Statistica Sinica **35** (2025), 1-21
doi:https://doi.org/10.5705/ss.202022.0336

# EMPIRICAL LIKELIHOOD INFERENCE OF VARIANCE COMPONENTS IN LINEAR MIXED-EFFECTS MODELS

Jingru Zhang[1], Wei Guo[2], Joanne S. Carpenter[3], Andrew Leroux[2,4], Kathleen R. Merikangas[2], Nicholas G. Martin[5], Ian B. Hickie[3], Haochang Shou[6] and Hongzhe Li*[6]

[1]*Fudan University,* [2]*National Institutes of Health,*
[3]*University of Sydney,* [4]*University of Colorado,*
[5]*QIMR Berghofer Medical Research Institute and* [6]*University of Pennsylvania*

*Abstract:* Linear mixed-effects models are widely used in analyzing repeated measures data, including clustered and longitudinal data, where inferences of both fixed effects and variance components are of interest. Unlike inference on fixed effect, which has been well studied, inference on the variance components is more challenging due to null value on the boundary and the unknown fixed effects as nuisance parameters. Existing methods require strong distributional assumptions on the random effects and random errors. In this paper, we develop empirical likelihood-based methods for the inference of the variance components in the presence of fixed effects. We derive a nonparametric version of the Wilks' theorem for the proposed empirical likelihood ratio statistics for variance components. We also develop an empirical likelihood test for multiple variance components related to a sequence of correlated outcomes. Simulation studies demonstrate that the proposed methods exhibit better type 1 error control than the likelihood-based or score test when the Gaussian distributional assumptions of the random effects are violated. We apply the methods to investigate the heritability of physical activity measured by actigraph device in the Australian Twin study and observe that such activity is heritable only in the quantile range from 0.375 to 0.514.

*Key words and phrases:* Boundary value, global test, heritability, nonparametric test, wearable device data.

## 1. Introduction

Longitudinal and clustered data commonly arise from observational studies or clinical trials, where subjects are measured repeatedly over time or within a cluster. The repeated measures within a subject or a cluster are often correlated. To analyze such data, linear mixed-effects (LME) models that incorporate both fixed and random effects are widely used. Many statistical methods have been developed for such linear mixed-effects models, especially methods for inference of the fixed effects. However, inference on the variance components is less

*Corresponding author. E-mail: hongzhe@upenn.edu