

Stochastic Restricted Maximum Likelihood Estimator in Logistic Regression Model

Varathan Nagarajah^{1,2}, Pushpakanthie Wijekoon³

¹Postgraduate Institute of Science, University of Peradeniya, Peradeniya, Sri Lanka

²Department of Mathematics and Statistics, University of Jaffna, Jaffna, Sri Lanka

³Department of Statistics and Computer Science, University of Peradeniya, Peradeniya, Sri Lanka

Email: varathan10@gmail.com, pushpaw@pdn.ac.lk

Received 2 November 2015; accepted 27 December 2015; published 30 December 2015

Copyright © 2015 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

In the presence of multicollinearity in logistic regression, the variance of the Maximum Likelihood Estimator (MLE) becomes inflated. Şiray *et al.* (2015) [1] proposed a restricted Liu estimator in logistic regression model with exact linear restrictions. However, there are some situations, where the linear restrictions are stochastic. In this paper, we propose a Stochastic Restricted Maximum Likelihood Estimator (SRMLE) for the logistic regression model with stochastic linear restrictions to overcome this issue. Moreover, a Monte Carlo simulation is conducted for comparing the performances of the MLE, Restricted Maximum Likelihood Estimator (RMLE), Ridge Type Logistic Estimator (LRE), Liu Type Logistic Estimator (LLE), and SRMLE for the logistic regression model by using Scalar Mean Squared Error (SMSE).

Keywords

Logistic Regression, Multicollinearity, Stochastic Restricted Maximum Likelihood Estimator, Scalar Mean Squared Error

1. Introduction

In many fields of study such as medicine and epidemiology, it is very important to predict a binary response variable, or to compute the probability of occurrence of an event, in terms of the values of a set of explanatory variables related to it. For example, the probability of suffering a heart attack is computed in terms of the levels of a set of risk factors such as cholesterol and blood pressure. The logistic regression model serves admirably this purpose and is the most used for these cases.

The general form of logistic regression model is

$$y_i = \pi_i + \varepsilon_i, \quad i = 1, \dots, n \tag{1}$$

which follows Bernoulli distribution with parameter π_i as

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

where x_i is the i^{th} row of X , which is an $n \times (p+1)$ data matrix with p explanatory variables and β is a $(p+1) \times 1$ vector of coefficients, ε_i is independent with mean zero and variance $\pi_i(1-\pi_i)$ of the response y_i . The maximum likelihood method is the most common estimation technique to estimate the parameter β , and the Maximum Likelihood Estimator (MLE) of β can be obtained as follows:

$$\hat{\beta}_{MLE} = C^{-1} X' \hat{W} Z, \tag{3}$$

where $C = X' \hat{W} X$; Z is the column vector with i^{th} element equals $\text{logit}(\hat{\pi}_i) + \frac{y_i - \hat{\pi}_i}{\hat{\pi}_i(1-\hat{\pi}_i)}$ and

$\hat{W} = \text{diag}[\hat{\pi}_i(1-\hat{\pi}_i)]$, which is an unbiased estimate of β . The covariance matrix of $\hat{\beta}_{MLE}$ is

$$\text{Cov}(\hat{\beta}_{MLE}) = \{X' \hat{W} X\}^{-1}. \tag{4}$$

As many authors have stated (Hosmer and Lemeshow (1989) [2] and Ryan (1997) [3], among others), the logistic regression model becomes unstable when there exists strong dependence among explanatory variables (multi-collinearity). For example, we suppose that the probability of a person surviving 10 or more extra years is modelled using three predictors Sex, Diastolic blood pressure and Body mass index. Since the response “whether the person surviving 10 or more extra years” is binary, the logistic regression model is appropriate for this problem. However, it is understood that the predictors Sex, Diastolic blood pressure and Body mass index may have some inter-relationship within each person. In this case, the estimation of the model parameters becomes inaccurate because of the need to invert near-singular information matrices. Consequently, the interpretation of the relationship between the response and each explanatory variable in terms of odds ratio may be erroneous. As a result, the estimates have large variances and large confidence intervals, which produce inefficient estimates.

To overcome the problem of multi-collinearity in the logistic regression, many estimators are proposed alternatives to the MLE. The most popular way to deal with this problem is called the Ridge Logistic Regression (RLR), which is first proposed by Schaffer *et al.* (1984) [4]. Later Principal Component Logistic Estimator (PCLE) by Aguilera *et al.* (2006) [5], the Modified Logistic Ridge Regression Estimator (MLRE) by Nja *et al.* (2013) [6], Liu Estimator by Mansson *et al.* (2012) [7], and Liu-type estimator by Inan and Erdogan (2013) [8] in logistic regression have been proposed.

An alternative technique to resolve the multi-collinearity problem is to consider parameter estimation with priori available linear restrictions on the unknown parameters, which may be exact or stochastic. That is, in some practical situations there exist different sets of prior information from different sources like past experience or long association of the experimenter with the experiment and similar kind of experiments conducted in the past. If the exact linear restrictions are available in addition to logistic regression model, many authors propose different estimators for the respective parameter β . Duffy and Santer (1989) [9] introduce a Restricted Maximum Likelihood Estimator (RMLE) by incorporating the exact linear restriction on the unknown parameters. Recently Şiray *et al.* (2015) [1] proposes a new estimator called Restricted Liu Estimator (RLE) by replacing MLE by RMLE in the logistic Liu estimator.

In this paper we propose a new estimator which is called as the Stochastic Restricted Maximum Likelihood Estimator (SRMLE) when the linear stochastic restrictions are available in addition to the logistic regression model. The rest of the paper is organized as follows. The proposed estimator and its asymptotic properties are given in Section 2. In Section 3, the mean square error matrix and the scalar mean square error for this new estimator are obtained. Section 4 describes some important existing estimators for the logistic regression models. Performance of the proposed estimator with respect to Scalar Mean Squared Error (SMSE) is compared with some existing estimators by performing a Monte Carlo simulation study in Section 5. The conclusion of the study is presented in Section 6.

2. The Proposed Estimator and its Asymptotic Properties

First consider the multiple linear regression model

$$y = X\beta + \varepsilon, \quad \varepsilon \sim N(0, \sigma^2 I), \quad (5)$$

where y is an $n \times 1$ observable random vector, X is an $n \times p$ known design matrix of rank p , β is a $p \times 1$ vector of unknown parameters and ε is an $n \times 1$ vector of disturbances.

The Ordinary Least Square Estimator (OLSE) of β is given by

$$\hat{\beta}_{\text{OLSE}} = S^{-1}X'y, \quad (6)$$

where $S = XX'$.

In addition to sample model (5), consider the following linear stochastic restriction on the parameter space β ;

$$r = R\beta + v; \quad E(v) = 0, \quad \text{Cov}(v) = \Omega > 0 \quad (7)$$

where r is an $m \times 1$ stochastic known vector, R is a $m \times p$ of full rank $m \leq p$ with known elements and v is an $(m \times 1)$ random vector of disturbances with mean $\mathbf{0}$ and dispersion matrix Ω , and Ω is assumed to be known $(m \times m)$ positive definite matrix. Further it is assumed that v is stochastically independent of ε , i.e. $E(\varepsilon v') = 0$.

The Restricted Ordinary Least Square Estimator (ROLSE) due to exact prior restriction (i.e. $v = 0$) in (7) is given by

$$\hat{\beta}_{\text{ROLSE}} = \hat{\beta}_{\text{OLSE}} + S^{-1}R'(RS^{-1}R')^{-1}(r - R\hat{\beta}_{\text{OLSE}}) \quad (8)$$

Theil and Goldberger (1961) [10] proposed the mixed regression estimator (ME) for the regression model (2.1) with the stochastic restricted prior information (7)

$$\hat{\beta}_{\text{ME}} = \hat{\beta}_{\text{OLSE}} + S^{-1}R'(\Omega + RS^{-1}R')^{-1}(r - R\hat{\beta}_{\text{OLSE}}) \quad (9)$$

Suppose that the following linear prior information is given in addition to the general logistic regression model (1)

$$h = H\beta + v; \quad E(v) = 0, \quad \text{Cov}(v) = \Psi \quad (10)$$

where h is an $(q \times 1)$ stochastic known vector, H is a $(q \times (p+1))$ of full rank $(q \leq p+1)$ known elements and v is an $(q \times 1)$ random vector of disturbances with mean $\mathbf{0}$ and dispersion matrix Ψ , and Ψ is assumed to be known $(q \times q)$ positive definite matrix. Further, it is assumed that v is stochastically independent of $\varepsilon^* = (\varepsilon_1, \varepsilon_2, \dots, \varepsilon_n)$, i.e. $E(\varepsilon^* v') = 0$.

