Bayesian Analysis for Multivariate Skew-Normal Simplex Mixed-Effects Models with Heterogeneous Dispersion

Xingde Duan[1], Shi Zhang[1], Wenzhuan Zhang[1](✉), and Xinli Miao[2]

[1] Guizhou University of Finance and Economics, Guiyang 550025, China
`zhangwenzhuan@mail.gufe.edu.cn`
[2] Chuxiong Normal University Chuxiong, Chuxiong 675000, China

Abstract. Continuous proportional data frequently appears in many areas of research, where proportional outcome are in the open interval $(0, 1)$. Simplex mixed-effects model is a powerful tool for modeling longitudinal continuous proportional data; however, the normality assumption of random effects in classic simplex mixed-effects model may be questionable in the analysis of skewed data. In this paper, we relax the normality assumption of random effects by specifying the random-effect distribution with the multivariate skew-normal distribution in mixed-effect model and simultaneously model the dispersion parameter (heterogeneity) in mixed-effect model. An efficient Markov chain Monte Carlo algorithm that combines the block Gibbs sampler, the Metropolis-Hastings algorithm and the data-augmentation technique is proposed for producing the joint Bayesian estimates of unknown parameters and random effects. The Deviance Information Criterion (DIC), as a popular model comparison criterion, is employed to select better model. The proposed methodology is illustrated by several simulation studies and a real example.

Keywords: Simplex distribution · Gibbs sampler · Metropolis–Hastings algorithm · Proportional data · Model selection

1 Introduction

Various random-effects models and marginal models based on simplex distribution are two classes of popular tool for modeling longitudinal continuous proportional data; therefore, statistical inference on these complex simplex models have been drawing much attention in the last two decades. For example, Song and

---

This research was supported in part by grants from the National Natural Science Foundation of China (11501073), the Science and Technology Foundation of Guizhou Province ([2020]1Y009), the talents for scientific research project of Guizhou University of Finance and Economics (2018YJ104).

H. Song and D. Jiang (Eds.): SIMUtools 2020, LNICST 369, pp. 313–326, 2021.
https://doi.org/10.1007/978-3-030-72792-5_27

Tan [12] and Song et al. [11] proposed the simplex marginal model with the constant and varying dispersion parameter under the GEE framework, respectively; Qiu et al. [9] proposed a simplex generalized linear mixed model by using the PQL/REML inference; Zhang and Wei [16] developed the stochastic approximation (SA) algorithm for simplex distribution nonlinear mixed models; Zhao et al. [17] studied Bayesian analysis of simplex distribution nonlinear mixed models based on MCMC algorithm; Bandyopadhyay et al. [3] proposed the augmented general proportion density (GPD) random effects model for clustered proportion data; Bonat et al. [4] investigated a class of simplex mixed models based on likelihood analysis. However, the random effects in all the above mentioned mixed-effects models follow the multivariate normal distribution. It is difficult for the normality assumption of random effects to capture the features of skewed and heavy-tailed data.

Recently, the skew-normal distribution is widely incorporated into various mixed effects models to relax the normality assumption of random effects and response variables. For example, Sahu et al. [10] proposed the Bayesian version of the multivariate skew-normal (SN) distributions; Arellano-Valle et al. [2] presented skew-normal linear mixed models by assuming the random effects and response variables to be a multivariate SN; Michaelis et al. [8] proposed a new class of multivariate distributional regression with skewed responses and skewed random effects by introducing the skew-normal and skew-t distribution; Xing et al. [14] proposed a two-part mixed-effects model with logistic mixed-effects model on the occurrence of positive values and linear mixed-effects model with skew-t (ST) and skew-normal (SN) distributions on the continuous positive values; Han et al. [6] proposed a class of skew-normal nonlinear mixed-effects joint models in the presence of covariates measured with errors to analyze complex longitudinal data; Zhang et al. [15] proposed the quantile regression-based partially linear mixed-effects joint models for longitudinal data with multiple features. However, to the best of our knowledge, there is little work done for Bayesian analysis of simplex mixed-effects model with random effects following the multivariate skew-normal distribution. Therefore, a novel Bayesian approach to simplex mixed-effects model is developed by specifying the random-effect distribution with the multivariate skew-normal distribution.

On the other hand, modelling dispersion parameter has been considered by different authors using various heterogeneous dispersion model. For example, see Artes and Jørgensen [1], Song et al. [11], Duan et al. [5]. In particular, Duan et al. [5] recommended the usage of modeling heterogeneous dispersion in a semiparametric simplex regression model when the homogeneous dispersion assumption is uncertain. Therefore, we introduce a new model for longitudinal continuous proportional data by incorporating multivariate skew-normal random effects into simplex regression models with heterogeneous dispersion. Also, an efficient Markov chain Monte Carlo algorithm that combines the block Gibbs sampler, the Metropolis-Hastings algorithm and the data-augmentation technique is developed.

2 Model

2.1 The Multivariate Skew-Normal Distribution

In this section, we first introduce the following notations. Let $N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ stand for a k-variate normal distribution with k-dimensional mean vector $\boldsymbol{\mu}$ and $k \times k$ covariance matrix $\boldsymbol{\Sigma}$. In what follows, the probability density function and cumulative distribution function of $N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ valued at z will be denoted by $\phi_k(z|\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\Phi_k(z|\boldsymbol{\mu}, \boldsymbol{\Sigma})$; respectively. For abbreviation, we let $\operatorname{diag}(a_1, \ldots, a_k)$ denote a diagonal matrix with elements $a_1, \ldots, a_k$ and I_k denote a $k \times k$ identity matrix. In this work, we implement the Bayesian version of the multivariate SN distributions proposed by Sahu et al. [10], and the corresponding probability density function is given as follow.

Definition 1. A k-dimensional random vector $\mathbf{Z}$ is subject to a k-variate skew-normal distribution, if its probability density function can be expressed in the following form

$$\begin{aligned} f(\mathbf{z}|\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta}) = {} & 2^k \phi_k(\mathbf{z}|\boldsymbol{\mu}, \boldsymbol{\Sigma} + \boldsymbol{\Delta}\boldsymbol{\Delta}^T) \\ & \times \Phi_k(\boldsymbol{\Delta}^T(\boldsymbol{\Sigma} + \boldsymbol{\Delta}\boldsymbol{\Delta}^T)^{-1}(\mathbf{z} - \boldsymbol{\mu})|0, (I_k + \boldsymbol{\Delta}^T \boldsymbol{\Sigma}^{-1} \boldsymbol{\Delta})^{-1}), \end{aligned} \tag{1}$$

where $\boldsymbol{\mu} \in \mathbb{R}^k$ is the location vector, $\boldsymbol{\Sigma}$ is the $k \times k$ scale positive-definite matrix, $\boldsymbol{\Delta} = \operatorname{diag}(\boldsymbol{\delta})$ with diagonal elements $\boldsymbol{\delta} = (\delta_1, \ldots \delta_k)^T \in \mathbb{R}^k$ is the $k \times k$ diagonal skewness matrix, $\mathbf{z}$ is a real k-dimensional vector.

