ESTIMATION OF THE BINARY LOGISTIC REGRESSION MODEL PARAMETER USING BOOTSTRAP RE-SAMPLING

R. LI, J. ZHOU and L. WANG

College of Data Science and Information Engineering, Guizhou Minzu University, Guizhou, Guiyang, 550025, China;
1904492559@qq.com

Abstract— In this paper, the non-parametric bootstrap and non-parametric Bayesian bootstrap methods are applied for parameter estimation in the binary logistic regression model. A real data study and a simulation study are conducted to compare the Non-parametric bootstrap, Non-parametric Bayesian bootstrap and the maximum likelihood methods. Study results shows that three methods are all effective ways for parameter estimation in the binary logistic regression model. In small sample case, the non-parametric Bayesian bootstrap method performs relatively better than the non-parametric bootstrap and the maximum likelihood method for parameter estimation in the binary logistic regression model.

Keywords— Non-parametric bootstrap; Non-parametric Bayesian bootstrap; Logistic Regression; Confidence Interval; Parameter Estimation.

I. INTRODUCTION

Logistic regression is a widely used method in modeling categorical dependent variable. Suppose dependent variable y follows a Bernoulli distribution and y_i denotes the i^{th} value of y, the binary logistic regression model can be described as follows:

$$y_i \sim Be(\pi_i), \quad i = 1,2,\ldots n \tag{1}$$

$$\pi_i = \frac{\exp(\beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip})}{1 + \exp(\beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip})} \quad i = 1,2,\ldots n \tag{2}$$

where $(1, x_{i1}, \ldots, x_{ip})$ is the i^{th} row of the $n \times (p+1)$ design matrix X with p independent variables, β_j, $j = 0, \ldots, p$ are unknown parameters. Through logit transformation, expression (2) can also be written in linear form as

$$logit(\pi_i) = \ln\frac{\pi_i}{1-\pi_i} = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip} \quad i = 1,2,\ldots n \tag{3}$$

For logistic regression model, an important problem is how to estimate the unknown parameters β_j. The most common method of estimating β_j is the maximum likelihood (ML) method. To combat the multicollinearity problem, Schaefer (1984) proposed a ridge estimator in logistic regression model and studied its performance under the mean squared error criterion. Hossain (2004) used Non-parametric Bootstrap Method to estimate parameters in multiple logistic regression models. Ariffin and Midi (2012) introduced two robust bootstrap methods for parameter estimation for logistic regression models, namely the diagnostic logistic before bootstrap and the weighted logistic bootstrap with probability. Rubin (1981) introduced the Bayesian bootstrap method according to the bootstrap method proposed by Efron (1979). Månsson and Shukur (2011) proposed a new ridge estimator in logistic regression model to combat multicollinearity. Adjei and Karim (2016) applied the parameter bootstrap, the non-parameter bootstrap (NB) and the ML methods in logistic regression model for parameter estimation and gave a detailed numerical study (Salgado *et al.*, 2017).

Motivated by Adjei and Karim (2016), we will apply the non-parametric Bayesian bootstrap (NBB) and Non-parametric bootstrap methods to estimate coefficient parameters in logistic regression model. The rest of the paper is organized as follows. In Section II, NB and NBB methods in logistic regression model are given in detail. In Section III and Section IV, a numerical study and a Monte Carlo simulation study is provided respectively. Finally, main results are given in Section V (Khoshbin *et al.*, 2017).

II. METHODOLOGY

Bootstrap re-sampling is a method of supplementing data by itself. Parametric bootstrap is used when the population distribution is known. On the other hand, semi-parametric bootstrap is usually applied when only some information on the population distribution is known. While non-parametric bootstrap re-sampling is used when the population distribution is unknown.

A. Non-parametric Bootstrap

The advantages of non-parametric bootstrap re-sampling have been clearly showed for data with unknown distribution. Efron (1979) states that the bootstrap is able to provide trustworthy answers despite of unfavorable circumstances. In the simplest setting a random sample is available and the non-parametric estimate is the empirical distribution function, while a parametric model with a parameter of fixed dimension is replaced by its maximum likelihood estimate (Davison *et al.*, 2003).

The algorithm of NB in the logistic regression model is given below:

Step 1: Make a new data set for binary response with covariate(s) (x, y) from group data.

Step 2: Draw bootstrap sample by sampling the pairs with replacements from new the data set $(x,y)_b^* = ((x_1,y_1)^*,\dots,(x_n,y_n)^*)$ for $(b = 1,2,\dots,B)$.

