

The Performance of Robust Methods in Logistic Regression Model

Idriss Abdelmajid Idriss Ahmed, Weihu Cheng

College of Applied Science, Department of Statistics, Beijing University of Technology, Beijing, China
Email: Idrissabdo59@gmail.com, Chengweihu@bjut.edu.cn

How to cite this paper: Ahmed, I.A.I. and Cheng, W.H. (2020) The Performance of Robust Methods in Logistic Regression Model. *Open Journal of Statistics*, 10, 127-138.
<https://doi.org/10.4236/ojs.2020.101010>

Received: January 22, 2020

Accepted: February 25, 2020

Published: February 28, 2020

Copyright © 2020 by author(s) and Scientific Research Publishing Inc.
This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).
<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Logistic regression is the most important tool for data analysis in various fields. The classical approach for estimating parameters is the maximum likelihood estimation, a disadvantage of this method is high sensitivity to outlying observations. Robust estimators for logistic regression are alternative techniques due to their robustness. This paper presents a new class of robust techniques for logistic regression. They are weighted maximum likelihood estimators which are considered as Mallows-type estimator. Moreover, we compare the performance of these techniques with classical maximum likelihood and some existing robust estimators. The results are illustrated depending on a simulation study and real datasets. The new estimators showed the best performance relative to other estimators.

Keywords

Logistic Regression, Maximum Likelihood Estimator, Robust Estimation, Outlier, Weighted Maximum Likelihood

1. Introduction

Logistic regression is a proper analysis method to model the data and explain the relationship between the binary response variable and explanatory variables. The maximum likelihood estimator is a common technique of parameter estimation in the binary regression model. Unfortunately, this method does not resistant against atypical observations in data. To handle this problem, many robust estimators an alternative to MLE have been proposed. [1] developed a diagnostic measurement of outlying observations and they showed that in the logistic regression, the MLE was very sensitively to outlying observations (see also [2]). [3] discussed different types of M-estimators for binary regression, these estimates

belong to the Mallows-type based on leverage down weight. [4] derived a robust estimator based on a modified median estimator for the logistic regression model and they also studied a Wald-type test statistic for the logistic regression model. [5] developed projection estimators for the GLM which are very robust but their computation is extremely complex. [6] defined a robust estimator based on the quasi-likelihood, which replaced the least squares estimator (l_2 norm) by the least absolute deviation estimator (L_1 norm) in the definition of quasi-likelihood. [7] proposed a natural class of robust estimator and testing procedures for binomial models and Poisson models, which are based on a concept of quasi-likelihood estimator proposed by [8]. [9] studied the breakdown of the maximum likelihood estimator in the logistic model. [10] suggested a highly robust and consistent estimator. [11] presented a stable and fast algorithm to compute the M-estimator introduced by [10]. [12] introduced a fast algorithm based on breakdown points of the trimmed likelihood for the generalized linear model. Another class of the robust estimator is the fisher-consistent estimators proposed by [10]. [14] studied a robust resistant estimator and this estimator based on the misclassification model. [15] generalized optimally bounded score functions studied by [16] for linear models to the logistic model.

In this article we investigate the use of weight functions introduced by [17] as a weight function for Mallows type (weighted maximum likelihood estimator) to obtain a robust estimation for logistic regression, in addition, to compare their performance with classical maximum likelihood estimator and some existing robust methods by means of simulation study and real data sets.

The maximum likelihood estimator for the logistic regression model is given in Section 2. In Section 3, we state a review and describe some of the existing robust techniques. We explain the performance of the estimators based on the results of a simulation study and real data in Section 4. The conclusion is given in Section 5.

2. Maximum Likelihood of Logistic Regression

Suppose the binary response variable y_i takes values (0,1), these numerical values represented the negative response and the positive response respectively. The mean of this variable will be the proportion of positive responses. If p is the proportion of the observations with an outcome of 1, then $1 - p$ is the probability of an outcome of 0. The predictor variables $x_i \in R^p$, $i = 1, 2, \dots, n$, the probability of positive response variable $\mu_i = p(y_i = 1 | x_i)$ is linked to predictor variables by the mean of a link function $g(\mu) = X\beta$, such that $g^{-1}(X\beta)$ is the logit link function which transforms the covariate values in the interval (0,1).

