Süleyman Demirel Üniversitesi
Fen Bilimleri Enstitüsü Dergisi
Cilt 22, Sayı 1, 53-59, 2018

Süleyman Demirel University
Journal of Natural and Applied Sciences
Volume 22, Issue 1, 53-59, 2018

DOI: 10.19113/sdufbed.71595

Performance of a New Restricted Biased Estimator in Logistic Regression

Yasin ASAR[*1]

[1]Necmettin Erbakan University, Faculty of Science, Department of Mathematics-Computer Sciences, 42090, Konya

(Alınış / Received: 26.07.2017, Kabul / Accepted: 08.11.2017, Online Yayınlanma / Published Online: 14.12.2017)

Keywords
Estimation,
Liu-type estimator,
MLE,
MSE,
Multicollinearity,
Monte Carlo simulation.

Abstract: It is known that the variance of the maximum likelihood estimator (MLE) inflates when the explanatory variables are correlated. This situation is called the multicollinearity problem. As a result, the estimations of the model may not be trustful. Therefore, this paper introduces a new restricted estimator (RLTE) that may be applied to get rid of the multicollinearity when the parameters lie in some linear subspace in logistic regression. The mean squared errors (MSE) and the matrix mean squared errors (MMSE) of the estimators considered in this paper are given. A Monte Carlo experiment is designed to evaluate the performances of the proposed estimator, the restricted MLE (RMLE), MLE and Liu-type estimator (LTE). The criterion of performance is chosen to be MSE. Moreover, a real data example is presented. According to the results, proposed estimator has better performance than MLE, RMLE and LTE.

Lojistik Regresyon’da Yeni Bir Kısıtlı Yanlı Tahmin Edicinin Performansı

Anahtar Kelimeler
Tahmin,
Liu-tipi tahmin edici,
MLE,
MSE,
Çoklu-bağlantı,
Monte Carlo simülasyonu.

Özet: Açıklayıcı değişkenler ilişkili olduğunda en çok olabilirlik tahmincisinin (MLE) varyansının şiştiği bilinmektedir. Bu durum çoklu-bağlantı problemi olarak adlandırılır. Sonuç olarak, modelin tahminleri güvenilir olmayabilir. Bu nedenle, bu makale, lojistik regresyon modelinde, düşük boyutlu veriler için ($n > p$), çoklu-bağlantı problemini gidermek için parametrelerin bir alt uzayda olduğu durumda uygulanabilecek yeni kısıtlı bir tahminciyi ortaya koymaktadır. Bu makalede ele alınan tahmin edicilerin hata kareler ortalamaları (MSE) ve matris MSE’leri(MMSE) verilmiştir. Monte Carlo deneyi, önerilen tahmincinin (RLTE), kısıtlı en çok olabilirlik tahmincisinin (RMLE), MLE ve Liu tipi tahmincisinin (LTE) performanslarını değerlendirmek üzere tasarlanmıştır. Ayrıca gerçek veri üzerinde bir örnek gösterilmiştir. Performans değerlendirme kriteri olarak MSE seçilmiştir. Sonuçlara göre, yeni tahmincinin MLE, RMLE ve LTE’den daha iyi performansı vardır.

1. Introduction

The binary logistic regression model has become the popular method of analysis in the situation that the outcome variable is discrete or dichotomous. Although, its original acceptance is important in the field of epidemiologic researches, this method has become a commonly employed method in applied sciences such as engineering, health policy, biomedical research, business and finance, criminology, ecology, linguistics and biology [9].
In the analysis of a dichotomous dependent variable, lots of distribution functions are used, see [5]. However, the logistic distribution being an extremely flexible and easily used function and providing clinically meaningful interpretation, it has become the popular distribution in this research area [9].
Now, consider the following binary logistic regression model with intercept where the dependent variable is distributed as Bernoulli $Be(\pi)$ such that

$$\pi = \frac{e^{X\beta}}{1+e^{X\beta}} \tag{1}$$

where X is the $n\times(p+1)$ design matrix such that $(n > p)$, $\beta = [\beta_0, \beta_1, \ldots \beta_p]^{'}$ is the $(p+1)\times 1$ coefficient vector and p is the number of explanatory variables. In order to estimate the coefficient vector β, the following log-likelihood function is needed to be maximized

$$L(\beta) = \sum_{i=1}^{N} y_i \log(\pi_i) + \sum_{i=1}^{N} (1-y_i)\log(1-\pi_i). \tag{2}$$

