A Bayesian Shared Parameter Model for Analysing Longitudinal Skewed Responses with Nonignorable Dropout

M. Ganjali[1,2,*] and T. Baghfalaki[1]

[1]*Department of Statistics, Shahid Beheshti University, Tehran, Iran*

[2]*School of Biological Science, Institute for Research in Fundamental Sciences (IPM), Iran*

Abstract: When the nature of a data set comes from a skew distribution, the use of usual Gaussian mixed effect model can be unreliable. In recent years, skew-normal mixed effect models have been used frequently for longitudinal data modeling in many biomedical studies. These models are flexible for considering skewness of the longitudinal data. In this paper, a shared parameter model is considered for simultaneously analysing nonignorable missingness and skew longitudinal outcomes. A Bayesian approach using Markov Chain Monte Carlo is adopted for parameter estimation. Some simulation studies are performed to investigate the performance of the proposed methods. The proposed methods are applied for analyzing an AIDS data set, where CD4 count measurements are gathered as longitudinal outcomes. In these data CD4 counts measurements are severely skew. In application section, different structures of skew-normal distribution assumptions for random effects and errors are considered where deviance information criterion is used for model comparison.

Keywords: Bayesian approach, Longitudinal data, Markov Chain Monte Carlo, Missingness mechanism, Nonignorable missing data, Random effects model.

1. INTRODUCTION

The normality assumption is usually used for analysing real data sets, but this assumption maybe unrealistic when there is strong skewness or heavy-tails in the data set. In practice, most of the real data sets are skewed or heavy-tailed, for example, CD4 count measurements in AIDS data set (see application section) are extremely skew. Therefore, development of statistical models using some flexible distribution families for analyzing this kind of data set is necessary.

Skew-normal distribution family [1] is a flexible family which includes normal one as a special case. Recently, some developments of various statistical applications using univariate and multivariate forms of skew-normal distribution are published. For examples vide [2-10]. Also, [11-15] discussed the use of skew-normal distribution for analyzing skew-normal data with missing values.

In longitudinal studies, subjects are measured repeatedly through time and missing values are a common problem in these studies. According to taxonomy of missing data mechanisms by [16], the missing process is categorized to be missing completely at random (MCAR) if the missingness mechanism is independent of both unobserved and observed data. It is missing at random (MAR) if conditional on observed data, missingness is independent of the unobserved measurements and it is not missing at random (NMAR) if it is neither MCAR nor MAR.

Models which can be used for analysing missing data are selection model, pattern-mixture model and shared parameter model. Selection model factorizes joint distribution of missing mechanism and responses as a marginal density of measurement process and a missingness process conditional on the outcomes. Pattern mixture model uses a marginal distribution of missingness process and a conditional distribution of the measurement process given the missingness process. Finally, shared parameter model shares one or two random components in models of responses and missingness mechanisms. Some references to such models are [17-22]. In this paper, we will use a shared parameter model for skew responses and will use a Bayesian approach for parameter estimation. Various versions of the multivariate skew-normal distribution were considered and used in literature [23-25, 4].

[26] proposed the multivariate skew-normal distribution through the following probability density function:

$$f(y|\mu,\Sigma,\Delta) = 2^k \phi_k(y|\mu,\Sigma+\Delta\Delta')$$

$$\times\Phi_k(\Delta'(\Sigma+\Delta\Delta')^{-1}(y-\mu)\,|\,0,(I_k+\Delta'\Sigma^{-1}\Delta)^{-1})$$

*Address correspondence to this author at the Department of Statistics, Shahid Beheshti University, Tehran, Iran; Tel: 0098 21 29902915;
Fax: 0098 21 22431649; E-mail: m-ganjali@sbu.ac.ir

E-ISSN: 1929-6029/14

where $\phi_k(y|\mu,\Sigma)$ and $\Phi_k(y|\mu,\Sigma)$ are the probability density function and cumulative distribution function of the $N_k(\mu,\Sigma)$, respectively. We denote this by $Y\sim SN_k(\mu,\Sigma,\Delta)$. In this form, $\Delta=diag(\delta)$ and δ is a vector of skewness parameter for variable Y. Note that for $\Delta=0$, the multivariate skew-normal distribution reduces to the usual symmetric multivariate normal $N_k(\mu,\Sigma)$ distribution. Its stochastic representation is given by: $Y\overset{d}{=}\Delta|X_0|+X_1$, where, $X_0\sim N_k(0,I_T)$ $X_1\sim N_k(\mu,\Sigma)$, X_0 and X_1 are independent, $\overset{d}{=}$ means "distributed as" and $|X_0|$ is the vector of the absolute values of the components of the vector X_0. This form can be used to derived many of properties of the distribution.