From now on, the aforementioned distribution can be denoted by $\mathbf{Z} \sim SN_k(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta})$. Clearly, the multivariate normal distribution is a special distribution of multivariate SN distribution, if $\boldsymbol{\Delta} = \mathbf{0}$. Following Arellano-Valle et al. [2], we introduce the following propositions.

Proposition 1. If $\mathbf{Z} \sim SN_k(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta})$, then $\mathbf{Z} = \boldsymbol{\Delta}|\mathbf{Z_0}| + \mathbf{Z}_1$, where $\mathbf{Z_0} \sim N_k(\mathbf{0}, I_k)$ and $\mathbf{Z}_1 \sim N_k(\boldsymbol{\mu}, \boldsymbol{\Sigma})$ and $\mathbf{Z_0}$ and $\mathbf{Z_1}$ are independent.

A proof of this proposition has been given in Sahu et al. [10] and Arellano-Valle et al. [2]. This proposition represents the stochastic representation which can readily generate the random variables from the multivariate SN distribution. In addition, this proposition has the following hierarchical structure: $\mathbf{Z}|\mathbf{Z}_0 = \mathbf{z_0} \sim N_k(\boldsymbol{\mu} + \boldsymbol{\Delta}|\mathbf{z_0}|, \boldsymbol{\Sigma})$ and $\mathbf{Z_0} \sim N_k(\mathbf{0}, I_k)$. Under the Bayesian framework, this hierarchical structure readily facilitates the Markov chain Monte Carlo algorithm in the skew-normal distribution.

Proposition 2. If $\mathbf{Z} \sim SN_k(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\Delta})$, then $E(\mathbf{Z}) = \boldsymbol{\mu} + \sqrt{\frac{2}{\pi}}\boldsymbol{\delta}$ and $\operatorname{var}(\mathbf{Z}) = \boldsymbol{\Sigma} + \left(1 - \frac{2}{\pi}\right)\boldsymbol{\Delta}^2$.

Similarly, the proof of this proposition refer to Arellano-Valle et al. [2]. In this work, we assume the location parameter $\boldsymbol{\mu} = -\sqrt{\frac{2}{\pi}}\boldsymbol{\delta}$ in order to construct a zero mean vector with respect to multivariate skew-normal distribution on random effects in the following mixed-effect model.

2.2 Model Formulation

To study a longitudinal proportional data based on regression-type technique, we assume that y_{ij} is proportional outcome of the ith patient measured at time $t_{ij}(i = 1, 2, \ldots, n,\ j = 1, \ldots, n_i)$. For each ith patient, $y_{i1}, y_{i2}, \ldots, y_{in_i}$ given $\mathbf{b_i}$, which is a $k \times 1$ vector of random effects characterized by the ith patient, are conditionally independent and each $y_{ij}|\mathbf{b_i}$ is assumed to be a simplex distribution with conditional means $\mu_{ij} = \mathrm{E}(Y_{ij}|\mathbf{b_i}) \in (0, 1)$ and dispersion parameter $\sigma_{ij}^2 > 0$:

$$\begin{cases} p(y_{ij}|\mathbf{b_i}, \sigma_{ij}^2, \beta, \gamma) = \left[2\pi\sigma_{ij}^2\{y_{ij}(1-y_{ij})\}^3\right]^{-1/2} \exp\left\{-\frac{1}{2\sigma_{ij}^2} d(y_{ij}; \mu_{ij})\right\} \\ \mathrm{logit}(\mu_{ij}) = \mathbf{x_{ij}^T}\boldsymbol{\beta} + \mathbf{z_{ij}^T}\mathbf{b_i}, \\ \log(\sigma_{ij}^2) = \mathbf{w_{ij}^T}\boldsymbol{\gamma}, \\ \mathbf{b_i} \sim SN_k\left(-\sqrt{2/\pi}\boldsymbol{\delta}, \boldsymbol{\Sigma}, \boldsymbol{\Delta}\right) \end{cases} \tag{2}$$

where $0 < y_{ij} < 1$, $d(y_{ij}; \mu_{ij}) = \frac{(y_{ij}-\mu_{ij})^2}{y_{ij}(1-y_{ij})\mu_{ij}^2(1-\mu_{ij})^2}$, $\mathrm{logit}(\mu_{ij}) = \log\left(\frac{\mu_{ij}}{1-\mu_{ij}}\right)$, $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown parameters for the fixed effects in the mean model, $\boldsymbol{\gamma}$ is a $q \times 1$ vector of unknown parameters to be estimated in the dispersion model, $\mathbf{x_{ij}}$, $\mathbf{z_{ij}}$ and $\mathbf{w_{ij}}$ are the $p \times 1$, $k \times 1$ and $q \times 1$ vector of design explanatory variables. Also, $\boldsymbol{\Sigma}$, $\boldsymbol{\delta}$ and $\boldsymbol{\Delta}$ are given by the Definition 1. Here, $\mathbf{w_{ij}}$ may comprise some or all of the variables in $\mathbf{x_{ij}}$. Thus, the model defined in (2) is referred to as a multivariate skew-norml simplex mixed-effects models with heterogeneous dispersion. In what follows, we will denote by $y_{ij}|\mathbf{b_i} \sim \mathrm{S}^{-1}(\mu_{ij}, \sigma_{ij}^2)$, if $y_{ij}|\mathbf{b_i}$ follows a simplex distribution with conditional means μ_{ij} and dispersion parameter σ_{ij}^2. For the convenience of the implementation of the MCMC algorithms, it can be seen from Proposition 1 and Proposition 2 that $\mathbf{b_i} \sim SN_k\left(-\sqrt{2/\pi}\delta, \mathbf{\Sigma}, \mathbf{\Delta}\right)$ in model (2) can derived as the simple hierarchical structure:

$$\mathbf{b_i}|\mathbf{u_i} \sim N_k(-\sqrt{2/\pi}\boldsymbol{\delta} + \boldsymbol{\Delta}\mathbf{u_i}, \boldsymbol{\Sigma}), \mathbf{u_i} \sim N_k(0, I_k)\mathrm{I}\{\mathbf{u_i} > 0\}, \tag{3}$$

where $\mathbf{u_i}$ is $k \times 1$ vector of unobserved latent variable, $\mathrm{I}\{\cdot\}$ represents an indicator function.