The log-likelihood function can be maximized by differentiating it with respect to β and setting the obtained expression called likelihood equations equal to zero. The likelihood equations are given as follows:

$$\sum_{i=1}^{N} (y_i - \pi_i) = 0 \tag{3}$$

[*] *Corresponding author: yasar@konya.edu.tr*

$$\sum_{i=1}^{N} x_i (y_i - \pi_i) = 0. \tag{4}$$

Since the Equations (3) and (4) are nonlinear in β, one may use the iteratively weighted least squares (IWLS) algorithm. Therefore, maximum likelihood estimator (MLE) of β can be obtained using IWLS algorithm where which is given as follows [21]

$$\widehat{\beta}^{t+1} = \widehat{\beta}^{t} + \left(X'\widehat{W}^{t}X\right)^{-1} X' \left(y - \widehat{\pi}^{t}\right) \tag{5}$$

where $\widehat{\pi}^t$ is the estimated values of $\widehat{\pi}$ using $\widehat{\beta}^t$ and $\widehat{W}^t = diag\left(\widehat{\pi}^t_j \left(1-\widehat{\pi}^t_j\right)\right)$ such that $\widehat{\pi}^t_j$ is the jth element of $\widehat{\pi}^t$. In the final step of the algorithm, one gets the maximum likelihood estimator as

$$\widehat{\beta}_{MLE} = S^{-1} X' \widehat{W} \widehat{z} \tag{6}$$

where $S = X'\widehat{W}X$, $\widehat{z}' = (z_1, z_2, ..., z_n)$ with $\eta_j = x_j'\beta$ and $\widehat{z}_j = \widehat{\eta}_j + (y_j - \widehat{\pi}_j) / (\partial\eta_j / \partial\widehat{\pi}_j)$.

The weighted sum of squares can be minimized approximately by using the MLE. However, this estimator becomes unstable when the regressor variables are correlated. This problem is called multicollinearity. Thus, due to the high variance and very low t-ratios, the estimations of MLE are no more trustful. This is because the matrix S becomes ill-conditioned when there is multicollinearity.

There are some solutions to this ill-conditioning problem. Ridge regression which is firstly defined by [7] for the linear model is a very popular method. The ridge estimator has been generalized to binary logistic regression by [22] successfully as follows:

$$\widehat{\beta}_{LR} = (S + kI)^{-1} X' \widehat{W} \widehat{z} \tag{7}$$

where $k > 0$ and I is the $(p+1) \times (p+1)$ identity matrix. The authors applied the ridge estimators defined by [7] and [8] in logistic regression. Recently, a number of logistic ridge estimators have been applied and investigated by [16]. Moreover, see the following studies for different characterizations of this method in different models: [19], [23] and [24].

Another solution to the problem is to use Liu estimator defined by [12]. Logistic version of this estimator is defined by [15]. The authors showed that logistic Liu estimator has a better performance than MLE according to mean squared error (MSE) criterion. Since the logistic Liu estimator uses shrinkage parameter, its length becomes smaller than the length of MLE.

The logistic Liu-type estimator (LTE) defined by [10] can also be used as a solution to the problem. There are two parameters used in this estimator which seems to be a combination of Liu estimator and ridge estimator. LTE was defined as follows:

$$\widehat{\beta}_{LTE} = (S + kI)^{-1} (S - dI)\, \widehat{\beta}_{MLE} \tag{8}$$

where $k > 0$ and $-\infty < d < \infty$. Different methods to select the parameters (k,d) used in LTE are proposed by [3].

In statistical research, there may be prior information regarding the variables considered in the statistical analysis. Such kind of information my arise from different sources such as past experience or being an expert of the area etc. (see [18]). Therefore, in this paper, we also consider imposing some restrictions on the parameter space of the coefficient vector.

The purpose of this paper is to propose a restricted estimator by imposing restrictions on LTE and make a comparison between the estimators considered in this study and the new restricted Liu-type estimator (RLTE) by designing a Monte Carlo simulation study and a real data application. The organization of the paper is as follows: In Section 2, derivation of the proposed estimator is considered and MSE characteristics of listed estimators are given. Moreover, the optimal shrinkage parameters of the new estimator are obtained. In Section 3, the details of Monte Carlo simulation are demonstrated, a discussion regrading results of simulation are provided. and a real-life application is performed to show the benefits of the new method. Finally, a brief summary and conclusion are given in Section 4.

2. Theory and Method

2.1. Definition of the new estimators

Consider the following restrictions on the parameter space of the coefficient vector β

$$H\beta = h \tag{9}$$

where H is a matrix of order $q \times (p+1)$ of known elements and h is a vector of known elements of order $q \times 1$. The restricted MLE (RMLE) is proposed by [6] by imposing restrictions on the log-likelihood function (2). Therefore, the following objective function should be maximized

$$L(\beta, \lambda) = L(\beta) + \lambda'(H\beta - h) \tag{10}$$

where λ is a vector of Lagrangian multipliers. A Newton-Raphson method can be applied to find the solution ([11], [20]). One can compute the derivatives of (10) with respect to β and λ as follows

