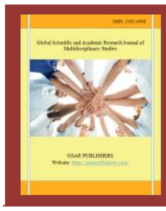

Global Scientific and Academic Research Journal of Multidisciplinary Studies
ISSN: 2583-4088 (Online)
Frequency: Monthly
Published By GSAR Publishers
Journal Homepage Link- https://gsarpublishers.com/journals-gsarjms-home/

COMPARISON OF MLE, LASSO, AND LIU ESTIMATOR METHODS TO OVERCOME MULTICOLLINEARITY IN MULTINOMIAL LOGISTIC REGRESSION: SIMULATION STUDY

BY

Siska Ayuditya Saputri[1*], Netti Herawati[1], Agus Sutrisno[1], Nusyirwan[1], Khoirin Nisa[1]

[1]Department of Mathematics, University of Lampung

Article History

Received: 09/05/2024
Accepted: 18/05/2024
Published: 20/05/2024

Vol – 3 Issue –5

PP: - 08-12

Abstract

The purpose of this study is to compare the performance of the MLE, LASSO, and Liu Estimator, and MLE methods in dealing with multicollinearity using simulated data with n=50,75,150, and 300 in multinomial logistic model (p=6) with $\rho = 0{,}3$ and 0,99. The best model was compered using AIC, MSE, BIC.is the best based zn the MSE, SE, AIC, BIC values. The result showed that LASSO and Liu Estimator methods were able to overcome partial multicollinearity in 3 independent variables and full multicollinearity 6 independent variables much better than MLE method. This result is based on MSE, AIC, and BIC values of LASSO and Liu Estimator which are much smaller than those of MLE.

Keywords*- LASSO, Liu Estimator, MLE, multicollinearity*

Introduction

Multicollinearity refers to a condition where two or more independent variables in a logistic regression model have a strong and correlated relationship with each other. After multicollinearity appears, it can complicate data interpretation and make the logistic regression model unfavorable. The main problem that arises is the instability of the parameter estimates, where the variance of the estimates becomes very large (Montgomery & Peck, 1992). When in analyzing regression we want to make conclusion, the presence of multicollinearity makes it a serious problem that needs to be addressed. Therefore, it is important to find the most suitable method to deal with multicollinearity. There are many methods to overcome this multicollinearity and even eliminate the multicollinearity, suggested by Jollife (2002). Of the many existing methods, the LASSO method is one of them, because it can shrink the regression coefficient to near zero or even exactly zero. The next method is Liu Estimator, which plays the same role as LASSO in overcoming multicollinearity whose estimator serves as a bias shrinking and generalization direct estimator. The multinomial logistic regression model requires that there is no multicollinearity between independent variables. If this happens, the parameter estimation will be poor. This study will compare which method is better between MLE, LASSO, and Liu Estimator in terms of overcoming multicollinearity in multinomial logistic regression based on the smallest MSE, SE, AIC, BIC values.

MATERIALS AND METHODS

Data

In this study, we simulated data with n = 50, 75, 150, and 300 containing full multicollinearity in 6 explanatory variables with (ρ = 0.99) and partial multicollinearity in 3 explanatory variables (ρ = 0.3) using the R package with 100 iterations. The explanatory variables are generated from Monte Carlo simulations:

$$\boldsymbol{X_p = (1-\rho^2)^{\frac{1}{2}} Z_{ij} + \rho Z_{i(p+1)}} \qquad (1)$$

Where ρ is determined in order to get a high correlation between the 6 explanatory variables. The dependent variable is derived from multinomial logistic regression probabilities for each category:

$$\pi_0(x) = \frac{1}{1 + exp\ g_1(x) + exp\ g_2(x) + exp\ g_3(x)} \qquad (2)$$

$$\pi_1(x) = \frac{exp\ g_1(x)}{1 + exp\ g_1(x) + exp\ g_2(x) + exp\ g_3(x)} \qquad (3)$$

$$\pi_2(x) = \frac{exp\ g_2(x)}{1 + exp\ g_1(x) + exp\ g_2(x) + exp\ g_3(x)} \qquad (4)$$

$$\pi_3(x) = \frac{exp\ g_3(x)}{1 + exp\ g_1(x) + exp\ g_2(x) + exp\ g_3(x)} \qquad (5)$$

*Corresponding Author: **Siska Ayuditya Saputri**

To see the value of multicollinearity in the data set, the variance variance inflation factor (VIF) is checked.