3 Bayesian Analysis of Mixed-Effects Model

Let $\mathbf{Y} = \{\mathbf{y_{ij}} : i = 1, \cdots, n, j = 1, \ldots, n_i\}$, $\mathbf{X} = \{\mathbf{x_{ij}} : i = 1, \cdots, n, j = 1, \ldots, n_i\}$, $\mathbf{Z} = \{\mathbf{z_{ij}} : i = 1, \cdots, n, j = 1, \ldots, n_i\}$, $\mathbf{W} = \{\mathbf{w_{ij}} : i = 1, \cdots, n, j = 1, \ldots, n_i\}$, $\mathbf{y_i} = \{y_{i1}, \ldots, y_{in_i}\}$, $\mathbf{x_i} = \{x_{i1}, \ldots, x_{in_i}\}$, $\mathbf{z_i} = \{z_{i1}, \ldots, z_{in_i}\}$, $\mathbf{w_i} = \{w_{i1}, \ldots, w_{in_i}\}$, $\mathbf{b} = \{\mathbf{b_i} : i = 1, \cdots, n\}$, $\mathbf{u} = \{\mathbf{u_i} : i = 1, \cdots, n\}$, $\mathbf{1_k} = (1, \ldots, 1)^T$ and $\theta = (\boldsymbol{\beta}, \boldsymbol{\gamma}, \boldsymbol{\Sigma}, \boldsymbol{\delta})$. Clearly, it is difficult to generate the random sample from the marginal posterior distribution $p(\theta|\mathbf{Y}, \mathbf{X}, \mathbf{Z}, \mathbf{W})$ because of high-dimensional integral with respect of random effects $\mathbf{b}$ and latent variable $\mathbf{u}$ involved. To address this issue, we adopt the data-augmentation strategy to augment the observed data $(\mathbf{Y}, \mathbf{X}, \mathbf{Z}, \mathbf{W})$ with the unobserved data $\mathbf{b}$ and $\mathbf{u}$. Specifically, an efficient Markov chain Monte Carlo algorithm is used to generate random observations from the joint posterior distribution $p\,(\theta, \mathbf{b}, \mathbf{u}|\mathbf{Y}, \mathbf{X}, \mathbf{Z}, \mathbf{W})$, which is proportional to

$$\left\{\prod_{i=1}^{n}[\prod_{j=1}^{n_i} p\,(y_{ij}|\mathbf{x_{ij}},\mathbf{z_{ij}},\mathbf{w_{ij}}, \mathbf{b_i}, \mathbf{u_i}, \theta)]p\,(\mathbf{b_i}|\mathbf{u_i}, \theta)\, p\,(\mathbf{u_i}|\mathbf{u_i} > 0)\right\} p\,(\theta)\,, \qquad (4)$$

where $p\,(\theta)$ is the prior distribution for unknown parameter θ. Furthermore, the prior distribution $p\,(\theta)$ is specified as follows:

$$\begin{aligned} &p\,(\boldsymbol{\beta}) \overset{D}{=} N_p\left(\boldsymbol{\beta}^{\mathbf{0}},\mathbf{H}_{\beta\mathbf{0}}\right), p\,(\boldsymbol{\gamma}) \overset{D}{=} N_q\left(\boldsymbol{\gamma}^{\mathbf{0}},\mathbf{H}_{\gamma\mathbf{0}}\right), p\,(\boldsymbol{\delta}) \overset{D}{=} N_k\left(\boldsymbol{\delta}^{\mathbf{0}},\mathbf{H}_{\delta\mathbf{0}}\right), \\ &p\,(\boldsymbol{\Sigma}) \overset{D}{=} IW_k\,(\rho_0, \mathbf{R_0})\,, \end{aligned} \qquad (5)$$

where $\boldsymbol{\beta}^{\mathbf{0}}$, $\mathbf{H}_{\beta\mathbf{0}}$, $\boldsymbol{\gamma}^{\mathbf{0}}$, $\mathbf{H}_{\gamma\mathbf{0}}$, $\boldsymbol{\delta}^{\mathbf{0}}$, $\mathbf{H}_{\delta\mathbf{0}}$, ρ_0, $\mathbf{R_0}$ are hyperparameters whose values are given based on prior information; and IW_k denotes the k-dimensional inverse Wishart distribution.

3.1 Full Conditional Distributions

In this section, we derive the full conditional distributions of interest in the MCMC sampling procedures as follows.

A tedious calculation gives that the full conditional distribution for latent variable $\mathbf{u_i}$ can be written as

$$\begin{aligned} &p\,(\mathbf{u_i}|\mathbf{y_i},\mathbf{x_i},\mathbf{z_i},\mathbf{w_i},\mathbf{b_i}, \theta) \propto \\ &\exp\left\{-\tfrac{1}{2}\mathbf{u_i^T u_i} - \tfrac{1}{2}(\mathbf{b_i} - \boldsymbol{\Delta}(\mathbf{u_i} - \sqrt{2/\pi}\cdot\mathbf{1_k}))^T \varSigma^{-1}(\mathbf{b_i} - \boldsymbol{\Delta}(\mathbf{u_i} - \sqrt{2/\pi}\cdot\mathbf{1_k}))\right\} \mathrm{I}(\mathbf{u_i} > 0), \end{aligned}$$

which gives rise to

$$\mathbf{u_i}|\mathbf{y_i},\mathbf{x_i},\ \mathbf{z_i},\mathbf{w_i},\mathbf{b_i}, \theta \sim N_k(\mathbf{A}_{ui}\mathbf{a}_{ui}, \mathbf{A}_{ui})\mathrm{I}(\mathbf{u_i} > 0),\ i = 1, 2, \ldots n \qquad (6)$$

where $\mathbf{A}_{ui} = (I_k + \boldsymbol{\Delta}\boldsymbol{\Sigma}^{-1}\boldsymbol{\Delta})^{-1}$, $\mathbf{a}_{ui} = \sqrt{2/\pi}\boldsymbol{\Delta}\boldsymbol{\Sigma}^{-1}\boldsymbol{\delta} + \boldsymbol{\Delta}\boldsymbol{\Sigma}^{-1}\mathbf{b}_i$.

The full conditional distribution for random effect $\mathbf{b_i}$ is proportional to

$$\begin{aligned} &p\,(\mathbf{b_i}|\mathbf{y_i},\mathbf{x_i},\mathbf{z_i},\mathbf{w_i},\mathbf{u_i}, \theta) \propto \\ &\exp\left\{-\tfrac{1}{2}[\sum_{j=1}^{n_i}\frac{d(y_{ij};\mu_{ij})}{\sigma_{ij}^2} + (\mathbf{b}_i - \boldsymbol{\Delta}(\mathbf{u}_i - \sqrt{2/\pi}.\mathbf{1}_k))^T \varSigma^{-1}(\mathbf{b}_i - \boldsymbol{\Delta}(\mathbf{u}_i - \sqrt{2/\pi}.\mathbf{1}_k))]\right\}. \end{aligned} \qquad (7)$$

The full conditional distribution for scale positive-definite matrix $\boldsymbol{\Sigma}$ is proportional to

$$\begin{aligned} &p\,(\boldsymbol{\Sigma}|\boldsymbol{Y}, \boldsymbol{X}, \boldsymbol{Z}, \boldsymbol{W}, \boldsymbol{b}, \boldsymbol{u}, \theta) \propto |\boldsymbol{\Sigma}|^{-(n+k+\rho_0+1)/2} \\ &\exp\left\{-\tfrac{1}{2}\mathrm{tr}\left[\varSigma^{-1}(\mathbf{R}_0 + \sum_{i=1}^{n}(\mathbf{b}_i - \boldsymbol{\Delta}(\mathbf{u}_i - \sqrt{2/\pi}.\mathbf{1}_k))(\mathbf{b}_i - \boldsymbol{\Delta}(\mathbf{u}_i - \sqrt{2/\pi}.\mathbf{1}_k))^T)\right]\right\} \end{aligned}$$

.