$$\frac{\partial L(\beta,\lambda)}{\partial \beta} = X'(y - \pi) + H'\lambda,$$

$$\frac{\partial^2 L(\beta,\lambda)}{\partial \beta \partial \beta'} = -\left(X'WX\right)$$

and

$$\frac{\partial L(\beta,\lambda)}{\partial \lambda} = H\beta - h.$$

Now, the t^{th} step of Newton-Raphson method is given by

$$\begin{aligned} \widehat{\beta}^{t+1}_{*} &= \widehat{\beta}^{t} + \left(X'\widehat{W}^{t}X\right)^{-1} \left[X'\left(y - \widehat{\pi}^{t}\right) + H'\lambda\right] \\ &= \widehat{\beta}^{t} + \left(X'\widehat{W}^{t}X\right)^{-1} X'\left(y - \widehat{\pi}^{t}\right) \\ &\quad + \left(X'\widehat{W}^{t}X\right)^{-1} H'\lambda \end{aligned} \tag{11}$$

where $\widehat{\beta}^t + \left(X'\widehat{W}^t X\right)^{-1} X'\left(y-\widehat{\pi}^t\right)$ is in the form of MLE and $\widehat{\beta}_*$ is the solution vector. Thus, Equation (11) becomes

$$\widehat{\beta}_*^{t+1} = \widehat{\beta}^{t+1} + \left(X'\widehat{W}^t X\right)^{-1} H'\lambda \tag{12}$$

where $\widehat{\beta}^{t+1} = \widehat{\beta}^t + \left(X'\widehat{W}^t X\right)^{-1} X'\left(y-\widehat{\pi}^t\right)$. Now, multiplying both sides of (12) by H and using $H\widehat{\beta}_*^{t+1} = h$,

$$h = H\widehat{\beta}^{t+1} + H\left(X'\widehat{W}^t X\right)^{-1} H'\lambda$$

is obtained. Then, one gets

$$h - H\widehat{\beta}^{t+1} = H\left(X'\widehat{W}^t X\right)^{-1} H'\lambda.$$

Since $H\left(X'\widehat{W}^t X\right)^{-1} H'$ is a positive definite matrix, λ can be estimated as

$$\widehat{\lambda} = \left[H\left(X'\widehat{W}^t X\right)^{-1} H'\right]^{-1}\left(h - H\widehat{\beta}^{t+1}\right). \tag{13}$$

Now, using (13) in (12), it is obtained that

$$\begin{aligned}\widehat{\beta}_*^{t+1} &= \widehat{\beta}^{t+1} + \left(X'\widehat{W}^t X\right)^{-1} H'\left[H\left(X'\widehat{W}^t X\right)^{-1} H'\right]^{-1} \\ &\quad \times \left(h - H\widehat{\beta}^{t+1}\right). \end{aligned} \tag{14}$$

Therefore, letting $\widehat{\beta}_* = \widehat{\beta}_{RMLE}$, in the the final step of this weighted procedure RMLE can be obtained as

$$\widehat{\beta}_{RMLE} = \widehat{\beta}_{MLE} - S^{-1}H'\left[HS^{-1}H'\right]^{-1}\left(H\widehat{\beta}_{MLE} - h\right). \tag{15}$$

Now, following [6], [13] and [20], the following penalized log-likelihood function is considered:

$$L(\beta,k,d,\lambda) = L(\beta) - \frac{1}{2}\|\sqrt{k}\beta + \frac{d}{\sqrt{k}}\widehat{\beta}_{MLE}\|^2 + \lambda'(H\beta - h) \tag{16}$$

where $\|\beta\|$ is the norm of β. Taking the derivatives of (16) with respect to β and λ, the followings are obtained

$$\frac{\partial L(\beta,k,d,\lambda)}{\partial\beta} = X'(y-\pi) - k\beta - d\widehat{\beta}_{MLE} + H'\lambda,$$

$$\frac{\partial^2 L(\beta,k,d,\lambda)}{\partial\beta\partial\beta'} = -\left(X'WX + kI\right)$$

and

$$\frac{\partial L(\beta,k,d,\lambda)}{\partial\lambda} = H\beta - h.$$

Similarly, the t^{th} step of Newton-Raphson method is given by

$$\begin{aligned}\widehat{\beta}_{**}^{t+1} &= \widehat{\beta}^t + \left(X'\widehat{W}^t X + kI\right)^{-1} \\ &\quad \times \left[X'\left(y-\widehat{\pi}^t\right) - k\beta - d\widehat{\beta}_{MLE} + H'\lambda\right]_{\beta=\widehat{\beta}^t}\end{aligned} \tag{17}$$

where $\widehat{\beta}_{**}$ is the solution the problem. Now, rearranging the terms of (17), it becomes