Duffy and Santner (1989) [9] proposed the Restricted Maximum Likelihood Estimator (RMLE) for the logistic regression model (1) with the exact prior restriction (i.e. $v = 0$) in (10)

$$\hat{\beta}_{\text{RMLE}} = \hat{\beta}_{\text{MLE}} + C^{-1}H'(HC^{-1}H')^{-1}(h - H\hat{\beta}_{\text{MLE}}) \quad (11)$$

Following RMLE in (11) and the Mixed Estimator (ME) in (9) in the Linear Regression Model, we propose a new estimator which is named as the Stochastic Restricted Maximum Likelihood Estimator (SRMLE) when the linear stochastic restriction (10) is available in addition to the logistic regression model (1).

$$\hat{\beta}_{\text{SRMLE}} = \hat{\beta}_{\text{MLE}} + C^{-1}H'(\Psi + HC^{-1}H')^{-1}(h - H\hat{\beta}_{\text{MLE}}) \quad (12)$$

Asymptotic Properties of SRMLE:

The $\hat{\beta}_{\text{SRMLE}}$ is asymptotically unbiased.

$$E(\hat{\beta}_{\text{SRMLE}}) = \beta \quad (13)$$

The asymptotic covariance matrix of SRMLE equals

$$\text{Var}(\hat{\beta}_{\text{SRMLE}}) = C^{-1} - C^{-1}H'(\Psi + HC^{-1}H')^{-1}HC^{-1} \tag{14}$$

3. Mean Square Error Matrix Comparisons

To compare different estimators with respect to the same parameter vector β in the regression model, one can use the well known Mean Square Error (MSE) Matrix (MSE) and/or Scalar Mean Square Error (SMSE) criteria.

$$\text{MSE}(\hat{\beta}, \beta) = E\left[(\hat{\beta} - \beta)(\hat{\beta} - \beta)'\right] = D(\hat{\beta}) + B(\hat{\beta})B'(\hat{\beta}) \tag{15}$$

where $D(\hat{\beta})$ is the dispersion matrix, and $B(\hat{\beta}) = E(\hat{\beta}) - \beta$ denotes the bias vector.

The Scalar Mean Square Error (SMSE) of the estimator $\hat{\beta}$ can be defined as

$$\text{SMSE}(\hat{\beta}, \beta) = \text{trace}\left[\text{MSE}(\hat{\beta}, \beta)\right] \tag{16}$$

For two given estimators $\hat{\beta}_1$ and $\hat{\beta}_2$, the estimator $\hat{\beta}_2$ is said to be superior to $\hat{\beta}_1$ under the MSE criterion if and only if

$$M(\hat{\beta}_1, \hat{\beta}_2) = \text{MSE}(\hat{\beta}_1, \beta) - \text{MSE}(\hat{\beta}_2, \beta) \geq 0. \tag{17}$$

The MSE and SMSE of the proposed estimator SRMLE is

$$\begin{aligned} \text{MSE}(\hat{\beta}_{\text{SRMLE}}) &= D(\hat{\beta}_{\text{SRMLE}}) + B(\hat{\beta}_{\text{SRMLE}})B'(\hat{\beta}_{\text{SRMLE}}) \\ &= C^{-1} - C^{-1}H'(\Psi + HC^{-1}H')^{-1}HC^{-1} \end{aligned} \tag{18}$$

$$\text{SMSE}(\hat{\beta}_{\text{SRMLE}}) = \text{trace}\left[C^{-1} - C^{-1}H'(\Psi + HC^{-1}H')^{-1}HC^{-1}\right] \tag{19}$$

$$\begin{aligned} \text{Gain in Efficiency} &= \text{MSE}(\hat{\beta}_{\text{MLE}}) - \text{MSE}(\hat{\beta}_{\text{SRMLE}}) = D(\hat{\beta}_{\text{MLE}}) - D(\hat{\beta}_{\text{SRMLE}}) \\ &= C^{-1}H'(\Psi + HC^{-1}H')^{-1}HC^{-1} \end{aligned} \tag{20}$$

Note that the difference given in (20) is non-negative definite. Thus by the MSE criteria it follows that $\hat{\beta}_{\text{SRMLE}}$ has smaller Mean square error than $\hat{\beta}_{\text{MLE}}$.

4. Some Existing Logistic Estimators

To examine the performance of the proposed estimator SRMLE over some existing estimators, the following estimators are considered.

1) Logistic Ridge Estimator

Schaefer *et al.* (1984) [4] proposed a ridge estimator for the logistic regression model (1).

$$\begin{aligned} \hat{\beta}_{\text{LRE}} &= (X'WX + kI)^{-1}X'WX\hat{\beta}_{\text{MLE}} \\ &= (C + kI)^{-1}C\hat{\beta}_{\text{MLE}} = Z_k\hat{\beta}_{\text{MLE}} \end{aligned} \tag{21}$$

where $k > 0$ is the ridge parameter and $Z_k = (C + kI)^{-1}C$.

The asymptotic MSE and SMSE of $\hat{\beta}_{\text{LRE}}$,

$$\begin{aligned} \text{MSE}(\hat{\beta}_{\text{LRE}}) &= E\left[(\hat{\beta}_{\text{LRE}} - \beta)(\hat{\beta}_{\text{LRE}} - \beta)'\right] = D(\hat{\beta}_{\text{LRE}}) + B(\hat{\beta}_{\text{LRE}})B'(\hat{\beta}_{\text{LRE}}) \\ &= Z_kC^{-1}Z_k' + (Z_k\beta - \beta)(Z_k\beta - \beta)' = (C + kI)^{-1}C(C + kI)^{-1} + \delta_1\delta_1' \end{aligned} \tag{22}$$

where $\delta_1 = Z_k\beta - \beta$.

$$\begin{aligned}
 \text{SMSE}(\hat{\beta}_{LRE}) &= \text{trace} \left\{ E \left[(\hat{\beta}_{LRE} - \beta)(\hat{\beta}_{LRE} - \beta)' \right] \right\} \\
 &= \text{trace} \left[(\hat{\beta}_{MLE} - \beta)' Z_k' Z_k (\hat{\beta}_{MLE} - \beta) \right] + (Z_k \beta - \beta)' (Z_k \beta - \beta) \\
 &= \text{trace} [C^{-1} Z_k' Z_k] + k^2 \beta' (C + kI)^{-2} \beta
 \end{aligned} \tag{23}$$

2) Logistic Liu Estimator

Following Liu (1993) [11], Urgan and Tez (2008) [12], Mansson *et al.* (2012) [7] examined the Liu Estimator for logistic regression model, which is defined as

$$\begin{aligned}
 \hat{\beta}_{LLE} &= (C + I)^{-1} (C + dI) \hat{\beta}_{MLE} \\
 &= Z_d \hat{\beta}_{MLE}
 \end{aligned} \tag{24}$$

where $0 < d < 1$ is a parameter and $Z_d = (C + I)^{-1} (C + dI)$.

The asymptotic MSE and SMSE of $\hat{\beta}_{LLE}$,

$$\begin{aligned}
 \text{MSE}(\hat{\beta}_{LLE}) &= E \left[(\hat{\beta}_{LLE} - \beta)(\hat{\beta}_{LLE} - \beta)' \right] \\
 &= D(\hat{\beta}_{LLE}) + B(\hat{\beta}_{LLE}) B'(\hat{\beta}_{LLE}) \\
 &= Z_d C^{-1} Z_d' + (Z_d \beta - \beta)(Z_d \beta - \beta)' \\
 &= \left\{ (C + I)^{-1} (C + dI) \right\} C^{-1} \left\{ (C + I)^{-1} (C + dI) \right\}' + \delta_2 \delta_2' \\
 &= (C + I)^{-1} (C + dI) (I + dC^{-1}) (C + I)^{-1} + \delta_2 \delta_2'
 \end{aligned} \tag{25}$$

where $\delta_2 = Z_d \beta - \beta$.

$$\begin{aligned}
 \text{SMSE}(\hat{\beta}_{LLE}) &= \text{trace} \left\{ E \left[(\hat{\beta}_{LLE} - \beta)(\hat{\beta}_{LLE} - \beta)' \right] \right\} \\
 &= \text{trace} \left[(\hat{\beta}_{MLE} - \beta)' Z_d' Z_d (\hat{\beta}_{MLE} - \beta) \right] + (Z_d \beta - \beta)' (Z_d \beta - \beta) \\
 &= \text{trace} [C^{-1} Z_d' Z_d] + k^2 \beta' (C + kI)^{-2} \beta
 \end{aligned} \tag{26}$$

3) Restricted MLE

As we mentioned in Section 2, Duffy and Santner (1989) [9] proposed the Restricted Maximum Likelihood Estimator (RMLE) for the logistic regression model (1) with the exact prior restriction (*i.e.* $\nu = 0$) in (10).

$$\hat{\beta}_{RMLE} = \hat{\beta}_{MLE} + C^{-1} H' (HC^{-1} H')^{-1} (h - H \hat{\beta}_{MLE}) \tag{27}$$

The asymptotic MSE and SMSE of $\hat{\beta}_{RMLE}$,

$$\text{MSE}(\hat{\beta}_{RMLE}) = A + \delta_3 \delta_3' \tag{28}$$

$$\text{SMSE}(\hat{\beta}_{RMLE}) = \text{trace}(A + \delta_3 \delta_3') \tag{29}$$

where $A = D(\hat{\beta}_{RMLE}) = C^{-1} - C^{-1} H' (HC^{-1} H')^{-1} H C^{-1}$

and $\delta_3 = \text{Bias}(\hat{\beta}_{RMLE}) = C^{-1} H' (HC^{-1} H')^{-1} (h - H \hat{\beta}_{MLE})$

