mathematics

MDPI

Article

Inference and Local Influence Assessment in a Multifactor Skew-Normal Linear Mixed Model

Zeinolabedin Najafi [1], Karim Zare [1,*,†], Mohammad Reza Mahmoudi [2,†], Soheil Shokri [3] and Amir Mosavi [4,5,6,7]

1 Department of Statistics, Marvdasht Branch, Islamic Azad University, Marvdasht 73711-13119, Iran
2 Department of Statistics, Faculty of Science, Fasa University, Fasa 74616-86131, Iran
3 Department of Statistics, Lahijan Branch, Islamic Azad University, Lahijan 44169-39515, Iran
4 Faculty of Civil Engineering, Technische Universität Dresden, 01069 Dresden, Germany
5 John von Neumann Faculty of Informatics, Obuda University, 1034 Budapest, Hungary
6 Institute of Information Society, University of Public Service, 1083 Budapest, Hungary
7 Institute of Information Engineering, Automation and Mathematics, Slovak University of Technology in Bratislava, 81243 Bratislava, Slovakia
* Correspondence: karim.zare@iau.ac.ir
† These authors contributed equally to this work.

Abstract: This work considers a multifactor linear mixed model under heteroscedasticity in random-effect factors and the skew-normal errors for modeling the correlated datasets. We implement an expectation–maximization (EM) algorithm to achieve the maximum likelihood estimates using conditional distributions of the skew-normal distribution. The EM algorithm is also implemented to extend the local influence approach under three model perturbation schemes in this model. Furthermore, a Monte Carlo simulation is conducted to evaluate the efficiency of the estimators. Finally, a real data set is used to make an illustrative comparison among the following four scenarios: normal/skew-normal errors and heteroscedasticity/homoscedasticity in random-effect factors. The empirical studies show our methodology can improve the estimates when the model errors follow from a skew-normal distribution. In addition, the local influence analysis indicates that our model can decrease the effects of anomalous observations in comparison to normal ones.

Keywords: EM algorithm; expectation–maximization algorithm; heteroscedasticity; Monte Carlo simulation; random effects; skew-normal; variance components; applied mathematics; Linear mixed models

MSC: 62F10; 62J05

Citation: Najafi, Z.; Zare, K.; Mahmoudi, M.R.; Shokri, S.; Mosavi, A. Inference and Local Influence Assessment in a Multifactor Skew-Normal Linear Mixed Model. *Mathematics* **2022**, *10*, 2820. https://doi.org/10.3390/math10152820

Academic Editor: Christophe Chesneau

Received: 6 July 2022
Accepted: 4 August 2022
Published: 8 August 2022

1. Introduction

Linear mixed models (LMMs) are useful for the statistical analysis of correlated datasets such as longitudinal data. For simplicity, it is usually assumed that both random effects and random errors follow a normal distribution. Under these restrictions, there are several proposals for estimating LMM parameters in the literature; among these, one can refer to Harvill [1], Fellner [2], Khuri et al. [3] and Wu et al. [4]. For example, Harvill [1] and Fellner [2] obtained the maximum likelihood (ML) estimates of parameters in a multifactor normal LMM under heteroscedasticity of random-effect factors. They showed the estimates, in addition to being consistent, were asymptotically normally distributed. However, as pointed out by Zhong and Davidian [5], using these estimation methods may cause invalid statistical inferences when the data are asymmetric. Therefore, many authors have criticized the common use of the normality assumption (see, e.g., [6–10]).

From a practical perspective, the most frequently used method to achieve normality is to apply a transformation on the variables. Although such methods may provide