$$\begin{aligned}\widehat{\beta}_{**}^{t+1} &= \left(X'\widehat{W}^t X + kI\right)^{-1} X'\widehat{W}^t X\left(\widehat{\beta}^t + \left(X'\widehat{W}^t X\right)^{-1} X'\left(y-\widehat{\pi}^t\right)\right) \\ &\quad + \left(X'\widehat{W}^t X + kI\right)^{-1}\left[H'\lambda - d\widehat{\beta}^{t+1}\right] \\ &= \left(X'\widehat{W}^t X + kI\right)^{-1} X'\widehat{W}^t X\widehat{\beta}^{t+1} - d\left(X'\widehat{W}^t X + kI\right)^{-1}\widehat{\beta}^{t+1} \\ &\quad + \left(X'\widehat{W}^t X + kI\right)^{-1} H'\lambda \\ &= \left(X'\widehat{W}^t X + kI\right)^{-1}\left(X'\widehat{W}^t X - dI\right)\widehat{\beta}^{t+1} \\ &\quad + \left(X'\widehat{W}^t X + kI\right)^{-1} H'\lambda \qquad (18) \\ &= \left(I - (k+d)\left(X'\widehat{W}^t X + kI\right)^{-1}\right)\widehat{\beta}^{t+1} + \left(X'\widehat{W}^t X + kI\right)^{-1} H'\lambda \\ &= \widehat{\beta}^{t+1} + \left(X'\widehat{W}^t X + kI\right)^{-1}\left[H'\lambda - (k+d)\widehat{\beta}^{t+1}\right] \qquad (19)\end{aligned}$$

where $\widehat{\beta}^{t+1} = \widehat{\beta}^t + \left(X'\widehat{W}^t X\right)^{-1} X'\left(y-\widehat{\pi}^t\right)$ is the MLE at the $(t+1)^{th}$ step. Multiplying both sides of (19) and using $H\widehat{\beta}_{**}^{t+1} = h$, one obtains

$$h = H\widehat{\beta}^{t+1} + H\left(X'\widehat{W}^t X + kI\right)^{-1}\left(H'\lambda - (k+d)\widehat{\beta}^{t+1}\right)$$

Then, after some algebra, the estimator of λ becomes

$$\widehat{\lambda} = \left(H\widehat{S}_k^{-1}H'\right)^{-1}\left[h - H\widehat{S}_k^{-1}\widehat{S}_d\widehat{\beta}^{t+1}\right] \tag{20}$$

where $\widehat{S}_k^{-1} = \left(X'\widehat{W}^t X + kI\right)^{-1}$, $\widehat{S}_d = \left(X'\widehat{W}^t X - dI\right)$. Substituting (20) in (18), it is obtained that

$$\widehat{\beta}_{**}^{t+1} = \widehat{S}_k^{-1}\widehat{S}_d\widehat{\beta}^{t+1} + \widehat{S}_k^{-1}H'\left(H\widehat{S}_k^{-1}H'\right)^{-1}\left[h - H\widehat{S}_k^{-1}\widehat{S}_d\widehat{\beta}^{t+1}\right] \tag{21}$$

At the final iteration, letting $\widehat{\beta}_{**} = \widehat{\beta}_{RLTE}$, the proposed estimator RLTE is given by

$$\widehat{\beta}_{RLTE} = \widehat{\beta}_{LTE} - S_k^{-1}H'\left[HS_k^{-1}H'\right]^{-1}\left(H\widehat{\beta}_{LTE} - h\right) \tag{22}$$

where $S_k = S + kI$. There is also an alternative expression of RLTE as follows

$$\widehat{\beta}_{RLTE} = M_k S_d\widehat{\beta}_{MLE} + S_k^{-1}H'\left[HS_k^{-1}H'\right]^{-1}h \tag{23}$$

where $M_k = S_k^{-1} - S_k^{-1}H'\left[HS_k^{-1}H'\right]^{-1}HS_k^{-1}$ and $S_d = S - dI$.

2.2. MSE characteristics of estimators

Containing all relevant information of an estimator MMSE and MSE functions are used in the literature to make comparisons between estimators. MMSE and MSE of an estimator $\tilde{\beta}$ are defined respectively by

$$MMSE\left(\tilde{\beta}\right) = E\left[\left(\tilde{\beta} - \beta\right)\left(\tilde{\beta} - \beta\right)'\right], \tag{24}$$

$$MSE\left(\tilde{\beta}\right) = tr\left(MMSE\left(\tilde{\beta}\right)\right) = E\left[\left(\tilde{\beta} - \beta\right)\left(\tilde{\beta} - \beta\right)'\right] \tag{25}$$

where tr is the trace of a matrix.

