.
- With a few exceptions, the table also shows that increasing the sample size n decreases the SMSE of the estimators. If the correlation between the explanatory variables is high, increasing the sample size has a favorable effect on the estimators.

- Furthermore, by evaluating the performances of the estimators based on the degree of correlation, it was seen that SMSE values of the estimators increase as the degree of correlation increases in general.
- Also, when considering the performances of the estimators according to the degree of correlation $\rho$, it can be concluded that SMSE values of the estimators increases as the number of explanatory variable increases at a $\rho$ level.

### 3.3 Numerical Example

[24], were the first to use the dataset by applying it to a logistic model to investigate the effects of temperature, pH, and soluble solids content on the response of Alicyclobacillus growth likelihood in apple juice. The matrix's eigenvalues are 13464.7990, 1715.9257, 56.5515, and 3.5445. As a result, the condition index (C.I) is 61.6342, indicating that multicollinearity exists in the model. Table 6 (Appendix) shows the estimated regression coefficient values from each estimator, as well as the accompanying mean squared error.

When there is multicollinearity, the ML estimator performs the least well, as expected. The choice of the biasing parameters k and d determines the efficiency of biased estimators. All of the proposed estimators performed admirably, and one of them has the minimum mean square error, which corresponds to the simulation outcome.

## 4 Conclusion

The proposed logistic new two-parameter estimator in this paper to alleviate the multicollinearity problem in a logistic regression model. In addition, this new estimator outperformed the current ones in terms of MSE. Both the theoretical technique and the Monte Carlo simulation study were used to evaluate the estimators' performance. The degree of correlation, sample size, and number of explanatory variables were all varied in the Monte Carlo experiment design. The results revealed that the estimators' performance was strongly dependent on these criteria. Finally, to demonstrate the efficiency of the proposed estimator, pena dataset was used and the results corresponded to some extent with those of the simulation study. The outcomes of this study are recommended to practitioners and applied researchers who employ a logistic regression model with correlated explanatory variables in their varied domains and research efforts.

*References:*

[1] Abonazel M. R., Rasha A. F.,( 2018), *Liu-Type Multinomial Logistic Estimator Liu-Type Multinomial Logistic Estimator*. Sankhya B 81(2): 203-225 (2019).

[2] Asar Y., (2016), Liu-type logistic estimators with optimal shrinkage parameter. *J. Modern Appl. Statist. Methods*, 15, 738–751

[3] Asar Y., ( 2017), Some New Methods to Solve Multicollinearity in Logistic Regression. *Communications in Statistics-Simulation and Computation*, 46(4):2576–86.

[4] Asar Y., Genç A., ( 2017), Two-parameter ridge estimator in the binary logistic regression. *Commun Stat Simul Comput.* 46(9): 7088-7099

[5] Asar Y., Wu J., (2019), An Improved and Efficient Biased Estimation Technique in Logistic Regression Model. *Communications in Statistics-Theory and Methods*, 49(9):2237–2252.

[6] Awwad F. A., Odeniyi K.A., Dawoud I., Yahya Z., Abonazel M.R, Kibria B.M, Eldin E.T. (2022), New Two-Parameter Estimators for the Logistic Regression Model with Multicollinearity. *WSEAS Transactions on Mathematics*, 21,403–414, https://doi.org/10.37394/23206.2022.21.48.

[7] Ertan E., Kadri U A, (2020), Communications in Statistics - Theory and Methods A New Liu-Type Estimator in Binary Logistic Regression Models." *Communications in Statistics - Theory and Methods*, 0(0):1–25.

[8] Oladapo O.J., Alabi O.O., Ayinde K., (2023), Performance of Some Dawoud-Kibra Estimators for Logistic Regression Model: Application to Pena data set. *Equations*. 3:130-139

[9] Schaefer R.L., Roi L.D., Wolfe R.A., (1984), A ridge logistic estimator, *Commu- nications in Statistics - Theory and Methods*, 13(1): 99–113.

[10] Ozkale M.,( 2016), Iterative Algorithms of Biased Estimation Methods in Binary Logistic Regression. *Statistical*, vol. 57 (4):991–1016.

[11] Schaefer R. L., ( 1986), Alternative regression collinear estimators in logistic when the data are col- linear. *Journal of Statistical Computation and Simulation*, 25 (1–2):75–91.

[12] Hoerl A.E., Kennar R.W., Baldwin K.F., (1975) Ridge regression: Some simulation. Commun. Stat. *Theory Methods*. 4: 105–123

[13] Kibria B.M.G., Lukman A.F. (2020) A new ridge type estimator for the linear regression

model: Simulations and applications. *Hindawi Scientifica*. 2020:1-16

[14] Lukman A.F., Kibria B.M.G., Nziku C.K., Amin M.; Adewuyi E.T.,Farghali, R.,(2023), K-L: Estimator: Dealing with Multicollinearity in the Logistic Regression Model. *Mathematics*, 11, 340.