Step 3: For each $b = 1,2,\dots,B$ estimate the bootstrap sample statistics $\hat{\beta}_1^*,\dots,\hat{\beta}_B^*$ where $\hat{\beta}_B^* = t((x,y)_b^*) = (\hat{\beta}_{b0}^*,\hat{\beta}_{b1}^*)$ by refitting model (3).

Step 4: The distribution of $\hat{\beta}_b^*$ around the original estimate β is similar to the sampling distribution of the β around the population parameter θ. Here, we assume that β is normally distribution, which is often approximately the case for statistics in sufficiently large samples.

For convenience of following discussion, we denote the estimated values of the mean, bias and standard error of the non-parametric bootstrap estimator $\hat{\beta}_{NB}^*$ as

$$\bar{\hat{\beta}}_{NB}^* = \frac{1}{B}\sum_{b=1}^{B}\hat{\beta}_b^* \quad (4)$$

$$bias(\hat{\beta}_{NB}^*) = E(\hat{\beta}_{NB}^*) - \beta = \bar{\hat{\beta}}_{NB}^* - \beta \quad (5)$$

$$SE(\hat{\beta}_{NB}^*) = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}\left(\hat{\beta}_b^* - \bar{\hat{\beta}}_{NB}^*\right)^2} \quad (6)$$

Based on the non-parametric bootstrap estimator $\hat{\beta}_{NB}^*$, we can three types of the confidence interval on the model parameters.

Assuming the distribution of the bootstrapped statistic is approximately normal and symmetric, we get the normal interval as

$$CI = \left(\beta - bias(\hat{\beta}_{NB}^*)\right) \pm z_{1-\alpha/2}SE(\hat{\beta}_{NB}^*) \quad (7)$$

where $z_{1-\alpha/2}$ is the $1-\alpha/2$ quantile of the standard normal distribution.

In this case, the sample estimator is an unbiased estimator of the population estimator (Banjanovic *et al.,* 2016).

Assuming the distribution of the bootstrapped statistic is approximately symmetric and approximately normal, then the percentile interval can be given as

$$\hat{\beta}_{NB(lower)}^* < CI < \hat{\beta}_{NB(upper)}^* \quad (8)$$

where $lower = [(B+1)\alpha/2]$, $upper = [(B+1)(1-\alpha/2)]$, and the square brackets indicate rounding to the nearest integer.

If there are no assumptions, we can get the confidence interval on the model parameters called as BCa interval

$$\hat{\beta}_{NB(lower^*)}^* < CI < \hat{\beta}_{NB(upper^*)}^* \quad (9)$$

where $lower^* = [Ba_1]$, $upper^* = [Ba_2]$ and

$$a_1 = \Phi\left[z + \frac{z - z_{1-\alpha/2}}{1 - \hat{a}(z - z_{1-\alpha/2})}\right] \quad (10)$$

$$a_2 = \Phi\left[z + \frac{z + z_{1-\alpha/2}}{1 - \hat{a}(z + z_{1-\alpha/2})}\right] \quad (11)$$

$$\hat{a} = \frac{\sum_{i=1}^{n}\left(\bar{\beta} - \beta_{(-i)}\right)^3}{6\left[\sum_{i=1}^{n}\left(\bar{\beta} - \beta_{(-i)}\right)^2\right]^{3/2}} \quad (12)$$

$$z = \Phi^{-1}\left[\frac{\#_{b=1}^{B}(\hat{\beta}_b^* < \beta)}{B+1}\right] \quad (13)$$

Use the same notation as Fox and Weisberg (2012), $\beta_{(-i)}$ represents the value of β produced when the i^{th} observation is deleted from the sample,

$$\bar{\beta} = \sum_{i=1}^{n}\beta_{(-i)}/n.$$

B. Non-parametric Bayesian Bootstrap

Rubin (1981) asserts that the Bayesian Bootstrap is a natural bayesian analogue of the bootstrap, and shows that operationally they are similar. The Bayesian Bootstrap Method is to reduce the sample data repetition rate on the original sample, and thus to make data improvement methods for the samples generated by the bootstrap method. In fact, the Bayesian Bootstrap Method is a weighted average, the weights are derived from the random numbers generated by the Dirichlet Distribution.