Mean Squared Error Comparisons

• **SRMLE versus LRE**

$$\begin{aligned}
 &MSE(\hat{\beta}_{LRE}) - MSE(\hat{\beta}_{SRMLE}) \\
 &= \left\{ D(\hat{\beta}_{LRE}) - D(\hat{\beta}_{SRMLE}) \right\} + \left\{ B(\hat{\beta}_{LRE})B'(\hat{\beta}_{LRE}) - B(\hat{\beta}_{SRMLE})B'(\hat{\beta}_{SRMLE}) \right\} \\
 &= \left\{ (C+kI)^{-1}C(C+kI)^{-1} \right\} - \left\{ C^{-1} - C^{-1}H'(\Psi + HC^{-1}H')^{-1}HC^{-1} \right\} + \delta_1\delta_1' \\
 &= \left\{ (C+kI)^{-1}C(C+kI)^{-1} \right\} - (C + H'\Psi^{-1}H)^{-1} + \delta_1\delta_1' \\
 &= \left\{ (C+kI)^{-1}C(C+kI)^{-1} + \delta_1\delta_1' \right\} - (C + H'\Psi^{-1}H)^{-1} \\
 &= (M_1 - N_1)
 \end{aligned} \tag{30}$$

where $M_1 = (C+kI)^{-1}C(C+kI)^{-1} + \delta_1\delta_1'$ and $N_1 = (C + H'\Psi^{-1}H)^{-1}$. One can obviously say that

$(C+kI)^{-1}C(C+kI)^{-1}$ and N_1 are positive definite and $\delta_1\delta_1'$ is non-negative definite matrices. Further by **Theorem 1** (see **Appendix 1**), it is clear that M_1 is positive definite matrix. By **Lemma 1** (see **Appendix 1**), if $\lambda_{\max}(N_1M_1^{-1}) < 1$, where $\lambda_{\max}(N_1M_1^{-1})$ is the largest eigen value of $N_1M_1^{-1}$ then $M_1 - N_1$ is a positive definite matrix. Based on the above arguments, the following theorem can be stated.

Theorem 4.1. The estimator SRMLE is superior to LRE if and only if $\lambda_{\max}(N_1M_1^{-1}) < 1$.

• **SRMLE Versus LLE**

$$\begin{aligned}
 &MSE(\hat{\beta}_{LLE}) - MSE(\hat{\beta}_{SRMLE}) \\
 &= \left\{ D(\hat{\beta}_{LLE}) - D(\hat{\beta}_{SRMLE}) \right\} + \left\{ B(\hat{\beta}_{LLE})B'(\hat{\beta}_{LLE}) - B(\hat{\beta}_{SRMLE})B'(\hat{\beta}_{SRMLE}) \right\} \\
 &= \left\{ (C+I)^{-1}(C+dI)(I+dC^{-1})(C+I)^{-1} \right\} - \left\{ C^{-1} - C^{-1}H'(\Psi + HC^{-1}H')^{-1}HC^{-1} \right\} + \delta_2\delta_2' \\
 &= \left\{ (C+I)^{-1}(C+dI)(I+dC^{-1})(C+I)^{-1} \right\} - (C + H'\Psi^{-1}H)^{-1} + \delta_2\delta_2' \\
 &= \left\{ (C+I)^{-1}(C+dI)(I+dC^{-1})(C+I)^{-1} + \delta_2\delta_2' \right\} - (C + H'\Psi^{-1}H)^{-1} \\
 &= (M_2 - N_2)
 \end{aligned} \tag{31}$$

where $M_2 = \left\{ (C+I)^{-1}(C+dI)(I+dC^{-1})(C+I)^{-1} + \delta_2\delta_2' \right\}$ and $N_2 = (C + H'\Psi^{-1}H)^{-1}$. One can obviously say that $(C+I)^{-1}(C+dI)(I+dC^{-1})(C+I)^{-1}$ and N_2 are positive definite and $\delta_2\delta_2'$ is non-negative definite matrices. Further by **Theorem 1** (see **Appendix 1**), it is clear that M_2 is positive definite matrix. By **Lemma 1** (see **Appendix 1**), if $\lambda_{\max}(N_2M_2^{-1}) < 1$, where $\lambda_{\max}(N_2M_2^{-1})$ is the the largest eigen value of $N_2M_2^{-1}$ then $M_2 - N_2$ is a positive definite matrix. Based on the above arguments, the following theorem can be stated.

Theorem 4.2. The estimator SRMLE is superior to LLE if and only if $\lambda_{\max}(N_2M_2^{-1}) < 1$.

• **SRMLE versus RMLE**

$$\begin{aligned}
 &MSE(\hat{\beta}_{RMLE}) - MSE(\hat{\beta}_{SRMLE}) \\
 &= \left\{ D(\hat{\beta}_{RMLE}) - D(\hat{\beta}_{SRMLE}) \right\} + \left\{ B(\hat{\beta}_{RMLE})B'(\hat{\beta}_{RMLE}) - B(\hat{\beta}_{SRMLE})B'(\hat{\beta}_{SRMLE}) \right\} \\
 &= \left\{ C^{-1} - C^{-1}H'(HC^{-1}H')^{-1}HC^{-1} \right\} - \left\{ C^{-1} - C^{-1}H'(\Psi + HC^{-1}H')^{-1}HC^{-1} \right\} + \delta_3\delta_3' \\
 &= \left\{ C^{-1}H'(\Psi + HC^{-1}H')^{-1}HC^{-1} - C^{-1}H'(HC^{-1}H')^{-1}HC^{-1} \right\} + \delta_3\delta_3' \\
 &= \left\{ C^{-1}H'(\Psi + HC^{-1}H')^{-1}HC^{-1} + \delta_3\delta_3' \right\} - \left\{ C^{-1}H'(HC^{-1}H')^{-1}HC^{-1} \right\} = (M_3 - N_3)
 \end{aligned} \tag{32}$$

where $M_3 = \left\{ C^{-1}H'(\Psi + HC^{-1}H')^{-1}HC^{-1} + \delta_3\delta_3' \right\}$ and $N_3 = \left\{ C^{-1}H'(HC^{-1}H')^{-1}HC^{-1} \right\}$. One can obviously say that $C^{-1}H'(\Psi + HC^{-1}H')^{-1}HC^{-1}$ and N_3 are positive definite and $\delta_3\delta_3'$ is non-negative definite matrices. Further by **Theorem 1** (see **Appendix 1**), it is clear that M_3 is positive definite matrix. By **Lemma 1** (see **Appendix 1**), if $\lambda_{\max}(N_3M_3^{-1}) < 1$, where $\lambda_{\max}(N_3M_3^{-1})$ is the the largest eigen value of $N_3M_3^{-1}$ then $M_3 - N_3$ is a positive definite matrix. Based on the above arguments, the following theorem can be stated.

Theorem 4.3. The estimator SRMLE is superior to RMLE if and only if $\lambda_{\max}(N_3M_3^{-1}) < 1$.

Based on the above results one can say that the new estimator SRMLE is superior to the other estimators with respect to the mean squared error matrix sense under certain conditions. To check the superiority of the estimators numerically, we then consider a simulation study in the next section.

5. A Simulation Study

A Monte Carlo simulation is done to illustrate the performance of the new estimator SRMLE over the MLE, RMLE, LRE, and LLE by means of Scalar Mean Square Error (SMSE). Following McDonald and Galarneau (1975) [13] the data are generated as follows:

$$x_{ij} = (1 - \rho^2)^{1/2} z_{ij} + \rho z_{i,p+1}, \quad i = 1, 2, \dots, n, \quad j = 1, 2, \dots, p \tag{33}$$

where z_{ij} are pseudo- random numbers from standardized normal distribution and ρ^2 represents the correlation between any two explanatory variables. Four explanatory variables are generated using (33). We considered four different values of ρ corresponding to 0.70, 0.80, 0.90 and 0.99. Further four different values of n corresponding to 20, 40, 50, and 100 are considered. The dependent variable y_i in (1) is obtained from the Bernoulli (π_i) distribution where $\pi_i = \frac{\exp(x_i'\beta)}{1 + \exp(x_i'\beta)}$. The parameter values of $\beta_1, \beta_2, \dots, \beta_p$ are chosen so that

$$\sum_{j=1}^p \beta_j^2 \text{ and } \beta_1 = \beta_2 = \dots = \beta_p.$$

Moreover, for the restriction, we choose

$$H = \begin{pmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{pmatrix}, \quad h = \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix} \text{ and } \Psi = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{34}$$

Further for the ridge parameter k and the Liu parameter d , some selected values are chosen so that $0 \leq k \leq 1$ and $0 \leq d \leq 1$.