Method

LASSO

According to Tibshirani (1996), LASSO coefficient estimation uses quadratic programming with inequality constraints. The LASSO estimates are obtained from the following equation:

$$\hat{\beta}^{LASSO} = \underset{\beta}{argmin} \sum_{i=1}^{n} \left(y_i - \beta_0 - \sum_{k=1}^{p} \beta_k x_{ik}\right)^2 \quad (6)$$

Conditional on $\sum_{k=1}^{p} |\hat{\beta}_k^{LASSO}| \leq t$, where t is a tuning parameter that controls the shrinkage of the LASSO coefficients by $t \geq 0$. So that if $0 < t < \sum_{k=1}^{p} |\hat{\beta}_k^{0}|$, then:

$$\hat{\beta}^{LASSO} = \underset{\beta}{argmin} \sum_{i=1}^{n} \left\{\frac{1}{2}\left(y_i - \beta_0 - \sum_{k=1}^{p} \beta_k x_{ik}\right)^2 + \lambda \sum_{k=1}^{p} |\hat{\beta}_k|\right\} \quad (7)$$

According to Hastie, et al. (2008), the LASSO method in logistic regression is applied by adding a LASSO penalty to the log-likelihood function to estimate the independent variable. Therefore, the logistic regression estimator can be obtained by maximizing the log-likelihood function. Logistic regression estimator for multinomial response variable with LASSO penalty in the following equation:

$$\hat{\beta}^{LASSO} = argmax\left\{l(\hat{\beta}) - \lambda \sum_{j=1}^{J} \sum_{k=1}^{p} |\hat{\beta}_{jk}|\right\} \quad (8)$$

Liu Estimator

Liu's method is a method in logistic regression that function as a biased shrinkage estimator and a method whose generalization is direct. Liu (1993) proposed the use of another estimator where the parameters obtained can serve as a linear function of the depreciation parameter d. According to Qasim, et al (2019), determining the value of d in the Liu's estimator is much simpler. Hoerl & Kennard (1970) suggested the following for the value of d in Liu's method:

$$d_1 = max\left[0, \frac{\hat{a}_{j\,max}^2 - 1}{\frac{1}{\lambda_{j\,max}} + a_{max}^2}\right] \quad (9)$$

Furthermore, for the next value of d proposed and based on the concepts in Kibria (2003):

$$d_2 = max\left[0, median \frac{\hat{a}_j^2 - 1}{\frac{1}{\lambda_j} + \hat{a}_j^2}\right] \quad (10)$$

$$d_3 = max\left[0, \frac{1}{p} \sum_{j}^{J} \left(\frac{\hat{a}_j^2 - 1}{\frac{1}{\lambda_j} + \hat{a}_j^2}\right)\right] \quad (11)$$

The following estimator is proposed where quantiles other than the median are used, and this approach was successfully applied by Khalaf & Shukur (2005):

$$d_4 = max\left[0, \max \frac{\hat{a}_j^2 - 1}{\frac{1}{\lambda_j} + \hat{a}_j^2}\right] \quad (12)$$

Maximum Likelihood Estimation

Besides LASSO and Liu Estimator that can estimate logistic regression parameters, another method is Maximum Likelihood Estimation (MLE). Because this study uses multinomial logistic regression, according to Kleinbaum & Klein (2010), the response variable is also multinomial, so the likelihood function is adjusted to the multinomial distribution. According to Hosmer & Lemeshow (2000), the parameter chosen is the one that maximizes the likelihood function. The general form of the Maximum Likelihood method parameter estimation function is: $L(\beta) \prod_{i=1}^{n} [\pi_0(x_i)^{y_{0i}} \pi_1(x_i)^{y_{1i}} \ldots \pi_{j-1}(x_i)^{y_{(j-1)i}}]$ **(13)** dimana $\beta = (\beta_{j0}, \beta_{j1}, \ldots, \beta_{jp})$.