The algorithm for NBB in the Logistic Regression Model is:

Step 1: Repeat step1~step3 of NB algorithm to get parameter estimate value $\hat{\beta}_{NB}^* = (\hat{\beta}_1^*,\hat{\beta}_2^*,\dots,\hat{\beta}_B^*)$.

Step 2: Generate random variables $u_1,u_2,\dots,u_{n-1}$ from (0,1) uniformly distributed, ordering them from small to large $u_{(1)},u_{(2)},\dots,u_{(n-1)}$, and calculating the gaps between two adjacent numbers $g_i = u_{(i)} - u_{(i-1)}$, $i = 1,2,\dots,n-1$, let $u_{(0)} = 0$, $u_{(n)} = 1$. Where g_i obeys the Dirichlet distribution and satisfies the formula $\sum_{i=1}^{n} g_i = 1$.

Step 3: According to the value of Step 1 and Step 2, we can get the required one non-parametric Bayesian Bootstrap Statistics for $\hat{\beta}^* = \sum_{b=1}^{B} g_b\,\hat{\beta}_b^*$.

Step 4: Repeat step 2-3 B times, for each $b = 1,2,\cdots,B$ estimate the non-parametric Bayesian Bootstrap Statistics $\hat{\beta}_1^*,\hat{\beta}_2^*,\cdots,\hat{\beta}_B^*$, that is $\hat{\beta}_{NBB}^* = (\hat{\beta}_1^*,\hat{\beta}_2^*,\cdots,\hat{\beta}_B^*)$.

Also, the confidence interval on the model parameters based on Non-parametric Bayesian Bootstrap Method is suggested as

$$CI = \bar{\hat{\beta}}_{NBB}^* \pm z_{1-\alpha/2}SE(\hat{\beta}_{NBB}^*) \quad (14)$$

where

$$\bar{\hat{\beta}}_{NBB}^* = \frac{1}{B}\sum_{b=1}^{B}\hat{\beta}_b^* \quad (15)$$

Table 1. Estimated values of parameters based on ML, NB and NBB methods.

Method	$\hat{\beta}_0$		$\hat{\beta}_1$		$\exp(\hat{\beta}_1)$
	Value	s.e	Value	s.e	
ML	-1.42829	0.22584	0.08454	0.00788	1.08822
NB	-1.42829	0.22144	0.08454	0.00814	1.08822
NBB	-1.43983	0.21905	0.08524	0.00791	1.08898

$$SE(\hat{\beta}_{NBB}^*) = \sqrt{\frac{1}{B-1}\sum_{b=1}^{B}\left(\hat{\beta}_b^* - \bar{\hat{\beta}}_{NBB}^*\right)^2} \quad (16)$$

III.A REAL DATA STUDY

In this section, we will apply the NB, NBB and ML methods to the serological data of Hepatitis A from Bulgaria once used by Adjei and Karim (2016). The data includes an independent variable, that is, the information about the age group of one year on the subject, and the response variable is a set of binary variables of serum properties. The number of seropositive is the number of infected by hepatitis A, the covariate is age and $\pi_i = \pi(\text{age})$ is the proportion of seropositive. According to section 1, the binary logistic regression model can be written as

$$logit[\pi_i] = \beta_0 + \beta_1 age_i, i = 1,2,\cdots,n \quad (17)$$

where β_1 is the effect of age on the log odds of infection, $n = 83$.

In this section and next section, the significance level $\alpha$is set to 0.05 and all computations are accomplished by R3.5.1.

A. Estimated values of parameters

From Table 1, we can see that the estimated values of parameters based on ML, NB and NBB methods are very close. The standard error of the estimator derived from the NBB method is relatively smaller than two other estimators. This means the NBB method performs relatively well under the standard error criterion.

It is known that the odds ratio equals $\exp(\hat{\beta}_1)$ in the binary logistic regression model. It can be seen that the odds ratio of the NBB method is the largest, indicating that the number of infected hepatitis A is 1.08898 times compared to the uninfected population as the age increases each year. The odds ratio of the ML and NB methods are almost identical and both are slightly smaller than the odds ratio of the NBB methods. This implies the NBB method is slightly well than the ML and NB methods in odds ratio mean.

To compare the reliability of the ML, NB and NBB methods, we choose the 95% confidence interval (CI) based on ML as a reference interval and calculate the frequencies of the estimated values of NB and NBB methods in the given interval.

