Bias In The Maximum Likelihood Estimation Of Parameters Of Nonlinear Regression Models

B. Mahaboob, B.Venkateswarlu, C. Narayana, M. Sivaiah

Abstract: Nonlinear model building has become an important tool in Predictive Analysis and Forecasting Theory. MLE is a phenomenon in which one can obtain unknown coefficients of a distribution by optimizing a likelihood function. Maximum likelihood estimate is the vector in parameter space which optimizes the likelihood function. This research article throws a light on the BIAS in the MLE of unknown coefficients of statistical models which are not linear. In addition to this a test for the linearity of regression has been proposed. If the ML function possesses derivatives one can apply first derivative test to obtain optimum values. But in some situations the equations of first degree of ML function are to be solved in explicit manner. For example in linear statistical model OLS estimator optimize the ML function. In vast number of cases advanced numerical techniques should be implemented in order to get ML function. As the application of ML technique is both flexible and intuitive this technique has become an indispensable tool in statistical inference.

Index Terms: Bias, MLE(Maximum Likelihood Estimation), Multivariate Normal distribution, Variance-Covariance Matrix, Residual Vector, Linearity of Regression, Rank correlation coefficient.

1. INTRODUCTION

AMathematician or Statistician suggests a model $M(\beta)$ for describing some hidden relationships where $\beta's$ are the unknown coefficients which tell the characteristics of such a connection. The method of estimation of parameters is a procedure by which the Mathematician or Statistician traces the best parameter choice β for observed entries. Usually β is used to encoding the best features of the model M. An estimate of parameter β say $\hat{\beta}$ accepts the validation of the model (i.e., verifying if the derived model fits the given entries) and forecasting future outcomes. Most renowned techniques in the estimation of parameters are MLE,, NLLSE,,MOM and MINIMAX Estimation. Gauss M. Corriero et al.[1] in their paper derived most general formulae of biases of order two in ML estimates of the regression, dispersion and precision parameters in nonlinear statistical models with t- distributed errors. In 2003, Mukesh Sharma et al.[2] in their paper, presented the maximum likelihood method for regression analyses of censored data below detection limit for nonlinear models and there proposed ML method has been translated into an equivalent least squares method. Man Chen et al. in 2016[3], in their paper proposed analytic bias -corrected MLEs to reduce the biases of the MLEs of two parameter Weibull distributions. R D Cook et al. [4], in their paper, investigated the biases of the residuals and maximum likelihood parameter estimates from normal nonlinear regression models. If the ML function possesses derivatives on can apply first derivative test to obtain optimum values. David E. Giles et.al [2013] depicted the bias of order two of the ML estimators of parameters of Lomax distribution (aka Pareto II) which is most useful in vast number of fields for finite sample sizes. Besides they shown that the bias is positive.

- *B. Mahaboob, Department of Mathematics, Koneru Lakshmaih Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh. Email id: mahaboob@kluniversity.in.*
- *B.Venkateswarlu, Department of Mathematics, Vellore institute of Technology, Vellore, Tamilnadu. Email id: venkatesh.reddy@vit.ac.in.*
- *C. Narayana, Department of Mathematics, Sri Harsha institute of P.G Studies, Nellore, Andhra Pradesh. Email id: nareva.nlr@gmail.com.*
- *M. Sivaiah, Department of Mathematics, NBKR Arts & Science College, Vidyanagar, Andhra Pradesh.*

2 BIAS IN THE MLE OF PARAMETERS OF NONLINEAR REGRESSION MODELS

A set of l nonlinear regression models can be specified by

$$Z_{ab} = g_a\left(Y_a, \psi\right) + \in_{ab} \quad (1)$$

Where $a = 1, 2, \ldots\ldots l$ and $b = 1, 2, \ldots\ldots, m$. Here $\in_{ab}$ stand for random errors following MVND satisfy in $E\left(\in_{ab}\right) = 0, E(\in_{ab}, \in_{pv}) = 0$, $\psi = \left(\psi_1, \psi_2, \ldots\ldots, \psi_q\right)'_{qx1}$ vectors of unknown parameters. The forms of the l models $g_a\left(Y_a, \psi\right)$ are known. $X_b = \left\{X_{pb}\right\}$ is a $v \times 1$ vector of the entries observed for the v (assume) free quantities in the both experiment $Z_b = \left\{Z_{ab}\right\}$ is a $l \times 1$ vector of the entries observed for the $'l'$ responses in the both investigations unless $b = v$ when $E(\in_{ab}, \in_{pb}) = \sigma_{ap}^{b}$ and hence the vectors of entries in the both investigations has the Var-Covar-Matrix

$$\sum{}_b = \left(\sigma_{ap}^{(p)}\right]$$

(2)