[15] Lukman A.F., Ayinde K., Binuomote S., Clement O.A., ( 2019a), Modified ridge-type estimator to combat multicollinearity: Application to chemical data. *Journal of Chemometrics*, 33: e3125.

[16] Lukman A. F., Adewuy E.T., Onate A.C., Ayinde K.,( 2020), A Modified Ridge-Type Logistic Estimator." Iranian Journal of Science and Technology, *Transactions A: Science*, 44(2):437–43.

[17] Aslam M., Ahmad S.,( 2020), The Modified Liu-Ridge-Type Estimator : A New Class of Biased Estimators to Address Multicollinearity. Communications in Statistics - *Simulation and Computation*, 51(11):6591-6609.

[18] Farebrother R. W.,(1976), Further results on the mean square error of ridge regression. *Journal of the Royal Statistical Society: Series B (Methodological)*, 38 (3):248–250

[19] Trenkler G., Toutenburg H.,(1990), Mean squared error matrix comparisons between biased estimators—An overview of recent results. *Stat. Pap.* 31, 165–179.

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

[21] McDonald G. C., Galarneau, D. I., (1975), A monte carlo evaluation of some ridge-type estimators," *Journal of the American Statistical Association*, 70(350): 407–416.

[22] Kibria B.M.G., (2003), Performance of some new ridge regression estimators, Communications in Statistics, *Simulation and Computation*, 32(2), 419–435.

[23] Newhouse J. P., Oman S. D., (1971), An evaluation ofridge estimators. *Rand Corporation*(P-716-PR), 1-16.

[24] Pena W., Massaguer P., Zuniga A., Saraiva S.H., (2011) Modeling the growth limit of Alicyclobacillus acidoterrestris CRA7152 in apple juice: effect of pH, Brix, temperature and nisin concentration. *Journal of Food Processing and Preservation*, 35 (4):509-517.

**Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)**

The authors equally contributed in the present research, at all stages from the formulation of the problem to the final findings and solution.

**Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself**

No funding was received for conducting this study.

**Conflict of Interest**

The authors have no conflicts of interest to declare.