Clearly, the full conditional distribution for $\boldsymbol{\Sigma}$ has the following inverse Wishart distribution:

$$\begin{aligned}&\boldsymbol{\Sigma}|\boldsymbol{Y},\boldsymbol{X},\boldsymbol{Z},\boldsymbol{W},\boldsymbol{b},\boldsymbol{u},\boldsymbol{\beta},\boldsymbol{\gamma},\boldsymbol{\delta}\sim\\&IW_k\left(\mathbf{R}_0+\sum_{i=1}^{n}(\mathbf{b}_i-\boldsymbol{\Delta}(\mathbf{u}_i-\sqrt{2/\pi}.\mathbf{1}_k))(\mathbf{b}_i-\boldsymbol{\Delta}(\mathbf{u}_i-\sqrt{2/\pi}.\mathbf{1}_k))^T, n+\rho_0\right).\end{aligned} \tag{8}$$

The full conditional distribution for parameters $\boldsymbol{\beta}$ in the mean model is proportional to

$$p\left(\boldsymbol{\beta}|\mathbf{Y},\mathbf{X},\mathbf{Z},\mathbf{W},\mathbf{b},\mathbf{u},\gamma,\boldsymbol{\Sigma},\boldsymbol{\delta}\right)\propto\exp\left\{-\frac{1}{2}\left[\sum_{i=1}^{n}\sum_{j=1}^{n_i}\frac{d\left(y_{ij};\mu_{ij}\right)}{\sigma_{ij}^2}+\left(\boldsymbol{\beta}-\boldsymbol{\beta}^{\mathbf{0}}\right)^T\mathbf{H}_{\boldsymbol{\beta}\mathbf{0}}^{-\mathbf{1}}\left(\boldsymbol{\beta}-\boldsymbol{\beta}^{\mathbf{0}}\right)\right]\right\}. \tag{9}$$

Clearly, the conditional distribution for skewness parameter $\boldsymbol{\delta}$ can be written as

$$\begin{aligned}&p\left(\boldsymbol{\delta}|\mathbf{Y},\mathbf{X},\mathbf{Z},\mathbf{W},\mathbf{b},\mathbf{u},\gamma,\boldsymbol{\Sigma},\boldsymbol{\beta}\right)\propto\\&\exp\left\{-\tfrac{1}{2}\left[\left(\boldsymbol{\delta}-\boldsymbol{\delta}^{\mathbf{0}}\right)^T\mathbf{H}_{\boldsymbol{\delta}\mathbf{0}}^{-\mathbf{1}}\left(\boldsymbol{\delta}-\boldsymbol{\delta}^{\mathbf{0}}\right)+\sum_{i=1}^{n}\left(\mathbf{b}_i-\boldsymbol{\Delta}\left(\mathbf{u}_i-\sqrt{2/\pi}.\mathbf{1}_k\right)\right)^T\Sigma^{-1}\left(\mathbf{b}_i-\boldsymbol{\Delta}\left(\mathbf{u}_i-\sqrt{2/\pi}.\mathbf{1}_k\right)\right)\right]\right\}\end{aligned}$$

which yields the multivariate normal distribution as follows:

$$\boldsymbol{\delta}|\mathbf{Y},\mathbf{X},\mathbf{Z},\mathbf{W},\mathbf{b},\mathbf{u},\boldsymbol{\gamma},\boldsymbol{\Sigma},\boldsymbol{\beta}\sim N_k(\mathbf{A}_\delta\mathbf{a}_\delta,\mathbf{A}_\delta), \tag{10}$$

where $\mathbf{A}_\delta = \left[\mathbf{H}_{\delta 0}^{-1}+\sum_{i=1}^{n}(\sqrt{2/\pi}\mathbf{I_k}-\text{diag}(\mathbf{u_i}))^T\boldsymbol{\Sigma}^{-1}(\sqrt{2/\pi}\mathbf{I_k}-\text{diag}(\mathbf{u_i}))\right]^{-1}$ and $\mathbf{a}_{\boldsymbol{\delta}}=\mathbf{H}_{\delta 0}^{-1}\delta_{\mathbf{0}}-\sum_{i=1}^{n}(\sqrt{2/\pi}\mathbf{I_k}-\text{diag}(\mathbf{u_i}))^T\boldsymbol{\Sigma}^{-1}\mathbf{b_i})$.

The full conditional distribution for parameter γ in the dispersion model is proportional to

$$\begin{aligned}&p\left(\gamma|\mathbf{Y},\mathbf{X},\mathbf{Z},\mathbf{W},\mathbf{b},\mathbf{u},\boldsymbol{\beta},\boldsymbol{\Sigma},\boldsymbol{\delta}\right)\propto\\&\exp\left\{-\tfrac{1}{2}\left[\sum_{i=1}^{n}\sum_{j=1}^{n_i}\log\sigma_{ij}^2+\sum_{i=1}^{n}\sum_{j=1}^{n_i}\tfrac{d(y_{ij};\mu_{ij})}{\sigma_{ij}^2}+\left(\boldsymbol{\gamma}-\boldsymbol{\gamma}^{\mathbf{0}}\right)^T\mathbf{H}_{\boldsymbol{\gamma}\mathbf{0}}^{-\mathbf{1}}\left(\boldsymbol{\gamma}-\boldsymbol{\gamma}^{\mathbf{0}}\right)\right]\right\}.\end{aligned} \tag{11}$$

Clearly, it is straightforward to generate random observations from the standard and familiar conditional distributions given in (6), (8) and (10); however, it is rather difficult to draw random observations from the unfamiliar and complicated conditional distributions given in (7), (9) and (11). Therefore, we will implement the Metropolis-Hastings (MH) algorithm to deal with these sampling difficulties as follows.

3.2 MH Algorithm

For the full conditional distribution $p\left(\mathbf{b_i}|\mathbf{y_i},\mathbf{x_i},\mathbf{z_i},\mathbf{w_i},\mathbf{u_i},\theta\right)$, $p(\boldsymbol{\beta}|\mathbf{Y},\mathbf{X},\mathbf{Z},\mathbf{W},\mathbf{b},\mathbf{u},\boldsymbol{\gamma},\boldsymbol{\Sigma},\boldsymbol{\delta})$ and $p\left(\gamma|\mathbf{Y},\mathbf{X},\mathbf{Z},\mathbf{W},\mathbf{b},\mathbf{u},\boldsymbol{\beta},\boldsymbol{\Sigma},\boldsymbol{\delta}\right)$, we choose the normal distribution and multivariate normal distribution as the proposal distribution.

