



Proposal of a bootstrap procedure using measures of influence in non-linear regression models with outliers

Larissa Ribeiro de Andrade¹, Marcelo Angelo Cirillo^{1*} and Luiz Alberto Beijo²

¹Departamento de Ciências Exatas, Universidade Federal de Lavras, Cx. Postal 3037, Lavras, Minas Gerais, Brazil. ²Departamento de Ciências Exatas, Universidade Federal de Alfenas, Alfenas, Minas Gerais, Brazil. *Author for correspondence. E-mail: macuflla@dex.ufla.br

ABSTRACT. The bootstrap method is generally performed by presupposing that each sample unit would show the same probability of being re-sampled. However, when a sample with outliers is taken into account, the empirical distribution generated by this method may be influenced, or rather, it may not accurately represent the original sample. Current study proposes a bootstrap algorithm that allows the use of measures of influence in the calculation of re-sampling probabilities. The method was reproduced in simulation scenarios taking into account the logistic growth curve model and the CovRatio measurement to evaluate the impact of an influential observation in the determinacy of the matrix of the co-variance of parameter estimates. In most cases, bias estimates were reduced. Consequently, the method is suitable to be used in non-linear models and allows the researcher to apply other measures for better bias reductions.

Keywords: CovRatio, accuracy, precision, Monte Carlo.

Proposta de um procedimento bootstrap utilizando medidas de influência em modelos de regressão não lineares na presença de outliers

RESUMO. Em geral o método *bootstrap* é realizado supondo que cada unidade amostral apresente a mesma probabilidade de ser reamostrada. Contudo, ao considerar uma amostra que apresente *outliers*, a distribuição empírica gerada através da execução desse método pode ser influenciada, no sentido de não representar fielmente a amostra original. Tendo por base este problema, o objetivo desse trabalho consistiu em propor um algoritmo *bootstrap* que permita utilizar medidas de influência no cálculo das probabilidades de reamostragem. Com este propósito, a ilustração desse método foi feita em alguns cenários de simulação, considerando o modelo não linear de crescimento logístico e a medida CovRatio, utilizada para avaliar o impacto de uma observação influente no determinante da matriz de covariância das estimativas dos parâmetros. Observou-se que na maioria dos casos as estimativas dos vieses foram reduzidas. Concluiu-se que o método é adequado de ser utilizado em modelos não lineares, permitindo ao pesquisador aplicar outras medidas de tal forma a proporcionar melhor redução do viés.

Palavras-chave: CovRatio, acurácia, precisão, Monte Carlo.

Introduction

When inferential methods in data analysis have to be applied, the researcher constantly faces atypical observations that are generally interpreted as outliers. The consequence of such observations consists in a breach of assumptions and/or the construction of statistical tests. It may be said that when outliers occur in a sample, the researcher must be very careful in the interpretation of results, since the latter may have been impaired.

An alternative consists in the application of robust inferential methods in the occurrence of such issues. While taking into consideration samples from binomial populations within this context, Silva and Cirillo (2010) carried out a study related to the performance of an estimator for the binomial

proportion by different concentration of outliers in the sample. Wood et al. (2005) used the same approach and proposed two estimators differentiated by the arithmetic and weighted mean of the observed proportions. When the estimators' variances were compared, the authors concluded that the recommendation of an estimator will be given in different situations, characterized by the distribution of proportions and the number of trials (n) performed.

It is worth mentioning that the use of bootstrap methods is specific to each problem. Field and Welsh (2007) discussed the effect of different ways of modeling clustered data and found that the consistency of variance estimates for a bootstrap method depended on the choice of model with the residual bootstrap. Nankervis (2005) conducted a

study of computer algorithms using double bootstrap to estimate intervals of confidence. It was a study of the robustness of statistical methods and model adjustments by outlier observations (CARONI; BILLOR, 2007; FILZMOSER et al., 2008; JACKSON; CHEN, 2004). In this context, Cirillo et al. (2006) used the bootstrap method and studied the characteristics of Levene's test with regard to the multivariate approach, known in the literature as a robust test to the breach in the supposition of normality.