The experiment is replicated 3000 times by generating new pseudo-random numbers and the estimated SMSE is obtained as

$$\begin{aligned} \hat{\text{SMSE}}(\hat{\beta}^*) &= \text{Mean} \left\{ \text{tr} \left[\text{MSE}(\hat{\beta}, \beta) \right] \right\} = \text{Mean} \left\{ \text{tr} \left[E \left[(\hat{\beta} - \beta)(\hat{\beta} - \beta)' \right] \right] \right\} \\ &= \text{Mean} \left\{ E \left[\text{tr} \left[(\hat{\beta} - \beta)(\hat{\beta} - \beta)' \right] \right] \right\} = \text{Mean} \left\{ E \left[\text{tr} \left[(\hat{\beta} - \beta)'(\hat{\beta} - \beta) \right] \right] \right\} \\ &= \text{Mean} \left\{ (\hat{\beta} - \beta)'(\hat{\beta} - \beta) \right\} = \frac{1}{3000} \sum_{nsim=1}^{3000} (\hat{\beta} - \beta)'(\hat{\beta} - \beta) \end{aligned} \tag{35}$$

The simulation results are listed in **Tables A1-A16 (Appendix 3)** and also displayed in **Figures A1-A4 (Appendix 2)**. From **Figures A1-A4**, it can be noticed that in general increase in degree of correlation between two explanatory variables ρ inflates the estimated SMSE of all the estimators and increase in sample size n declines the estimated SMSE of all the estimators. Further, the new estimator SRMLE has smaller SMSE compared to MLE with respect to all the values of ρ and n . However, when $n = 20$ and $0.3 \leq k, d \leq 0.5$,

SRMLE performs better compared to the estimators LRE, and LLE. From **Table A17 (Appendix 3)**, it is clear that when k and d are small LLE is better than other estimators in the MSE sense, and LRE is better when k and d are large. For moderate k and d values the proposed estimator is good, but this will change with the n and ρ values. Therefore we then analyse the estimators LRE, LLE and SRMLE further by using different k and d values and the results are summarized in **Table A18** and **Table A19 (Appendix 3)**. According to these results it is clear that SRMLE is even superior to LRE and LLE for certain values of k and d .

6. Concluding Remarks

In this research, we introduced the Stochastic Restricted Maximum Likelihood Estimator (SRMLE) for logistic regression model when the linear stochastic restriction was available. The performances of the SRMLE over MLE, LRE, RMLE, and LLE in logistic regression model were investigated by performing a Monte Carlo simulation study. The research had been done by considering different degree of correlations, different numbers of observations and different values of parameters k , d . It was noted that the SMSE of the MLE was inflated when the multicollinearity was presented and it was severe particularly for small samples. The simulation results showed that the proposed estimator SRMLE had smaller SMSE than the estimator MLE with respect to all the values of n and ρ . Further it was noted that the proposed estimator SRMLE was superior over the estimators LLE and LRE for some k and d values related to different ρ and n .

Acknowledgements

We thank the editor and the referee for their comments and suggestions, and the Postgraduate Institute of Science, University of Peradeniya, Sri Lanka for providing necessary facilities to complete this research.

References

- [1] Şiray, G.U., Toker, S. and Kaçiranlar, S. (2015) On the Restricted Liu Estimator in Logistic Regression Model. *Communications in Statistics—Simulation and Computation*, **44**, 217-232. <http://dx.doi.org/10.1080/03610918.2013.771742>
- [2] Hosmer, D.W. and Lemeshow, S. (1989) *Applied Logistic Regression*. Wiley, New York.
- [3] Ryan, T.P. (1997) *Modern Regression Methods*. Wiley, New York.
- [4] Schaefer, R.L., Roi, L.D. and Wolfe, R.A. (1984) A Ridge Logistic Estimator. *Communications in Statistics—Theory and Methods*, **13**, 99-113. <http://dx.doi.org/10.1080/03610928408828664>
- [5] Aguilera, A.M., Escabias, M. and Valderrama, M.J. (2006) Using Principal Components for Estimating Logistic Regression with High-Dimensional Multicollinear Data. *Computational Statistics & Data Analysis*, **50**, 1905-1924. <http://dx.doi.org/10.1016/j.csda.2005.03.011>
- [6] Nja, M.E., Ogoke, U.P. and Nduka, E.C. (2013) The Logistic Regression Model with a Modified Weight Function. *Journal of Statistical and Econometric Method*, **2**, 161-171.
- [7] Mansson, G., Kibria, B.M.G. and Shukur, G. (2012) On Liu Estimators for the Logit Regression Model. The Royal Institute of Technology, Centre of Excellence for Science and Innovation Studies (CESIS), Paper No. 259. <http://dx.doi.org/10.1016/j.econmod.2011.11.015>
- [8] Inan, D. and Erdogan, B.E. (2013) Liu-Type Logistic Estimator. *Communications in Statistics—Simulation and Computation*, **42**, 1578-1586. <http://dx.doi.org/10.1080/03610918.2012.667480>
- [9] Duffy, D.E. and Santner, T.J. (1989) On the Small Sample Prosperities of Norm-Restricted Maximum Likelihood Estimators for Logistic Regression Models. *Communications in Statistics—Theory and Methods*, **18**, 959-980. <http://dx.doi.org/10.1080/03610928908829944>
- [10] Theil, H. and Goldberger, A.S. (1961) On Pure and Mixed Estimation in ECONOMICS. *International Economic Review*, **2**, 65-77. <http://dx.doi.org/10.2307/2525589>
- [11] Liu, K. (1993) A New Class of Biased Estimate in Linear Regression. *Communications in Statistics—Theory and Methods*, **22**, 393-402. <http://dx.doi.org/10.1080/03610929308831027>
- [12] Urgan, N.N. and Tez, M. (2008) Liu Estimator in Logistic Regression When the Data Are Collinear. *International Conference on Continuous Optimization and Knowledge-Based Technologies*, Linthuania, Selected Papers, Vilnius, 323-327.
- [13] McDonald, G.C. and Galarneau, D.I. (1975) A Monte Carlo Evaluation of Some Ridge-Type Estimators. *Journal of the*

American Statistical Association, **70**, 407-416. <http://dx.doi.org/10.1080/01621459.1975.10479882>

- [14] Rao, C.R. and Toutenburg, H. (1995) *Linear Models: Least Squares and Alternatives*. 2nd Edition, Springer-Verlag, New York, Inc.
- [15] Rao, C.R., Toutenburg, H., Shalabh and Heumann, C. (2008) *Linear Models and Generalizations*. Springer, Berlin.

Appendix 1

Theorem 1. Let $A: n \times n$ and $B: n \times n$ such that $A > 0$ and $B \geq 0$. Then $A + B \geq 0$. (Rao and Toutenburg, 1995) [14].

Lemma 1. Let the two $n \times n$ matrices $M > 0$, $N \geq 0$, then $M > N$ if and only if $\lambda_{\max}(NM^{-1}) < 1$. (Rao et al., 2008) [15].

Appendix 2

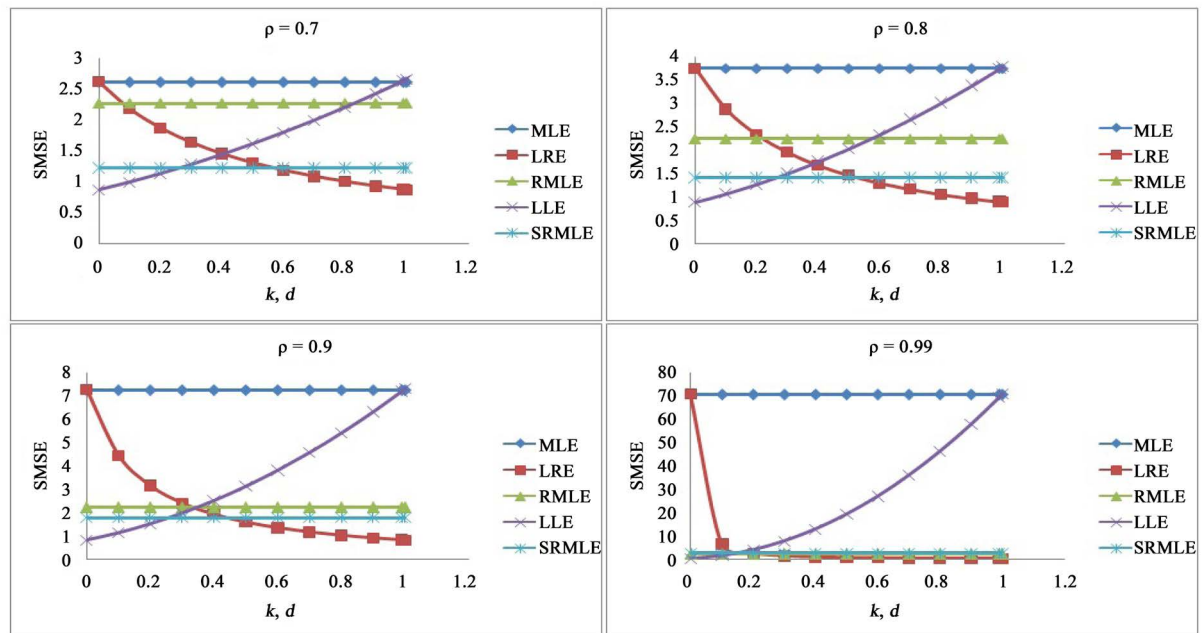


Figure A1. Estimated SMSE values for MLE, LRE, RMLE, LLE and SRMLE for $n = 20$.