We can write the multiple logistic regression model by:

$$p(Y = 1 | X = x_i) = F(x_i^T \beta) = \frac{\exp(x_i^T \beta)}{1 + \exp(x_i^T \beta)}, i = 1, \dots, n, \quad (1)$$

where $X = (1, x_1, \dots, x_p)$ are the values of the predictor variables and

$\beta^T = (\beta_0, \beta_1, \dots, \beta_p)$ is the vector of unknown parameters. The binary regression model can be defined by:

$$\eta_i = x_i^T \beta,$$

where η_i is a linear predictor and also known as transformation function,

where $\eta_i = \ln\left(\frac{\mu_i}{1-\mu_i}\right)$ this transformation is known as a logit link function.

There are another two transformation functions used in practice for modeling binomial and Bernoulli data:

- The probit function $g(\mu_i) = \eta_i = \phi^{-1}(\mu_i)$,
- The complementary log-log function $g(\mu_i) = \eta_i = \ln(-\ln(1-\mu_i))$.

In this article, we focus on a logit function as a link function. The classical maximum likelihood estimator is used to estimating the vector of unknown parameters β . Suppose that the response variables y_i distributed according to the Bernoulli distribution and the probability distribution for the i^{th} observation is given by:

$$f(y_i) = \mu_i^{y_i} (1 - \mu_i)^{1-y_i}, i = 1, 2, \dots, n,$$

and each observation y_i takes the value 1 with probability μ_i or the value 0 with probability $(1 - \mu_i)$. The likelihood function is given by:

$$l(\beta; y_1, y_2, \dots, y_n) = \prod_{i=1}^n f(y_i) = \prod_{i=1}^n \mu_i^{y_i} (1 - \mu_i)^{1-y_i}, \quad (2)$$

then, we take a log-likelihood of above formula:

$$l(\beta; y_1, y_2, \dots, y_n) = \ln \prod_{i=1}^n f(y_i) = \sum_{i=1}^n \left[y_i \ln\left(\frac{\mu_i}{1-\mu_i}\right) \right] + \sum_{i=1}^n \ln(1 - \mu_i),$$

where $\eta_i = \ln[\mu_i(1-\mu_i)] = x_i^T \beta$ and $1 - \mu_i = [1 + \exp(x_i^T \beta)]^{-1}$, then, the log-likelihood can be written as:

$$\begin{aligned} \ln l(\beta; Y) &= \sum_{i=1}^n \ln y_i x_i^T \beta - \sum_{i=1}^n \ln [1 + \exp(x_i^T \beta)] \\ &= \beta^T XY - \sum_{i=1}^n \ln [1 + \exp(x_i^T \beta)]. \end{aligned} \quad (3)$$

In design experiments, we have repeated observations or trials at each level of the explanatory variables (x). Let η_i be a number of the trials at each level of the predictor and y_i be the number of 1's observed at the i^{th} observations with $n = n_1, n_2, \dots, n_m$. Then, the log-likelihood is given by:

$$\ln l(\beta; Y) = \beta^T XY - \sum_{i=1}^m n_i \ln [1 + \exp(x_i^T \beta)], \quad (4)$$

however, the likelihood function can be maximized by differentiating it with respect to β :

$$\frac{\partial \ln l(\beta, Y)}{\partial \beta} = X^T Y - \sum_{i=1}^m \left[\frac{n_i}{1 + e^{x_i^T \beta}} \right] e^{x_i^T \beta} X_i,$$

where, $e^{x_i^T \beta} / (1 + e^{x_i^T \beta}) = 1 / (1 + e^{-x_i^T \beta}) = \mu_i$, we have

$$\partial \ln(\beta; Y) = X^T Y - \sum_i^m n_i \mu_i X_i,$$

where $n_i \mu_i$, represents the mean of the binomial variable, we can write above equation in matrix notation as $X^T(Y - \mu)$, where