When the use of the bootstrap method in linear models was considered, Roberts and Martin (2006) used the bootstrap approach to obtain the critical points from the studentized residuals that detected the influence of outliers in linear regression issues in adjusting regression models while taking into consideration the non-normal error distributions.

Within a similar context, bootstrap methods for the adjustment of robust regression by the least trimmed squares (LST) method were developed by Salibian-Barrera and Zamar (2002) and by Van Aelst and Willems (2002). In order to reduce the computational effort, Willems and Van Aelst (2005) proposed a more efficient bootstrap algorithm with more accurate estimates. In the case of LMS method, Jiménez-Gamero et al. (2004) proposed a method called reduced bootstrap for the median.

In heteroscedastic models, Flachaire (2005) compared re-sampling methods that took into account different heteroscedastic residual structures in the re-sampling process with paired bootstrap. The author concluded that, taking into account the heteroscedastic structure, some asymptotic tests had a better performance. New bootstrap estimators for heteroscedastic models may be found in Cribari-Neto and Gois (2002), Cribari-Neto and Soares (2003).

In terms of predictive models, Austin and Tu (2004) provided significant considerations in the use of algorithms in different situations illustrated by multi-co-linearity, model selection methods and statistical adjustment.

In the case of non-linear model studies, the researcher is constantly faced with situations of inability to obtain analytical solutions or to fix assumptions that would make feasible the derivation process. Due to this situation, the obtainment of accurate estimates becomes more serious when the sample to be submitted to the application of the bootstrap method provides a significant amount of outliers.

The usual methods, known in the literature as parametric and nonparametric bootstrap (DAVISON; HINKLEY, 1997), may provide

unreliable estimates since the probability of a sample unit be randomly selected to compose a bootstrap sample follows a uniform discrete distribution. In other words, all sampling units have the same probability of being selected. The application of these methods will certainly generate sub-samples that may have more outliers than the original sample (LOK; LEE, 2011). Salibian-Barrera et al. (2009) developed a bootstrap method for robust estimators which is computationally faster and more resistant to outliers than the classical bootstrap. This fast and robust bootstrap method is, under reasonable regularity conditions, asymptotically consistent.

Regarding to the application of the bootstrap method and its problems to the presence of outliers, the attainment of estimates of nonlinear model parameters, for small samples, provided a field of research on the use of computing resources that reduced the computational effort.

Current research proposes a bootstrap algorithm that allows the use of influence measures in the calculation of re-sampling probabilities. The method will be illustrated by some simulation scenarios according to the logistic growth curve model and the CovRatio measure used to evaluate the impact of an influential observation in determining the covariance matrix of estimated parameters.

Material and methods

Methodology is described according to the following steps: (i) Monte Carlo Simulation of the logistic model; (ii) bootstrap procedure with the incorporation of CovRatio influence measure and (iii) evaluation of the accuracy and precision of estimates of the logistic growth curve model parameter.

(i) Monte Carlo simulation of the logistic growth curve model

The logistic model studied in current research has been defined below (1)

$$y_i = \alpha / \{1 + \exp[\beta(x_i - y)]\} + \varepsilon_i \quad (i=1, \dots, n) \quad (1)$$

where sample sizes are specified in $n = 50$ and 150 , x_i , $i = 1, \dots, n$, a co-variate with fixed effect. The parameter α referred to the upper asymptote; β corresponded to the curve intercept and γ indicated the average growth rate. Finally, ε_i referred to the i^{th} residue generated by a contaminated normal distribution, with the following distributions:

$$\varepsilon_i \sim (1-\delta) \times N(0,1) + (\delta) \times \text{Beta}(6,2); \quad (2)$$

$$\varepsilon_i \sim (1-\delta) \times N(0,1) + (\delta) \times \text{Beta}(6,6); \quad (3)$$

$$\varepsilon_i \sim (1-\delta) \times N(0,1) + (\delta) \times \text{Beta}(2,6); \quad (4)$$

where δ is the outlier percentage in the sample, previously specified at 5 and 10%. It should be noted that, with the above specific probabilities, the residues distributed by normal distribution indicated the reference population, whereas the residues generated by Beta distribution corresponded to outliers. Contaminated normal distribution was thus characterized. The parameters defined in Beta distribution, described in the expressions (2-4), were arbitrarily set, such that the distribution of residues to which outliers were generated, would present varying degrees of asymmetry (CASSEL et al., 1999), as Figure 1 shows.