# APPENDIX

Table 1. Estimated MSE when n=150 and 200

| n | rho | p | MLE | RIDGE | K-L | MRT | LMLRT | LMLRT1 | LMLRT2 | LMLRT3 | LMLRT4 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 150 | 0.8 | 4 | 0.5796 | 0.2965 | 0.4216 | 0.5427 | 0.4716 | ***0.2044*** | 0.2967 | 0.3944 | 0.3397 |
| | | 5 | 0.7947 | 0.3311 | 0.5387 | 0.7043 | 0.5705 | ***0.2212*** | 0.3599 | 0.5039 | 0.4355 |
| | | 6 | 1.0624 | 0.3954 | 0.7031 | 0.8850 | 0.7144 | ***0.2334*** | 0.4106 | 0.5768 | 0.5037 |
| | | 8 | 1.6231 | 0.5331 | 1.0604 | 1.1444 | 0.9507 | ***0.2940*** | 0.4559 | 0.6504 | 0.5787 |
| | 0.9 | 4 | 1.1705 | 0.4714 | 0.7588 | 0.9453 | 0.7521 | ***0.2514*** | 0.3329 | 0.4261 | 0.3754 |
| | | 5 | 1.6978 | 0.5657 | 1.0170 | 1.2033 | 0.9370 | ***0.3190*** | 0.3757 | 0.4948 | 0.4331 |
| | | 6 | 2.2822 | 0.6924 | 1.3432 | 1.4718 | 1.1720 | ***0.3500*** | 0.3965 | 0.5338 | 0.4669 |
| | | 8 | 3.4482 | 0.9488 | 2.0761 | 1.9938 | 1.6654 | 0.4490 | ***0.4326*** | 0.6164 | 0.5357 |
| | 0.95 | 4 | 2.3606 | 0.7823 | 1.3538 | 1.5374 | 1.2089 | 0.4010 | ***0.3757*** | 0.3956 | 0.3915 |
| | | 5 | 3.5776 | 1.0351 | 1.9423 | 1.9905 | 1.6049 | 0.6208 | ***0.4235*** | 0.4530 | 0.4358 |
| | | 6 | 4.7369 | 1.2658 | 2.6113 | 2.4294 | 2.0442 | 0.6434 | ***0.4022*** | 0.4685 | 0.4356 |
| | | 8 | 7.4971 | 1.8423 | 4.3192 | 3.5441 | 3.1483 | 0.8237 | ***0.4339*** | 0.5511 | 0.4912 |
| | 0.99 | 4 | 13.4932 | 3.7645 | 6.5523 | 5.6816 | 5.5197 | 2.5290 | 1.0759 | ***0.5724*** | 0.8919 |
| | | 5 | 20.5679 | 5.1738 | 10.4741 | 8.0181 | 7.7544 | 3.8675 | 1.0016 | ***0.4885*** | 0.7423 |
| | | 6 | 25.8467 | 6.1231 | 13.6637 | 9.8644 | 9.8294 | 3.1077 | 0.8624 | ***0.4473*** | 0.6325 |
| | | 8 | 40.9413 | 9.1203 | 22.9815 | 15.3945 | 15.9127 | 4.1403 | 0.8473 | ***0.4474*** | 0.5691 |
| 200 | 0.8 | 4 | 0.3916 | 0.2184 | 0.2952 | 0.4140 | 0.3615 | **0.1576** | 0.2268 | 0.3143 | 0.2658 |
| | | 5 | 0.6437 | 0.3082 | 0.4689 | 0.6211 | 0.5277 | **0.2117** | 0.3031 | 0.4284 | 0.3704 |
| | | 6 | 0.8897 | 0.3619 | 0.6054 | 0.7843 | 0.6308 | **0.2098** | 0.3527 | 0.5149 | 0.4488 |
| | | 8 | 1.2845 | 0.4573 | 0.8585 | 1.0079 | 0.8316 | **0.2633** | 0.4165 | 0.6162 | 0.5435 |
| | 0.9 | 4 | 0.7981 | 0.3506 | 0.5322 | 0.7260 | 0.5731 | **0.2040** | 0.2848 | 0.3793 | 0.3270 |
| | | 5 | 1.3254 | 0.5094 | 0.8678 | 1.0556 | 0.8483 | **0.2840** | 0.3402 | 0.4536 | 0.3977 |
| | | 6 | 1.8711 | 0.6049 | 1.1251 | 1.2971 | 1.0190 | **0.3243** | 0.3706 | 0.5057 | 0.4407 |
| | | 8 | 2.6269 | 0.7650 | 1.6083 | 1.6800 | 1.3579 | **0.3798** | 0.4179 | 0.6143 | 0.5319 |
| | 0.95 | 4 | 1.6061 | 0.5568 | 0.9327 | 1.1685 | 0.8844 | **0.2906** | 0.3395 | 0.3938 | 0.3700 |
| | | 5 | 2.7185 | 0.8755 | 1.5883 | 1.7183 | 1.3897 | 0.4987 | **0.3769** | 0.4328 | 0.4052 |
| | | 6 | 3.8387 | 1.0614 | 2.1348 | 2.1099 | 1.7154 | 0.5555 | **0.3869** | 0.4594 | 0.4237 |
| | | 8 | 5.5954 | 1.4089 | 3.2041 | 2.8579 | 2.4266 | 0.6454 | **0.3890** | 0.5358 | 0.4684 |
| | 0.99 | 4 | 8.8079 | 2.3613 | 4.2401 | 3.8125 | 3.4847 | 1.4688 | 0.6567 | ***0.4210*** | 0.5863 |
| | | 5 | 14.3758 | 3.8963 | 7.4887 | 6.0546 | 5.5617 | 2.9135 | 0.7786 | ***0.4370*** | 0.6326 |
| | | 6 | 21.4196 | 5.2274 | 11.4633 | 8.2711 | 8.0812 | 2.6524 | 0.9027 | ***0.4347*** | 0.6548 |
| | | 8 | 31.9423 | 7.0465 | 17.9255 | 12.4407 | 12.6170 | 3.0267 | 0.6826 | ***0.4110*** | 0.4858 |