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_m \end{bmatrix}.$$

As a result, the MLE estimator is typically done by solving the score equation:

$$X^T(Y - \mu) = 0. \quad (5)$$

Equation (5) is nonlinear in β one may use the iteratively weighted least squares (IWLS) algorithm. The method of iteratively weighted least squares is used to solve certain optimization problems, in logistic regression model the (IWLS) is used to find the maximum likelihood estimates with objective function of the form of:

$$\arg \min_{\beta} \sum_{i=1}^n |y_i - f_i(\beta)|^p, \quad (6)$$

by an iterative method in which each step involves solving a weighted least squares problems of the form:

$$\beta^{(k+1)} = \arg \min_{\beta} \sum_{i=1}^n W_i(\beta^k) |Y_i - f_i(\beta)|^2 = \arg \min_{\beta} \sum_{i=1}^n W_i(k) |Y_i - X_i \beta|^2,$$

where $W(\beta^k)$ is the diagonal matrix of weights, usually will all elements set initially to $W_i(0) = 1$. Let use rewrite β as a matrix form:

$$\left(X^T W(\beta^k) X \right)^{-1} X^T W(\beta^k) X \beta = \left(X^T W(\beta^k) X \right)^{-1} X^T W(\beta^k) \tilde{X},$$

where $W(\tilde{X}) = X \beta$ is the vector of linear predictor η_i , in the other hand the Newton method can be factorized as:

$$\begin{aligned} \beta^{(k+1)} &= \left(X^T W(\beta^k) X \right)^{-1} X^T W(\beta^k) \left(\tilde{X}(\beta^k) + \tilde{Y}(\beta^k) \right) \\ &= \left(X^T W(\beta^k) X \right)^{-1} X^T W(\beta^k) Z(\beta^k), \end{aligned}$$

with a new vector $Z(\beta) = \eta_i(\beta) + (Y_i - \mu_i(\beta)g(\mu_i(\beta)))_i$. That is $\beta^{(k+1)}$ is the solution of a weighted least square problem with weight $Z^{(k)}$, response vector $Z^{(k)}$ and explanatory variable $X^{(k)}$.

3. Robust Estimators in Logistic Regression

An outlier is an observation deviated from the other values in data and produces the large residuals. In the logistic regression model, an outlier can be occurred in the response variables as well as in the predictor variables or in both. In the binary regression model, all the response variables y_i are binary, takes the numerical values 0 or 1, therefore, an outlier in the response variable can only occur as a transposition $0 \rightarrow 1$ or $1 \rightarrow 0$ discussed by [14]. An error in response

variables is also well-known as a misclassification error or residual outlier. Extreme observation in explanatory variables is known as a leverage point or leverage outlier: there are two types of leverage point, good and bad. A good leverage point occurs when $y=0$ with a small value of $p(y=1|x_i)$ or when $y=1$ with a large value of $p(y=1|x_i)$, and vice versa for a bad leverage point. The classical maximum likelihood estimation can be influenced by leverage points and misclassification in the response variables, studied by [13] and [14]. To solve this problem, there are many robust estimators proposed for GLM, specifically, for the logistic and Poisson models. For instance, the Mallows-type technique of [2] and we can also cite works of ([2] [3] [7] [10] [13]).

In this article we proposed a new class of robust techniques for logistic regression, they are weighted maximum likelihood estimators, where the weight depends on the weight functions introduced by [17] as a weight of explanatory variables in Mallows-type estimator. In addition, we compare the performance of these techniques with classical maximum likelihood, Mallows-type estimator and unbiased bounded-influence estimator, in the presence of outliers.

