

# Empirical Likelihood-Based Method for LME Models

**Haiyan Su**

*Montclair State University, USA*

## INTRODUCTION

Longitudinal data are often encountered in biomedical, epidemiological, social science and business studies. Repeated measurements are made on subjects over time and responses within a subject are usually correlated. The interest of longitudinal study is usually the growth curve over time. Linear mixed-effects models (LME) are widely used for longitudinal data because it allows both the population level fixed effects and individual level random effects to appear linearly in the model linearly in the model (Pinheiro & Bates, 2000). In modeling the fixed effects and random effects coefficients, the normality assumption of the random error terms and the random effect terms are required. Violation of this assumption may yield invalid statistical inference. Since the normality assumption is too restrictive as pointed out by Zhang and Davidian (2001), broader classes of distributions for the mixed models have been proposed by Verbeke and Lesaffre (1997), Zhang and Davidian (2001), Chen et al., (2002) among others. Ma and Genton (2004) relaxed the normality assumption by using semiparametric generalized skew-elliptical distributions. Zhou and He (2008) replaced the normal distribution with skew-t distributions. Lachos et al., (2009) proposed a Bayesian approach with asymmetric heavy tailed distribution. The generalized estimation equation (GEE) (Liang & Zeger, 1986) approach relaxed this assumption by focusing on population-level effects using the quasi-likelihood method. GEE is a very popular approach researcher use when normality assumption is violated for continuous

responses and it is supported by most available software such as SAS and R .

In this paper, we develop an empirical likelihood-based method to get inference for population level parameter of interest in linear mixed-effects models when normality assumption is violated. Empirical likelihood was introduced by Owen (1988, 1990). It amounts to computing the parametric likelihood of a general multinomial distribution which has its atoms at all observed data points. Empirical likelihood provides a good alternative among the nonparametric methods that can be used to make statistical inference when the normality assumption is violated. The advantage of empirical likelihood compared to the bootstrap method and the jackknife method arises because it is a nonparametric method of inference based on a data-driven likelihood ratio function. A thorough summary of the advantages of empirical likelihood over its competitors is given in Hall and La Scala (1990). In this study, we suggest doing a projection to erase nuisance parameters for dimension reduction and then using the profile empirical likelihood principle. We then apply the Bartlett correction method to obtain the adjusted confidence interval. The paper is organized as follows: We give the main results including the model formulation, the Bartlett correction procedures and proofs of the theoretical results in the “Main Focus” section. Then follows results from simulation studies by comparing the proposed method with normal approximation method and the GEE method and a real application study. We then give the conclusion of the paper and the future trends.

## MAIN FOCUS

### Model Formulation

The general form of a LME model is  $Y_i = X_i\beta + Z_i b_i + \varepsilon_i$ ,  $i = 1, \dots, n$ , where  $b_i \sim N(0, D)$ ,  $\varepsilon_i \sim N(0, R_i)$ .  $\beta$  is a  $p \times 1$  vector of fixed regression coefficient and  $b_i$  is a  $k \times 1$  vector of random coefficient for subject  $i$ .  $Y_i = (Y_{i1}, Y_{i2}, \dots, Y_{im_i})^T$  is the response variable for subject  $i$ .  $m_i$  is the number of measurements for each subject,  $X_i$  is a  $m_i \times p$  covariance matrix for fixed effects and  $Z_i$  is a  $m_i \times k$  covariance matrix of random effects for subject  $i$ .  $\varepsilon_i$  is a  $m_i \times 1$  vector of the error term which is independent of  $b_i$ .  $D$  is an unknown covariance matrix of the random effects  $b_i$ . We assume the covariance matrix  $R_i = \sigma^2 \text{Im}_i$ , where  $\sigma^2$  is an unknown parameter and  $\text{Im}_i$  is a  $m_i \times m_i$  identity matrix. Let  $N = \sum m_i$  denote the total number of observations, and then the above LME model can be written in the matrix notation:

$$Y = X\beta + Zb + \varepsilon, b \sim N(0, G), \varepsilon \sim N(0, R) \quad (1)$$

where  $Y = (Y_1^T, \dots, Y_n^T)^T$  is a  $N \times 1$  vector,  $X = (X_1^T, \dots, X_n^T)^T$  is a  $N \times p$  matrix of rank  $p$ ,  $Z = \text{diag}(Z_1, \dots, Z_n)$  is a  $N \times r$  matrix of rank  $r = nk$ ,  $b = (b_1^T, \dots, b_n^T)^T$ ,  $\varepsilon = (\varepsilon_1^T, \dots, \varepsilon_n^T)^T$ ,  $b \sim N(0, G)$ ,  $G = \text{diag}(D, \dots, D)$ . For the above linear mixed-effects models, our interest are to estimate  $\beta$  or some sub-vector of  $\beta$ . Random effect  $b$  and random error variance  $\sigma^2$  are treated as nuisance parameters. In order to apply our proposed method for linear mixed-effects models, we reformulate the linear mixed-effects model into the linear model formulation first. Denote  $\Sigma_i = \sigma^2 \text{Im}_i + Z_i D Z_i^T$ ,  $\Sigma = \text{diag}(\Sigma_1, \dots, \Sigma_n) = \sigma^2 I_N + Z G Z^T = \sigma^2 I_N + \text{diag}(Z_1 D Z_1^T, \dots, Z_n D Z_n^T)$ ,  $\varepsilon^* = Zb + \varepsilon$ . The above LME model can be written as a simple linear model