In order to evaluate the estimates of the biases in the MLE of the parameters of nonlinear regression models, the following are defined

$\in=\left(\in_1^1,\in_2^1,\ldots\ldots,\in_m^1\right)'$ and $Z=\left(z_1^1,z_2^1,\ldots\ldots,z_m^1\right)'$

$g(\psi)=\{g_1(Y_1,\psi),\ldots\ldots\ldots\ldots,g_l(Y_1,\psi),g_1(Y_2,\psi),\ldots,g_l(Y_m,\psi)\}$

Where $\in$ is the $ml\times1$ vector of ml residuals. z is the $ml\times1$ vector observations and $g(\psi)$ are mk forecasts of ψ By the maximum likelihood or the generalized least squares estimation one can minimize the expression $\left(z-g(\psi)\right)'\sum^{-1}\left(z-g(\psi)\right)$ with respect to ψ . Where is the Var-Covar-Matrix for all ml entries and is block diagonal matrix as every one of the m investigations are taken to be statistically independent. First order condition gives

$$G\left(\hat{\psi}\right)'=\sum^{-1}\left[z-g\left(\hat{\psi}\right)\right]=0$$

(3)

Here $G(\psi)$ represents $ml\times q$ matrix of first order derivatives of $g(\psi)$, the models for every response in every investigation w.r.t every one of the q-parameters, and those differential coefficients are found out at ψ. Let $\Delta=\hat{\psi}-\psi$ represent the discrepancy in the parameter estimate $\hat{\psi}$ following (2.4). Now vector residual is expressed as $E=Z-g\left(\hat{\psi}\right)=\in-G(\psi)\Delta-\frac{1}{2}H\Delta$

(4)

Here the models are taken to be sufficiently denoted by Taylor power series expansion in exponents of Δ shortened after terms of order 2 and G is the $ml\times q$ matrix

$$H=\{J_{11}\Delta,\ldots.,J_{1l}\Delta,J_{21}\Delta,\ldots..,J_{ml}\Delta\}^1$$

(5)

Where each J_{ba} is a $q\times q$ matrix of order two differential coefficients of the a^{th} response model w.r.t to its q unknown constants at the constraints of the both investigations obtained at ψ . Further up to this order of correction

$$G\left(\hat{\psi}\right)=G(\psi)+H \qquad (6)$$

Suppose that $\Delta=B\in+N$ (7)

Here N stands for second order polynomial of the random error variables and

$N=\left\{\in' C_1\in,\in' C_2\in,\ldots\ldots..,\in' C_q\in\right\}'$ and B is $q\times ml$ and $C's$ are $ml\times ml$ matrices and N is $q\times1$ vector. It remains to find the matrices A and B .This supposed structure has the essential principle that when random error variables have null variance the expectation of the discrepancy i.e., the bias vanishes, from (2.8) the bias is obtained by

$$E(\Delta)=E(N) \qquad (8)$$

Put (4) and (6) in (2) to get

$$(G+H)'\sum^{-1}\left(\in-G\Delta-\frac{1}{2}H\Delta\right)=0 \qquad (9)$$

Where $G=G(\varphi)$. Now (7) into (9) and equalizing coefficients to 0 and hence this expression is identically 0 to order 2 in the random error variables $\in$. For the term linear in $\in$, on can obtain

$$G'\sum^{-1}\left(\in-GB\in\right)=0 \qquad (10)$$

$B=UG^{-1}\sum^{-1}$ (11)

Here $U=(G^1\sum^{-1}G)^{-1}$ (12)

For the second order term in $\in$ one can get

$$G^1\sum^{-1}\left(-GN-\frac{1}{2}KB\in\right)+K^1\sum^{-1}\left(\in-GB\in\right)=0 \qquad (13)$$