It should be underscored that residues may be generated by asymmetrical normal distributions (AZZALINI; CAPITONIO, 2003; FANG et al., 1990), following the aim of current investigation.

The parametric rates in the model simulation (1) were arbitrarily defined by $\alpha = 1$, $\beta = 1$ and $\gamma = 1.5$. Thus, for each sample generated, the least squares estimates were obtained by Gauss-Newton iterative method (MAZUCHELI; ACHCAR, 2002), with the following configurations: maximum number of iterations set at $r = 1000$, and convergence criterion set at $\xi = 1e^{-100}$. Keeping these specifications in mind, 500 Monte Carlo simulations were performed for each situation within the combination of sample size (n), outlier percentage in the sample (δ) and distribution of residues (Figure 1). However, only K accomplishments that showed the convergence achieved in the Monte Carlo process and in the bootstrap procedure described in Section (ii), were taken into account.

ii) Bootstrap procedure with the incorporation of CovRatio influence measure

The incorporation of the CovRatio influence measure in bootstrap procedure for each situation was initially performed in two steps:

(a) Obtaining Hessians matrices, represented by X on the solution of the least squares by Gauss-Newton's method;

(b) Obtaining estimates of ordinary least squares, excluding the intercept and specifying the linearized model by $Y = X + \varepsilon$. It should be emphasized that the term linearized model in current assay does not deal with a transformation that makes the model linear; it is used as a symbolic notation with regard to the general linear model.

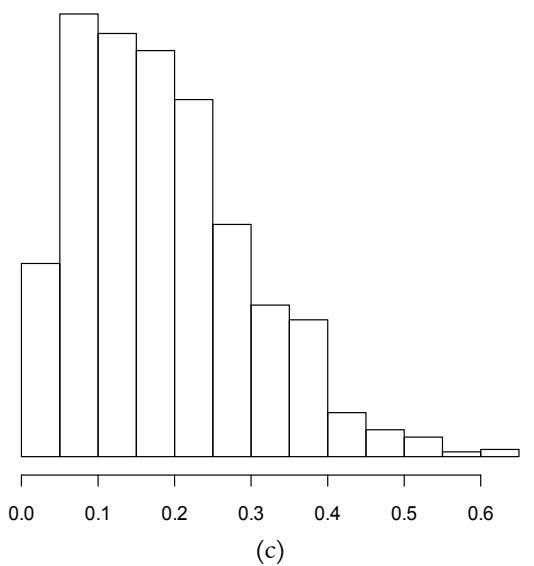
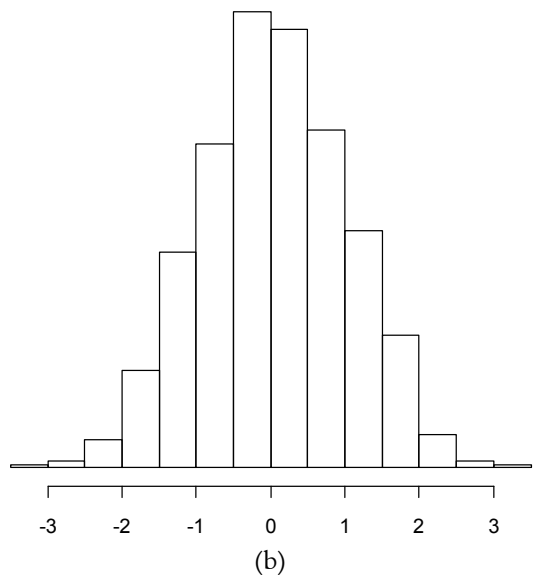
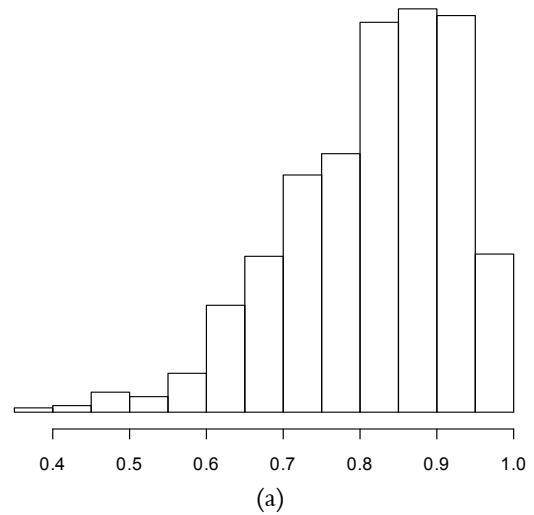


Figure 1. Distributions of residues used in the generation of outliers (a) Beta (6.2); (b) Beta (6.6) and (c) Beta (2.6).