RESULTS AND DISCUSSION

The simulated data in this study has 6 explanatory variables, and there are 2 versions, namely, first, only 3 independent variables contain multicollinearity, and second, all independent variables contain multicollinearity. Multicollinearity is checked through VIF and correlation values. The results of the analysis of simulated data with n = 50, 75, 150, and 300, and the ρ is 0.3 and 0.99 which causes the simulated data to contain high multicollinearity. In this study, LASSO and Liu Estimator can overcome the high multicollinearity problem, while MLE can only overcome multicollinearity when 3 independent variables contain multicollinearity when n=300 and 6 independent variables contain multicollinearity when n=150.

The analysis result in Tables 1 to 4 show that LASSO and Liu Estimator appear to have smaller MSE, AIC, and BIC values than MLE for the 3 independent variables multicollinearity when the samples are 50 and 75, and 150, but when the sample is 300, the MSE value of MLE is smaller than Liu Estimator.

Table 1: MSE, AIC, and BIC values of LASSO, MLE, Liu Estimator Methods for n=50

$n = 50$		
MLE	MSE	25726.44
	AIC	37.3399
	BIC	50.7241
LASSO	MSE	0.158605
	AIC	29.8598
	BIC	45.1311
Liu Estimator	MSE	0.0523
	AIC	29.3399
	BIC	44.6361

Table 2: MSE, AIC, and BIC values of LASSO, MLE, Liu Estimator Methods for n=75

$n = 75$		
MLE	MSE	2605.016
	AIC	54.4793
	BIC	70.7017
LASSO	MSE	0.07816
	AIC	46.9703
	BIC	65.4974

*Corresponding Author: **Siska Ayuditya Saputri**

Liu Estimator	MSE	0.0530
	AIC	46.4793
	BIC	65.0192

Table 3: MSE, AIC, and BIC values of LASSO, MLE, Liu Estimator Methods for n=150

$n = 150$		
MLE	MSE	0.09758
	AIC	104.4535
	BIC	125.5279
LASSO	MSE	0.03319
	AIC	96.9519
	BIC	121.0623
Liu Estimator	MSE	0.0510
	AIC	96.4534
	BIC	120.5385

Table 4: MSE, AIC, and BIC values of LASSO, MLE, Liu Estimator Methods for n=300

$n = 300$		
MLE	MSE	0.04052
	AIC	186.9717
	BIC	212.8982
LASSO	MSE	0.02194
	AIC	179.4671
	BIC	209.0698
Liu Estimator	MSE	0.0510
	AIC	178.9717
	BIC	208.602

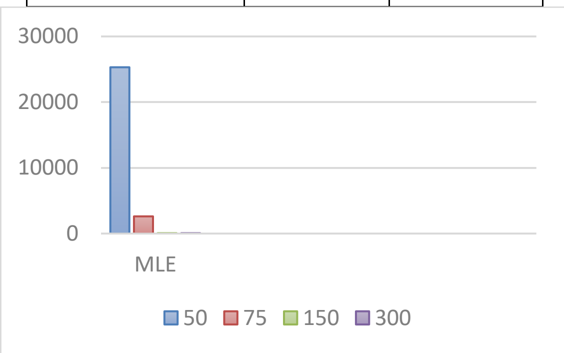

Figure 1: Graph of MSE value of MLE method when 3 independent variables contain multicollinearity

