Journal of Multivariate Analysis 136 (2015) 1–11

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Estimation in skew-normal linear mixed measurement error models

Ameneh Kheradmandi, Abdolrahman Rasekh *

Department of Statistics, Shahid Chamran University, Ahvaz, Iran

ARTICLE INFO

Article history:
Received 3 July 2014
Available online 19 December 2014

AMS 2008 subject classifications:
62J05
62F10

Keywords:
Maximum likelihood
EM algorithm
Structural model
Skewness

ABSTRACT

In this paper we define a class of skew-normal linear mixed measurement error models. This class provides a useful generalization of normal linear mixed models with measurement error in fixed effects variables. It is assumed that the random effects, model errors and measurement errors follow a skew-normal distribution, extending usual symmetric normal model in order to avoid data transformation. We find the likelihood function of the observed data, which can be maximized by using standard optimization techniques. Next, an EM-type algorithm is proposed for estimating the parameters that seems to provide some advantages over a direct maximization of the likelihood. Finally, we propose results of a simulation study and an example of real data for illustration.

1. Introduction

Linear mixed effects models are the most common statistical tools for analyzing repeated measurement data and in particular, longitudinal data in biomedical, agricultural, environmental and also in economics and social sciences. In the models, independent variables are often measured with non-negligible errors (Davidian and Giltinan, [16]). Hence considerable interest has been focused on the study of the estimation of parameters in measurement error models. Main references on the subject include Armstrong [4], Fuller [19], Stefanski and Carroll [25], Cheng and Van Ness [13], Zhong et al. [29], Carroll et al. [12] and Zare et al. [27]. Let y_i be an observed continues response. We assume the following linear mixed measurement error model as:

$$\begin{aligned} y_i &= \alpha + \beta x_i + b_i z_i + e_i, \\ X_i &= x_i + u_i, \quad i = 1, \ldots, n \end{aligned} \tag{1}$$

where α and β are fixed parameters and b_i's are random effects with $b_i \overset{\text{i.i.d.}}{\sim} N(0, \sigma_b^2)$. In this model X_i is the observed value of the x_i with measurement error u_i. Furthermore, it is assumed that $(e_i, u_i)' \overset{\text{i.i.d.}}{\sim} N_2((0,0)', \text{diag}\{\sigma_e^2, \sigma_u^2\})$, with i.i.d. means independent and identically distributed. In the structural model, it is also assumed that $x_i \overset{\text{i.i.d.}}{\sim} N(\mu_x, \sigma_x^2)$, and $\{x, e, u, b\}$ are mutually independent (see, Fuller [19]). There are several proposals of estimation for mixed effects models, for example see, Davidian and Giltinan [15,16], Karcher and Wang [23], Demidenko and Stukel [17] and Vonesh et al. [26]. Zare et al. [27] studied the estimation problem for the functional mixed measurement error model. Under normality assumption, they

* Corresponding author.
E-mail addresses: amene.kheradmandi@gmail.com (A. Kheradmandi), rasekh_a@scu.ac.ir (A. Rasekh).

http://dx.doi.org/10.1016/j.jmva.2014.12.007

applied the corrected score method of Nakamura [24] to obtain the estimators of the parameters. Cui et al. [14] derived the moments of estimators for the parameters in mixed effects model with error in variables.

On the other hand, the normality (symmetry) assumption is a routine but possibly restrictive assumption for different statistical models including the linear mixed measurement error models. In recent years, considerable interest has focused on the models relaxing normality assumption and incorporating asymmetry. Azzalini [6] introduces skew-normal distribution, extending usual normal model in order to avoid data transformation. The univariate skew-normal density function with location parameter μ, scale parameter σ^2 and skewness parameter λ, is defined by

$$f(x;\mu,\sigma^2,\lambda)=2\phi\left(\frac{x-\mu}{\sigma}\right)\Phi\left(\lambda\frac{x-\mu}{\sigma}\right),\quad x,\mu,\lambda\in R,\ \sigma>0$$

where $\phi(.)$ and $\Phi(.)$ denote the probability density function and cumulative distribution function, respectively, of the normal distribution. A random variable $Z=\frac{x-\mu}{\sigma}$ follows a standard skew-normal distribution with $\mu=0$ and $\sigma^2=1$, which is denoted by $SN(\lambda)$. The skew-normal distribution has the following properties:

(i) $E(X)=\mu+\sqrt{\frac{2}{\pi}}\frac{\lambda}{\sqrt{1+\lambda^2}}$.
(ii) $\mathrm{Var}(X)=(1-\frac{2\lambda}{\pi(1+\lambda^2)})\sigma^2$.
(iii) $\nu=\frac{1}{2}(4-\pi)(\frac{E^2(X)}{\mathrm{Var}(X)})^{\frac{3}{2}}$ and $\kappa=2(\pi-3)(\frac{E^2(X)}{\mathrm{Var}(X)})^2$, where ν and κ are asymmetry and kurtosis indexes, respectively.
(iv) If $\lambda=0$ then $X\sim N(\mu,\sigma^2)$.
(v) As pointed out by Henze [21], if $Z\sim SN(\lambda)$ then

$$Z\overset{d}{=}\delta|Z_0|+(1-\delta^2)^{\frac{1}{2}}Z_1. \tag{2}$$

where $Z_0\sim N(0,1)$ and $Z_1\sim N(0,1)$ are independent variables, $\delta=\frac{\lambda}{\sqrt{(1+\lambda^2)}}$ and $\overset{d}{=}$ means "distributed as".

Other properties of this distribution and its variations have been discussed by several authors including Azzalini [7], Henze [21], Azzalini and Capitanio [9], Azzalini and Dalla Valla [10], Arnold and Beaver [5] and Azzalini [8]. Arellano-Valle et al. [3] define a class of skew-normal measurement error models for a linear regression model. Skew-normal linear mixed models are introduced in Arellano-Valle et al. [1] and Bolfarine et al. [11] consider influence diagnostics for this model. In this paper, we define skew-normal linear mixed measurement error model that follows by replacing the normality assumptions by the assumptions that the random terms e_i and u_i, the random effect b_i and the independent variable x_i have the skew-normal distribution. We obtain the likelihood function of the observed data $\boldsymbol{Z}_i=(y_i,X_i),\ i=1,\ldots,n$. In addition, we consider some special cases where $\lambda_e=\lambda_u=\lambda_b=0$ or $\lambda_e=\lambda_u=\lambda_x=0$ in detail. We present an EM algorithm, which can overcome some difficulties detected by using direct maximization of the likelihood function, especially in terms of robustness with respect to starting values. The plan of the paper is as follows: In Section 2, we derive the marginal density of the observed data $\boldsymbol{Z}_i$ by integrating out the unobserved variables x_i and b_i. In Section 3, we present an EM-type algorithm and a simulation study is given in Section 4. To illustrate the usefulness of the proposed methods, an application to a real data set is reported in Section 5. Finally, Section 6 is dedicated to the concluding remarks.

2. The skew-normal linear mixed measurement error model

To consider a structural skew-normal mixed measurement error model, under the model defined by (1), we assume that:

$$\begin{aligned}&e_i\overset{\text{i.i.d.}}{\sim}SN(0,\sigma_e^2,\lambda_e),\qquad u_i\overset{\text{i.i.d.}}{\sim}SN(0,\sigma_u^2,\lambda_u),\\&x_i\overset{\text{i.i.d.}}{\sim}SN(0,\sigma_x^2,\lambda_x),\qquad b_i\overset{\text{i.i.d.}}{\sim}SN(0,\sigma_b^2,\lambda_b),\quad i=1,\ldots,n\end{aligned} \tag{3}$$

with $\{x,e,u,b\}$ are independent. Leading, under above model, we have

$$\begin{aligned}&y_i|x_i,b_i\sim SN(\alpha+\beta x_i+b_iz_i,\sigma_e^2,\lambda_e),\\&X_i|x_i,b_i\sim SN(x_i,\sigma_u^2,\lambda_u).\end{aligned}$$

In the following, we drop the subscript i in a sample unit to simplify notation. Hence, the conditional distribution of (y,X) given x and b can be computed by

$$f(y,X|x,b)=\frac{2^2}{\sigma_e\sigma_u}\phi\left(\frac{y-\alpha-\beta x-bz}{\sigma_e}\right)\phi\left(\frac{X-x}{\sigma_u}\right)\Phi\left(\lambda_e\frac{y-\alpha-\beta x-bz}{\sigma_e}\right)\Phi\left(\lambda_u\frac{X-x}{\sigma_u}\right).$$

Furthermore,

$$f(x,b)=\frac{2^2}{\sigma_x\sigma_b}\phi\left(\frac{x-\mu_x}{\sigma_x}\right)\phi\left(\frac{b}{\sigma_b}\right)\Phi\left(\lambda_x\frac{x-\mu_x}{\sigma_x}\right)\Phi\left(\lambda_b\frac{b}{\sigma_b}\right).$$