The procedure for implementation of MH algorithm is presented as follows. Given the current value $\mathbf{b}_i^{(t)}$, $\boldsymbol{\beta}^{(t)}$ and $\boldsymbol{\gamma}^{(t)}$, some new candidates $\mathbf{b}_i^*$, $\boldsymbol{\beta}^*$ and $\boldsymbol{\gamma}^*$ generated from $N_k(\mathbf{b}_i^{(t)}, \sigma_{b_i}^2 \Omega_{b_i})$, $N\left(\boldsymbol{\beta}^{(t)}, \sigma_\beta^2 \Omega_\beta\right)$ and $N\left(\boldsymbol{\gamma}^{(t)}, \sigma_\gamma^2 \Omega_\gamma\right)$ can calculate the following accepted probability

$$\min\left\{1, \frac{p\left(\mathbf{b}_i^* | \mathbf{y_i}, \mathbf{x_i}, \mathbf{z_i}, \mathbf{w_i}, \mathbf{u_i}, \theta\right)}{p\left(\mathbf{b}_i^{(t)} | \mathbf{y_i}, \mathbf{x_i}, \mathbf{z_i}, \mathbf{w_i}, \mathbf{u_i}, \theta\right)}\right\},$$

$$\min\left\{1, \frac{p\left(\boldsymbol{\beta}^* | \mathbf{Y}, \mathbf{X}, \mathbf{Z}, \mathbf{W}, \mathbf{b}, \mathbf{u}, \gamma, \boldsymbol{\Sigma}, \boldsymbol{\delta}\right)}{p\left(\boldsymbol{\beta}^{(t)} | \mathbf{Y}, \mathbf{X}, \mathbf{Z}, \mathbf{W}, \mathbf{b}, \mathbf{u}, \gamma, \boldsymbol{\Sigma}, \boldsymbol{\delta}\right)}\right\},$$

$$\min\left\{1, \frac{p\left(\boldsymbol{\gamma}^* | \mathbf{Y}, \mathbf{X}, \mathbf{Z}, \mathbf{W}, \mathbf{b}, \mathbf{u}, \boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{\delta}\right)}{p\left(\boldsymbol{\gamma}^{(t)} | \mathbf{Y}, \mathbf{X}, \mathbf{Z}, \mathbf{W}, \mathbf{b}, \mathbf{u}, \boldsymbol{\beta}, \boldsymbol{\Sigma}, \boldsymbol{\delta}\right)}\right\},$$

where

$$\Omega_{b_i}^{-1} = \sum_{j=1}^{n_i} \left[3\mu_{ij}\left(1-\mu_{ij}\right) + \frac{1}{\sigma_{ij}^2 \mu_{ij}\left(1-\mu_{ij}\right)}\right] z_{ij} z_{ij}^T + \boldsymbol{\Sigma}^{-1},$$

$$\Omega_\beta^{-1} = \sum_{i=1}^{n} \sum_{j=1}^{n_i} \left[3\mu_{ij}\left(1-\mu_{ij}\right) + \frac{1}{\sigma_{ij}^2 \mu_{ij}\left(1-\mu_{ij}\right)}\right] x_{ij} x_{ij}^T + \mathbf{H}_{\boldsymbol{\beta}\mathbf{0}}^{\mathbf{-1}},$$

$$\Omega_\gamma^{-1} = \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n_i} w_{ij} w_{ij}^T + \mathbf{H}_{\boldsymbol{\gamma}\mathbf{0}}^{\mathbf{-1}},$$

and $\sigma_{b_i}^2$, σ_β^2 and σ_γ^2 are referred to as the tuned variance coefficients.

3.3 Bayesian Model Selection

Deviance Information Criterion (DIC) proposed by Spiegelhalter et al. [13] is a popular criteria for model selection under the Bayesian framework. It is easily to calculate the estimator for DIC in the complex model by using MCMC algorithm . Thus, it follows from the idea of Spiegelhalter et al. [13] that the DIC under our proposed model is defined as

$$\mathrm{DIC} = \overline{D(\theta)} + p_D$$

where $\overline{D(\theta)} = -2\mathrm{E}_{\theta,\mathbf{b},\mathbf{u}}\left\{\log p\left(\mathbf{Y}, \mathbf{b}, \mathbf{u} | \theta, \mathbf{X}, \mathbf{Z}, \mathbf{W}\right) | \mathbf{Y}, \mathbf{X}, \mathbf{Z}, \mathbf{W}\right\}$ is the posterior mean of deviance $D(\theta) = -2\mathrm{E}_{\mathbf{b},\mathbf{u}}\left\{\log p\left(\mathbf{Y}, \mathbf{b}, \mathbf{u} | \theta, \mathbf{X}, \mathbf{Z}, \mathbf{W}\right) | \mathbf{Y}, \mathbf{X}, \mathbf{Z}, \mathbf{W}\right\}$, $p_D = \overline{D(\theta)} - D\left(\widehat{\theta}\right)$ is a Bayesian measure of model complexity, $\widehat{\theta}$ is the posterior mean of θ . Then, DIC can be simplied as

$$\begin{aligned}
\mathrm{DIC} &= 2\overline{D(\theta)} - D\left(\widehat{\theta}\right) \\
&= -4\mathrm{E}_{\theta,\mathbf{b},\mathbf{u}}\left\{\log p\left(\mathbf{Y}, \mathbf{b}, \mathbf{u} | \theta, \mathbf{X}, \mathbf{Z}, \mathbf{W}\right) | \mathbf{Y}, \mathbf{X}, \mathbf{Z}, \mathbf{W}\right\} \\
&\quad + 2\mathrm{E}_{\mathbf{b},\mathbf{u}}\left\{\log p\left(\mathbf{Y}, \mathbf{b}, \mathbf{u} | \widehat{\theta}, \mathbf{X}, \mathbf{Z}, \mathbf{W}\right) | \mathbf{Y}, \mathbf{X}, \mathbf{Z}, \mathbf{W}\right\}.
\end{aligned}$$

Clearly, the aforementioned expression for DIC involve the complicated expectation algorithm; therefore, we use Monte Carlo method to approximate the estimator for DIC as follows:

$$\widehat{DIC} = -\frac{4}{T}\sum_{t=1}^{T}\log p\left(\mathbf{Y}, \mathbf{b}^{(t)}, \mathbf{u}^{(t)}|\theta^{(t)}, \mathbf{X}, \mathbf{Z}, \mathbf{W}\right)$$
$$+\frac{2}{T}\sum_{t=1}^{T}\log p\left(\mathbf{Y}, \mathbf{b}^{(t)}, \mathbf{u}^{(t)}|\widehat{\theta}, \mathbf{X}, \mathbf{Z}, \mathbf{W}\right).$$