In this subsection, the MMSE and MSE functions of the estimators are obtained. To obtained these functions, firstly, covariance matrices and bias vectors of estimators are computed. Firstly, since MLE is asymptotically unbiased, the covariance matrix, MMSE and MSE of MLE are given as follows (see, [16])

$$Cov\left(\widehat{\beta}_{MLE}\right) = S^{-1}, \quad (26)$$

$$MMSE(\widehat{\beta}_{MLE}) = S^{-1} \quad (27)$$

$$MSE(\widehat{\beta}_{MLE}) = \sum_{j=1}^{p+1} \frac{1}{\lambda_j} \quad (28)$$

where λ_i's are the eigenvalues of the matrix S.
RMLE has the following theoretical properties (see [1]):

$$\begin{aligned} Cov\left(\widehat{\beta}_{RMLE}\right) &= S^{-1} - S^{-1}H'\left[HS^{-1}H'\right]^{-1}HS^{-1} \\ &= M, \quad (29) \\ Bias\left(\widehat{\beta}_{RMLE}\right) &= -S^{-1}H'\left[HS^{-1}H'\right]^{-1}(H\beta - h) \\ &= -\delta, \quad (30) \\ MMSE\left(\widehat{\beta}_{RMLE}\right) &= M + \delta\delta', \quad (31) \\ MSE\left(\widehat{\beta}_{RMLE}\right) &= \sum_{j=1}^{p+1}\left[m_{jj} + \delta_j^2\right] \quad (32) \end{aligned}$$

where $Cov(\eta)$ is the covariance matrix and $Bias(\eta)$ is the bias of the vector η, m_{jj} is the j^{th} diagonal of $V'MV$ and δ_j is the j^{th} component of $V'\delta$ such that the columns of V are the eigenvectors of S.
The bias and covariance of LTE are presented by

$$Bias\left(\widehat{\beta}_{LTE}\right) = -(d+k)S_k^{-1}\beta \quad (33)$$

and

$$Cov\left(\widehat{\beta}_{LTE}\right) = S_k^{-1}S_d\Lambda^{-1}S_dS_k^{-1}. \quad (34)$$

In [1], MMSE and MSE of LTE are respectively given as

$$MMSE\left(\widehat{\beta}_{LTE}\right) = S_k^{-1}S_dS^{-1}S_dS_k^{-1} + (d+k)^2 S_k^{-1}\beta\beta'S_k^{-1}, \quad (35)$$

and

$$MSE\left(\widehat{\beta}_{LTE}\right) = \sum_{j=1}^{p+1}\left(\frac{(\lambda_j - d)^2}{\lambda_j(\lambda_j+k)^2}\right) + \sum_{j=1}^{p}\left(\frac{(d+k)^2\alpha_j^2}{(\lambda_j+k)^2}\right) \quad (36)$$

where α_j is the j^{th} component of $V'\beta$.
Using the alternative definition of RLTE, we can compute MMSE and MSE of RLTE as the following

$$Cov\left(\widehat{\beta}_{RLTE}\right) = M_kS_dS^{-1}S_dM_k, \quad (37)$$

$$Bias\left(\widehat{\beta}_{RLTE}\right) = -(d+k)M_k\beta, \quad (38)$$

$$MMSE\left(\widehat{\beta}_{RLTE}\right) = M_kS_dS^{-1}S_dM_k + (d+k)^2 M_k\beta\beta'M_k, \quad (39)$$

$$MSE\left(\widehat{\beta}_{RLTE}\right) = \sum_{j=1}^{p+1}\left[\frac{(\lambda_j-d)^2}{\lambda_j}m(k)_{jj}^2 + (d+k)^2\alpha_j^2 m(k)_{jj}^2\right] \quad (40)$$

where $m(k)_{jj}$ is the j^{th} diagonal of $V'M_kV$, λ_j is the j^{th} eigenvalue of S.
Since the values of k and d are not known in real data, it is useful to compare the estimators for some definite values of these parameters. Therefore, we obtain the MSE differences between the estimators.
Using the equations (28) and (40), we compute the difference

$$\begin{aligned} \Delta_1 &= MSE\left(\widehat{\beta}_{MLE}\right) - MSE\left(\widehat{\beta}_{RLTE}\right) \\ &= \sum_{j=1}^{p+1}\frac{1}{\lambda_j} - \left[\frac{(\lambda_j-d)^2}{\lambda_j}m(k)_{jj}^2 + (d+k)^2\alpha_j^2 m(k)_{jj}^2\right] \\ &= \sum_{j=1}^{p+1}\frac{1}{\lambda_j}\left(1 - m(k)_{jj}^2\left[(\lambda_j-d)^2 + (d+k)^2\lambda_j\alpha_j^2\right]\right). \end{aligned} \quad (41)$$