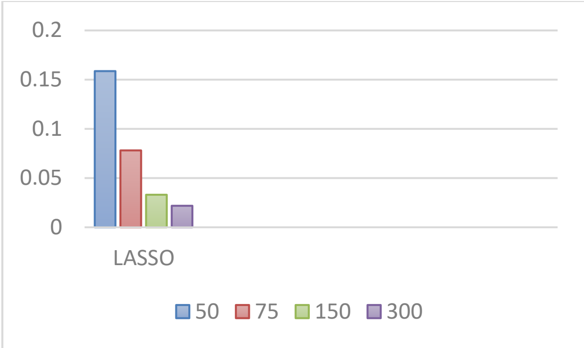

Figure 2: Graph of MSE value of LASSO method when 3 independent variables contain multicollinearity

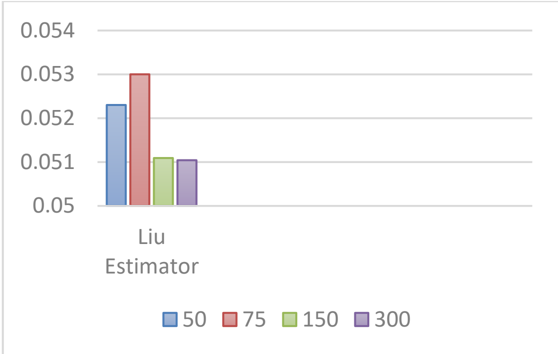

Figure 3: Graph of MSE value of Liu Estimator method when 3 independent variables contain multicollinearity

Figures 1, 2, and 3 show the graphical comparison of MSE values for MLE, LASSO, and Liu Estimator. Which result in that LASSO and Liu Estimator have smaller MSE values than MLE.

The analysis results in Tables 5 to 8 show that LASSO and Liu Estimator have smaller SE, MSE, AIC, and BIC values than MLE for the 6 explanatory variables that contain multicollinearity.

Table 5: MSE, AIC, and BIC values of LASSO, MLE, Liu Estimator Methods for n=50

$n = 50$		
MLE	MSE	72447.83
	AIC	32.1249
	BIC	45.5090
LASSO	MSE	0.9790
	AIC	24.6447
	BIC	39.9160
Liu Estimator	MSE	0.0588
	AIC	24.1249
	BIC	39.4211

Table 6: MSE, AIC, and BIC values of LASSO, MLE, Liu Estimator Methods for n=75

$n = 75$		
MLE	MSE	1602.587
	AIC	43.6510
	BIC	59.8734

LASSO	MSE	1.1947
	AIC	36.1421
	BIC	54.6692
Liu Estimator	MSE	0.0590
	AIC	35.6510
	BIC	54.1909

Table 7: MSE, AIC, and BIC values of LASSO, MLE, Liu Estimator Methods for n=150

$n = 150$		
MLE	MSE	0.5220
	AIC	74.8202
	BIC	95.8946
LASSO	MSE	0.10102
	AIC	67.3186
	BIC	91.4290
Liu Estimator	MSE	0.05370
	AIC	66.8202
	BIC	90.9053

Table 8: MSE, AIC, and BIC values of LASSO, MLE, Liu Estimator Methods for n=300

$n = 300$		
MLE	MSE	0.1603
	AIC	135.9484
	BIC	161.8749
LASSO	MSE	0.0715
	AIC	128.4438
	BIC	158.0465
Liu Estimator	MSE	0.05374
	AIC	127.9484
	BIC	157.5786

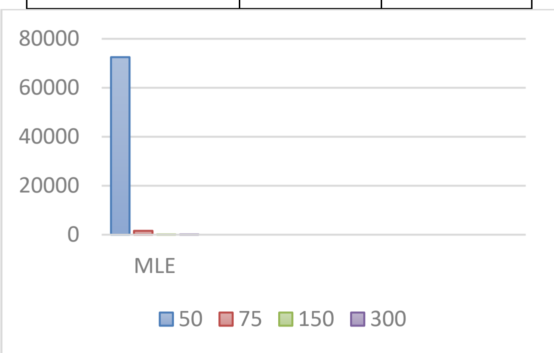

Figure 4: Graph of MSE value of MLE method when 6 independent variables contain multicollinearity