4 Numerical Examples

4.1 Simulation Study

In this simulation study, we suppose that the longitudinal proportional data $y_{ij}(i = 1, 2, \ldots, n,\ j = 1, \ldots, n_i)$ given skew-normal random effect $\mathbf{b_i}$ follows the simplex distribution, $y_{ij}|\mathbf{b_i} \sim \overset{-1}{\mathrm{S}}(\mu_{ij}, \sigma^2_{ij})$. The structure of the mean model and the dispersion model are given by

$$\text{logit}(\mu_{ij}) = x_{1ij}\beta_1 + x_{2ij}\beta_2 + b_i$$

$$\log(\sigma^2_{ij}) = \gamma_1 + x_{2ij}\gamma_2,$$

where discrete covariate x_{1ij} is randomly drawn as $-1, 0, 1$, continuous covariate x_{2ij} is simulated from the uniform distribution $U[0, 1]$, the true values are set as $\beta = (\beta_1, \beta_2)^T = (0.5, 0.5)^T$ and $\gamma = (\gamma_1, \gamma_2)^T = (3, 2)^T$. Also, the random effect $\mathbf{b_i}$ is distributed as the skew-normal distribution, $\mathbf{b_i} \sim SN\left(-\sqrt{2/\pi}\delta, \Sigma, \Delta\right)$, where the true value of the skewness parameter δ is 3 and variance component Σ is 0.6.

In order to investigate the effect of different prior inputs on the Bayesian estimate for parameter, three types of prior information for β and γ are specified by

Type I: $\boldsymbol{\beta^0} = (0.5, 0.5)^T$, $\mathbf{H_{\beta 0}} = 0.25I_2$, $\boldsymbol{\gamma^0} = (3, 2)^T$, $\mathbf{H_{\gamma 0}} = 0.25I_2$, $\boldsymbol{\delta^0} \sim N(0, 1)$, $\mathbf{H_{\delta 0}} = 100$, $\mathbf{R_0} = 2$ and $\rho_0 = 5$, where I_2 stands for 2×2 identity matrix. This scenario represents a good prior information.

Type II: $\boldsymbol{\beta^0} = 1.5 \times (0.5, 0.5)^T$, $\mathbf{H_{\beta 0}} = 10I_2$, $\boldsymbol{\gamma^0} = 1.5 \times (3, 2)^T$, $\mathbf{H_{\gamma 0}} = 10I_2$; whilst the other hyperparameter were set to be the same as those given in Type I. This scenario shows an inaccurate prior information.

Type III: $\boldsymbol{\beta^0} = (0, 0)^T$, $\mathbf{H_{\beta 0}} = 100I_2$, $\boldsymbol{\gamma^0} = (0, 0)^T$, $\mathbf{H_{\gamma 0}} = 100I_2$. Similarly, the other hyperparameter were set to be the same as those given in Type I. This scenario regards a non-informative prior information.

For the above simulated data sets, we used the preceding proposed hybrid algorithm combining the Gibbs sampler and the Metropolis Hastings algorithm to evaluate the Bayesian estimates of unknown parameter and random effects based on different sample size $n = 50$ and $n = 100$. To monitor convergence of the proposed algorithm, we randomly chose and plotted a mixing process of three Markov chains for unknown parameters via three different starting values in a few test runs. Figure 1 indicated that the traces of three Markov chains for parameters β and γ with three different starting values mix well under type III prior input. Another useful tool for convergence diagnosis is the "estimated potential scale reduction (EPSR)" values obtained from three parallel chains of all the parameters generated with three different starting values in a few test replications. We observed that the EPSR values are less than 1.2 after about 1000 iterations in all test cases and plotted the EPSR values for all unknown parameters against iterations in a randomly selected case in Fig. 2. To be conservative, for each setting, 5000 observations collected after 5000 burn-ins in producing the Bayesian estimate based on 100 replications were reported in Table 1, where 'EST' denotes the mean of the estimates based on 100 replications, 'SD' is the standard deviation of the estimates based on 100 replications, 'RMS' is the root mean square between the estimates based on 100 replications and its true value. Table 1 showed that (i) Bayesian estimates for parameters β and γ are relatively close to the corresponding true values under various prior inputs with different sample size, which implies that the estimates are not sensitive to the prior inputs and sample size in our considered simulation studies; (ii) the skewness parameter δ gave some large deviation with small sample size; whereas the performance of Bayesian estimate for δ is satisfactory as the sample size goes larger; (iii) the values of 'SD' and 'RMS' are quite close, which implies that the estimated standard deviation is rather reliable when the Bias is small.

4.2 A Real Example

We reanalysed the longitudinal proportional data for a prospective ophthalmology study taken from Meyers et al. [7]. The prospective ophthalmology data were analyzed by different authors using various approaches. To compare our proposed approach with the existing approach, we simultaneously considered a normal simplex mixed-effects models with heterogeneous dispersion; that is, $\mathbf{b_i} \sim N(0, \varSigma)$. Thus, we called $\mathbf{b_i} \sim N(0, \varSigma)$ as the normal approach; whereas $\mathbf{b_i} \sim SN_k\left(-\sqrt{2/\pi}\delta, \mathbf{\Sigma}, \mathbf{\Delta}\right)$ was called as the skewed normal approach in the remainder of this paper. This prospective ophthalmology study described that three gas concentration levels of C_3F_8 were injected into the 31 patients' eye before surgery and percentage of remained gas volume in all patients were recorded at followed-up three to eight times over a 3-month period. We considered the following several variables such as percentage of remained gas volume, gas concentration of C_3F_8, and time for 31 patients. Our scientific interest of this study is to link percentage of remained gas volume with three initial gas concentration levels of C_3F_8 and time while accounting for which covariates lead to heterogeneity. Therefore, following the work of Song et al. [11], the structure of the mean parameter and the dispersion parameter in simplex distribution are modeled by

Table 1. Summary statistics of the estimates in the simulation study.

Type	Parameters	n=50			n=100		
		EST	SD	RMS	EST	SD	RMS
I	β_1	0.4930	0.0938	0.0935	0.4906	0.0571	0.0576
	β_2	0.5011	0.1570	0.1562	0.4848	0.1277	0.1279
	γ_1	2.9988	0.1177	0.1171	3.0028	0.0741	0.0738
	γ_2	1.9618	0.1305	0.1354	1.9713	0.1054	0.1087
	Σ	0.6783	0.2686	0.2785	0.6103	0.2525	0.2515
	δ	2.8102	0.7924	0.8109	3.0093	0.3534	0.3517
II	β_1	0.4882	0.0877	0.0881	0.5024	0.0604	0.0602
	β_2	0.5024	0.2189	0.2178	0.5079	0.1404	0.1399
	γ_1	2.9982	0.1245	0.1239	3.0083	0.0763	0.0764
	γ_2	2.0013	0.1718	0.1709	2.0036	0.1237	0.1232
	Σ	0.6771	0.2645	0.2742	0.6391	0.3218	0.3225
	δ	2.6431	1.0444	1.0987	2.9275	0.4292	0.4332
III	β_1	0.5092	0.1015	0.1014	0.5006	0.0523	0.0520
	β_2	0.5039	0.2047	0.2037	0.4942	0.1534	0.1527
	γ_1	3.0077	0.1161	0.1158	3.0191	0.0803	0.0821
	γ_2	2.0232	0.1696	0.1703	1.9999	0.1165	0.1159
	Σ	0.6617	0.2704	0.2761	0.5838	0.2276	0.2270
	δ	2.7506	0.8660	0.8970	3.0025	0.3035	0.3020