Similarly, using (32) and (40), we can also compute

$$\begin{aligned} \Delta_2 &= \Delta_2 = MSE\left(\widehat{\beta}_{RMLE}\right) - MSE\left(\widehat{\beta}_{RLTE}\right) \\ &= \sum_{j=1}^{p+1}\left[m_{jj}+\delta_j^2\right] - \left[\frac{(\lambda_j-d)^2}{\lambda_j}m(k)_{jj}^2 + (d+k)^2\alpha_j^2 m(k)_{jj}^2\right] \\ &= \sum_{j=1}^{p+1}\left(m_{jj} - m(k)_{jj}^2\left(\frac{(\lambda_j-d)^2}{\lambda_j} - (d+k)^2\alpha_j^2\right) + \delta_j^2\right). \end{aligned} \quad (42)$$

Finally, using (36) and (40), the following difference is obtained

$$\begin{aligned} \Delta_3 &= MSE\left(\widehat{\beta}_{LTE}\right) - MSE\left(\widehat{\beta}_{RLTE}\right) \\ &= \sum_{j=1}^{p+1}\left(\frac{(\lambda_j-d)^2}{\lambda_j(\lambda_j+k)^2} + \frac{(d+k)^2\alpha_j^2}{(\lambda_j+k)^2}\right) \\ &\quad - \left[\frac{(\lambda_j-d)^2}{\lambda_j}m(k)_{jj}^2 + (d+k)^2\alpha_j^2 m(k)_{jj}^2\right] \\ &= \sum_{j=1}^{p+1}\left(\frac{1}{(\lambda_j+k)^2} - m(k)_{jj}^2\right)\left(\frac{(\lambda_j-d)^2}{\lambda_j} + (d+k)^2\alpha_j^2\right) \end{aligned} \quad (43)$$

If the differences Δ_1, Δ_2 and Δ_3 can be showed that they are positive, then it means that RLTE is superior to the others. However, we skip the detailed theoretical comparisons and refer to [2] for similar comparisons. Therefore, we design a Monte Carlo experiment to compare the estimators in Section 3.

2.3. How to choose k and d

This subsection presents how to choose the biasing parameters used in RLTE. Since the MSE functions are quadratic functions of the parameter d and nonlinear functions of the parameter k, fixing the value of k, the optimal values of the parameter d can be obtained. In order to find the optimal parameter, fixing the value of the parameter k, it is sufficient to minimize the MSE function given in the past subsection by differentiating the MSE functions according to d and solving the resultant expression for d. Since, the optimal way of choosing the value of the parameter k cannot be obtained, the value of k is computed by using $k = \frac{p+1}{\widehat{\beta}'_{MLE}\widehat{\beta}_{MLE}}$ due to [22].

$$\frac{\partial MSE\left(\hat{\beta}_{RLTE}\right)}{\partial d} = 2\sum_{j=1}^{p+1}\left[\frac{(\lambda_j - d)}{\lambda_j} m(k)_{jj}^2 + (d+k)\,\alpha_j^2 m(k)_{jj}^2\right] = 0$$

the optimal parameter d_{RLTE} is computed as follows:

$$d_{RLTE} = \frac{\sum_{j=1}^{p+1}\left[m(k)_{jj}^2\left(1 - k\alpha_j^2\right)\right]}{\sum_{j=1}^{p+1}\left[m(k)_{jj}^2\left(\alpha_j^2 + \frac{1}{\lambda_j}\right)\right]}. \tag{44}$$

3. Numerical Experiments

In order to evaluate the performances of listed estimators, a Monte Carlo simulation experiment is conducted. Details and results of the simulation are presented in this section.

3.1. Details of the simulation

In a simulation study, defining important factors in designing the simulation is crucial. The main effective factor of this study is degree of correlation ρ among independent variables. In the experiment, strength of correlation ρ varies such that $\rho = 0.90, 0.99$ and 0.999. The sample size and number of regressor variables, being crucial factors, are varied as in many researches, for example see [14], [15], [16] and [25].
The following equation is used to produce the dataset having different strengths of correlation:

$$x_{ij} = \left(1-\rho^2\right)^{1/2} z_{ij} + \rho z_{ip} \tag{45}$$

where $i = 1,2,\ldots,n$, $j = 1,2,\ldots,p$ and z_{ij} is a random number produced using standard normal distribution. The response variable is also is generated from the Bernoulli distribution $Be(\pi_i)$ where

$$\pi_i = \frac{e^{x_i\beta}}{1+e^{x_i\beta}} \tag{46}$$

where x_i is the i^{th} row of data matrix X.
To impose some restrictions on the parameter space, following [17], the following restriction matrices are chosen for $p = 4$ and $p = 8$ respectively:

$$H_4 = \begin{bmatrix} 1 & 0 & -2 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}$$

$$H_8 = \begin{bmatrix} 1 & 0 & -2 & 1 & -3 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & -3 & 1 & -2 & 1 \end{bmatrix}$$

and $h = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ for both cases.
Performances of estimators are investigated by the simulated MSEs which are computed by the following equation:

$$MSE\left(\tilde{\beta}\right) = \sum_{r=1}^{5000}\left[\frac{\left(\tilde{\beta}-\beta\right)_r^{\prime}\left(\tilde{\beta}-\beta\right)_r}{5000}\right] \tag{47}$$

where $\left(\tilde{\beta}-\beta\right)_r$ is the difference for each estimator considered in this study at the rth step of simulation. Finally, we estimate the parameters of LTE following [1] and the parameters of RLTE are estimated by the proposed method.