When $Y\sim SN_k(\mu,\Sigma,\Delta)$, then $E[Y]=\mu+\sqrt{\frac{2}{\pi}}\delta$ and $var(Y)=\Sigma+(1-\frac{2}{\pi})\Delta^2$.

In his paper, we have developed skew-normal linear mixed effect model for considering missingness mechanism in a shared parameter model framework. We have used Bayesian approach and the available WinBUGS software [27] for implementation of the model, where the Bayesian criterion DIC (deviance information criterion) has been used for model comparison. We have used some simulation studies for investigating the performance of the proposed method for analysing skew responses with possibility of nonignorable missingness. As an illustrative example, we have reanalyzed an AIDS data with CD4 count measurement as response variable which is severely skew.

This paper is organized as follows: Section 2 includes the model, notations and concepts of skew-normal shared parameter model. Section 3 includes the MCMC scheme for the proposed model. Section 4 presents some simulation studies to provide some viewpoints about using the proposed method. Section 5 includes an application section. In this section after exploratory analysis of the data set, the data are analysed using proposed methodology. The last section includes some conclusions.

2. NOTATION AND MODELS

We specify a mixed effect model and a logistic model for the longitudinal responses and missingness mechanism, respectively, such that the two models share some random effects. Let $Y_i=(Y_{i1},Y_{i2},...Y_{im_i})'$ denote the vector of longitudinal measurements for the i^{th} individual at times $t_{i1},t_{i2},...,t_{im_i}$, also let Y_i be partitioned into two subvectors: Y_i^{obs} containing the observed components of Y_i and Y_i^{mis} containing the missed components of Y_i.

We assume a skew-normal linear mixed effects model with the following hierarchical model for the longitudinal responses:

$$Y_i|b_i,\beta_1,\Psi,\Delta_e\overset{ind.}{\sim}SN_{m_i}(X_{1i}\beta_1+Z_ib_i,\Psi,\Delta_e) \quad (1)$$

$$b_i|D,\Delta_b\overset{iid.}{\sim}SN_q(0,D,\Delta_b),i=1,2,...,n.$$

where Y_i is a $m_i\times 1$ vector of response for the i^{th}, $i=1,2,...,n$, individual X_{1i} with dimension $m_i\times p_1$ is a design matrix corresponding to the fixed effects, β_1 with dimension $p_1\times 1$ is a vector of regression coefficients, Z_i with dimension $m_i\times q$ is the design matrix corresponding to the $q\times 1$ random effects vector b_i, and Δ_e and Δ_b are diagonal matrices of error skewness parameter and random effect skewness parameters with elements $\delta_{e1},\delta_{e2},...,\delta_{em_i}$ and $\delta_{b1},\delta_{b2},...,\delta_{bq}$, respectively. Also, Ψ and D are scale matrices of error and random effects, respectively. For simplicity and identifiability of parameters, in the reminder of the paper, we assume $\Psi=\sigma_eI_{m_i}$ and $\Delta_e=\delta_eI_q$.

Let R_{ij} be an indicator variable for observing the i^{th} individual's outcome at j^{th} follow-up time, i.e. $R_{ij}=1$, if an outcome is observed and $R_{ij}=0$ if an outcome is missed. Using the shared parameter model framework, we assume the following model for the probability of a subject having missing value at time point t_{ij}:

$$\eta\Big(Prob(R_{ij}=1|b_i,\beta_2,\gamma)\Big)=x_{2ij}'\beta_2+\gamma'b_i,$$

where η is a link function for missingness mechanism such as probit, logit or complementary log-log. x_{2ij} is a p_2-dimensional vector of covariate, β_2 is a $p_2\times 1$