This is due to the replacement of the planning matrix by the Hessian matrix in (a). Thus, assuming the linearized models in (b), the influence measure $F_i, i = 1, \dots, n$, represented by CovRatio, used as a criterion for calculating the weights for each sampling observation, is imposed on the bootstrap procedure, following the algorithm below:

1st - Consider a sample defined by set $D = \{(y_1, x_1, \varepsilon_1), \dots, (y_n, x_n, \varepsilon_n)\}$ where each element has been generated on a linearized model.

2nd - Calculate the influence measure $F_i, i=1,2,\dots,n$, related to CovRatio.

3rd - Assign weight w_i to each observation, according to the rule

(a) If $F_i \geq |r_i|$ then $w_i = |F_i|$

(b) If $F_i \leq |r_i|$ then $w_i = |r_i|$

where r_i indicates the i^{th} studentized residuals of linearized non-linear model.

4th - Calculate the probabilities of re-sampling by

$$p_i = 1 - \frac{w_i}{\sum_{i=1}^n w_i}, \quad i=1, \dots, n.$$

5th - Assign to each element of set D the weights p_1, \dots, p_n , as below:

$$D = \{(y_1, x_1, \varepsilon_1), (y_2, x_2, \varepsilon_2), \dots, (y_n, x_n, \varepsilon_n)\}$$

After the execution of this algorithm, the subsets formed from $B=500$ re-sample with replacement, performed on D (first step), are represented by

$$D_b = \{(y_1^*, x_1^*, \varepsilon_1^*), (y_2^*, x_2^*, \varepsilon_2^*), \dots, (y_n^*, x_n^*, \varepsilon_n^*)\}$$

with $b=1, \dots, B$.

Ordinary least squares estimates were obtained for each subset D_b . At the end of this procedure, B subsets were considered within the generation of empirical distribution of each parameter.

iii) Accuracy and precision evaluation of estimates of logistic growth curve model parameter

When the estimates of the logistic growth curve model parameters were obtained, an accuracy and precision analysis was undertaken for the validation of the bootstrap procedure proposed in Section (ii). Monte Carlo estimates (MC) and bootstrap with influence measure were the two approaches proposed, with estimates corrected by Monte Carlo bias (BMIC).

Bias calculation was performed by assuming K samples, interpreted as the number of ‘valid’ samples, or rather, the simulated sample showed a

convergence of the Gauss-Newton iterative method in the Monte Carlo simulation and in bootstrap procedure. Consequently, for each ‘valid’ sample, the computed estimates of logistic model (1) produced an empirical distribution, by which the bias could be calculated, according to expressions (5) and (6).

$$Bias(\widehat{MC})_j = \frac{1}{K} \left| \sum_{k=1}^K (\hat{\theta}_{mjk} - \theta_j) \right|, \tag{5}$$

$$Bias(\widehat{BMIC})_j = \frac{1}{K} \left| \sum_{k=1}^K (\hat{\theta}_{bjk} - \theta_j) \right| - Bias(\widehat{MC})_j, \tag{6}$$

where θ_j is j^{th} ($j = 1, 2, 3$) parameter of the parametric vector $\theta' = [\alpha \ \beta \ \gamma]$; $\hat{\theta}_{mjk}$ and $\hat{\theta}_{bjk}$

respectively referred to Monte Carlo and bootstrap estimate of the j^{th} parameter specified in θ' . In terms

of accuracy, the standard deviation of estimates obtained through the Monte Carlo method (MC) and bootstrap (BMIC) was computed, following expressions (7) and (8).