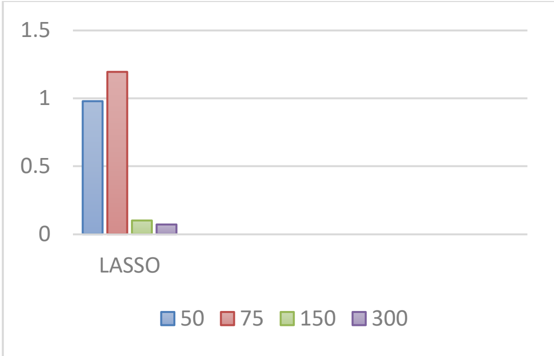

Figure 5: Graph of MSE value of LASSO method when 6 independent variables contain multicollinearity

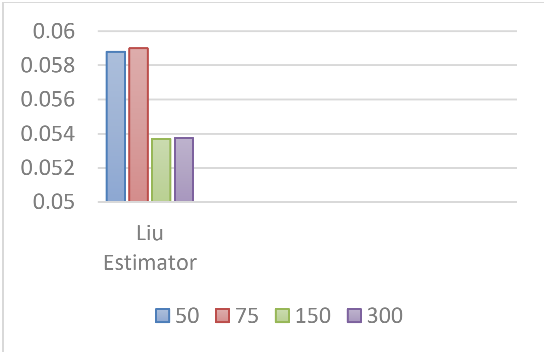

Figure 6: Graph of MSE value of Liu Estimator method when 6 independent variables contain multicollinearity

Figures 4, 5, and 6 show the graphical comparison of MSE values for MLE, LASSO, and Liu Estimator. Which result in that LASSO and Liu Estimator have smaller MSE values than MLE.

Table 9: MSE, AIC, and BIC values of LASSO, MLE, Liu Estimator Methods for n=150

Multicollinearity		MLE	LASSO	Liu Estimator
x_1, x_2, x_3	n = 50	25726.44	0.158605	0.0523
	n = 75	2605.016	0.07816	0.0530
	n = 150	0.09758	0.03319	0.0510
	n = 300	0.04052	0.02194	0.0510
$x_1, x_2, x_3 x_4, x_5, x_6$	n = 50	72447.83	0.9790	0.0588
	n = 75	1602.587	1.1947	0.0590
	n = 150	0.5220	0.10102	0.05370
	n = 30	0.1603	0.0715	0.05374

	0			

To determine the best method, we used the Mean Square Error (MSE), Akaike Information Criterion (AIC), and Bayesian Information Criterion (BIC) of the models obtained using the three methods studied which can be seen in tables 1 to 8. Liu Estimator has the smallest AIC and BIC values compared to LASSO and MLE. It can also be seen that LASSO and Liu Estimator have smaller MSE values than MLE, which means that the best methods are LASSO and Liu Estimator. Table 9 shows that LASSO and Liu Estimator are able to overcome multicollinearity because they have the smallest MSE value compared to MLE.

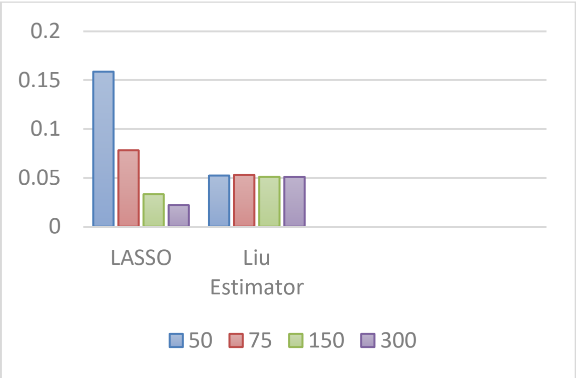

Figure 7: Graph of MSE values in LASSO and Liu Estimator when 3 independent variables are correlated