vector of fixed effects, b_i is the same subject-specific random effects defined in the longitudinal model (1) and γ is a vector of parameters for random effects. The model can be viewed as an improvement of the shared parameter model (SPM) proposed in the context of longitudinal data with nonignorable dropout [28-30]. Therefore, if $R_i=(R_{i1},R_{i2},...,R_{im_i})'$ is the random vector of missing data status, then

$$f(r_i|b_i,\beta_2,\gamma)=\prod_{k=1}^{m_i} f(r_{ik}|b_i,\beta_2,\gamma).$$

In the Bayesian SPM the regression parameter γ is of particular interest because it controls how the missingness effects the measurement process. If $\gamma=0$, the missing data are ignorable and the usual Bayesian mixed effect model can be used, for analysing the data. In this case we do not need any missing data adjustment. Also, note that if $\gamma\neq 0$, the degree of non-ignorability depends on its components values, i.e. a large value of each component of γ causes more non-ignorability.

Biometrical researchers have an attractive appeal for the shared parameter model, because they generally believe that there may be some latent, yet to be measured, quantity underlying a person's susceptibility to both cognitive decline and missing data due to adverse outcome. Also, the shared parameter model does not explicitly assume that missingness depends on the unobserved outcome, instead it depends on a latent variable that is inherent in all outcomes from the same subject.

In the next section, we have considered MCMC scheme for implementation of the above described model.

3. THE MCMC SCHEME

For MCMC computation, we use the stochastic representation of the multivariate skew-normal distribution. Therefore, the model can be represented hierarchically by the following representation:

$$Y_i|b_i,\beta_1,\sigma_e^2,\delta_e,u_{ei}\overset{ind.}{\sim}N_{m_i}(X_{1i}\beta_1+Z_i b_i+\delta_e u_{ei},\sigma_e^2 I_{m_i}),$$

$$U_{ei}\overset{ind.}{\sim}N_{m_i}(0,I_{m_i})I_{\{u_{ei}>0\}},$$

and,

$$b_i|D,\Delta_b,u_{bi}\overset{ind.}{\sim}N_q(\Delta_b u_{bi},D),$$

$$U_{bi}\overset{ind.}{\sim}N_q(0,I_{m_i})I_{\{u_{bi}>0\}},\ i=1,2,...,n.$$

where $I_{\{u_i>0\}}$ means that all components of vector u_i are larger than 0.

Implementation of MCMC methods such as Gibbs sampling needs to specify prior distributions. The full model specification is given by:

$$Y_i|b_i,\beta_1,\sigma_e^2,\delta_e,u_{ei}\overset{ind.}{\sim}N_{m_i}(X_{1i}\beta_1+Z_i b_i+\delta_e u_{ei},\sigma_e^2 I), \quad (2)$$

$$U_{ei}\overset{ind.}{\sim}N_{m_i}(0,I_{m_i})I_{\{u_{ei}>0\}},$$

$$R_{ij}|b_i,\beta_2,\gamma\overset{ind.}{\sim}Ber(\eta^{-1}(x'_{2ij}\beta_2+\gamma' b_i)),$$

$$b_i|D,\Delta_b,u_{bi}\overset{ind.}{\sim}N_q(\Delta_b u_{bi},D),$$

$$U_{bi}\overset{ind.}{\sim}N_q(0,I_{m_i})I_{\{u_{bi}>0\}},$$

$$\beta_1\sim N_{p_1}(\mu_{\beta_1},\Sigma_{\beta_1}),\quad \beta_2\sim N_{p_2}(\mu_{\beta_2},\Sigma_{\beta_2}),$$

$$\gamma\sim N_q(\mu_\gamma,\Sigma_\gamma),\quad \sigma_e^2\sim IG(\alpha_e,\tau_e),\quad D\sim IW(\Psi_D,\tau_D),$$

$$\delta_e\sim N(\mu_{\delta_e},\kappa_{\delta_e}),\quad \delta_b\sim N_q(\mu_{\delta_b},\kappa_{\delta_b}),$$

where $Ber(.)$, $IG(.,.)$ and $IW(.,.)$ denote Bernoulli distribution, inverse gamma distribution and inverse Wishart distribution, respectively. We have used logit link for missingness mechanism. Also, the hyperparameters of these priors are selected such that they lead to the low-informative prior distributions. The full conditional distributions of hierarchical model (2) can be found in appendix.