As shown in Table 3, the proportion of the NBB method in the confidence interval of MLE is slightly bigger than the proportion of the MLE and NB methods, and the MLE and NB are approximately equal, indicating that the NBB method has the slightly higher reliability among the three methods.

Table 2. Proportion of the estimated values in the 95% confidence interval based on ML.

Method	$\hat{\beta}_0$		$\hat{\beta}_1$	
	95%CI based on ML	Proportion	95% CI based on ML	Proportion
NB	(-1.8830, -0.9958)	95.58%	(0.0699,0.1009)	94.36%
NBB	(-1.8830, -0.9958)	95.56%	(0.0699,0.1009)	95.02%
ML	(-1.8830, -0.9958)	95.00%	(0.0699,0.1009)	95.00%

Table 3. Confidence intervals and confidence interval length by ML, NB and NBB.

Method		$\hat{\beta}_0$		$\hat{\beta}_1$	
		95%CI	CI length	95%CI	CI length
MLE		(-1.8830, -0.9958)	0.887	(0.0699,0.1009)	0.0310
	Normal	(-1.8450, -0.9770)	0.868	(0.0676,0.0995)	0.0319
NB	Percentile	(-1.9050, -1.0350)	0.870	(0.0709,0.1030)	0.0321
	BCa	(-1.8610, -1.0040)	0.857	(0.0688,0.1003)	0.0315
NBB		(-1.8828, -1.0288)	0.854	(0.0702,0.1013)	0.0311

B. Confidence intervals of parameters

Next, we give the confidence intervals derived from the MLE, NB (normal, percentile, BCa) and NBB in Table 3. The accuracy of the three methods is judged by the length of the confidence interval, From Table 3, we can see that length of the confidence interval of NBB is almost the same as the length of the confidence interval of MLE and NB. For the NB method, normal and percentile are reasonably similar to each other, and with respect to normal and percentile, the length of the confidence interval of the BCa method is the shortest, so the results of the BCa method appear to be more precise.

IV.MONTE CARLO SIMULATION

To further compare the performance of the ML, NB and NBB methods, we are to perform a Monte Carlo simulation study.

A. The design of the experiment

The dependent variable of the logistic regression is generated using pseudo-random numbers from the $Be(\pi_i)$ distribution and one explanatory variable are generated as$x_i \sim N(40,24), i = 1, \cdots, n$. $\pi_i$is calculated as follows:

$$\pi_i = \frac{e^{\beta_0+\beta_1 x_i}}{1 + e^{\beta_0+\beta_1 x_i}} \tag{18}$$

To compare the results with Section 3,$\beta_0$and β_1 is set to -1.43 and 0.08 respectively which are derive from the real data.

Four different sample sizes for , such as 60, 80, 100, and 150 are considered. The number of re-sampling times of NB and NBB is 5000 times. All simulation results are summarized in Tables 4-6.

B. Estimated values of parameters

It can be seen from Table 4 that the estimated values based on ML, NB and NBB are very close in all cases. When $n = 80$, the estimated values are the closest to the true value. This is reasonable for the true values of $\beta_0$and β_1 are derived from the real data with sample size 83.

On the whole, there is no significant difference on the standard error of the ML, NB, and NBB three methods. In detail, it can be seen from Table 4 that bold digit is the smallest in each row. When the sample size is not big, the standard error of the NBB method is the smallest, the NB method is the second smallest, and the ML method is the biggest. This means that the NBB method performs slightly well than the NB method and the ML method when the sample size is not big.

In the same way used in Section 3, we choose the 95% confidence interval (CI) based on ML as a reference interval and calculate the frequencies of the estimated values of NB and NBB methods in the given interval to compare the reliability of the ML, NB and NBB methods.

It can be seen from Table 5 that when $n = 60,80$, the estimated values of the NBB method in the given confidence interval bigger than that of the NB method and the ML method. From the above analysis, it can be concluded that when the sample size is small, the NBB method is the most reliable among the three methods.

C. Confidence intervals of parameters

Finally, we discuss the accuracy of the ML, NB and NBB methods in the confidence interval length mean. From Table 6, we can draw conclusions that in the confidence interval estimation of the logistic regression model, when the sample size is small, the NBB method has a more accurate confidence interval, followed by the NB method, and finally the MLE method. When the sample size is large, the accuracy of the three methods tends to be similar.

Table 4. Estimated values of parameters based on ML, NB and NBB methods.