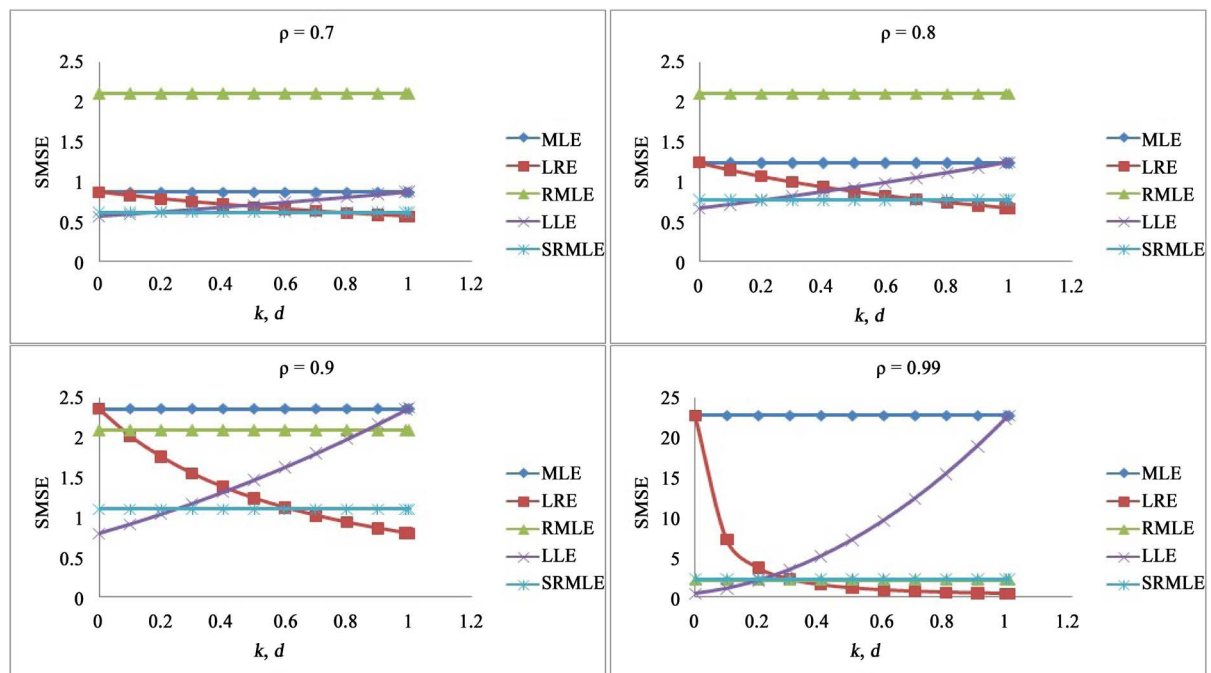


Figure A2. Estimated SMSE values for MLE, LRE, RMLE, LLE and SRMLE for $n = 50$.

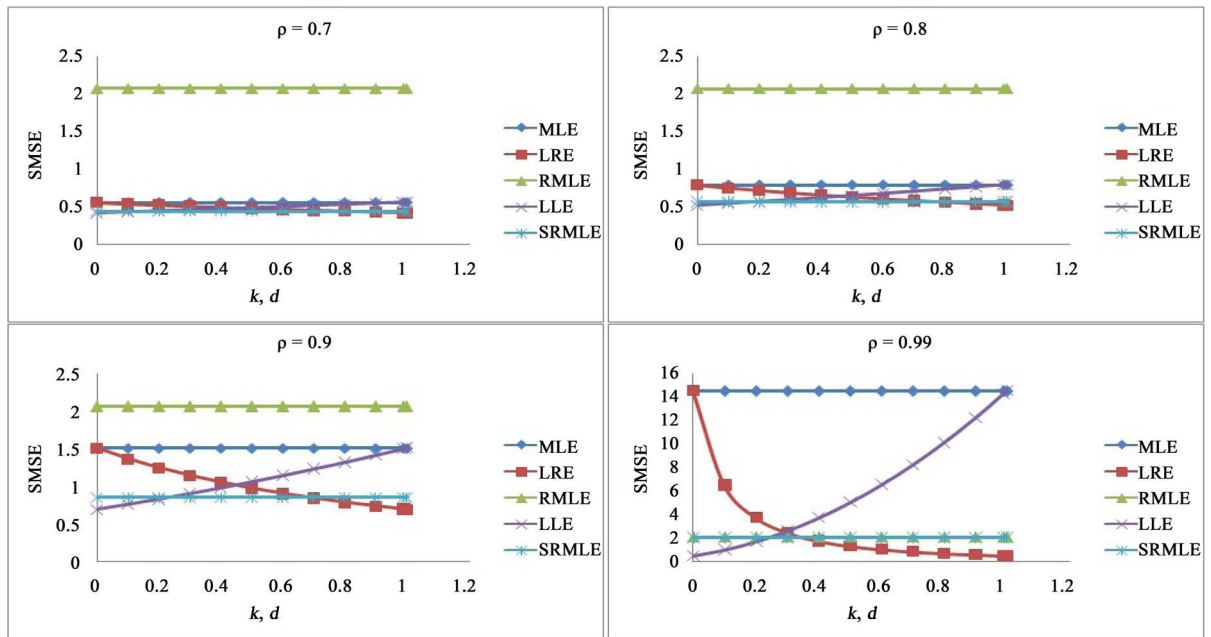


Figure A3. Estimated SMSE values for MLE, LRE, RMLE, LLE and SRMLE for $n = 75$.

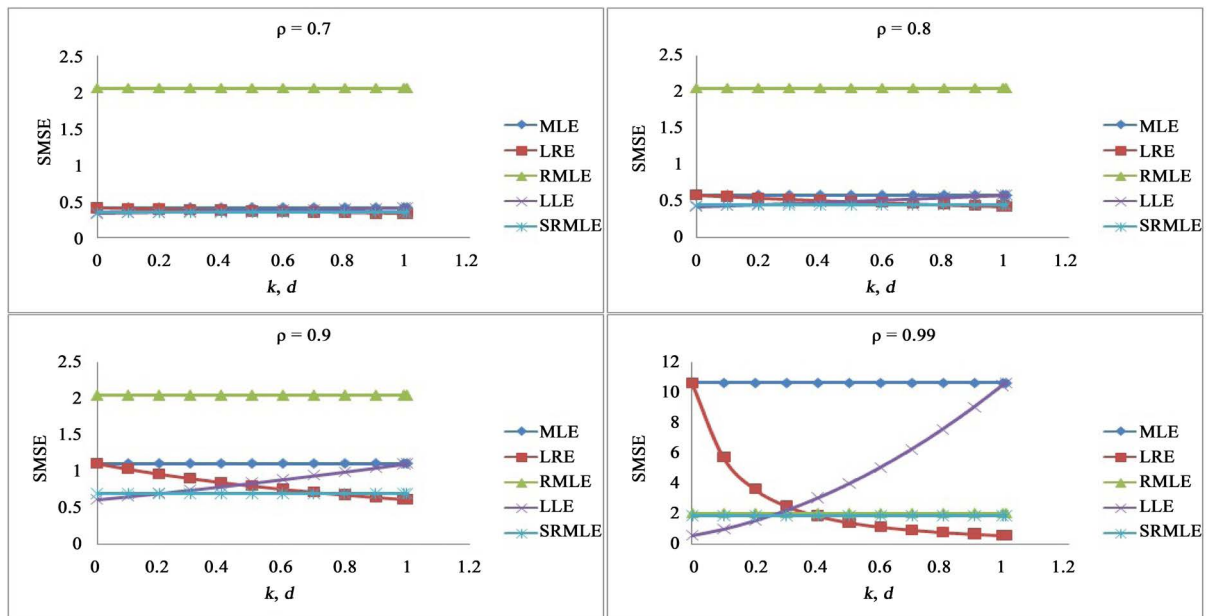


Figure A4. Estimated SMSE values for MLE, LRE, RMLE, LLE and SRMLE for $n = 100$.

Appendix 3

Table A1. The estimated MSE values for different k, d when $n = 20$ and $\rho = 0.70$.