Where K is the $ml\times q$ and is given by

$K=\{J_{11}B\in,\ldots\ldots.,J_{1l}B\in,J_{21}B\in,\ldots\ldots.,J_{ml}B\in\}^1$ by taking expectations in (13) one can get

$$G^1\sum^{-1}\left(-GE(N)-\frac{1}{2}W\right)=0 \qquad (14)$$

Where W is an $ml\times1$ vector which is given by $W=\left\{tr(J_{11}U),tr(J_{12}U),\ldots\ldots,tr(J_{ml}U)\right\}^1$

it can be easily seen that

$$E\left(K^1\sum^{-1}\left(\in-GB\in\right)\right)=0 \qquad (15)$$

From the bias in the MLE of properties of nonlinear models is given by

$$E(\Delta)=-\frac{1}{2}UG^1\sum^{-1}W$$

(16)

Where U and W are given by (12) and (16). In practice the parameter estimates $\hat{\psi}$ must be used instead of unknown parameters ψ . Put $G_{ml\times q}=\left\{G_1^1,G_2^1,\ldots\ldots..,G_m^1\right\}$. Where each G_b^1 and $l\times q$ matrix of first order derivatives of each of the K responses with respect to each of the q-parameters, evaluate at the conditions of the both experiment and at parameter values ψ . Now the bias can be written as

$$E(\psi) = -\frac{1}{2}\left[\sum_{t=1}^{m} G_t^1 \sum\nolimits_t^{-1} G_t\right]^{-1} \sum_{b=1}^{m} G_b^1 \sum\nolimits_b^{-1} \begin{bmatrix} tr\left\{\left(\sum_{c=1}^{m} G_c^1 \sum\nolimits_c^{-1} G_c\right)^{-1} J_{b1}\right\} \\ . \\ . \\ . \\ tr\left\{\left(\sum_{c=1}^{m} G_c^1 \sum\nolimits_c^{-1} G_c\right)^{-1} J_{bl}\right\} \end{bmatrix}$$

(17)

It is essential to identify that for linear models all the J's vanish and the biases also vanish. Also even if just one response is nonlinear in the parameters then in general all the parameter estimates are biased.

The term $U = \left(\sum_{t=1}^{n} G_t^1 \sum\nolimits_t^{-1} G_t\right)^{-1}$ is in fact the VAR-COVAR Matrix of the estimates of parameters and determinant is the "Generalized variance" of the estimates of parameters.

3 A TEST FOR LINEARITY OF REGRESSION

Choose a model specified by $Z_j = \gamma + \delta g(Y_j) + \in_j,\ j = 1, 2, ...m$. Where γ, δ unknown parameters are Y_j 's are fixed constants.

Assumptions:

B_1 : The errors $\in_j$'s are continuous, B_2 : the errors $\in_j$'s are independently, symmetrically and continuously distributed around the same median.

The testing hypothesis are $H_0 : g(Y) = Y$ □ $H_1 : g$ is a nonlinear convex or concave function. Rename the (Y, Z) pairs such that $Y_1 < Y_2 < ...... < Y_m$

Assume that $m = zp$ and if m is odd delete the middle pair $\left(Y_{(m+1)(0.5)}, Z_{(m+1)(0.5)}\right)$

Evaluate the p sample slopes as

$$\hat{\delta}_t = \frac{Z_{p+t} - Z_t}{Y_{p+t} - Y_t} \qquad (18)$$

The test statistic is given by

$$D = \sum_{j=1}^{p-1} \sum_{k=j+1}^{p} \phi_{jk},\ where\ \phi_{jk} = \begin{cases} 1, if\ \hat{\delta}_j < \hat{\delta}_k \\ -1, if\ \hat{\delta}_j > \hat{\delta}_k \end{cases}$$

For significantly large values of D accept H_0. This is equivalent either to subject H_0 or accept H_0 for this significantly large or small values of Kendall's rank coefficient $\rho = \frac{2D}{p(p-1)}$ and this is calculated for the sequences $(1, 2, ..., p)$ and $\left(\hat{\delta}_1, \hat{\delta}_2, ..., \hat{\delta}_p\right)$. If H_0 is true the successive slopes $\hat{\delta}_1, \hat{\delta}_2, ..., \hat{\delta}_p$ cannot show an appreciate trend where it H_1 is true the slopes should tend to increase or decrease. One can point out that (i) Under either $B_1 or B_2$ the $\hat{\delta}$'s are independent and (ii) Under H_0 and either $B_1 or B_2$ $P\left(\hat{\gamma}_j > \hat{\gamma}_k\right) = P\left(\hat{\gamma}_j < \hat{\gamma}_k\right) = \frac{1}{2}$ for all $j \neq k$.