3.1. Conditionally Unbiased Bounded-Influence Estimator (CUBIF)

In the CUBIF estimator, the weights of controlling atypical observations depend on the response variables and the predictor variables, this estimator is also known as the Scheppe class estimator introduced by [2]. The idea of this method is to minimize a measure of efficiency based on the asymptotic variance-covariance matrix to bound the measure of infinitesimal sensitivity. The M-estimators are the solution of the form of $\sum \psi(y_i, x_i, \beta) = 0$, such that $E((y, x, \beta) | x_i) = 0$, where y_i represents the i^{th} response variable, x_i represents the i^{th} explanatory variables, β is a vector of unknown parameters and ψ is a known $(p \times 1)$ functions that does not depend on i or n . We can write the optimal function of ψ as follows:

$$\psi(y, x, \beta, B) = W(\beta, y, x, b, B) \left[y - g(\beta^T x) - c \left(\beta^T x, \frac{b}{h(x, B)} \right) x \right], \quad (7)$$

where B is a variance covariance matrix, b is bounded on the measure of infinitesimal sensitivity and $h(x, B) = (x^T B^{-1} x)^{1/2}$ is a leverage measure. The function $c \left(\frac{B^T x, b}{h(x, B)} \right)$ is a bias correction term with corrected residual given by:

$$r(y, x, \beta, b, B) = y_i - g(\beta^T x) - c \left(\beta^T x, \frac{b}{h(x, B)} \right). \quad (8)$$

The weights function in the form of $W(\beta, y, x, B) = W_b r((y, x, \beta) h(x, B))$, where W_b represent the Huber weights function given by $W_b = \min \left\{ 1, \frac{b}{|x|} \right\}$. The weight function W downweights observations with a high leverage point

and large corrected residual making M-estimator to have bonded influence.

3.2. Mallows Type Class (Mallows)

[2] proposed Mallows-type leverage dependent weight estimator, this estimator minimizes the weighted log-likelihood function, where the weight depends on the explanatory variables. [3] discussed more deeply on Mallows-type estimator and suggested a simple way to make the maximum likelihood estimator more robust by downweighting the atypical observation in the predictor variables. The leverage of observation x can be measured by the following:

$$h_n(x) = \left((x - \hat{\mu}_n)' \hat{\Sigma}_n^{-1} (x - \hat{\mu}_n) \right)^{1/2}, \quad (9)$$

where $\hat{\mu}_n$ represents a robust location estimator, and $\hat{\Sigma}_n$ represents a robust variance-covariance matrix of the continuous covariates (x_1, x_2, \dots, x_n) . The initial robust scale and location estimator of continuous $\hat{\Sigma}$ and $\hat{\mu}$, can be calculated by using minimum covariance determinant (MCD) approach. The Mallows type estimator for logistic regression can be obtained by a solution of the form of:

$$\sum_{i=1}^n w_i \left[y_i \log(\Pi_i(\beta)) + (1 - y_i) \log(1 - \Pi_i(\beta)) \right], \quad (10)$$

where $w_i = W(h_n(x_i))$, W is a non-increasing function such that $W(u)$ is bounded. [3] suggested choosing W depends on a constant $c > 0$.

$$W(u) = \left(1 - \frac{u^2}{c_2} \right)^3 I(|u| \leq c),$$

this estimate is called the weighted maximum likelihood estimate (Mallows-type estimator) and the influence function of WMLE is given by:

$$If(y, x, \beta) = M^{-1}(y - f(\beta'x))xW(h(x)), \quad (11)$$

where $h(x) = \left((x - \mu)' \Sigma^{-1} (x - \mu) \right)^{1/2}$ with μ and Σ are the limit values of $\hat{\mu}$ and $\hat{\Sigma}$ and

$$M = E(W(x))F^{-1}(\beta x)xx'.$$

3.3. Weighted Maximum Likelihood Estimator (WMLE)

Similar to the strategy used in constructing the Mallows-type estimator, we proposed a new class of robust techniques, they are the weighted maximum likelihood estimators, with weight depends on the weight functions introduced by [17]. First, compute the initial location and scatter estimators of the explanatory variables $\hat{\mu}^{(0)}$ and $\hat{\Sigma}^{(0)}$ respectively. Then, calculate the squared Mahalanobis distances of the explanatory variables which can be defined as:

$$m^2 = \left(x_i - \hat{\mu}^{(0)} \right)' \left(\hat{\Sigma}^{(0)} \right)^{-1} \left(x_i - \hat{\mu}^{(0)} \right).$$