$$D_{MC} = \sqrt{\frac{\sum_{k=1}^K (\hat{\theta}_{mjk} - \theta_j)^2}{K - 1}} \tag{7}$$

$$D_{BMIC} = \sqrt{\frac{\sum_{k=1}^K (\hat{\theta}_{bjk} - \theta_j)^2}{K - 1}} \tag{8}$$

Results and discussion

Model accuracy of the estimates of the BMIC bootstrap method with bias correction

Results described in Table 1 correspond to the study of accuracy of the estimates of the logistic growth curve model parameters obtained by the Monte Carlo and the bootstrap methods (Section ii) proposed in current research. This was done by relative bias in the two approaches, namely, Monte Carlo simulations (MC) and bootstrap re-sample, respectively, with measures of influence and estimates corrected by Monte Carlo bias (BMIC).

Owing to results in Table 1, with different degrees of symmetry for the residue, BMIC in the two outlier percentage, $\delta = 5$ and 10%, showed an improvement in estimate accuracy relating to parameter α , which represents the upper asymptote.

Specifically with regard to the symmetric distribution of residues Beta (6.6), in agreement with results by Cook et al. (1986) concerning the

Table 1. Results relating to bias on estimates of logistic growth curve model parameter obtained by MC and BMIC methods due to different sample sizes (n), percentage of outliers in the sample (δ) and distribution of residues with different degrees of symmetry: left (Beta (6.2)); right (Beta (2.6)) and symmetric (Beta (6.6)).

		Upper asymptote (α)					
Method	n	$\delta = 5\%$			$\delta = 10\%$		
		Beta (6.6)	Beta (6.2)	Beta (2.6)	Beta (6.6)	Beta (6.2)	Beta (2.6)
MC	50	0.2506	0.1823	0.2007	0.1131	0.1539	0.1588
	150	0.1090	0.0997	0.0624	0.0671	0.0715	0.0710
BMIC	50	0.0285	0.0167	0.0208	0.0327	0.0878	0.0539
	150	0.0065	0.0376	0.0507	0.0440	0.0618	0.0208
		Intercept of model (β)					
Method	n	$\delta = 5\%$			$\delta = 10\%$		
		Beta (6.6)	Beta (6.2)	Beta (2.6)	Beta (6.6)	Beta (6.2)	Beta (2.6)
MC	50	0.1854	0.5729	0.4607	0.2838	0.3124	0.2129
	150	0.6067	1.0903	0.6270	0.7730	0.8470	1.1090
BMIC	50	0.2908	0.0913	0.1778	0.3765	0.1782	0.1594
	150	0.0137	0.5854	0.1527	0.3514	0.2736	0.6737
		Average growth rate (γ)					
Method	n	$\delta = 5\%$			$\delta = 10\%$		
		Beta (6.6)	Beta (6.2)	Beta (2.6)	Beta (6.6)	Beta (6.2)	Beta (2.6)
MC	50	0.7949	0.7836	0.8138	0.8020	0.8019	0.8028
	150	0.8122	0.8548	0.8419	0.8449	0.8684	0.8308
BMIC	50	1.8578	0.6463	1.8239	0.7527	1.3738	2.5804
	150	0.3760	2.0029	0.4386	1.9275	1.6052	0.1141

MC = Monte Carlo method. BMIC = bootstrap method with measure of influence corrected by Monte Carlo bias.

biases of maximum likelihood estimates in non-linear regression models, the authors concluded that low bias rate might be a merely consequence of the position of the co-variate sample space. Since results were based on a few simulation scenarios, no strong statistical evidence exists that would allow broader conclusions on BMIC performance.

In situations involving the asymmetric distribution of residuals (Beta (2.6)) and Beta (6.2)), evidence exists that the breach in conditions of regularity of nonlinear models have contributed to high bias results, which may be mainly observed in small samples (n = 50). Consequently, the BMIC method provided a reduction in bias estimate and the method may be recommended for such situations.

It is worth noting that results obtained for large samples (n = 150), even in limitation, have been adequate in certain specific situations. This is especially true for results of Beta (2.6) distribution by different amounts of outliers ($\delta = 5\%$) and ($\delta = 10\%$) in which the rates were below 0.01 and consistent with recommendation by Box (1971) who considered it a reasonable rate relative bias equal to or less than 0.01.