Many Bayesian criteria have been proposed in the literature, we consider DIC, deviance information criterion [27] for model comparison. Let Θ and $Z=(z_1,...,z_N)'$ be the entire model parameters and data, respectively. Define $D(\Theta)=-2\ln f(z|\Theta)$ $=-2\sum_{i=1}^{N}\ln f(z_i|\Theta)$ where, $f(z_i|\Theta)$ is marginal distribution of z_i, then $E\left[D(\Theta)\right]$ is a measure of fit and can be approximated by using the MCMC output in a Monte Carlo integration. This index is given by

$\bar{D} = \frac{1}{m}\sum_{k=1}^{m} D(\Theta^{(k)})$. The DIC criterion is given by $\widehat{DIC} = \bar{D} + \hat{p}_D$, where, p is the number of parameters and N is the total number of observations. The $\hat{p}_D$ is the effective number of parameters [27], and is defined as $p_D = E\left[D(\Theta)\right] - D\left(E\left[\Theta\right]\right)$. The term $D\left(E\left[\Theta\right]\right)$ is the deviance of posterior mean obtained when considering the mean values of the generated posterior mean of the model parameters, which is estimated by $\widehat{D} = D\left(\frac{1}{m}\sum_{k=1}^{m}\Theta^{(k)}\right)$. The smaller is the DIC, the better is the fit of the model.

4. SIMULATION STUDIES

To investigate the performance of the proposed methodology, we conducted some imulation studies. A simulation study for investigating performance of the skew-normal within subject error is considered. Also, some simulation studies, for investigating the performance of the skew-normal random effects model are considered. A simulation study for comparing the results of the proposed method with those of a method which ignores the missing mechanism even when the missingness is truly nonignorable is also considered. The last simulation is considered for investigating the performance of the proposed method when the outcomes are actually normal.

4.1. The Skew-Normal within Subject Error

In this simulation study, we generate 500 samples with sample size $n = 200$ with two rates of missingness 15% and 30%. We simulate data according to the following linear mixed effect model:

$$y_{ij} = \beta_{10} + \beta_{11}t_j + \beta_{12}x_i + b_{1i} + b_{2i}t_j + \varepsilon_{ij}, \quad (3)$$

where $t_j = j$, for $j = 1,2,\cdots,5$, and x_i is generated from Bernoulli distribution with success probability 0.2. In this model, $\beta_{10} = 10$, $\beta_{11} = -3$ and $\beta_{12} = -2$. Also, we considered the following missingness mechanism:

$$R_{ij} \sim Ber(p_{ij}), \quad (4)$$

where,

$$logit(p_{ij}) = \beta_{20} + \beta_{21}x_i + \beta_{22}b_{1i} + \beta_{23}b_{2i}. \quad (5)$$

In this model, $\varepsilon_{ij} \sim SN(0,\sigma_e^2,\delta_e)$, where $\sigma_e^2 = 1$ and $\delta_e = 4$, also, $b_i = (b_{1i}, b_{2i})' \sim N_2(0,D)$, such that $D = \begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$.

We have analyzed the simulated data under normal and skew-normal distribution assumptions. The results of this simulation study are given in Table **1**. This table contains estimated value of parameters, standard errors, relative biases and root of mean square errors. The two later criteria are defined as

$$Rel.Bias(\theta) = \frac{1}{N}\sum_{i=1}^{N}(\frac{\widehat{\theta}_i}{\theta} - 1), \quad MSE(\theta) = \frac{1}{N}\sum_{i=1}^{N}(\widehat{\theta}_i - \theta)^2$$

where $\widehat{\theta}_i$ is the estimated value of θ for the i^{th} sample. In this table, the header "Skew-normal Scenario" and "Normal Scenario" identify the underlying model. The results show that the rate of missingness is an effective measures, such that the less the rate of missingness, the smaller are the relative biases and the smaller are the root of MSEs. The results of this table show that using normal distribution assumption when one has to use the skew-normal distribution leads to biased estimates. The regression coefficients of normal model are estimated without considerable bias, but intercept and variance estimates are severely bias. Also, the relative biases and root of MSEs in skew-normal model are near zero, this means that the proposed model possess consistency properties.