	$k, d = 0.0$	$k, d = 0.1$	$k, d = 0.2$	$k, d = 0.3$	$k, d = 0.4$	$k, d = 0.5$	$k, d = 0.6$	$k, d = 0.7$	$k, d = 0.8$	$k, d = 0.9$	$k, d = 0.99$	$k, d = 1.0$
MLE	2.6097	2.6097	2.6097	2.6097	2.6097	2.6097	2.6097	2.6097	2.6097	2.6097	2.6097	2.6097
LRE	2.6097	2.1774	1.8688	1.6361	1.4543	1.3084	1.1892	1.0901	1.0068	0.9361	0.8812	0.8755
RMLE	2.2682	2.2682	2.2682	2.2682	2.2682	2.2682	2.2682	2.2682	2.2682	2.2682	2.2682	2.2682
LLE	0.8755	0.9995	1.1355	1.2835	1.4435	1.6156	1.7997	1.9958	2.2039	2.4240	2.6325	2.6562
SRMLE	1.2274	1.2274	1.2274	1.2274	1.2274	1.2274	1.2274	1.2274	1.2274	1.2274	1.2274	1.2274

Table A2. The estimated MSE values for different k, d when $n = 20$ and $\rho = 0.80$

	$k, d = 0.0$	$k, d = 0.1$	$k, d = 0.2$	$k, d = 0.3$	$k, d = 0.4$	$k, d = 0.5$	$k, d = 0.6$	$k, d = 0.7$	$k, d = 0.8$	$k, d = 0.9$	$k, d = 0.99$	$k, d = 1.0$
MLE	3.7509	3.7509	3.7509	3.7509	3.7509	3.7509	3.7509	3.7509	3.7509	3.7509	3.7509	3.7509
LRE	3.7509	2.8786	2.3312	1.9525	1.6750	1.4633	1.2971	1.1637	1.0548	0.9646	0.8960	0.8890
RMLE	2.2452	2.2452	2.2452	2.2452	2.2452	2.2452	2.2452	2.2452	2.2452	2.2452	2.2452	2.2452
LLE	0.8890	1.0689	1.2733	1.5023	1.7558	2.0340	2.3367	2.6640	3.0158	3.3923	3.7521	3.7933
SRMLE	1.4179	1.4179	1.4179	1.4179	1.4179	1.4179	1.4179	1.4179	1.4179	1.4179	1.4179	1.4179

Table A3. The estimated MSE values for different k, d when $n = 20$ and $\rho = 0.90$.

	$k, d = 0.0$	$k, d = 0.1$	$k, d = 0.2$	$k, d = 0.3$	$k, d = 0.4$	$k, d = 0.5$	$k, d = 0.6$	$k, d = 0.7$	$k, d = 0.8$	$k, d = 0.9$	$k, d = 0.99$	$k, d = 1.0$
MLE	7.2447	7.2447	7.2447	7.2447	7.2447	7.2447	7.2447	7.2447	7.2447	7.2447	7.2447	7.2447
LRE	7.2447	4.4635	3.1455	2.4005	1.9193	1.5859	1.3436	1.1611	1.0199	0.9083	0.8267	0.8186
RMLE	2.2263	2.2263	2.2263	2.2263	2.2263	2.2263	2.2263	2.2263	2.2263	2.2263	2.2263	2.2263
LLE	0.8186	1.1287	1.5135	1.9731	2.5075	3.1165	3.8003	4.5589	5.3922	6.3002	7.1813	7.2829
SRMLE	1.7693	1.7693	1.7693	1.7693	1.7693	1.7693	1.7693	1.7693	1.7693	1.7693	1.7693	1.7693

Table A4. The estimated MSE values for different k, d when $n = 20$ and $\rho = 0.99$.

	$k, d=0.0$	$k, d=0.1$	$k, d=0.2$	$k, d=0.3$	$k, d=0.4$	$k, d=0.5$	$k, d=0.6$	$k, d=0.7$	$k, d=0.8$	$k, d=0.9$	$k, d=0.99$	$k, d=1.0$
MLE	70.5890	70.5890	70.5890	70.5890	70.5890	70.5890	70.5890	70.5890	70.5890	70.5890	70.5890	70.5890
LRE	70.5890	6.5671	2.6098	1.4620	0.9711	0.7153	0.5650	0.4692	0.4045	0.3589	0.3288	0.3259
RMLE	2.2118	2.2118	2.2118	2.2118	2.2118	2.2118	2.2118	2.2118	2.2118	2.2118	2.2118	2.2118
LLE	0.3259	1.5179	4.0071	7.7935	12.8770	19.2580	26.9360	35.9120	46.1840	57.7540	69.2758	70.6209
SRMLE	2.5410	2.5410	2.5410	2.5410	2.5410	2.5410	2.5410	2.5410	2.5410	2.5410	2.5410	2.5410

Table A5. The estimated MSE values for different k, d when $n = 50$ and $\rho = 0.70$.

	$k, d = 0.0$	$k, d = 0.1$	$k, d = 0.2$	$k, d = 0.3$	$k, d = 0.4$	$k, d = 0.5$	$k, d = 0.6$	$k, d = 0.7$	$k, d = 0.8$	$k, d = 0.9$	$k, d = 0.99$	$k, d = 1.0$
MLE	0.8648	0.8648	0.8648	0.8648	0.8648	0.8648	0.8648	0.8648	0.8648	0.8648	0.8648	0.8648
LRE	0.8648	0.8210	0.7811	0.7447	0.7113	0.6807	0.6525	0.6266	0.6025	0.5803	0.5617	0.5597
RMLE	2.1057	2.1057	2.1057	2.1057	2.1057	2.1057	2.1057	2.1057	2.1057	2.1057	2.1057	2.1057
LLE	0.5597	0.5875	0.6161	0.6454	0.6756	0.7065	0.7383	0.7708	0.8042	0.8383	0.8697	0.8732
SRMLE	0.6141	0.6141	0.6141	0.6141	0.6141	0.6141	0.6141	0.6141	0.6141	0.6141	0.6141	0.6141

Table A6. The estimated MSE values for different k, d when $n = 50$ and $\rho = 0.80$.

	$k, d = 0.0$	$k, d = 0.1$	$k, d = 0.2$	$k, d = 0.3$	$k, d = 0.4$	$k, d = 0.5$	$k, d = 0.6$	$k, d = 0.7$	$k, d = 0.8$	$k, d = 0.9$	$k, d = 0.99$	$k, d = 1.0$
MLE	1.2320	1.2320	1.2320	1.2320	1.2320	1.2320	1.2320	1.2320	1.2320	1.2320	1.2320	1.2320
LRE	1.2320	1.1399	1.0593	0.9882	0.9252	0.8690	0.8186	0.7733	0.7324	0.6953	0.6648	0.6600
RMLE	2.0967	2.0967	2.0967	2.0967	2.0967	2.0967	2.0967	2.0967	2.0967	2.0967	2.0967	2.0967
LLE	0.6615	0.7101	0.7607	0.8134	0.8682	0.9250	0.9839	1.0448	1.1077	1.1728	1.2330	1.2398
SRMLE	0.7675	0.7675	0.7675	0.7675	0.7675	0.7675	0.7675	0.7675	0.7675	0.7675	0.7675	0.7675

Table A7. The estimated MSE values for different k, d when $n = 50$ and $\rho = 0.90$

	$k, d = 0.0$	$k, d = 0.1$	$k, d = 0.2$	$k, d = 0.3$	$k, d = 0.4$	$k, d = 0.5$	$k, d = 0.6$	$k, d = 0.7$	$k, d = 0.8$	$k, d = 0.9$	$k, d = 0.99$	$k, d = 1.0$
MLE	2.3557	2.3557	2.3557	2.3557	2.3557	2.3557	2.3557	2.3557	2.3557	2.3557	2.3557	2.3557
LRE	2.3557	2.0182	1.7556	1.5460	1.3754	1.2343	1.1161	1.0158	0.9301	0.8560	0.7976	0.7916
RMLE	2.0892	2.0892	2.0892	2.0892	2.0892	2.0892	2.0892	2.0892	2.0892	2.0892	2.0892	2.0892
LLE	0.7916	0.9067	1.0313	1.1651	1.3082	1.4607	1.6226	1.7937	1.9742	2.1640	2.3428	2.3631
SRMLE	1.0961	1.0961	1.0961	1.0961	1.0961	1.0961	1.0961	1.0961	1.0961	1.0961	1.0961	1.0961

Table A8. The estimated MSE values for different k, d when $n = 50$ and $\rho = 0.99$.

	$k, d = 0.0$	$k, d = 0.1$	$k, d = 0.2$	$k, d = 0.3$	$k, d = 0.4$	$k, d = 0.5$	$k, d = 0.6$	$k, d = 0.7$	$k, d = 0.8$	$k, d = 0.9$	$k, d = 0.99$	$k, d = 1.0$
MLE	22.7202	22.7202	22.7202	22.7202	22.7202	22.7202	22.7202	22.7202	22.7202	22.7202	22.7202	22.7202
LRE	22.7202	7.1920	3.6283	2.2157	1.5085	1.1030	0.8486	0.6784	0.5590	0.4719	0.4124	0.4066
RMLE	2.0835	2.0835	2.0835	2.0835	2.0835	2.0835	2.0835	2.0835	2.0835	2.0835	2.0835	2.0835
LLE	0.4066	1.0428	2.0336	3.3790	5.0791	7.1337	9.5430	12.3070	15.4255	18.8990	22.3278	22.7265
SRMLE	2.1982	2.1982	2.1982	2.1982	2.1982	2.1982	2.1982	2.1982	2.1982	2.1982	2.1982	2.1982

Table A9. The estimated MSE values for different k, d when $n = 75$ and $\rho = 0.70$.