*Bold values show the smallest MSE*

Table 2. Estimated MSE when n=250 and 300

| n | Rho | p | MLE | RIDGE | K-L | MRT | LMLRT | LMLRT1 | LMLRT2 | LMLRT3 | LMLRT4 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 250 | 0.8 | 4 | 0.3389 | 0.1935 | 0.2578 | 0.3807 | 0.3324 | **0.1418** | 0.1886 | 0.2568 | 0.2180 |
| | | 5 | 0.4684 | 0.2425 | 0.3476 | 0.4971 | 0.4247 | **0.1730** | 0.2398 | 0.3515 | 0.2985 |
| | | 6 | 0.6067 | 0.2957 | 0.4498 | 0.5952 | 0.5135 | **0.1920** | 0.3020 | 0.4392 | 0.3898 |
| | | 8 | 0.9848 | 0.3936 | 0.7000 | 0.8458 | 0.7102 | **0.2284** | 0.3996 | 0.5842 | 0.5260 |
| | 0.9 | 4 | 0.6863 | 0.3133 | 0.4698 | 0.6695 | 0.5297 | **0.1810** | 0.2545 | 0.3427 | 0.2927 |
| | | 5 | 0.9953 | 0.4108 | 0.6715 | 0.8880 | 0.7126 | ***0.2406*** | 0.2787 | 0.3976 | 0.3355 |
| | | 6 | 1.2397 | 0.4749 | 0.8243 | 1.0068 | 0.8153 | ***0.2580*** | 0.3404 | 0.4846 | 0.4185 |
| | | 8 | 2.1362 | 0.6905 | 1.3751 | 1.4857 | 1.2214 | ***0.3383*** | 0.3934 | 0.5819 | 0.5009 |
| | 0.95 | 4 | 1.3855 | 0.5102 | 0.8360 | 1.0762 | 0.8106 | ***0.2619*** | 0.2930 | 0.3525 | 0.3227 |
| | | 5 | 1.9988 | 0.6814 | 1.2084 | 1.4231 | 1.1268 | 0.3881 | **0.3174** | 0.3880 | 0.3516 |
| | | 6 | 2.5455 | 0.8113 | 1.5298 | 1.6331 | 1.3237 | 0.4275 | **0.3272** | 0.4480 | 0.3876 |
| | | 8 | 4.4491 | 1.2178 | 2.6283 | 2.4524 | 2.0787 | 0.5884 | **0.3770** | 0.5257 | 0.4577 |
| | 0.99 | 4 | 7.0001 | 1.9259 | 3.4004 | 3.2123 | 2.7879 | 1.2674 | 0.6442 | 0.4158 | 0.5625 |
| | | 5 | 10.6428 | 2.8765 | 5.5976 | 4.7262 | 4.2751 | 1.8874 | 0.5906 | **0.3684** | 0.5067 |
| | | 6 | 14.0437 | 3.6615 | 7.6001 | 6.0468 | 5.5868 | 1.9008 | 0.6038 | **0.3948** | 0.4940 |
| | | 8 | 24.9821 | 5.8387 | 14.1309 | 10.0496 | 10.0759 | 2.6666 | 0.6122 | **0.4016** | 0.4662 |
| 300 | 0.8 | 4 | 0.2649 | 0.1670 | 0.2114 | 0.3068 | 0.2774 | **0.1230** | 0.1653 | 0.2255 | 0.1897 |
| | | 5 | 0.3884 | 0.2125 | 0.2968 | 0.4389 | 0.3799 | **0.1579** | 0.2263 | 0.3235 | 0.2777 |
| | | 6 | 0.5511 | 0.2701 | 0.4102 | 0.5654 | 0.4806 | **0.1760** | 0.2820 | 0.4144 | 0.3658 |
| | | 8 | 0.7246 | 0.3176 | 0.5298 | 0.6898 | 0.5839 | **0.1972** | 0.3601 | 0.5326 | 0.4905 |
| | 0.9 | 4 | 0.5477 | 0.2728 | 0.3924 | 0.5721 | 0.4639 | **0.1577** | 0.2152 | 0.2923 | 0.2479 |
| | | 5 | 0.8046 | 0.3458 | 0.5534 | 0.7687 | 0.6094 | **0.2105** | 0.2602 | 0.3839 | 0.3197 |
| | | 6 | 1.1421 | 0.4333 | 0.7531 | 0.9658 | 0.7626 | **0.2312** | 0.3041 | 0.4460 | 0.3781 |
| | | 8 | 1.5609 | 0.5351 | 1.0275 | 1.1955 | 0.9618 | **0.2736** | 0.3732 | 0.5850 | 0.4999 |
| | 0.95 | 4 | 1.1400 | 0.4545 | 0.7192 | 0.9632 | 0.7315 | **0.2346** | 0.2828 | 0.3384 | 0.3121 |
| | | 5 | 1.6083 | 0.5546 | 0.9672 | 1.1915 | 0.9114 | 0.3171 | **0.2934** | 0.3787 | 0.3348 |
| | | 6 | 2.4319 | 0.7680 | 1.4614 | 1.6153 | 1.2760 | 0.3733 | **0.3084** | 0.4055 | 0.3535 |
| | | 8 | 3.2922 | 0.9429 | 1.9936 | 1.9829 | 1.6212 | 0.4430 | **0.3514** | 0.5357 | 0.4547 |
| | 0.99 | 4 | 6.1537 | 1.7760 | 3.0902 | 2.9648 | 2.5481 | 1.0788 | 0.5825 | ***0.3722*** | 0.5061 |
| | | 5 | 8.5604 | 2.2965 | 4.4483 | 3.8770 | 3.4364 | 1.5782 | 0.4735 | **0.3269** | 0.4038 |
| | | 6 | 13.3230 | 3.4459 | 7.2296 | 5.7655 | 5.3185 | 1.6921 | 0.4935 | **0.3370** | 0.3976 |
| | | 8 | 18.7246 | 4.4697 | 10.6896 | 7.7337 | 7.5113 | 2.0036 | 0.5269 | **0.3790** | 0.4172 |