Table 1. The estimated MSE values when $p = 4$

n	ρ	MLE	LTE	RMLE	RLTE
50	0.9	1.8735	1.1048	1.2743	1.0358
	0.99	2.4023	1.1603	1.3037	1.0166
	0.999	10.0165	2.5475	4.9053	1.0518
200	0.9	1.7663	1.1719	1.2763	1.0697
	0.99	1.8914	1.0852	1.3221	1.0239
	0.999	3.3698	1.3687	1.7409	1.0212
500	0.9	1.7321	1.2350	1.2107	1.0887
	0.99	1.7794	1.0772	1.2390	1.0283
	0.999	2.3380	1.1683	1.4210	1.0190

Table 2. The estimated MSE values when $p = 8$

n	ρ	MLE	LTE	RMLE	RLTE
50	0.9	2.0697	1.1517	1.3692	1.0262
	0.99	3.1446	1.4788	2.3250	1.0239
	0.999	26.1147	8.1679	12.7852	1.1934
200	0.9	1.7703	1.1103	1.3542	1.0331
	0.99	2.1467	1.1837	1.6183	1.0189
	0.999	5.9215	2.4199	3.8429	1.0260
500	0.9	1.7423	1.1581	1.3700	1.0640
	0.99	1.8525	1.1041	1.3668	1.0189
	0.999	3.3166	1.5889	2.2829	1.0208

3.2. Results of the simulation study

In Tables 1 - 2, the simulated MSEs of listed estimators are reported for different values of n, ρ and p. According to the tables, MLE seems to have the highest MSE values and RLTE has the lowest MSE values for all situations. LTE shows a better performance than MLE and RMLE in almost all cases.
Moreover, it is observed that the MSE values of MLE increase as the degree of correlation increases. On the other hand, the MSE of LTE and RMLE increases with a few exceptions as the degree of correlation increases. The MSE of RLTE shows a degenerated pattern in this situation. As the sample size increases, the MSE of MLE decreases. However, there is no regular pattern in the other estimators for this situation. Moreover, if the number of explanatory variables increases, the MSE values increase generally.

3.3. A real data application

In this subsection a real data application is presented to show the usefulness of the new estimator. The data set is taken from [9] and it is called "Myopia Study". There 618 observations and 17 different explanatory variables. However, for illustrative purposes, only the following variables are considered in the analysis.

- LT: Lens thickness
- ACD: Anterior chamber depth
- SPHEQ: Spherical equivalent refraction
- AL: Axial length
- VCD: Vitreous chamber depth

The response variable is that whether a subject has a myopia or not (coded as 1 and 0 respectively). Since all the variables are in the same scale (mm), the design matrix is not centred and standardized. The correlation matrix is presented in Table 3. Moreover, the condition number of the matrix $X'\widehat{W}X$ is computed as 14426 which shows that there is a collinearity problem in the data [4].

Table 3. The correlation matrix of the Myopia data set

	SPHEQ	AL	ACD	LT	VCD
SPHEQ	1.0000	-0.3055	-0.2388	0.0727	-0.2471
AL	-0.3055	1.0000	0.4563	-0.3289	0.9419
ACD	-0.2388	0.4563	1.0000	-0.3393	0.1994
LT	0.0727	-0.3289	-0.3393	1.0000	-0.4516
VCD	-0.2471	0.9419	0.1994	-0.4516	1.0000

The restriction matrix $H_1 = \begin{bmatrix} 1 & -1 & -1 & 1 & 0 \end{bmatrix}$ with $h = 0$ is used in order to compare the variables with the opposite sign of correlation and the correlation between AL and VCD is 0.94. Moreover, another restriction matrix H_2 defined as follows

$$H_2 = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 \end{bmatrix}$$

with $h_2 = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}'$ is also used computing the MSE performances of the estimators. The estimated MSE values are reported in Table 4. According to Table 4, it is observed that RLTE has the lowest MSE value and the MSE of MLE is inflated due to multicollinearity.

Table 4. MSE values for different restriction matrices

Restriction	MLE	LTE	RMLE	RLTE
H_1	5253.2388	245.2170	6.2390	5.0788
H_2	5253.2388	245.2170	1.1929	0.1927

4. Conclusion

In this paper, a new restricted estimator is proposed in the logistic regression. Theoretical properties of the new estimator are investigated. Moreover, MMSE and MSE functions are obtained. By a Monte Carlo simulation, the estimators MLE, RMLE, LTE and RLTE are compared in the sense of simulated MSE values. According to the results of the simulation, it is concluded that the new estimator RLTE has better performance than the others especially when the degree of correlation is high and the sample size is low. Furthermore, an application of the mentioned methods are also applied to a real life example and RLTE has the least MSE value. Therefore, RLTE is a better alternative when the multicollinear situations are present in the data.