The weight function we proposed can be defined as: first weight:

$w_1 = (0.8 * m^2 + 0.2)$, where m^2 refers to squares Mahalanobis distances, then:

$$w_1 = \left(0.8 * (x_i - \hat{\mu}^{(0)})^T (\hat{\Sigma}^{(0)})^{-1} (x_i - \hat{\mu}^{(0)}) + 0.2 \right),$$

second weight: $w_2 = (0.8 * (m^2)^2 + 0.2)$, then, we can write in the form of:

$$w_2 = \left(0.8 * \left((x_i - \hat{\mu}^{(0)})^T (\hat{\Sigma}^{(0)})^{-1} (x_i - \hat{\mu}^{(0)}) \right)^2 + 0.2 \right).$$

Then, the weighted maximum likelihood estimators for logistic regression can be obtained by a solution of the form of:

$$\sum_{i=1}^n w_i \left[y_i \log(\Pi_i(\beta)) + (1 - y_i) \log(1 - \Pi_i(\beta)) \right]. \quad (12)$$

For these weights no observation is trimmed we used the modified algorithm for Mallows-type estimator of [3] for computation of the weighted maximum likelihood estimates.

4. Evaluation of the Robust Estimators

In order to examine the performance of the estimators, two approaches have been taken. The first includes simulated models for comparing the new techniques with the classical MLE, Mallows type estimator for [2] and [3]. In the second, we used real data sets of leukemia data.

4.1. Simulation Study

In this subsection, a simulation study was carried out to examine the performance of new robust techniques (WMLE_{w1}, WMLE_{w2}) and compare with MLE, conditionally unbiased bounded influence (CUBI) of [2] and the Mallows-type estimator (Mallows) of [3]. The weighted maximum likelihood estimator was computed using the modified algorithm for Mallows type estimator. The Mallows and CUBI were computed by the standard available in the robust package of R. The simulation study involves four models, these are an uncontaminated model (model 1), 5% of the data are contaminated (model 2), 10% moderate contaminated (model 3) and 20% extreme contaminated model (model 4). In the first scenario without contamination, we generated two predictor variables according to the standard normal distribution with mean zero and variance one, $x_1 \sim N(0,1)$ and $x_2 \sim N(0,1)$, with four sample sizes, $n = (100, 200, 300, 400)$. The large sample size was selected to guarantee the existence of the overlapping in each replication. The response variable is generated from the Bernoulli distribution with parameter equal to

$$\pi_i = \frac{\exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2)}{1 + \exp(\beta_0 + \beta_1 x_1 + \beta_2 x_2)}. \text{ The true parameters for the clean model setting}$$

as $\beta = (\beta_0, \beta_1, \beta_2)^T = (0.66, 0.24, 0.14)$. In the second scenario, 5% of the data are contaminated with the amount of deviating atypical observation in the

x -direction is taken as $\delta = 5$, the vector of true parameter for contaminated observations equals $\beta = (0.5, 1, 1)$. The predictor variables for contaminated models were generated according to normal distribution, $z_1 \sim N(0, 1)$ and $z_2 \sim N(0, 1)$. The third and four models are in the same way as the second scenario with a percentage of contamination equals 10% and 20% respectively and a quantity of deviating atypical observation in the x -direction equals (5). The new values of the predictor variables are denoted by x_1^* and x_2^* , where the response variable y_i^* for the contaminated model are generated from the following model equations:

$$y_i^* = \begin{cases} 0 & \text{if } \beta_o + \beta_1 x_1^* + \beta_2 x_2^* \geq 0. \\ 1 & \text{if } \beta_o + \beta_1 x_1^* + \beta_2 x_2^* < 0. \end{cases} \quad (13)$$