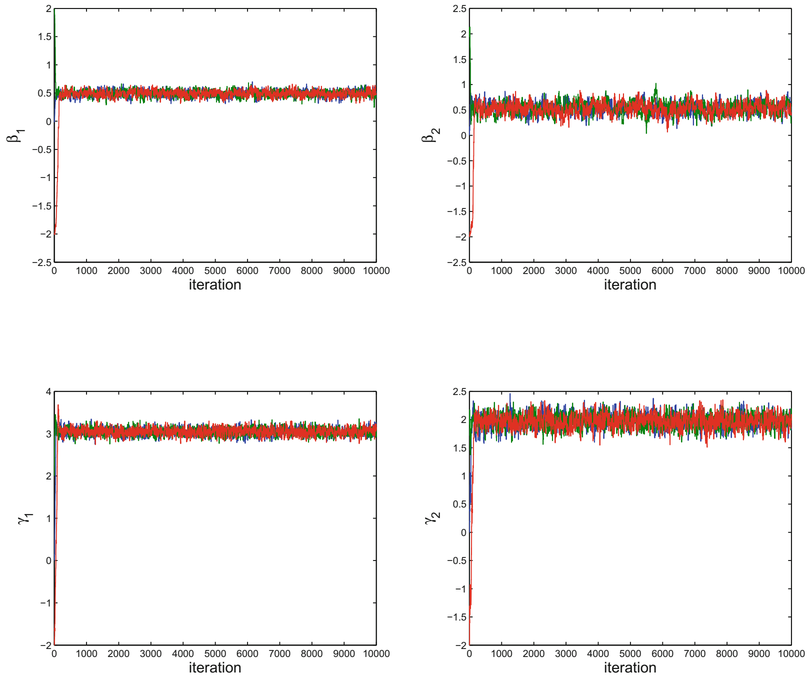

Fig. 1. Mixing process of three Markov chains for parameters β_1, β_2, γ_1 and γ_2 under prior III.

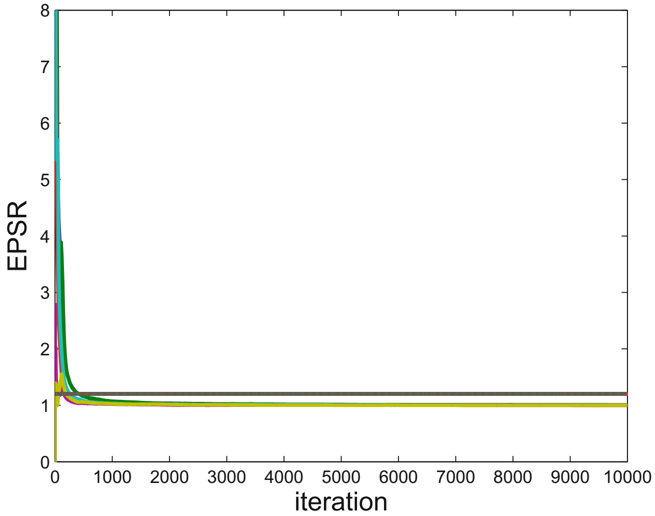

Fig. 2. EPSR values of all parameters against iterations in the simulation study.

$$\begin{cases} y_{ij}|\mathbf{b_i} \sim \mathrm{S}^{-1}(\mu_{ij}, \sigma_{ij}^2) \\ \mathrm{logit}(\mu_{ij}) = \beta_0 + \beta_1 \log(t_{ij}) + \beta_2 \log^2(t_{ij}) + \beta_3 x_{ij} + \mathbf{b_i}, \\ \log(\sigma_{ij}^2) = \gamma_0 + \gamma_1 \log(t_{ij}) + \gamma_2 x_{ij} \\ \mathbf{b_i} \sim SN_k\left(-\sqrt{2/\pi}\delta, \mathbf{\Sigma}, \mathbf{\Delta}\right) \end{cases} \tag{12}$$

where response y_{ij} denote the jth percentage of remained gas volume for the ith patient at follow-up day t_{ij}, the time covariate t_{ij} is the follow-up day after the C_3F_8 injection, and the discrete covariate x_{ij} are equal to -1, 0 and 1 by transforming the initial gas concentration levels of 15%, 20% and 25%, parameters $\boldsymbol{\delta}$, $\boldsymbol{\Sigma}$ and $\boldsymbol{\Delta}$ involved random effect $\mathbf{b_i}$ are specified by the Definition 1.

The model in Eq. (12) under the skewed normal approach together with the proceeding proposed Bayesian hybrid algorithm was fitted to the data set. To implement the hybrid algorithm, we set the hyperparameters with non-informative prior information in Eq. (5) to be $\boldsymbol{\beta^0} = \mathbf{0}$, $\mathbf{H_{\beta 0}} = 100I_4$, $\boldsymbol{\gamma^0} = \mathbf{0}$, $\mathbf{H_{\gamma 0}} = 100I_3$, $\boldsymbol{\delta^0} \sim N(0,1)$, $\mathbf{H_{\delta 0}} = 100$, $\mathbf{R_0} = 2$ and $\rho_0 = 5$. In the MH algorithm, we took the tuned variance parameters $\sigma_{b_i}^2 = 18$, $\sigma_\beta^2 = 2$ and $\sigma_\gamma^2 = 2$, which give rise to the average acceptance rates to be 31.8%, 30.5% and 30.7%, respectively. Also, we plotted the "estimated potential scale reduction (EPSR)" values against iterations in Fig. 3 to investigate the convergence of the proposed hybrid. It can be seen from Fig. 3 that the hybrid algorithm converges about 2500 iterations because EPSR values of all unknown parameters were less than 1.2 after about 2500 iterations. Thus, we collected 5000 observations after 5000 burn-ins in producing the Bayesian estimate under the normal approach and skewed normal approach. Bayesian estimates (EST), as well as their standard error (Stderr) and 95% highest posterior density interval (HPD) of the unknown parameters and the estimated value of DIC are presented in Table 2, which indicated the following: (i) the estimated DIC for skewed normal approach is less than those for normal approach and HPD interval of parameter $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are generally shorter than those for normal approach, which show that the performance of our proposed skewed normal approach is clearly better than those for normal approach in our considered example; (ii) it is necessary to consider the structure of the heterogeneous dispersion because both covariates of time and initial gas concentration levels are significant factors in the dispersion model; (iii) the proposed skewed normal approach can capture possible right-skewness of the percentage data of remained gas volume given the random effects.