4 CONCLUSIONS

In the above research work the BIAS in the MLE of parameters of nonlinear statistical models has been extensively discussed. Besides a test for linearity of regression has been proposed. In the context of future research one can propose a test for the specification of errors in statistical models which are nonlinear and can estimate a mixed general Cobb-Douglas type function with multiplicative and additive errors. Moreover the above propositions can be executed in estimating Seemingly Unrelated Nonlinear Regression Equations Model and Gompertz Growth type function considering additive error term.

REFERENCES

[1] Gauss M. Cordeiro, Klaus L.P. Vasuncellos and Maria Luuiza F. Santos (1998), "on the second order bias of parameter estimates in nonlinear regression models with student 't' errors', Journal of statistical computation and Simulation, Vol. 60, issue 4.

[2] Mukesh Sharma, Rinku Agarwal (2003), "Minimum Likelihood method for parameter estimation in nonlinear models with below detection data", Environmental and Ecological statistics, Vol.10, Issue4, and Pp: 445-454

[3] Man Chen, Zheng Zhang and Cencul (2016), "On the Bias of the maximum likelihood estimation of parameters of the Weibull distribution", Mathematical and computational Applications.

[4] Ioannis Kormidis (2014), "Bias in parametric estimation reduction and useful side effects", Wiley online library.

[5] Gauss M. Cordeiro, Silvia L.P. Ferrari, Mignel A. Uribe-opazo, Klaus L.P Vasconcellos (2000)," Corrected maximum likelihood estimation in a class of symmetric nonlinear regression models," Statistics and Probability letters 46(4), Pp: 317-328.

[6] R.D. Cook, C. L. Tsai, B. C. Wei (1986), "Bias in nonlinear regression", Biometrika, Vol.73, Issue (3), Pp: 615-623

[7] Halbert White (1981), "Consequences and detection of Mis-specified nonlinear regression models", Vol. 76, Pp: 413-433.

[8] James G. MacKinnon (1983),"Tests for model specification in the presence of alternative hypothesis. Some further results", Science Direct, Vol.21, issue (1), Pp: 53-70.

[9] A. Ronald Gallant (1975),"Seemingly unrelated nonlinear regressions, Journal of Economics, Vol. (3), and Pp: 35-50.

[10] Risto Heijmans, Heinz Neudeker (1998), "Estimation of the SURE model", Statistical papers, springer Vol. (39), issue (4), Pp: 423-430.

[11] E.J. Kontoghiorghes, M.R.B. Clarke (1995), “An alternative approach for the numerical solution of SURE models”, Computational statistics and Data analysis, Vol. (19), issue (4), Pp: 369-377.

[12] Popli G.K (2000), “A note on the instrumental variable estimators in the nonlinear models”, Journal of Quantitative Economics, 16, Pp: 31-36.

[13] Venkateswarlu .B et.al (2011), “On problems of statistical modelling”, International journal Engineering science and Technology, Vol. (3), issue (6), Pp: 4914-4920

[14] Hanumantha Rao, S., Vasanta Kumar, V., Srinivasa Rao, T., & Srinivasa Kumar, B. (2016). A two-phase unreliable M/Ek/1 queueing system with server startup, N-policy, delayed repair and state dependent arrival rates. Global Journal of Pure and Applied Mathematics, 12(6), 5387-5399.

[15] [15] Rao, S. H., Kumar, V. V., Rao, T. S., & Kumar, B. S. (2016). M/M/1 two-phase gated queuing system with unreliable server and state dependent arrivals. International Journal of Chemical Sciences, 14(3), 1742-1754.

[16] . Appa Rao, B. V. (2016). Minimum energy control of Δ-differentiable positive matrix sylvester dynamical systems. International Journal of Chemical Sciences, 14(2), 751-761

[17] Appa Rao, B. V,& Vara Prasad, K. A. S. N. (2016). Controllability and observability of Sylvester matrix dynamical systems on time scales. Kyung pook Mathematical Journal, 56(2), 529539.doi:10.5666/KMJ.2016.56.2.529.