The performance of these estimators is examined based on the Bias and mean squared error (MSE) for different scenarios. However, the estimator which has small Bias and MSE is a good one. In each scenario run included over 1000 repetitions. Therefore, the bias and mean squared error for each parameter are computed as follows:

$$\text{Biss} = \left| \frac{1}{1000} \sum_{i=1}^{1000} \beta_i - \beta \right|,$$

and

$$\text{MSE} = \left[\frac{1}{1000} \sum_{i=1}^n |\hat{\beta}_i - \beta|^2 \right].$$

4.2. Results from the Monte Carlo Simulation Study

Table 1 reports bias and mean squared errors of the five estimators for the contaminated data. The results indicate that the bias and MSE of MLE, Mallows, CUBIF estimators are fairly close to each other, both WMLE_{w1} and WMLE_{w2} estimators perform less compared to other estimators. It can observe that the bias and mean squared errors decrease when the sample size is increased. As can be seen from **Table 2** under 5% of the data was contaminated, the two new robust techniques WMLE_{w1} and WMLE_{w2} have overall the best performance among all compared estimators for different sample sizes.

The results of moderate in **Table 3** with 10% of the data are contaminated and extreme bad leverage point in **Table 4** with 20% of the data are contaminated demonstrated that our weighted maximum likelihood estimators WMLE_{w1} and WMLE_{w2} perform better than other estimators in the term of bias and mean squared errors. However, the classical maximum likelihood estimates perform poorly in the contaminated model due to the sensitivity of outliers. In summary, the two new estimators show the best performance among all compared techniques in contaminated data. Moreover, these new estimators have reasonable perform in clean data.

Table 1. Bias and mean squared errors of estimators for model 1.

Methods	$n = 100$		$n = 200$		$n = 300$		$n = 400$	
	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
MLE	0.1872	0.0881	0.1269	0.0413	0.1018	0.0264	0.0881	0.0195
MALLOWS	0.1839	0.0873	0.1269	0.0410	0.0967	0.0247	0.0861	0.0194
CUBIF	0.1853	0.0884	0.1290	0.0422	0.0992	0.0254	0.0855	0.0190
WMLE _{w₁}	0.3442	0.2947	0.3122	0.1302	0.0987	0.0858	0.2278	0.0742
WMLE _{w₂}	0.3240	0.2991	0.3318	0.1471	0.2018	0.1723	0.2464	0.0961

Table 2. Bias and mean squared errors of estimators for model 2.

Methods	$n = 100$		$n = 200$		$n = 300$		$n = 400$	
	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
MLE	0.8736	1.5627	0.8203	1.2564	0.7805	1.1229	0.7808	1.0993
MALLOWS	0.8742	1.5535	0.7988	1.2115	0.7835	1.1340	0.7947	1.1442
CUBIF	0.8799	1.5888	0.7811	1.1654	0.7754	1.1215	0.7728	1.0486
WMLE _{w₁}	0.3322	0.2800	0.0691	0.1492	0.5345	0.4209	0.4261	0.2946
WMLE _{w₂}	0.3288	0.2424	0.0165	0.1578	0.4666	0.3274	0.4301	0.3053

Table 3. Bias and mean squared errors of estimators for model 3.

Methods	$n = 100$		$n = 200$		$n = 300$		$n = 400$	
	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
MLE	0.9062	1.6799	0.8301	1.3019	0.7858	1.1394	0.7921	1.1371
MALLOWS	0.9077	1.7056	0.7981	1.2197	0.8041	1.1953	0.7969	.1502
CUBIF	0.9033	1.6877	0.8080	1.2462	0.7867	1.1581	0.7841	1.1219
WMLE _{w₁}	0.4358	0.3190	0.0604	0.1368	0.4961	0.3526	0.4606	0.3089
WMLE _{w₂}	0.4024	0.2718	0.0047	0.1350	0.4410	0.2812	0.4373	0.2805

Table 4. Bias and mean squared errors of estimators for model 4.