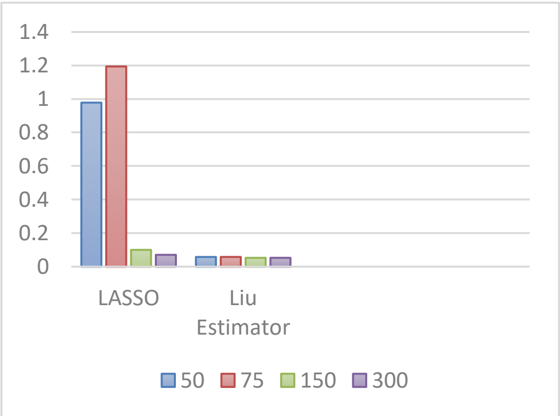

Figure 8: Graph of MSE values in LASSO and Liu Estimator when 6 independent variables are correlated

Based on Figure 7, when the 3 independent variables are correlated when the samples are 50 and 75, Liu Estimator has a smaller MSE value than LASSO, but when the samples are 150 and 300, LASSO has a smaller MSE value than Liu Estimator. Furthermore, in Table 8 on the data of 6 correlated independent variables, in all samples Liu Estimator outperforms LASSO because it has a smaller MSE value compared to LASSO. This means that Liu Estimator can overcome multicollinearity problems in all samples when 6 independent variables are correlated.

CONCLUSION

Based on the results of the above research, it can be concluded that:

1. *In all samples studied n = 50, 75, 150, and 300 when 3 independent variables contain multicollinearity and 6 independent variables contain multicollinearity, the LASSO and Liu Estimator methods are better than the MLE method to overcome multicollinearity.*
2. *When n = 50 and 75 with 3 independent variables contain multicollinearity, Liu Estimator method is better than LASSO method, but at n = 150 and 300 when 3 independent variables contain multicollinearity, LASSO method is better than Liu Estimator. Then at n = 50, 75, 150, and 300 when 6 independent variables contain multicollinearity, the Liu Estimator method is better used than LASSO because it is seen from the smaller MSE value of Liu Estimator compared to the LASSO method.*
3. *Based on the AIC and BIC values, the Liu Estimator method is better used to estimate the model compared to LASSO and MLE.*

REFERENCES

1. Hastie, T., Tibshirani, R., & Friedman, J. (2008). The Elements of Statistical Learning: Data Mining, Inference, and Prediction. (Second Edition). New York: Springer.
2. Hoerl, A.E. & Kennard, R.W. (1970). Ridge regression: biased estimation for nonorthogonal problems. Technometrics, 12, 55–67.
3. Jolliffe, I. T. (2002). Principal Component Analysis. (Second Edition). New York: Springer-Verlag.
4. Khalaf, G. & Shukur, G. (2005). Choosing Ridge Parameter for Regression Problems. Commun. Stat. Theory Methods, 34(5), 1177–1182.
5. Kibria, B.M.G. (2003). Performance of some new ridge regression estimators. Commun. Stat. Theory Methods, 32, 419–435.
6. Liu K. (1993). A new class of biased estimate in linear regression. Communications in Statistics, 22(2), 393-402.
7. Montgomery, D. C. & Peck, E. A. (1992). Introduction To Linear Regression Analysis. (Second Edition). New York: John Wiley & Sons.
8. Qasim, M., Amin, M., & Omer, T. (2019). Performance Of Some New Liu Parameters for the Linear Regression Model. Communication In Statistics Theory And Methods, 49(17), 4178-4196.
9. Tibshirani, R. (1996). Regression Shrinkage and Selection via LASSO. Journal of the Royal Statistical Society. Series B (Methodological), 58(1), 267-288.

*Corresponding Author: **Siska Ayuditya Saputri**