[18] Srinivasa Rao, T., Suresh Kumar, G., Vasavi, C., & Murty, M. S. N. (2016). Observability of fuzzy difference control systems. International Journal of Chemical Sciences, 14(4), 2516-2526.

[19] Vasavi, C., Suresh Kumar, G., & Murty, M. S. N. (2016). Fuzzy dynamic equations on time scales under generalized delta derivative via contractive-like mapping principles. Indian Journal of Science and Technology, 9(25) doi:10.17485/ijst/2016/v9i25/85323.

[20] Vasavi, C., Suresh Kumar, G., & Murty, M. S. N. (2016). Fuzzy hukuhara delta differential and applications to fuzzy dynamic equations on time scales. Journal of Uncertain Systems, 10(3), 163-180.

[21] Vasavi, C., Suresh Kumar, G., & Murty, M. S. N. (2016). Generalized differentiability and integrality for fuzzy set-valued functions on time scales. Soft Computing, 20(3), 1093-1104. doi:10.1007/s00500-014-1569-1

[22] Sobhanbabu, Y., Appa Rao, B. V., Srinivasa Rao, T., & Prasad, K. A. S. N. V. (2017). Reliability of matrix riccati type dynamical systems on time scales. International Journal of Civil Engineering and Technology, 8(12), 216-223

[23] Sobhanbabu, Y., Rao, B. V. A., Rao, T. S., & Prasad, K. A. S. N. V. (2017). Adaptive control and reliability on nabla settings. International Journal of Mechanical Engineering and Technology, 8(12), 316-324.

[24] Bindu, P., Sarala, Y., Srinivasa Rao, T., & Madhusudhana Rao, D. (2017). Complete left ideals in ternary semigroups. Proceedings of the Jangjeon Mathematical Society, 20(3), 347-351. doi:10.17777/pjms2017.20.3.347.

[25] Madhavi, R., Karri, R. R., Sankar, D. S., Nagesh, P., & Lakshminarayana, V. (2017). Nature inspired techniques to solve complex engineering problems. Journal of Industrial Pollution Control, 33(1), 1304-1311.

[26] Kumar, D., Rajyalakshmi, K., & Asadi, S. S. (2017). Digital marketing strategical role to promote technical education in andhra and telangana: An exploratory study. International Journal of Civil Engineering and Technology, 8(10), 197-206.

[27] . Kishore, G. N. V., Rao, K. P. R., & Bindu, V. M. L. H. (2017). Suzuki type unique common fixed point theorem in partial metric spaces by using (C): Condition with rational expressions. Afrika Matematika, 28(5-6), 793-803. Doi: 10.1007/s13370-017-0484-x.

[28] Vijay Prasad, S., Peter Praveen, J., Tiwari, A., Prasad, K., Bindu, P., Donthi, R., & Mahaboob, B. (2018). An application of LPP - graphical method for solving multi server queuing model. International Journal of Mechanical Engineering and Technology, 9(1066-1069), 1066-1069.

[29] Srinuvasa Rao, B., Kishore, G. N. V., Hari Krishna, Y., & Sarwar, M. (2018). Generalized (α,β) - rational contractions in ordered sb-metric spaces with applications. International Journal of Mechanical Engineering and Technology, 9(3), 272-282.

[30] .Kumar, P., & Keerthika, P. S. (2018). An inventory model with variable holding cost and partial backlogging under interval uncertainty: Global criteria method. International Journal of Mechanical Engineering and Technology, 9(11), 1567-1578.

[31] Rao, I. V. R., Mohammad, A. B., & Kumar, D. S. (2018). A novel image recognition method using multiple component resemblance score synthesis and genetic algorithm. Paper presented at the Journal of Physics: Conference Series, 1139(1) doi:10.1088/1742-6596/1139/1/012069.

[32] Nayagam, V. L. G., Jeevaraj, S., & Dhanasekaran, P. (2018). An improved ranking method for comparing trapezoidal intuitionistic fuzzy numbers and its applications to multicriteria decision making. Neural Computing and Applications, 30(2), 671-682. doi: 10.1007/s00521-016-2673-1

[33] Seetha, M. P., Sarala, Y., Jaya Lalitha, G., & Anjaneyulu, A. (2018). Ideal theory in a-ternary semi groups. International Journal of Civil Engineering and Technology, 9(3), 170-177.