Taking into consideration the results for curve intercept, represented by parameter β , it may be said that BMIC reduced bias in the two sample sizes within situations where the distribution of residuals was asymmetric. However, biases did not provide adequate rates, according to criterion by Box (1971).

With regard to the above criterion for results in current research, correction by Monte Carlo bias, imposed on the bias calculation of bootstrap estimates, was not efficient. An alternative to achieve

a more significant reduction in bias rates is the application of bias correction techniques, suggested by Cox and Snell (1968) and Mackinnon and Smith (1998). An example may be found in Cordeiro (2004), or rather, a proposal for the bias correction of maximum likelihood estimators in the class of nonlinear regression symmetric homoscedastic models.

Alternatively, other measures of influences, such as Dffits, Dfbetas and the use of a transformation in the variable response to minimize the effects of outliers, may be suggested. Rodrigues et al. (2010) suggested isotonic regression with different weights. Within the context of robust models, alternatives to normal errors have been proposed in the literature, with heavier tail distributions for residues with regard to normal distribution so that the influence of aberrant points could be reduced.

In terms of accuracy of estimates of parameter γ , interpreted as an average rate of growth, BMIC provided inconsistent results. This fact was notorious mainly for small samples. When small and even moderate sample sizes were taken into account, the formulas that calculated second-order biases were extremely useful to ensure improvement in the accuracy of the estimators.

Precision of estimates of the BMIC bootstrap method with bias correction

Following the methodology described in the specifications and evaluated for accuracy, Table 2 demonstrates the precision results.

When all the situations analyzed are taken into account, estimates of results for BMIC method, in the case of the parameter α , were imprecise.

Table 2. Results of standard deviations of the parameter estimates of logistic growth curve model obtained by MC and BMIC for different sample sizes (n), percentage of outliers in the sample (δ) and distribution of residues with different degrees of symmetry: left (Beta (6.2)); right (Beta (2.6)) and symmetric (Beta(6.6)).

		Upper asymptote (α)					
Method	N	$\delta = 5\%$			$\delta = 10\%$		
		Beta (6.6)	Beta (6.2)	Beta (2.6)	Beta (6.6)	Beta (6.2)	Beta (2.6)
MC	50	0.7130	0.2954	0.5303	0.2383	0.2718	0.3293
	150	0.2256	0.2427	0.1513	0.2172	0.1797	0.2270
BMIC	50	0.4138	0.4435	0.3293	0.4794	0.6221	0.5011
	150	0.2498	0.4691	0.3647	0.3977	0.6015	0.5075
		Intercept of the model (β)					
Method	n	$\delta = 5\%$			$\delta = 10\%$		
		Beta (6.6)	Beta (6.2)	Beta (2.6)	Beta (6.6)	Beta (6.2)	Beta (2.6)
MC	50	1.9535	2.4357	1.7542	1.6458	1.5706	1.1813
	150	1.7344	2.9820	1.6555	2.6015	1.1162	2.8225
BMIC	50	0.7019	0.7519	1.5189	1.6382	0.8350	0.7937
	150	0.5931	0.9043	1.0720	0.8608	0.6108	0.8357
		Average growth rate (γ)					
Method	n	$\delta = 5\%$			$\delta = 10\%$		
		Beta (6.6)	Beta (6.2)	Beta (2.6)	Beta (6.6)	Beta (6.2)	Beta (2.6)
MC	50	0.3057	0.6745	0.3107	0.3064	0.3147	0.3092
	150	0.5041	0.3429	0.2546	0.2955	0.2303	0.4360
BMIC	50	13.7013	10.6441	8.2448	10.0699	11.7297	8.3438
	150	34.3643	38.2027	41.7933	23.5487	46.1629	35.0767

MC = Monte Carlo method. BMIC = bootstrap method with measure of influence corrected by Monte Carlo bias.

One of the causes of imprecision might have been the lack of accuracy of results in Table 1. In this context, Meyer et al. (2006) indicated that the overestimation of the lack of precision was largely related to lack of model accuracy (high bias). The same author suggested that other related measures, such as the square root of the mean error of prediction, might be used since this measure corrected the lack of accuracy.