*Bold values show the smallest MSE*

Table 3. Estimated MSE when n=350 and 400

| n | Rho | p | MLE | RIDGE | K-L | MRT | LMLRT | LMLRT1 | LMLRT2 | LMLRT3 | LMLRT4 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 350 | 0.8 | 4 | 0.2581 | 0.1638 | 0.2062 | 0.3101 | 0.2791 | **0.1195** | 0.1460 | 0.1980 | 0.1675 |
| | | 5 | 0.3476 | 0.1993 | 0.2725 | 0.4034 | 0.3559 | **0.1479** | 0.1985 | 0.2932 | 0.2458 |
| | | 6 | 0.4382 | 0.2311 | 0.3347 | 0.4803 | 0.4146 | **0.1554** | 0.2586 | 0.3751 | 0.3359 |
| | | 8 | 0.6713 | 0.2942 | 0.4861 | 0.6713 | 0.5527 | **0.1828** | 0.3214 | 0.4882 | 0.4449 |
| | 0.9 | 4 | 0.5090 | 0.2586 | 0.3660 | 0.5561 | 0.4499 | **0.1526** | 0.1983 | 0.2779 | 0.2313 |
| | | 5 | 0.7422 | 0.3359 | 0.5240 | 0.7280 | 0.5877 | **0.2057** | 0.2528 | 0.3622 | 0.3078 |
| | | 6 | 0.9455 | 0.3824 | 0.6439 | 0.8653 | 0.6861 | **0.2079** | 0.2999 | 0.4458 | 0.3817 |
| | | 8 | 1.4387 | 0.4881 | 0.9229 | 1.1228 | 0.8761 | **0.2607** | 0.3243 | 0.5114 | 0.4329 |
| | 0.95 | 4 | 1.0199 | 0.4167 | 0.6491 | 0.9047 | 0.6812 | **0.2050** | 0.2634 | 0.3276 | 0.2962 |
| | | 5 | 1.5720 | 0.5825 | 0.9964 | 1.2132 | 0.9465 | 0.3307 | **0.2953** | 0.3772 | 0.3334 |
| | | 6 | 2.0095 | 0.6613 | 1.2276 | 1.4126 | 1.1008 | 0.3356 | **0.3084** | 0.4176 | 0.3612 |
| | | 8 | 3.0688 | 0.8594 | 1.8181 | 1.8755 | 1.4915 | 0.4094 | **0.3070** | 0.4803 | 0.4004 |
| | 0.99 | 4 | 5.0725 | 1.4502 | 2.5265 | 2.5055 | 2.0131 | 0.8873 | 0.5026 | **0.3410** | 0.4397 |
| | | 5 | 8.4149 | 2.3935 | 4.5686 | 3.9445 | 3.5038 | 1.5491 | 0.5498 | **0.3489** | 0.4533 |
| | | 6 | 10.9381 | 2.8483 | 5.9679 | 4.9225 | 4.5104 | 1.4040 | 0.4611 | **0.3466** | 0.4023 |
| | | 8 | 17.8678 | 4.1775 | 10.1576 | 7.3596 | 6.8396 | 1.8770 | 0.4332 | **0.3462** | 0.3629 |
| 400 | 0.8 | 4 | 0.2100 | 0.1393 | 0.1709 | 0.2580 | 0.2366 | ***0.1080*** | 0.1284 | 0.1654 | 0.1428 |
| | | 5 | 0.3063 | 0.1823 | 0.2416 | 0.3530 | 0.3136 | ***0.1436*** | 0.1740 | 0.2575 | 0.2120 |
| | | 6 | 0.3854 | 0.2061 | 0.2952 | 0.4480 | 0.3869 | ***0.1448*** | 0.2114 | 0.3218 | 0.2758 |
| | | 8 | 0.6195 | 0.2908 | 0.4640 | 0.6388 | 0.5410 | ***0.1806*** | 0.3069 | 0.4657 | 0.4324 |
| | 0.9 | 4 | 0.4183 | 0.2215 | 0.3051 | 0.4834 | 0.3968 | ***0.1389*** | 0.1793 | 0.2500 | 0.2088 |
| | | 5 | 0.6134 | 0.2840 | 0.4308 | 0.6367 | 0.5094 | ***0.1876*** | 0.2218 | 0.3293 | 0.2741 |
| | | 6 | 0.8099 | 0.3334 | 0.5548 | 0.7862 | 0.6184 | ***0.1846*** | 0.2700 | 0.4062 | 0.3444 |
| | | 8 | 1.3153 | 0.4837 | 0.8888 | 1.0858 | 0.8715 | ***0.2481*** | 0.3168 | 0.5043 | 0.4312 |
| | 0.95 | 4 | 0.8700 | 0.3646 | 0.5619 | 0.8122 | 0.6122 | ***0.1896*** | 0.2362 | 0.3053 | 0.2692 |
| | | 5 | 1.2468 | 0.4567 | 0.7705 | 1.0354 | 0.7841 | 0.2651 | ***0.2530*** | 0.3387 | 0.2949 |
| | | 6 | 1.6616 | 0.5416 | 1.0123 | 1.2581 | 0.9513 | ***0.2711*** | 0.2739 | 0.3920 | 0.3306 |
| | | 8 | 2.8084 | 0.8548 | 1.7388 | 1.7877 | 1.4577 | 0.4077 | ***0.2880*** | 0.4533 | 0.3757 |
| | 0.99 | 4 | 4.7594 | 1.3892 | 2.3907 | 2.4092 | 1.9403 | 0.9162 | 0.5456 | ***0.3730*** | 0.4850 |
| | | 5 | 6.9868 | 1.8994 | 3.6333 | 3.3088 | 2.8841 | 1.2590 | 0.4669 | ***0.3173*** | 0.3940 |
| | | 6 | 9.1182 | 2.2492 | 4.8795 | 4.1465 | 3.6618 | 1.0403 | 0.3773 | ***0.3209*** | 0.3464 |
| | | 8 | 15.1766 | 3.7466 | 8.8586 | 6.5905 | 6.2012 | 1.5898 | 0.3931 | ***0.3127*** | 0.3200 |