4.2. The Skew Random Effects

In this section, we have considered two different skew random effects models. We generate 500 samples with sample size $n = 200$, and we considered two rates of missingness 15% and 30%. The random effects are generated from gamma distribution and within-subject errors from normal as described in the following. The first model is

$$y_{ij} = \beta_{10} + \beta_{11}t_j + \beta_{12}x_i + b_i + \varepsilon_{ij},$$

$$R_{ij} \sim Ber(p_{ij}); \quad logit(p_{ij}) = \beta_{20} + \beta_{21}x_i + \beta_{22}b_{1i}.$$

where $b_{1i} \sim Gamma(4,1)$ and $\varepsilon_{ij} \sim N(0,1)$. The results of this simulation study are given in Table **2**. In this table, we report $E\left[\beta_{10} + b_{1i}\right]$ and $Var\left[b_{1i}\right]$.

The results of this table show that under skew-normal distribution assumption, the parameters are

Table 1: Results of simulation study for 500 samples with sample size $n=200$ generated by skew-normal within-subject errors and analyzed under two distribution assumptions, d_{11}, d_{12} and d_{22} are distinct elements of the matrix D. (Est.: posterior mean, SE: standard error, Rel. Bias: relative bias and RMSE: root of mean squared errors)

Model		Skew-normal Scenario						Normal scenario					
Rate of missingness		15%			30%			15%			30%		
Parameters	Real	Est. (SE)	Rel. Bias	RMSE	Est. (SE)	Rel. Bias	RMSE	Est. (SE)	Rel. Bias	RMSE	Est. (SE)	Rel. Bias	RMSE
β_{11}	10	10.020(0.203)	0.001	0.203	10.075(0.239)	0.007	0.246	13.188(0.122)	0.318	3.19	13.243(0.155)	0.324	3.247
β_{12}	-3	-2.982(0.145)	-0.005	0.145	-2.943(0.082)	-0.018	0.098	-2.977(0.058)	-0.007	0.062	-2.961(0.064)	-0.013	0.346
β_{13}	-2	-2006(0.458)	0.003	0.454	-2.021(0.369)	0.011	0.363	-1.925(0.241)	-0.037	0.248	-1.986(0.359)	-0.006	0.285
d_{11}	1	0.893(0.155)	-0.106	0.157	0.860(0.191)	-0.139	0.204	0.753(0.164)	-0.246	0.295	0.756(0.154)	-0.243	0.239
d_{12}	0.5	0.432(0.087)	-0.156	0.154	0.329(0.124)	-0.341	0.209	0.368(0.128)	-0.263	0.183	0.281(0.103)	-0.436	0.155
d_{22}	1	0.938(0.074)	-0.061	0.096	0.872(0.092)	-0.127	0.156	0.939(0.085)	-0.061	0.104	0.862(0.073)	-0.138	-
δ_e	4	3.982(0.184	-0.004	0.184	4.002(0.181)	0.001	0.178	-	-	-	-	-	5.908
δ_e^2	1	1.045(0.343)	0.045	0.344	1.057(0.412)	0.057	0.409	6.926(0.285)	5.926	5.932	6.900(0.320)	5.9	

Table 2: Results of simulation study for 500 samples generated by univariate gamma random effects and analyzed under two distribution assumptions with n=200 sample size. (Est.: posterior mean, SE: standard error, Rel. Bias: relative bias and RMSE: root of mean square errors)