	$k, d = 0.0$	$k, d = 0.1$	$k, d = 0.2$	$k, d = 0.3$	$k, d = 0.4$	$k, d = 0.5$	$k, d = 0.6$	$k, d = 0.7$	$k, d = 0.8$	$k, d = 0.9$	$k, d = 0.99$	$k, d = 1.0$
MLE	0.5544	0.5544	0.5544	0.5544	0.5544	0.5544	0.5544	0.5544	0.5544	0.5544	0.5544	0.5544
LRE	0.5544	0.5368	0.5202	0.5046	0.4898	0.4758	0.4626	0.4501	0.4382	0.4271	0.4174	0.4164
RMLE	2.0701	2.0701	2.0701	2.0701	2.0701	2.0701	2.0701	2.0701	2.0701	2.0701	2.0701	2.0701
LLE	0.4164	0.4295	0.4429	0.4565	0.4703	0.4844	0.4987	0.5135	0.5280	0.5430	0.5567	0.5582
SRMLE	0.4368	0.4368	0.4368	0.4368	0.4368	0.4368	0.4368	0.4368	0.4368	0.4368	0.4368	0.4368

Table A10. The estimated MSE values for different k, d when $n = 75$ and $\rho = 0.80$.

	$k, d = 0.0$	$k, d = 0.1$	$k, d = 0.2$	$k, d = 0.3$	$k, d = 0.4$	$k, d = 0.5$	$k, d = 0.6$	$k, d = 0.7$	$k, d = 0.8$	$k, d = 0.9$	$k, d = 0.99$	$k, d = 1.0$
MLE	0.7833	0.7833	0.7833	0.7833	0.7833	0.7833	0.7833	0.7833	0.7833	0.7833	0.7833	0.7833
LRE	0.7833	0.7511	0.7169	0.6853	0.6561	0.6289	0.6037	0.5803	0.5584	0.5380	0.5208	0.5189
RMLE	2.0642	2.0642	2.0642	2.0642	2.0642	2.0642	2.0642	2.0642	2.0642	2.0642	2.0642	2.0642
LLE	0.5189	0.5433	0.5684	0.5941	0.6204	0.6474	0.6750	0.7032	0.7321	0.7617	0.7888	0.7918
SRMLE	0.5620	0.5620	0.5620	0.5620	0.5620	0.5620	0.5620	0.5620	0.5620	0.5620	0.5620	0.5620

Table A11. The estimated MSE values for different k, d when $n = 75$ and $\rho = 0.90$.

	$k, d = 0.0$	$k, d = 0.1$	$k, d = 0.2$	$k, d = 0.3$	$k, d = 0.4$	$k, d = 0.5$	$k, d = 0.6$	$k, d = 0.7$	$k, d = 0.8$	$k, d = 0.9$	$k, d = 0.99$	$k, d = 1.0$
MLE	1.5040	1.5040	1.5040	1.5040	1.5040	1.5040	1.5040	1.5040	1.5040	1.5040	1.5040	1.5040
LRE	1.5040	1.3650	1.2461	1.1434	1.0541	0.9756	0.9064	0.8451	0.7904	0.7414	0.7015	0.6973
RMLE	2.0593	2.0593	2.0593	2.0593	2.0593	2.0593	2.0593	2.0593	2.0593	2.0593	2.0593	2.0593
LLE	0.6973	0.7632	0.8324	0.9049	0.9809	1.0602	1.1429	1.2290	1.3184	1.4112	1.4976	1.5074
SRMLE	0.8479	0.8479	0.8479	0.8479	0.8479	0.8479	0.8479	0.8479	0.8479	0.8479	0.8479	0.8479

Table A12. The estimated MSE values for different k, d when $n = 75$ and $\rho = 0.99$.

	$k, d = 0.0$	$k, d = 0.1$	$k, d = 0.2$	$k, d = 0.3$	$k, d = 0.4$	$k, d = 0.5$	$k, d = 0.6$	$k, d = 0.7$	$k, d = 0.8$	$k, d = 0.9$	$k, d = 0.99$	$k, d = 1.0$
MLE	14.4762	14.4762	14.4762	14.4762	14.4762	14.4762	14.4762	14.4762	14.4762	14.4762	14.4762	14.4762
LRE	14.4762	6.4739	3.7327	2.4479	1.7388	1.3051	1.0201	0.8226	0.6801	0.5738	0.4997	0.4925
RMLE	2.0555	2.0555	2.0555	2.0555	2.0555	2.0555	2.0555	2.0555	2.0555	2.0555	2.0555	2.0555
LLE	0.4925	0.9960	1.6984	2.5998	3.7000	4.9992	6.4974	8.1944	10.0904	12.1850	14.2409	14.4792
SRMLE	2.0212	2.0212	2.0212	2.0212	2.0212	2.0212	2.0212	2.0212	2.0212	2.0212	2.0212	2.0212

Table A13. The estimated MSE values for different k, d when $n = 100$ and $\rho = 0.70$.

	$k, d = 0.0$	$k, d = 0.1$	$k, d = 0.2$	$k, d = 0.3$	$k, d = 0.4$	$k, d = 0.5$	$k, d = 0.6$	$k, d = 0.7$	$k, d = 0.8$	$k, d = 0.9$	$k, d = 0.99$	$k, d = 1.0$
MLE	0.4083	0.4083	0.4083	0.4083	0.4083	0.4083	0.4083	0.4083	0.4083	0.4083	0.4083	0.4083
LRE	0.4083	0.3989	0.3899	0.3812	0.3730	0.3651	0.3574	0.3502	0.3431	0.3364	0.3306	0.3300
RMLE	2.0524	2.0524	2.0524	2.0524	2.0524	2.0524	2.0524	2.0524	2.0524	2.0524	2.0524	2.0524
LLE	0.3300	0.3376	0.3453	0.3531	0.3611	0.3691	0.3772	0.3853	0.3936	0.4020	0.4096	0.4105
SRMLE	0.3400	0.3400	0.3400	0.3400	0.3400	0.3400	0.3400	0.3400	0.3400	0.3400	0.3400	0.3400

Table A14. The estimated MSE values for different k, d when $n = 100$ and $\rho = 0.80$.

	$k, d = 0.0$	$k, d = 0.1$	$k, d = 0.2$	$k, d = 0.3$	$k, d = 0.4$	$k, d = 0.5$	$k, d = 0.6$	$k, d = 0.7$	$k, d = 0.8$	$k, d = 0.9$	$k, d = 0.99$	$k, d = 1.0$
MLE	0.5801	0.5801	0.5801	0.5801	0.5801	0.5801	0.5801	0.5801	0.5801	0.5801	0.5801	0.5801
LRE	0.5801	0.5602	0.5413	0.5236	0.5068	0.4911	0.4760	0.4618	0.4484	0.4356	0.4274	0.4235
RMLE	2.0481	2.0481	2.0481	2.0481	2.0481	2.0481	2.0481	2.0481	2.0481	2.0481	2.0481	2.0481
LLE	0.4235	0.4381	0.4531	0.4682	0.4836	0.4993	0.5153	0.5316	0.5481	0.5650	0.5803	0.5821
SRMLE	0.4454	0.4454	0.4454	0.4454	0.4454	0.4454	0.4454	0.4454	0.4454	0.4454	0.4454	0.4454

Table A15. The estimated MSE values for different k, d when $n = 100$ and $\rho = 0.90$.

	$k, d = 0.0$	$k, d = 0.1$	$k, d = 0.2$	$k, d = 0.3$	$k, d = 0.4$	$k, d = 0.5$	$k, d = 0.6$	$k, d = 0.7$	$k, d = 0.8$	$k, d = 0.9$	$k, d = 0.99$	$k, d = 1.0$
MLE	1.1056	1.1056	1.1056	1.1056	1.1056	1.1056	1.1056	1.1056	1.1056	1.1056	1.1056	1.1056
LRE	1.1056	1.0302	0.9629	0.9025	0.8481	0.7989	0.7542	0.7135	0.6763	0.6422	0.6139	0.6109
RMLE	2.0444	2.0444	2.0444	2.0444	2.0444	2.0444	2.0444	2.0444	2.0444	2.0444	2.0444	2.0444
LLE	0.6109	0.6534	0.6975	0.7432	0.7905	0.8393	0.8898	0.9418	0.9955	1.0507	1.1017	1.1075
SRMLE	0.6967	0.6967	0.6967	0.6967	0.6967	0.6967	0.6967	0.6967	0.6967	0.6967	0.6967	0.6967

Table A16. The estimated MSE values for different k, d when $n = 100$ and $\rho = 0.99$.

	$k, d = 0.0$	$k, d = 0.1$	$k, d = 0.2$	$k, d = 0.3$	$k, d = 0.4$	$k, d = 0.5$	$k, d = 0.6$	$k, d = 0.7$	$k, d = 0.8$	$k, d = 0.9$	$k, d = 0.99$	$k, d = 1.0$
MLE	10.6280	10.6280	10.6280	10.6280	10.6280	10.6280	10.6280	10.6280	10.6280	10.6280	10.6280	10.6280
LRE	10.6280	5.7256	3.6179	2.5069	1.8466	1.4212	1.1308	0.9235	0.7703	0.6538	0.5714	0.5632
RMLE	2.0416	2.0416	2.0416	2.0416	2.0416	2.0416	2.0416	2.0416	2.0416	2.0416	2.0416	2.0416
LLE	0.5632	0.9868	1.5399	2.2227	3.0350	3.9769	5.0483	6.2494	7.5800	9.0402	10.4652	10.6300
SRMLE	1.8827	1.8827	1.8827	1.8827	1.8827	1.8827	1.8827	1.8827	1.8827	1.8827	1.8827	1.8827

Table A17. Summary of the Tables A1-A16.

