Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 116 (2017) 182–189

Procedia Computer Science

www.elsevier.com/locate/procedia

2nd International Conference on Computer Science and Computational Intelligence 2017, ICCSCI 2017, 13-14 October 2017, Bali, Indonesia

Hierarchical Linear Mixed Model for Poverty Analysis in Indonesia

Ro'fah Nur Rachmawati[a]*, Novi Hidayat Pusponegoro[b]

[a]*Statistics Department, Bina Nusantara University, Kebun Jeruk Raya No. 27, Jakarta 11530, Indonesia*
[b]*Statistics Department, Sekolah Tinggi Ilmu Statistik, Jl Otista No. 64C, Jakarta 13330, Indonesia*

Abstract

Hierarchical Linear Mixed Model (HLMM) is an extension of Linear Mixed Model with hierarchical levels of observation. HLMM allows researchers to model different types of covariance structure to describe data properly while in classical linear model the covariance structure defines in constant variance and correlation that hardly applicable for longitudinal data. This paper describes two levels HLMM which represents a single growth curves model. Level-1 presents growth shape to capture within-subject effect and level-2 presents growth parameters that characterized between-subject differences. We model the covariance structure of level-1 random effect to excavate individual growth performance and applied to longitudinal data from poverty data of 34 provinces in Indonesia. Different types of covariance structures are modeled using PROC MIXED in SAS system, produce that AR(1) is the alternative of constant covariance structure and ARH(1) as an alternative for non-constant variance structure based on -2RLL, AIC and BIC criteria.

© 2017 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the 2nd International Conference on Computer Science and Computational Intelligence 2017.

Keywords: Longitudinal data; R Covariance structure; GLMM; PROC MIXED; Growth curves.

1. Hierarchical Linear Mixed Model

General Linear Mixed Model (GLMM) is often used to analyze data in various field of social science, psychology, agriculture and health research. GLMM in formal terms, are an extension model of mixed-effects described by Rao (1965)[1] for growth curves and by Laird and Ware (1982)[2] for longitudinal data analysis. The GLMM

* Corresponding author. Tel.: +0-215-345-830; fax: +0-215-300-244.
E-mail address: rofah.nr@binus.ac.id

1877-0509 © 2017 The Authors. Published by Elsevier B.V.
Peer-review under responsibility of the scientific committee of the 2nd International Conference on Computer Science and Computational Intelligence 2017.
10.1016/j.procs.2017.10.027

is very distinct from classical linear model by its analysis of data which show non-constant correlation and variability[3]. It models variance and covariance (random effects), so that the parameters of covariance are employed to model data with specific characteristics.

A special case of GLMM when its residual has Gaussian distribution is Linear Mixed Model (LMM), and when the LMM has two or more levels of group/subject we called the model as Hierarchical Linear Mixed Model (HLMM). There are many advantages of HLMM compared to the classic linear model: (1) As one of the specification of GLMM, with HLMM we can characterized our data by modeling its random effect using various types of covariance structure. By modeling its random effect through the covariance structures, we can excavate the individual effect of group/subject or treatment, while in classical linear model the covariance structure is presume in very ordinary structure, i.e. the constant variance and no correlation between treatment and time. (2) In health research, psychology and agricultural research, we often observe a person or an object repeatedly at different times. The object that are used for repeated measure usually have high correlation at each time and for this reason, statistical inferences derived from the classic model application may be inadequate. HLMM allows us to model repeated measure with correlation that occurs overtimes. (3) Longitudinal data is widely used for developmental research in many areas of science, because it has opposite characteristics from time series data. First, unlike times series data that has the same interval observation for each subject, longitudinal data has flexibility for different interval observation or in other words the repeated measure do not have equally spaced. Second, when times series data must have complete observation, longitudinal data can have missing observation (value). The excellence of GLMM and so HLMM are, HLMM allows researchers to analyze data with missing observations, and adjust conditions in which observations are not equally spaced, and to model the covariance structure of the data rather than presume a certain structure, as in the case of classical inference statistics.

There are two specific GLMM usually used for model individual performance trends, these approaches are known as Hierarchical Linear Model (HLM) and Latent Growth Curve Modeling (LGCM)[3]. The HLM approach has great flexibility with regards to the data but only for a limited class of structures for the parameters, whereas the LGCM allows many choices of parameters for a limited set of data structure[4]. Growth curves are very interested to analyze because we can analyze simultaneously individual level effect/individual differences (or in statistical point of view we see it as within-subject) and group level effect (or between-subject)[5] that changes overtime. We saw many growth curves in developmental research such as Jaume et. al. (2010)[3] that use children weight growth curves, Daniel M, James A. D., and James D. M. (2008)[5] discuss growth curves and individual differences in statistical and computational point of view. Growth curves research develops using nonlinear behaves as stated in Kevin J. G. et. al. (2011)[6] and applicate the model for psychology area, and using the nonlinear concepts for parents-child development[7]. Growth curves using Curve Matching or called data-driven technique also develop to repair individual prediction of childhood growth[8]. Growth curves are usually represented by means of the two-level HLM, individual effects (within-subject) variation are defined at the first level and between-subject variation is modeled at the second level[9]. In other word, as stated in Van der Leeden (1998)[10]; Wu, Clopper, & Wooldridge (1999)[11] observations are the first-level units, and subjects are the second-level units.

Using the foundation of Harville (1977)[12] and Jaume A. et al. (2010)[3], stated that HLMM as the unification of two level (hierarchical) observation in a single model, this paper aims to: (1) Analyze HLMM as two levels observation in a single growth curves model that represent observation that changes overtime as longitudinal data. (2) Using Two levels of HLMM to analyze poverty data in 34 provinces in Indonesia, (3) Modeling the covariance structure of poverty data using different types of covariance structures such as Unstructured (UN), Compound Symmetric (CS), Heterogeneous Compound Symmetric (CSH), First-order Autoregressive (AR(1)), Heterogeneous First-order Autoregressive (ARH(1)) and First-order Autoregressive Moving Average ARMA(1,1), and (4) Choose the best covariance structures that fit poverty data properly using Negative 2-Residual Loglikelihood (-2RLL), Akaike Information Criteria (AIC) and Bayesian Information Criteria (BIC).

Download English Version:

https://daneshyari.com/en/article/4960442

Download Persian Version:

https://daneshyari.com/article/4960442

Daneshyari.com