With regard to the precision of estimates related to parameter β , the curve intercept, results in Table 2 showed that method BMIC reduced average deviation patterns in the two situations, characterized by different amounts of outliers contained in the sample, previously fixed at $\delta = 5$ and 10%. With regard to estimates regarding the average rate of growth, represented by parameter γ , results in Table 2 indicated that, within the evaluation of these scenarios, the BMIC method resulted in low precision estimates. This fact was detected in all situations evaluated, including larger samples ($n = 150$).

It is highly relevant to mention that Evans (1996) discussed the effect of the degree of asymmetry of two growth curves on the precision of estimates of least squares of their parameters by Student's t statistic. The author employed the 4-parameter logistic model developed by Stone (1980), with the fourth parameter as θ .

The parameter's function is given by relating the time that the inflection point is located on the curve as a function of upper asymptote represented by the model's parameter α . The author concluded that asymmetry had a more detrimental effect on the

variances of the parameters of the modified logistic model. In fact, the average growth rate was the most affected parameter.

Conclusion

The bootstrap procedure may be applicable to non-linear model fitting in a given sample with outliers. However, caution must be taken in the choice of measure influences as a calculation criterion to obtain re-sampling probabilities.

In terms of the application to the logistic growth curve model for distributions with outliers from asymmetric distributions, BMIC reduced bias by more accurate and precise estimates, specifically parameters α and β , which respectively represented the upper asymptote and the intercept.

References

- AUSTIN, P. C.; TU, J. V. Bootstrap methods for developing predictive models. **The American Statistician**, v. 58, n. 2, p. 131-137, 2004.
- AZZALINI, A.; CAPITONIO A. Distributions generated by perturbation of symmetry with emphasis on a multivariate skew t -distribution. **Journal of the Royal Statistical Society. Series B**, v. 65, n. 2, p. 367-389, 2003.
- BOX, M, J. Bias in nonlinear estimation. **Journal of the Royal Statistical Society: Serie B**, v. 33, n. 2, p. 171-201, 1971.
- CARONI, C.; BILLOR, N. Robust detection of multiple outliers in grouped multivariate data. **Journal of Applied Statistics**, v. 34, n. 10, p.1241-1250, 2007.
- CASSEL, C.; HACKL, P.; WESTLUND, A. H. Robustness of partial least-squares method for estimating latent variable quality structures. **Journal of Applied Statistics**, v. 26, n. 3 p. 435-446, 1999.