Methods	$n = 100$		$n = 200$		$n = 300$		$n = 400$	
	Bias	MSE	Bias	MSE	Bias	MSE	Bias	MSE
MLE	0.9997	2.1176	0.8496	1.4131	0.8194	1.2636	0.8145	1.2167
MALLOWS	0.9746	2.0663	0.8571	1.4281	0.8328	1.3034	1.8308	1.2621
CUBIF	1.0167	2.1856	0.8710	1.8890	0.8349	1.3201	0.8134	1.2106
WMLE _{w₁}	0.2511	0.1654	0.0352	0.1548	0.2393	0.1224	0.2494	0.1150
WMLE _{w₂}	0.2350	0.1556	0.0729	0.1510	0.1897	0.0925	0.2404	0.1013

4.3. Leukemia Data

The datasets analyzed here. This data includes 33 leukemia patients. Three variables were measured for each patient: Time, AG and WBC. The response vari-

able is a survival time patient in weeks, we coded into (1 = the patient survived more than 52 weeks, 0 = otherwise). The two explanatory variables are: WBC measured a white blood cell count of patient and AG is a binary variable (1 = present of morphologic characteristic of white blood cells, 0 = absent of morphologic characteristic of white blood cells) according to an identification method of atypical observation in the leukemia data, the observation number 17 looks like atypical. A logistic regression model was fitted using binary survival time y as the response variable and AG and WBC as the predictor variables. The estimators examined here are new weighted maximum likelihood estimates (WMLE_{w1}, WMLE_{w2}), MLE, MLE₁₇ (MLE₁₇ is the maximum likelihood estimator for clean data after excluding observation number 17), Mallows (Mallows type estimator) and CUBIF (conditionally unbiased bounded-influence function estimator).

It can be observed from **Table 5**, the MLE is very sensitively to influential observations. In addition, after deleting observation number 17 reduced the effect of WBC close to zero. The new WMLE estimators (WMLE_{w1}, WMLE_{w2}) are showed the best performance among all other estimators for the leukemia data. However, Mallows estimates are sensibly close to the MLE₁₇.

5. Conclusion

In this study, we introduced two new robust techniques of logistic regression, also known as weighted maximum likelihood estimators. In order to examine the performance of new techniques, we conducted simulation experiments under different scenarios and real datasets. The classical maximum likelihood estimates show the lack of robustness when outliers are present. Our simulation experiments for uncontaminated models demonstrated that the MLE, Mallows and CUBIF estimators are fairly perform close to each other, while, the new weighted techniques perform less compared to other estimators. In both simulation study under different contaminated scenarios and real datasets, the new proposed weighted maximum likelihood techniques showed the best performance among all compared estimators. The new techniques used here to construct robust estimators can also be extension to other generalized linear models like Poisson regression model and negative binomial model.

Table 5. The estimated parameters and standard errors for the leukemia data.

Methods	Intercept		WBC		AG	
	Est.	S.E.	Est.	S.E.	Est.	S.E.
MLE	-1.3073	0.8145	0.3717	0.0186	2.2617	0.9522
MLE32	0.2119	1.0830	-0.0002	0.0001	2.5580	1.2341
MALLOWS	0.1602	1.0697	-0.2245	0.0129	2.5252	1.2159
CUBIF	-1.4503	1.8504	-0.0527	0.0337	0.2202	0.9756
WMLE _{w1}	-0.0011	0.3992	0.0012	0.0070	1.4744	0.4769
WMLE _{w2}	-1.5486	0.4588	-0.0064	0.0066	1.3786	0.5481