*Bold values show the smallest MSE*

Table 4. Estimated MSE when n=450 and 500

| n | Rho | p | MLE | RIDGE | K-L | MRT | LMLRT | LMLRT1 | LMLRT2 | LMLRT3 | LMLRT4 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 450 | 0.8 | 4 | 0.2040 | 0.1380 | 0.1686 | 0.2534 | 0.2336 | **0.1034** | 0.1267 | 0.1613 | 0.1402 |
| | | 5 | 0.2551 | 0.1618 | 0.2082 | 0.3050 | 0.2782 | **0.1258** | 0.1682 | 0.2349 | 0.2038 |
| | | 6 | 0.3257 | 0.1897 | 0.2587 | 0.3861 | 0.3420 | **0.1396** | 0.1832 | 0.2894 | 0.2429 |
| | | 8 | 0.4912 | 0.2424 | 0.3736 | 0.5489 | 0.4667 | **0.1613** | 0.2572 | 0.4000 | 0.3684 |
| | 0.9 | 4 | 0.4087 | 0.2212 | 0.3043 | 0.4752 | 0.3939 | **0.1359** | 0.1671 | 0.2250 | 0.1895 |
| | | 5 | 0.5357 | 0.2652 | 0.3937 | 0.5891 | 0.4841 | **0.1729** | 0.2024 | 0.3046 | 0.2501 |
| | | 6 | 0.6857 | 0.3079 | 0.4860 | 0.7108 | 0.5677 | **0.1780** | 0.2490 | 0.3822 | 0.3228 |
| | | 8 | 1.0718 | 0.4040 | 0.7286 | 0.9683 | 0.7679 | **0.2160** | 0.3170 | 0.4946 | 0.4265 |
| | 0.95 | 4 | 0.8312 | 0.3547 | 0.5473 | 0.7902 | 0.6010 | **0.1776** | 0.2182 | 0.2820 | 0.2465 |
| | | 5 | 1.1256 | 0.4475 | 0.7379 | 0.9770 | 0.7526 | 0.2515 | **0.2483** | 0.3469 | 0.2946 |
| | | 6 | 1.4479 | 0.5093 | 0.9076 | 1.1506 | 0.8794 | ***0.2574*** | 0.2759 | 0.3838 | 0.3305 |
| | | 8 | 2.2363 | 0.6807 | 1.3747 | 1.5519 | 1.2165 | 0.3338 | **0.3333** | 0.5016 | 0.4281 |
| | 0.99 | 4 | 4.4099 | 1.3306 | 2.2883 | 2.3381 | 1.8639 | 0.8200 | 0.4504 | ***0.3245*** | 0.4093 |
| | | 5 | 6.1716 | 1.7307 | 3.3056 | 3.0440 | 2.6183 | 1.0627 | 0.4170 | ***0.3106*** | 0.3746 |
| | | 6 | 8.4557 | 2.2967 | 4.7267 | 3.9807 | 3.5050 | 1.0925 | 0.4256 | ***0.3238*** | 0.3746 |
| | | 8 | 12.6568 | 3.0710 | 7.3407 | 5.6040 | 5.1555 | 1.2914 | 0.3514 | ***0.3301*** | 0.3187 |
| 500 | 0.8 | 4 | 0.1664 | 0.1180 | 0.1394 | 0.2035 | 0.1911 | ***0.0972*** | 0.1070 | 0.1337 | 0.1170 |
| | | 5 | 0.2383 | 0.1541 | 0.1954 | 0.2911 | 0.2661 | ***0.1212*** | 0.1417 | 0.2043 | 0.1725 |
| | | 6 | 0.3013 | 0.1792 | 0.2413 | 0.3640 | 0.3240 | ***0.1333*** | 0.1726 | 0.2723 | 0.2294 |
| | | 8 | 0.4356 | 0.2281 | 0.3391 | 0.4996 | 0.4335 | ***0.1580*** | 0.2575 | 0.3792 | 0.3637 |
| | 0.9 | 4 | 0.3454 | 0.1981 | 0.2639 | 0.4112 | 0.3504 | ***0.1336*** | 0.1523 | 0.2134 | 0.1777 |
| | | 5 | 0.4817 | 0.2433 | 0.3544 | 0.5393 | 0.4431 | ***0.1622*** | 0.2003 | 0.2977 | 0.2491 |
| | | 6 | 0.6336 | 0.2969 | 0.4606 | 0.6624 | 0.5360 | ***0.1701*** | 0.2344 | 0.3643 | 0.3087 |
| | | 8 | 0.9258 | 0.3659 | 0.6447 | 0.8880 | 0.7077 | ***0.2020*** | 0.2902 | 0.4723 | 0.4069 |
| | 0.95 | 4 | 0.7259 | 0.3264 | 0.4919 | 0.7178 | 0.5544 | ***0.1668*** | 0.2020 | 0.2639 | 0.2288 |
| | | 5 | 0.9762 | 0.3849 | 0.6386 | 0.8954 | 0.6833 | ***0.2102*** | 0.2343 | 0.3346 | 0.2834 |
| | | 6 | 1.3067 | 0.4888 | 0.8494 | 1.0853 | 0.8447 | ***0.2408*** | 0.2696 | 0.3794 | 0.3280 |
| | | 8 | 1.9731 | 0.6212 | 1.2388 | 1.4510 | 1.1272 | 0.3050 | ***0.2900*** | 0.4667 | 0.3871 |
| | 0.99 | 4 | 3.9188 | 1.2080 | 2.0537 | 2.1456 | 1.7026 | 0.7127 | 0.4212 | ***0.2875*** | 0.3720 |
| | | 5 | 5.3339 | 1.5168 | 2.8969 | 2.7554 | 2.2945 | 0.9166 | 0.3769 | ***0.3013*** | 0.3522 |
| | | 6 | 7.2427 | 2.0039 | 4.0699 | 3.5550 | 3.0878 | 0.9687 | 0.3546 | ***0.2892*** | 0.3196 |
| | | 8 | 10.8929 | 2.6307 | 6.2643 | 4.9804 | 4.4708 | 1.1523 | 0.3385 | ***0.3377*** | 0.3251 |