Table 2. Summary statistics of the estimates in the real example.

Parameters	Skew-Normal approach			Normal approach		
	EST	SD	HPD interval	EST	SD	HPD interval
β_0	2.8540	0.5463	[1.7276, 3.8235]	2.7446	0.5123	[1.6772,3.6224]
β_1	0.3106	0.5983	[–0.7124, 1.7133]	0.3813	0.9255	[–0.5841, 1.8754]
β_2	–0.4273	0.1429	[–0.7321, –0.1594]	–0.4372	0.1519	[–0.7896, –0.2046]
β_2	0.7158	0.3312	[0.0997, 1.3596]	0.5823	0.3107	[0.0187, 1.2538]
γ_0	5.7944	0.2759	[5.2541, 6.2945]	5.8476	0.2726	[5.3070, 6.3328]
γ_1	–0.4789	0.1066	[–0.6859, –0.2835]	–0.4936	0.1105	[–0.7127, –0.2770]
γ_2	–0.5229	0.1750	[–0.8773, –0.1829]	–0.5177	0.1772	[–0.8382, –0.1724]
Σ	0.6494	0.4498	[0.1112, 1.5479]	1.3485	0.5194	[0.5052, 2.3720]
δ	1.0002	1.3245	[–1.8292, 2.7875]	–	–	–
DIC	–4143.5	–	–	–4119.5	–	–

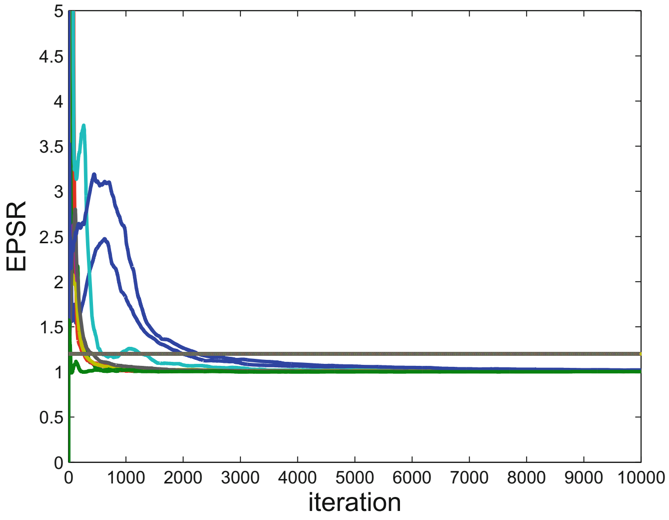

Fig. 3. EPSR values of all parameters against iterations in the real example.

5 Conclusion

This paper proposed a multivariate skew-norml simplex mixed-effects models with heterogeneous dispersion, obtained by incorporating skew-normal random effects into simplex regression model. To conduct Bayesian analysis for the proposed model, a hybrid algorithm combining the Gibbs sampler and the MH algorithm is used to simultaneously obtain Bayesian estimates of unknown parameters as well as their standard errors, random effect, model comparison criteria. Several simulation studies and a real example from a prospective ophthalmology study are used to illustrate the proposed methodologies. Empirical results show that our proposed skew-normal approach performs better than normal approach based on the estimate of HPD interval and DIC (see Table 2).

References

1. Artes, R., Jørgensen, B.: Longitudinal data estimating equations for dispersion models. Scand. J. Stat. **27**(2), 321–334 (2000)
2. Arellano-Valle, R.B., Bolfarine, H., Lachos, V.H.: Bayesian inference for skew-normal linear mixed models. J. Appl. Stat. **34**(6), 663–682 (2007)
3. Bandyopadhyay, D., Galvis, D.M., Lachos, V.H.: Augmented mixed models for clustered proportion data. Stat. Methods Med. Res. **26**(2), 880–897 (2017)
4. Bonat, W.H., Lopes, J.E., Shimakura, S.E., Ribeiro Jr, P.J.: Likelihood analysis for a class of simplex mixed models. Chil. J. Stat. **9**(2), 3–7 (2018)
5. Duan, X.D., Zhao, Y.Y., Tang, A.M.: A semiparametric Bayesian approach to simplex regression model with heterogeneous dispersion. Commun. Stat. Simul. Comput. **48**(8), 2487–2500 (2019)
6. Han, G., Huang, Y.X., Yuan, A.: Bayesian-frequentist hybrid approach for skew-normal nonlinear mixed-effects joint models in the presence of covariates measured with errors. Stat. Interface **11**(2), 223–236 (2018)
7. Meyers, S.M., Ambler, J.S., Tan, M., Werner, J.C., Huang, S.S.: Variation of perfluorpropane disapperance after vitrectomy. Retina **12**(4), 359–363 (1992)
8. Michaelis, P., Klein, N., Kneib, T.: Bayesian multivariate distributional regression with skewed responses and skewed random effects. J. Comput. Graph. Stat. **27**(3), 602–611 (2018)
9. Qiu, Z., Song, P.X.K., Tan, M.: Simplex mixed-effects models for longitudinal proportional data. Scand. J. Stat. **35**(4), 577–596 (2008)
10. Sahu, S.K., Dey, D.K., Branco, M.D.: A new class of multivariate skew distributions with applications to Bayesian regression models. Can. J. Stat. **31**(2), 129–150 (2003)
11. Song, P.X.K., Qiu, Z., Tan, M.: Modelling heterogeneous dispersion in marginal models for longitudinal proportional data. Biometrical J. **46**(5), 540–553 (2004)
12. Song, P.X.K., Tan, M.: Marginal models for longitudinal continuous proportional data. Biometrics **56**(2), 496–502 (2000)
13. Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Van der Linde, A.: Bayesian measures of model complexity and fit. J. Royal Stat. Soc. Ser. B **64**(2), 583–639 (2002)
14. Xing, D.Y., Huang, Y.X., Chen, H.N., Zhu, Y.L., Dagne, G.A., Baldwin, J.: Bayesian inference for two-part mixed-effects model using skew distributions, with application to longitudinal semicontinuous alcohol data. Stat. Methods Med. Res. **26**(4), 1838–1853 (2017)

15. Zhang, H.Z., Huang, Y.X., Wang, W., Chen, H.N., Langland-Orban, B.: Bayesian quantile regression-based partially linear mixed-effects joint models for longitudinal data with multiple features. Stat. Methods Med. Res. **28**(2), 569–588 (2019)
16. Zhang, W.Z., Wei, H.J.: Maximum likelihood estimation for simplex distribution nonlinear mixed models via the stochastic approximation algorithm. Rocky Mountain J. Math. **38**(5), 1863–1875 (2008)
17. Zhao, Y.Y., Xu, D.K., Duan, X.D., Dai, L.: Bayesian estimation of simplex distribution nonlinear mixed models for longitudinal data. Int. J. Appl. Math. Stat. **52**(3), 1–10 (2014)