References

[1] Asar, Y. 2017. Some new methods to solve multicollinearity in logistic regression. Communications in Statistics-Simulation and Computation, 46(4) 2576-2586.

[2] Arashi, M., Kibria, B. M. G., Norouzirad, M., Nadarajah, S. 2014. Improved preliminary test and Stein-rule Liu estimators for the ill-conditioned elliptical linear regression model. Journal Multivariate Analysis. 126 53–74.

[3] Asar, Y., Genç, A. 2016. New shrinkage parameters for the Liu-type logistic estimators. Communications in Statistics-Simulation and Computation, 45(3) 1094-1103.

[4] Belsley, D. A., Kuh, E., Welsch, R.E. 1980. Regression diagnostics: Identifying influential data and sources of collinearity. Wiley, New York.

[5] Cox, D. R., Snell, E. J. 1989. Analysis of Binary Data (2nd ed.): Chapman & Hall.

[6] Duffy, D. E., Santner, T. J. 1989. On the small sample properties of norm-restricted maximum likelihood estimators for logistic regression models. Communications in Statistics-Theory and Methods, 18(3) 959-980.

[7] Hoerl, A. E., Kennard, R. W. 1970. Ridge regression: Biased estimation for nonorthogonal problems. Technometrics, 12(1) 55-67.

[8] Hoerl, A. E., Kennard, R. W., Baldwin, K. F. 1975. Ridge regression: some simulations. Communications in Statistics-Theory and Methods, 4(2) 105-123.

[9] Hosmer, D.D., Lemeshow, S., Sturdivant, R.X. 2013. Applied Logistic Regression. Hoboken, New Jersey: John Wiley & Sons.

[10] İnan, D., Erdoğan, B. E. 2013. Liu-Type Logistic Estimator. Communications in Statistics-Simulation and Computation, 42(7) 1578-1586.

[11] LeCessie, S., VanHouwelingen, J.C. 1992. Ridge estimators in logistic regression. Applied Statistics 41(1) 191-201.

[12] Liu, K. 1993. A new class of blased estimate in linear regression. Communications in Statistics-Theory and Methods, 22(2) 393-402.

[13] Liu, K. 2003. Using Liu-type estimator to combat collinearity. Communications in Statistics-Theory and Methods, 32(5), 1009-1020.

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

[15] Månsson, K., Kibria, B. M. G., Shukur, G. 2012. On Liu estimators for the logit regression model. Economic Modelling, 29(4) 1483-1488.

[16] Månsson, K., and Shukur, G. 2011.On Ridge Parameters in Logistic Regression. Communications in Statistics-Theory and Methods, 40(18) 3366-3381.

[17] Månsson, K., Kibria, B. M. G., Shukur, G. 2016. A restricted Liu estimator for binary regression models and its application to an applied demand system. Journal of Applied Statistics, 43(6) 1119-1127.

[18] Rao, R.C., Shalabh, Toutenburg, H.S., Heumann, C.

2008. Linear models and generalizations, least squares and alternatives. ISBN 978-3-540-74226-5 3rd edition Springer Berlin Heidelberg New York.

[19] Roozbeh, M., Arashi, M. 2014. Feasible ridge estimator in seemingly unrelated semiparametric models. Communications in Statistics-Simulation and Computation, 43(10), 2593-2613.

[20] Özkale, M. R., Lemeshow, S., Sturdivant, R. 2017. Logistic regression diagnostics in ridge regression. Computational Statistics, https://doi.org/10.1007/s00180-017-0755-x.

[21] Saleh, A. M. E., Kibria, B. M. G. 2013. Improved ridge regression estimators for the logistic regression model. Computational Statistics, 28(6) 2519-2558.

[22] Schaefer, R. L., Roi, L. D., Wolfe, R. A. 1984. A Ridge Logistic Estimator. Communications in Statistics-Theory and Methods, 13(1) 99-113.

[23] Yüzbaşı, B., Ahmed, S. E., Güngör, M. 2017. Improved Penalty Strategies in Linear Regression Models, REVSTAT–Statistical Journal, 15(2), 251-276.

[24] Yüzbaşı, B., Ejaz Ahmed, S. 2016. Shrinkage and penalized estimation in semi-parametric models with multicollinear data. Journal of Statistical Computation and Simulation, 86(17), 3543-3561.

[25] Wu, J. 2016. Modified restricted Liu estimator in logistic regression model. Computational Statistics, 31(4) 1557-1567.