		Value			s.e		
		ML	NB	NBB	ML	NB	NBB
n=60	$\hat{\beta}_0$	-0.95355	-0.95355	-1.02052	0.67857	0.62564	**0.58936**
	$\hat{\beta}_1$	0.05568	0.05568	0.05923	0.01991	0.01766	**0.01603**
	$\exp(\hat{\beta}_1)$	1.05726	1.05726	1.06102	1.02011	1.01782	**1.01616**
n=80	$\hat{\beta}_0$	-1.29650	-1.29650	-1.37875	0.56235	0.55571	**0.51739**
	$\hat{\beta}_1$	0.08233	0.08233	0.08660	0.02039	0.01867	**0.01709**
	$\exp(\hat{\beta}_1)$	1.08581	1.08581	1.09046	1.02060	1.01885	**1.01724**
n=100	$\hat{\beta}_0$	-0.92258	-0.92258	-0.97127	0.46296	0.48759	**0.46263**
	$\hat{\beta}_1$	0.06647	0.06647	0.06895	0.01495	0.01483	**0.01380**
	$\exp(\hat{\beta}_1)$	1.06873	1.06873	1.07138	1.01506	1.01494	**1.01390**
n=150	$\hat{\beta}_0$	-1.45467	-1.45467	-1.5231	0.45318	0.43892	**0.42951**
	$\hat{\beta}_1$	0.08106	0.08106	0.08448	**0.01410**	0.01589	0.01538
	$\exp(\hat{\beta}_1)$	1.08444	1.08444	1.08820	**1.01420**	1.01602	1.01550

Table 5. Proportion of the estimated values in the 95% confidence interval based on ML.

Method		$\hat{\beta}_0$		$\hat{\beta}_1$	
		95%CI based on ML	Proportion	95%CI based on ML	Proportion
n=60	NB	(-2.4096,0.2877)	96.70%	(0.0217,0.1009)	97.18%
	NBB	(-2.4096,0.2877)	97.20%	(0.0217,0.1009)	98.38%
	MLE	(-2.4096,0.2877)	95.00%	(0.0217,0.1009)	95.00%
n=80	NB	(-2.4995, -0.2634)	95.68%	(0.0469,0.1279)	96.84%
	NBB	(-2.4995, -0.2634)	96.60%	(0.0469,0.1279)	98.02%
	MLE	(-2.4995, -0.2634)	95.00%	(0.0469,0.1279)	95.00%
n=100	NB	(-1.8815, -0.0467)	94.48%	(0.0396,0.0988)	95.74%
	NBB	(-1.8815, -0.0467)	94.82%	(0.0396,0.0988)	96.22%
	MLE	(-1.8815, -0.0467)	95.00%	(0.0396,0.0988)	95.00%
n=150	NB	(-2.4067, -0.6155)	95.58%	(0.0557,0.1114)	92.88%
	NBB	(-2.4067, -0.6155)	95.90%	(0.0557,0.1114)	93.40%
	MLE	(-2.4067, -0.6155)	95.00%	(0.0557,0.1114)	95.00%

Table 6. Confidence intervals and confidence interval length by ML, NB and NBB.

		ML	NB			NBB
			Normal	Percentile	BCa	
n=60	$\hat{\beta}_0$	(-2.4096,0.2877)	(-2.0986,0.3540)	(-2.378,0.0639)	(-2.226,0.1812)	(-2.2682,0.0534)
		2.6973	2.4526	2.4419	2.4072	2.3216
	$\hat{\beta}_1$	(0.0217,0.1009)	(0.0175,0.0868)	(0.0319,0.1002)	(0.0284,0.0931)	(0.0322,0.0943)
		0.0792	0.0693	0.0683	0.0647	0.0621
n=80	$\hat{\beta}_0$	(-2.4995,-0.2634)	(-2.289,-0.111)	(-2.586,-0.413)	(-2.373,-0.266)	(-2.4583,-0.4029)
		2.236	2.178	2.173	2.107	2.0554
	$\hat{\beta}_1$	(0.0469,0.1279)	(0.0411,0.1143)	(0.0557,0.1292)	(0.0501,0.1172)	(0.0560,0.1236)
		0.0810	0.0732	0.0735	0.0671	0.0676
n=100	$\hat{\beta}_0$	(-1.8815,-0.0467)	(-1.821,0.0902)	(-1.9996,-0.107)	(-1.8628,0.033)	(-1.897,-0.0885)
		1.8348	1.9112	1.8926	1.8958	1.8085
	$\hat{\beta}_1$	(0.0396,0.0988)	(0.0346,0.0928)	(0.0438,0.1018)	(0.0400,0.0950)	(0.0445,0.0985)
		0.0592	0.0582	0.058	0.055	0.054
n=150	$\hat{\beta}_0$	(-2.4067,-0.6155)	(-2.256, -0.535)	(-2.464, -0.724)	(-2.346, -0.644)	(-2.4027,-0.7018)
		1.7912	1.721	1.74	1.702	1.7009
	$\hat{\beta}_1$	(0.0557,0.1114)	(0.0470,0.1092)	(0.0585,0.1199)	(0.0551,0.1136)	(0.0570,0.1172)
		0.0557	0.0622	0.0614	0.0585	0.0602