$$Y = X\beta + \varepsilon^*, \varepsilon^* \sim N(0, \Sigma), \quad (2)$$

where the error terms  $\varepsilon^*$  are independent for different subjects, but not independent for the

observations within the same subject or cluster. To make the error terms independent, we need to apply some transformation. We multiply  $\Sigma^{-1/2}$  on both sides of Equation (2) and obtain

$$\Sigma^{-1/2}Y = \Sigma^{-1/2}X\beta + \Sigma^{-1/2}\varepsilon^*. \quad (3)$$

Now the error terms  $\Sigma^{-1/2}\varepsilon^*$  in Equation (3) are independent with each other. Thus we can apply the least square method to estimate  $\beta$ . The least squares function of  $\beta$  can be expressed by

$$L_\beta = (Y - X\beta)^T \Sigma^{-1} (Y - X\beta).$$

By differentiating  $L_\beta$  with respect to  $\beta$ , we have

$$\frac{\partial L_\beta}{\partial \beta} = X^T \Sigma^{-1} (Y - X\beta) = \sum_{i=1}^n X_i^T \Sigma_i^{-1} (Y_i - X_i \beta) = 0.$$

i.e.,  $\beta$  should satisfy the empirical estimating equation

$$E\{X^T \Sigma^{-1} (Y - X\beta)\} = 0. \quad (4)$$

Consequently, in the estimating procedure, we solve  $\hat{\beta}$  as the solution to

$$\frac{1}{n} \sum_{i=1}^n X_i^T \hat{\Sigma}_i^{-1} (Y_i - X_i \beta) = 0, \quad (5)$$

where  $\hat{\Sigma}_i$  is a consistent estimator of  $\Sigma_i$ . Based on Equation (5), we can define the empirical likelihood function for  $\beta$ . Let  $p_i$  be the probability assigned at  $(Y_i, X_i)$ . The empirical-likelihood ratio function for estimating  $\beta$  is given by

$$R_n(\beta) = \sup \left\{ \prod_{i=1}^n n p_i \mid \sum_{i=1}^n p_i X_i^T \hat{\Sigma}_i^{-1} (Y_i - X_i \beta) = 0, p_i \geq 0, \sum_{i=1}^n p_i = 1 \right\} \quad (6)$$

We show that under certain conditions,  $-2\log\{R_n(\beta)\}$  converges to a chi-square distribution asymptotically.

10 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

[www.igi-global.com/chapter/empirical-likelihood-based-method-for-lme-models/107284](http://www.igi-global.com/chapter/empirical-likelihood-based-method-for-lme-models/107284)

## Related Content

---

### Analyzing Social Media for Corporate Reputation Management: How Firms Can Improve Business Agility

Christoph Seebach, Roman Beckand Olga Denisova (2013). *International Journal of Business Intelligence Research* (pp. 50-66).

[www.irma-international.org/article/analyzing-social-media-for-corporate-reputation-management-/83478](http://www.irma-international.org/article/analyzing-social-media-for-corporate-reputation-management-/83478)

### Vintage Analytics and Data Warehouse Design

Ido Millet, Syed S. Andaleeband John L. Fazel (2014). *International Journal of Business Intelligence Research* (pp. 62-81).

[www.irma-international.org/article/vintage-analytics-and-data-warehouse-design/120052](http://www.irma-international.org/article/vintage-analytics-and-data-warehouse-design/120052)

### Utilizing Business Intelligence to Enhance Online Education at For-Profit and Non-Profit Institutions

Ehi Aimuwuand Sanjay Bapna (2013). *International Journal of Business Intelligence Research* (pp. 29-42).

[www.irma-international.org/article/utilizing-business-intelligence-enhance-online/76910](http://www.irma-international.org/article/utilizing-business-intelligence-enhance-online/76910)

### Chinese Cultural Values and Knowledge Sharing

Alan K.M. Auand Matthew C.H. Yeung (2016). *Business Intelligence: Concepts, Methodologies, Tools, and Applications* (pp. 1482-1490).

[www.irma-international.org/chapter/chinese-cultural-values-and-knowledge-sharing/142684](http://www.irma-international.org/chapter/chinese-cultural-values-and-knowledge-sharing/142684)

### Bankruptcy Prediction Using Principal Component Analysis Based Neural Network Models

G. A. Rekha Paiand G. A. Vijayalakshmi Pai (2010). *Business Intelligence in Economic Forecasting: Technologies and Techniques* (pp. 138-156).

[www.irma-international.org/chapter/bankruptcy-prediction-using-principal-component/44253](http://www.irma-international.org/chapter/bankruptcy-prediction-using-principal-component/44253)