Best Estimator	$n = 20$			
	$\rho = 0.7$	$\rho = 0.8$	$\rho = 0.9$	$\rho = 0.99$
LLE	$0.0 \leq k, d \leq 0.2$	$0.0 \leq k, d \leq 0.2$	$0.0 \leq k, d \leq 0.2$	$0.0 \leq k, d \leq 0.1$
SRMLE	$0.3 \leq k, d \leq 0.5$	$0.3 \leq k, d \leq 0.5$	$0.3 \leq k, d \leq 0.4$	$k, d = 0.2$
LRE	$0.6 \leq k, d \leq 1.0$	$0.6 \leq k, d \leq 1.0$	$0.5 \leq k, d \leq 1.0$	$0.3 \leq k, d \leq 1.0$
$n = 50$				
LLE	$0.0 \leq k, d \leq 0.1$	$0.0 \leq k, d \leq 0.2$	$0.0 \leq k, d \leq 0.2$	$0.0 \leq k, d \leq 0.2$
SRMLE	$0.2 \leq k, d \leq 0.7$	$0.3 \leq k, d \leq 0.7$	$0.3 \leq k, d \leq 0.6$	$k, d = 0.3$
LRE	$0.8 \leq k, d \leq 1.0$	$0.8 \leq k, d \leq 1.0$	$0.7 \leq k, d \leq 1.0$	$0.4 \leq k, d \leq 1.0$
$n = 75$				
LLE	$0.0 \leq k, d \leq 0.1$	$0.0 \leq k, d \leq 0.1$	$0.0 \leq k, d \leq 0.2$	$0.0 \leq k, d \leq 0.2$
SRMLE	$0.2 \leq k, d \leq 0.8$	$0.2 \leq k, d \leq 0.7$	$0.3 \leq k, d \leq 0.6$	$k, d = 0.3$
LRE	$0.9 \leq k, d \leq 1.0$	$0.8 \leq k, d \leq 1.0$	$0.7 \leq k, d \leq 1.0$	$0.4 \leq k, d \leq 1.0$
$n = 100$				
LLE	$0.0 \leq k, d \leq 0.1$	$0.0 \leq k, d \leq 0.1$	$0.0 \leq k, d \leq 0.1$	$0.0 \leq k, d \leq 0.2$
SRMLE	$0.2 \leq k, d \leq 0.8$	$0.2 \leq k, d \leq 0.8$	$0.2 \leq k, d \leq 0.7$	$k, d = 0.3$
LRE	$0.9 \leq k, d \leq 1.0$	$0.9 \leq k, d \leq 1.0$	$0.8 \leq k, d \leq 1.0$	$0.4 \leq k, d \leq 1.0$

Table A18. The best estimators and the corresponding k, d values when $n = 20$ and $n = 50$.

		$n = 20$	$n = 50$
$\rho = 0.7$	LLE	$0.1 \leq k \leq 0.6, 0.1 \leq d \leq 0.2$	$0.1 \leq k \leq 0.8, d = 0.1$
	LRE	$0.9 \leq k \leq 0.99, 0.1 \leq d \leq 0.2$	$0.9 \leq k \leq 0.99, d = 0.1$
	SRMLE	$0.6 \leq k \leq 0.99, 0.3 \leq d \leq 0.99$	$0.8 \leq k \leq 0.99, 0.2 \leq d \leq 0.99$
$\rho = 0.8$	LLE	$0.1 \leq k \leq 0.6, 0.1 \leq d \leq 0.2$	$0.1 \leq k \leq 0.8, 0.1 \leq d \leq 0.2$
	LRE	$0.8 \leq k \leq 0.99, 0.1 \leq d \leq 0.2$	$0.9 \leq k \leq 0.99, 0.1 \leq d \leq 0.2$
	SRMLE	$0.6 \leq k \leq 0.99, 0.3 \leq d \leq 0.99$	$0.8 \leq k \leq 0.99, 0.3 \leq d \leq 0.99$
$\rho = 0.9$	LLE	$0.1 \leq k \leq 0.5, 0.1 \leq d \leq 0.2$	$0.1 \leq k \leq 0.6, 0.1 \leq d \leq 0.2$
	LRE	$0.8 \leq k \leq 0.99, 0.1 \leq d \leq 0.2$	$0.1 \leq k \leq 0.6, 0.1 \leq d \leq 0.2$
	SRMLE	$0.5 \leq k \leq 0.99, 0.3 \leq d \leq 0.99$	$0.7 \leq k \leq 0.99, 0.3 \leq d \leq 0.99$
$\rho = 0.99$	LLE	$0.1 \leq k \leq 0.4, 0.3 \leq d \leq 0.99$	$0.1 \leq k \leq 0.6, 0.3 \leq d \leq 0.99$
	LRE	$0.1 \leq k \leq 0.2, d = 0.1$	$0.1 \leq k \leq 0.5, d = 0.1$
	SRMLE	$0.3 \leq k \leq 0.99, d = 0.1$	$0.6 \leq k \leq 0.99, d = 0.1$
$\rho = 0.99$	LLE	$0.2 \leq k \leq 0.99, 0.2 \leq d \leq 0.99$	$0.4 \leq k \leq 0.99, 0.2 \leq d \leq 0.99$
	LRE	$0.1 \leq k \leq 0.2, 0.2 \leq d \leq 0.99$	$0.1 \leq k \leq 0.3, 0.2 \leq d \leq 0.99$
	SRMLE		

Table A19. The best estimators and the corresponding k, d values when $n = 75$ and $n = 100$.

		$n = 75$	$n = 100$
$\rho = 0.7$	LLE	$0.1 \leq k \leq 0.8, d = 0.1$	$0.1 \leq k \leq 0.8, d = 0.1$
	LRE	$0.9 \leq k \leq 0.99, d = 0.1$	$0.9 \leq k \leq 0.99, 0.1 \leq d \leq 0.99$
	SRMLE	$0.8 \leq k \leq 0.99, 0.2 \leq d \leq 0.99$	$0.8 \leq k \leq 0.99, 0.2 \leq d \leq 0.99$
$\rho = 0.8$	LLE	$0.1 \leq k \leq 0.8, 0.2 \leq d \leq 0.99$	$0.1 \leq k \leq 0.8, 0.2 \leq d \leq 0.99$
	LRE	$0.1 \leq k \leq 0.8, d = 0.1$	$0.1 \leq k \leq 0.8, d = 0.1$
	SRMLE	$0.9 \leq k \leq 0.99, d = 0.1$	$0.9 \leq k \leq 0.99, 0.1 \leq d \leq 0.99$
$\rho = 0.9$	LLE	$0.8 \leq k \leq 0.99, 0.2 \leq d \leq 0.99$	$0.8 \leq k \leq 0.99, 0.2 \leq d \leq 0.99$
	LRE	$0.1 \leq k \leq 0.7, 0.2 \leq d \leq 0.99$	$0.1 \leq k \leq 0.8, 0.1 \leq d \leq 0.99$
	SRMLE	$0.1 \leq k \leq 0.7, 0.2 \leq d \leq 0.99$	$0.1 \leq k \leq 0.8, 0.1 \leq d \leq 0.99$
$\rho = 0.99$	LLE	$0.1 \leq k \leq 0.7, 0.1 \leq d \leq 0.2$	$0.1 \leq k \leq 0.8, d = 0.1$
	LRE	$0.9 \leq k \leq 0.99, 0.1 \leq d \leq 0.2$	$0.9 \leq k \leq 0.99, d = 0.1$
	SRMLE	$0.7 \leq k \leq 0.99, 0.3 \leq d \leq 0.99$	$0.8 \leq k \leq 0.99, 0.2 \leq d \leq 0.99$
$\rho = 0.99$	LLE	$0.1 \leq k \leq 0.6, 0.3 \leq d \leq 0.99$	$0.1 \leq k \leq 0.7, 0.2 \leq d \leq 0.99$
	LRE	$0.1 \leq k \leq 0.4, 0.1 \leq d \leq 0.2$	$0.1 \leq k \leq 0.4, 0.1 \leq d \leq 0.2$
	SRMLE	$0.5 \leq k \leq 0.99, 0.1 \leq d \leq 0.2$	$0.7 \leq k \leq 0.99, 0.1 \leq d \leq 0.2$
$\rho = 0.99$	LLE	$0.4 \leq k \leq 0.99, 0.3 \leq d \leq 0.99$	$0.4 \leq k \leq 0.99, 0.3 \leq d \leq 0.99$
	LRE	$0.1 \leq k \leq 0.3, 0.3 \leq d \leq 0.99$	$0.1 \leq k \leq 0.3, 0.3 \leq d \leq 0.99$
	SRMLE		