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] Pregibon, D. (1981) Logistic Regression Diagnostics. *The Annals of Statistics*, **4**, 705-724. <https://doi.org/10.1214/aos/1176345513>
- [2] Kunsch, H.R., Stefonski, L.A. and Carroll, R.J. (1989) Conditionally Unbiased Bounded-Influence Estimation in General Regression Models, with Applications to Generalized Linear Models. *Journal of the American Statistical Association*, **84**, 460-466. <https://doi.org/10.1080/01621459.1989.10478791>
- [3] Carroll, R.J. and Pederson, S. (1993) On Robustness in the Logistic Regression Model. *Journal of the Royal Statistical Society. Series B (Methodological)*, **55**, 693-706. <https://doi.org/10.1111/j.2517-6161.1993.tb01934.x>
- [4] Hobza, T., Martin, N. and Pardo, L. (2017) A Wald-Type Test Statistic Based on Robust Modified Median Estimator in Logistic Regression Models. *Journal of Statistical Computation and Simulation*, **87**, 2309-2333. <https://doi.org/10.1080/00949655.2017.1330414>
- [5] Bergesio, A. and Yohai, V.J. (2011) Projection Estimators for Generalized Linear Models. *Journal of the American Statistical Association*, **106**, 661-671. <https://doi.org/10.1198/jasa.2011.tm09774>
- [6] Morgenthaler, S. (1992) Least-Absolute-Deviations Fits for Generalized Linear Models. *Biometrika*, **79**, 747-754. <https://doi.org/10.1093/biomet/79.4.747>
- [7] Cantoni, E. and Ronchetti, E. (2001) Robust Inference for Generalized Linear Models. *Journal of the American Statistical Association*, **96**, 1022-1030. <https://doi.org/10.1198/016214501753209004>
- [8] Wedderburn, R.W.M. (1974) Quasi-Likelihood Functions, Generalized Linear Models, and the Gauss Newton Method. *Biometrika*, **61**, 439-447. <https://doi.org/10.1093/biomet/61.3.439>
- [9] Croux, C., Flandre, C. and Haesbroeck, G. (2002) The Breakdown Behavior of the Maximum Likelihood Estimator in the Logistic Regression Model. *Statistics and Probability Letters*, **60**, 377-386. [https://doi.org/10.1016/S0167-7152\(02\)00292-4](https://doi.org/10.1016/S0167-7152(02)00292-4)
- [10] Bianco, A.M. and Yohai, V.J. (1996) Robust Estimation in the Logistic Regression Model. In: Reider, H., Ed., *Robust Statistics, Data Analysis, and Computer Intensive Methods*, Wiley, New York, 17-34. https://doi.org/10.1007/978-1-4612-2380-1_2
- [11] Croux, C. and Haesbroeck, G. (2003) Implementing the Bianco and Yohai Estimator for Logistic Regression. *Computational Statistics and Data Analysis Journal*, **44**, 273-295. [https://doi.org/10.1016/S0167-9473\(03\)00042-2](https://doi.org/10.1016/S0167-9473(03)00042-2)
- [12] Muller, C.H. and Neykov, N. (1981) Breakdown Points of Trimmed Likelihood Estimators and Related Estimators in Generalized Linear Models. *Journal of Statistical Planning and Inference*, **116**, 503-519. [https://doi.org/10.1016/S0378-3758\(02\)00265-3](https://doi.org/10.1016/S0378-3758(02)00265-3)
- [13] Pregibon, D. (1982) Resistant Fits for Some Commonly Used Logistic Models with Medical Applications. *Biometrics*, **38**, 485-498. <https://doi.org/10.2307/2530463>
- [14] Copas, J.B. (1988) Binary Regression Models for Contaminated Data. *Journal of the Royal Statistical Society. Series B (Methodological)*, **50**, 225-265. <https://doi.org/10.1111/j.2517-6161.1988.tb01723.x>
- [15] Stefanski, L.A., Carroll, R.J. and Ruppert, D. (1986) Optimally Bounded Score

Functions for Generalized Linear Models with Applications to Logistic Regression. *Biometrika*, **73**, 413-424. <https://doi.org/10.2307/2336218>

- [16] Krasker, W.S. and Welsch, R.E. (1982) Efficient Bounded-Influence Regression Estimation. *Journal of the American Statistical Association*, **77**, 595-604. <https://doi.org/10.1080/01621459.1982.10477855>
- [17] Simeckova, M. (2005) Maximum Weighted Likelihood Estimator in Logistic Regression. WDS05 Proceedings of Contributed Papers, 144-148.