*Bold values show the smallest MSE*

Table 5. Estimated MSE when n=600 and 700

| n | Rho | P | MLE | RIDGE | K-L | MRT | LMLRT | LMLRT1 | LMLRT2 | LMLRT3 | LMLRT4 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| 600 | 0.8 | 4 | 0.1302 | 0.0984 | 0.1123 | 0.1602 | 0.1533 | ***0.0846*** | 0.0877 | 0.1029 | 0.0921 |
| | | 5 | 0.1964 | 0.1328 | 0.1637 | 0.2425 | 0.2250 | ***0.1080*** | 0.1174 | 0.1576 | 0.1318 |
| | | 6 | 0.2546 | 0.1613 | 0.2086 | 0.3110 | 0.2833 | ***0.1244*** | 0.1372 | 0.2086 | 0.1686 |
| | | 8 | 0.3610 | 0.1989 | 0.2846 | 0.4280 | 0.3740 | ***0.1448*** | 0.2008 | 0.3033 | 0.2886 |
| | 0.9 | 4 | 0.2596 | 0.1603 | 0.2045 | 0.3228 | 0.2827 | ***0.1188*** | 0.1316 | 0.1710 | 0.1475 |
| | | 5 | 0.4115 | 0.2214 | 0.3105 | 0.4801 | 0.4011 | ***0.1534*** | 0.1697 | 0.2472 | 0.2046 |
| | | 6 | 0.5261 | 0.2588 | 0.3883 | 0.5866 | 0.4839 | ***0.1608*** | 0.1924 | 0.3019 | 0.2524 |
| | | 8 | 0.7712 | 0.3212 | 0.5392 | 0.7580 | 0.6086 | ***0.1924*** | 0.2671 | 0.4217 | 0.3733 |
| | 0.95 | 4 | 0.5455 | 0.2672 | 0.3849 | 0.5839 | 0.4610 | ***0.1443*** | 0.1678 | 0.2271 | 0.1927 |
| | | 5 | 0.8251 | 0.3505 | 0.5532 | 0.7898 | 0.6123 | 0.2091 | ***0.1968*** | 0.2843 | 0.2387 |
| | | 6 | 1.1255 | 0.4373 | 0.7435 | 0.9973 | 0.7775 | ***0.2205*** | 0.2256 | 0.3424 | 0.2877 |
| | | 8 | 1.6924 | 0.5613 | 1.0796 | 1.2999 | 1.0213 | 0.2864 | ***0.2777*** | 0.4458 | 0.3723 |
| | 0.99 | 4 | 2.9281 | 0.9313 | 1.5569 | 1.7288 | 1.3243 | 0.5693 | ***0.3345*** | 0.2665 | 0.3141 |
| | | 5 | 4.3912 | 1.2577 | 2.3400 | 2.3302 | 1.9295 | 0.7678 | ***0.3399*** | 0.2842 | 0.3171 |
| | | 6 | 6.0241 | 1.6911 | 3.3996 | 3.1106 | 2.6350 | 0.8336 | ***0.3240*** | 0.2886 | 0.3008 |
| | | 8 | 9.3612 | 2.3173 | 5.4862 | 4.3095 | 3.8047 | 0.9851 | ***0.3152*** | 0.3219 | 0.3038 |
| 700 | 0.8 | 4 | 0.1163 | 0.0885 | 0.1006 | 0.1452 | 0.1395 | ***0.0744*** | 0.0756 | 0.0868 | 0.0776 |
| | | 5 | 0.1593 | 0.1117 | 0.1350 | 0.2052 | 0.1926 | ***0.0918*** | 0.0978 | 0.1255 | 0.1070 |
| | | 6 | 0.2184 | 0.1433 | 0.1812 | 0.2763 | 0.2535 | ***0.1119*** | 0.1178 | 0.1824 | 0.1441 |