Specifically, it is easy to know that when $n = 60,80$, the confidence interval length of the NBB method is the shortest, the second shortest is NB method, and the MLE method has the longest confidence interval; when $n = 100$, the confidence interval length of the NBB method is the shortest, and confidence interval length of the MLE method is almost the same as the NB method; when $n = 150$,the confidence interval of the three methods is almost the same length. Among them, in the NB method, the confidence interval length of the BCa method is better than the normal and percentile methods, and the confidence interval lengths of the normal and percentile methods are reasonably similar.

V. CONCLUSIONS

In this paper, the NB and NBB methods are applied to estimate the coefficient parameters in the binary logistic regression model. A numerical example and simulation study shows that three methods are effective ways for parameter estimation.

For the estimator based on the NBB method is relatively superior to that of the ML and NB in the standard error sense and the reliability sense. Also, in practical meaning the odds ration derived from the NBB method performs relatively well in small sample case.

Also, the confidence interval derived from the NBB method is relatively better than the ML method and NB method in accuracy sense for small sample.

ACKNOWLEDGEMENTS

This work was supported by Guizhou Provincial Department of Education (Grant No: 20161111047), Guizhou Science and Technology Department (Grant No: Qian Science [2017]1083), Guizhou Provincial Education Department (Grant No: Qian Science co-JG word LKM [2015]014) and High-level Innovative Talents Project of Guizhou Province.

REFERENCES

Adjei, I.A. and R. Karim, "An application of Bootstrapping in Logistic Regression Model", *Open Access Library Journal*, **3(9)** (2016).

Ariffin, S.B. and H. Midi, "Robust bootstrap methods in logistic regression model", *International Conference on Statistics in Science, Business, and Engineering*, Langkawi, Malaysia, 1-6 (2012).

Banjanovic, E.S. and J.W. Osborne, "Confidence intervals for effect sizes: applying bootstrap resampling", *Practical Assessment, Research & Evaluation*, **21(5)**, 1-20 (2016).

Davison, A.C., D.V. Hinkley and G.A. Young, "Recent developments in bootstrap methodology", *Statistical Science*, **18(2)**, 141-157 (2003).

Efron, B. "Bootstrap methods: Another look at the Jackknife". *Ann. Statist*. **7(1)**, 1-26 (1979).

Fox, J. and S. Weisberg, "Bootstrapping regression models in R: An appendix to an R Companion to Applied Regression", *SAGE Publications, Thousand Oaks*, (2012).

Hossain, A., "Nonparametric Bootstrapping for Multiple Logistic Regression Model using R", *BRAC University Journal*, **1(2)**, 109-113 (2004).

Khoshbin, S. and K. Seyyedi. "Removal of acid red 1 dye pollutant from contaminated waters by electrocoagulation method using a recirculating tubular reactor", *Latin American Applied Research*, **47(3-4)**, 101-105 (2017).

Månsson, K. and G. Shukur, "On ridge parameters in logistic regression", *Communications in Statistic-Theory Methods*, **40(18)**, 3366-3381 (2011).

Rubin, D.B, "The Bayesian Bootstrap", *Ann. Stat.*, **9(1)**, 130-134 (1981).

Salgado, R., W. Bucheli and G.C. Arteaga, "Electrochemistry as an alternative for the obtaining of conductive polymer nano-threads", *Latin American Applied Research*, **47(3)**, 89-94 (2017).

Schaefer, R.L., L.D. Roi and R.A. Wolfe, "A ridge logistic estimator", *Communications in Statistic-Theory Methods*, **13(1)**, 99-113 (1984).

Received: December 15th 2017
Accepted: June 30th 2018
Recommended by Guest Editor
Juan Luis García Guirao