- CIRILLO, M. A.; FERREIRA, D. F.; SÁFADI, T. Estudo do poder e tamanho do teste de Levene multivariado via simulação Monte Carlo e bootstrap. **Acta Scientiarum. Technology**, v. 28, n. 1, p. 105-112, 2006.
- COOK, R.; TSAI, C. L.; WEI, B. Bias in nonlinear regression. **Biometrika**, v. 73, n. 3, p. 615-623, 1986.
- CORDEIRO, G. M. Corrected LR tests in symmetric nonlinear regression models. **Journal of Statistical Computation and Simulation**, v. 74, n. 8, p. 609-620, 2004.
- COX, D. R.; SNELL, E. A. General definition of residuals. **Journal of the Royal Statistical Society: Series B**, v. 30, n. 2, p. 248-275, 1968.
- CRIBARI-NETO, F.; GOIS, M. C. A. Uma análise Monte Carlo do desempenho de matrizes de covariância sob heterocedasticidade de forma desconhecida. **Revista Brasileira de Economia**, v. 56, n. 2, p. 309-334, 2002.
- CRIBARI-NETO, F.; SOARES, A. C. N. Inferência em modelos heterocedásticos. **Revista Brasileira de Economia**, v. 57, n. 2, p. 319-335, 2003.
- DAVISON, A. C.; HINKLEY, D. V. **Bootstrap methods and their application**. New York: Cambridge University Press, 1997.
- EVANS, M. Aspects of maximum likelihood estimation of asymmetric growth curves. **Journal of Applied Statistics**, v. 23, n. 5, p. 467-492, 1996.
- FANG, K. T.; KOTZ, S.; NG, K. W. **Symmetric multivariate and related distributions**. London: Chapman and Hall, 1990.
- FIELD, A. C.; WELSH, A. H. Bootstrapping clustered data. **Journal of the Royal Statistical Society: Serie B**, v. 69, n. 3, p. 369-390, 2007.
- FILZMOSER, P.; MARONNA, R.; WERNER, M. Outlier identification in high dimensions. **Computational Statistics and Data Analysis**, v. 52, n. 3, p. 1694-1711, 2008.
- FLACHAIRE, E. Bootstrapping heteroskedastic regression models: wild bootstrap vs. pairs bootstrap. **Computational Statistics and Data Analysis**, v. 49, n. 3, p. 361-376, 2005.
- JACKSON, D. A.; CHEN, Y. Robust principal component analysis and outlier detection with ecological data. **Environmetrics**, v. 15, n. 2, p. 129-139, 2004.
- JIMÉNEZ-GAMERO, M. D.; MUNOZ-GARCIA, J.; PINO-MEJÍAS, R. Reduced bootstrap for the median. **Statistica Sinica**, v. 14, n. 4, p. 1179-1198, 2004.
- LOK, W. S.; LEE, S. M. S. Robustness diagnosis for bootstrap Inference. **Journal of Computational and Graphical Statistics**, v. 20, n. 2, p. 448-460, 2011.
- MACKINNON, J. G.; SMITH, A. A. Approximation bias correction in econometrics. **Journal of Econometrics**, v. 85, n. 2, p. 205-230, 1998.
- MAZUCHELLI, J.; ACHCAR, J. A. Algumas considerações em regressão não linear. **Acta Scientiarum. Technology**, v. 24, n. 6, p. 1761-1770, 2002.
- MEYER, P. M.; MACHADO, P. F.; COLDEBELLA, A.; CASSOLI, L. D.; COELHO, K. O.; RODRIGUES, P. H. M. Validação de modelos de predição de nitrogênio uréico no leite, estimando-se o consumo individual pelo consumo no rebanho. **Acta Scientiarum. Animal Sciences**, v. 28, n. 1, p. 73-79, 2006.
- NANKERVIS, J. C. Computational algorithms for double bootstrap confidence intervals. **Computational Statistics and Data Analysis**, v. 49, n. 2, p. 461-475, 2005.
- ROBERTS, S.; MARTIN, M. A. An evaluation of bootstrap methods for outlier detection in least squares regression. **Journal of Applied Statistics**, v. 33, n. 7, p. 703-720, 2006.
- RODRIGUES, A.; CHAVES, L. M.; SILVA, F. F.; ZEVIANI, W. M. Utilização da regressão isotônica em estudos de curvas de crescimento. **Revista Brasileira de Biometria**, v. 28, n. 4, p. 85-101, 2010.
- SALIBIAN-BARRERA, M.; ZAMAR, R. H. Bootstrapping robust estimates of regression. **The Annals of Statistics**, v. 30, n. 2, p. 556-582, 2002.
- SALIBIAN-BARREIRA, M.; VAN AELST, S.; WILLEMS, G. Fast and robust bootstrap. **Statistical Methods and Applications**, v. 17, n. 1, p. 41-71, 2009.
- SILVA, A. M.; CIRILLO, M. A. C. Estudo por simulação Monte Carlo de um estimador robusto utilizado na inferência de um modelo binomial contaminado. **Acta Scientiarum. Technology**, v. 32, n. 3, p. 303-307, 2010.
- STONE, R. Sigmoids. **Bulletin of Applied Statistics**, v. 7, n. 1, p. 59-119, 1980.
- VAN AELST, S.; WILLEMS, G. Fast and Robust bootstrap for multivariate inference: the R package FRB. **Journal of Statistical Software**, v. 53, n. 3, p. 1-32, 2013.
- WOOD, G. R.; LAI, C. D.; QIAO, C. G. Estimation of a proportion using several independent samples of binomial mixtures. **The Australian and New Zealand Journal of Statistics**, v. 47, n. 4, p. 441-448, 2005.
- WILLEMS, G.; VAN AELST, S. Fast and robust bootstrap for LTS. **Computational Statistics and Data Analysis**, v. 48, n. 4, p. 703-715, 2005.

Received on June 12, 2012.

Accepted on October 10, 2012.

License information: This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.