| | | 8 | 0.3198 | 0.1824 | 0.2553 | 0.3947 | 0.3481 | ***0.1373*** | 0.1782 | 0.2930 | 0.2598 |
| | 0.9 | 4 | 0.2382 | 0.1482 | 0.1882 | 0.3077 | 0.2698 | ***0.1039*** | 0.1057 | 0.1327 | 0.1148 |
| | | 5 | 0.3440 | 0.1909 | 0.2643 | 0.4307 | 0.3636 | ***0.1351*** | 0.1375 | 0.2080 | 0.1676 |
| | | 6 | 0.4678 | 0.2400 | 0.3494 | 0.5334 | 0.4436 | ***0.1571*** | 0.1893 | 0.2840 | 0.2420 |
| | | 8 | 0.6900 | 0.2995 | 0.4934 | 0.7345 | 0.5934 | ***0.1816*** | 0.2371 | 0.3947 | 0.3418 |
| | 0.95 | 4 | 0.4825 | 0.2342 | 0.3371 | 0.5466 | 0.4254 | ***0.1269*** | ***0.1436*** | 0.1910 | 0.1628 |
| | | 5 | 0.7220 | 0.3121 | 0.4948 | 0.7388 | 0.5703 | 0.1785 | ***0.1771*** | 0.2554 | 0.2146 |
| | | 6 | 0.9731 | 0.3855 | 0.6383 | 0.8930 | 0.6914 | ***0.2130*** | 0.2187 | 0.3136 | 0.2673 |
| | | 8 | 1.4495 | 0.4925 | 0.9285 | 1.1954 | 0.9229 | ***0.2531*** | 0.2600 | 0.4276 | 0.3554 |
| | 0.99 | 4 | 2.4882 | 0.7486 | 1.2836 | 1.5509 | 1.1374 | 0.3994 | 0.2936 | ***0.2573*** | 0.2848 |
| | | 5 | 3.9217 | 1.1402 | 2.1542 | 2.2228 | 1.7608 | 0.6480 | 0.2723 | ***0.2668*** | 0.2717 |
| | | 6 | 5.2754 | 1.4670 | 2.9738 | 2.8039 | 2.2832 | 0.7118 | 0.3232 | ***0.2963*** | 0.3098 |
| | | 8 | 7.8739 | 1.9249 | 4.5594 | 3.8268 | 3.2857 | 0.8187 | 0.2656 | ***0.3063*** | 0.2772 |

*Bold values show the smallest MSE*

Table 6. Estimated Coefficients and Scalar Mean Squared Error

| | $\hat{\beta}_1$ | $\hat{\beta}_2$ | $\hat{\beta}_3$ | $\hat{\beta}_4$ | SMSE |
|---|---|---|---|---|---|
| MLE | 1.0834 | -0.0543 | 0.0529 | -0.4255 | 0.30047 |
| RIDGE | 0.0010 | -0.0163 | 0.0055 | -0.0010 | 0.21956 |
| K-L | 0.8445 | -0.0508 | 0.0547 | -0.3576 | 0.20019 |
| MRT | 0.8922 | -0.0515 | 0.0543 | -0.3710 | 0.21677 |
| LMLRT | 0.0031 | -0.0277 | 0.0155 | -0.0060 | 0.18516 |
| LMLRT1 | 0.0115 | -0.0345 | 0.0279 | -0.0294 | ***0.17859*** |
| LMLRT2 | 0.0057 | -0.0317 | 0.0211 | -0.0140 | 0.18184 |
| LMLRT3 | 0.0112 | -0.0345 | 0.0276 | -0.0287 | 0.17872 |
| LMLRT4 | 0.0113 | -0.0345 | 0.0277 | -0.0291 | 0.17866 |