Model		Skew-normal Scenario						Normal scenario					
Rate of missingness		15%			30%			15%			30%		
parameters	real	Est. (SE)	Rel. Bias	RMSE	Est. (SE)	Rel. Bias	RMSE	Est.(SE)	Rel. Bias	RMSE	Est. (SE)	Rel. Bias	RMSE
$E[\beta_{10}+b_{i1}]$	14	14.024(0.172	0.002	1.72	14.007(0.183)	0.001	0.183	13.976(0.188)	-0.002	0.188	13.934(0.221)	-0.004	0.217
β_{12}	-3	-3.001(0.004)	0.001	0.005	-2.999(0.007)	-0.001	0.006	-2.998(0.004)	-0.001	0.004	-2.998(0.006)	-0.001	0.006
β_{13}	-2	-2.043(0.242)	0.022	0.244	-1.992(0.278)	-0.003	0.278	-1.975(0.298)	-0.012	0.297	-2.074(0.432)	0.036	0.41
σ_e^2	1	1.001(0.048)	0.001	0.047	1.002(0.069)	0.002	0.068	1.001(0.059)	0.001	0.058	1.015(0.057)	0.014	0.055
$Var[b_{i1}]$	4	4.073(0.523)	0.018	0.524	3.983(0.515)	-0.004	0.513	1.014(0.057)	-0.746	2.985	1.101(0.089)	1.001	3.102
δ_b	-	3.063(0.564)	-	-	3.108(0.276)	-	-	-	-	-	-	-	-

Table 3: Results of simulation study for 500 samples generated by bivariate gamma random effects and analyzed under two distribution assumptions with n=200 sample size. (Est.: posterior mean, SE: standard error, Rel. Bias: relative bias and RMSE: root of mean square errors), also *shows that relative bias can not be computed as the real value is 0.000

Model		Skew-normal Scenario						Normal scenario					
Rate of missingness		15%			30%			15%			30%		
parameters	real	Est. (SE)	Rel. Bias	RMSE	Est. (SE)	Rel. Bias	RMSE	Est. (SE)	Rel. Bias	RMSE	Est. (SE)	Rel. Bias	RMSE
$E[\beta_{10}+b_{1i}]$	12	12.047(0.153)	0.003	0.158	12.118(0.137)	0.001	0.137	11.974(0.147)	-0.002	0.148	11.981(0.150)	-0.001	0.149
$E[\beta_{10}+b_{2i}]$	12	11.977(0.215)	-0.002	0.215	11.705(0.211)	-0.024	0.361		-	-	-	-	-
β_{12}	-3	-2.578(0.117)	-0.14	0.437	-2.425(0.116)	-0.191	0.586	-1.206(0.135)	-0.597	1.798	-1.977(0.132)	-0.341	1.03
β_{13}	-2	-1.998(0.199)	-0.001	0.197	-2.002(0.195)	0.001	0.193	-2.019(0.247)	0.009	0.245	-2.003(0.234)	0.002	0.232
$\sigma_{\varepsilon_e^2}$	1	0.975(0.062)	-0.024	0.066	0.987(0.071)	-0.012	0.071	1.056(0.082)	0.055	0.098	1.708(0.404)	0.707	0.813
$Var[b_{1i}]$	2	2.082(0.296)	0.041	0.304	1.876(0.339)	-0.061	0.358	1.517(0.315)	-0.241	0.574	0.968(0.343)	-0.515	1.086
$cov[b_{1i}, b_{2i}]$	0	-0.009(0.021)	-	0.023	-0.013(0.027)	-	0.03	-0.082(0.137)	-	0.159	0.004(0.106)	-	0.105
$Var[b_{2i}]$	2	1.861(0.235)	-0.069	0.271	1.279(0.183)	-0.36	0.743	1.058(0.149)	-0.471	0.952	0.384(0.065)	-0.807	1.616
δ_{b_1}	-	2.145(0.200)	0.072	0.245	2.003(0.232)	0.001	0.23	-	-	-	-	-	-
δ_{b_2}	-	2.057	0.028	0.174	1.611(0.172)	-0.194	0.424	-	-	-	-	-	-

Table 4: Results of simulation study for 500 samples generated by skew-normal within-subject errors and gamma random effects and analyzed under two distribution assumptions with n=200 sample size. (Est.: posterior mean, SE: standard error, Rel. Bias: relative bias and RMSE: root of mean square error), also *shows that relative bias can not be computed as the real value is 0.000

Model		Skew-normal Scenario						Normal scenario					
Rate of missingness		15%			30%			15%			30%		
parameters	real	Est. (SE)	Rel. Bias	RMSE	Est. (SE)	Rel. Bias	RMSE	Est. (SE)	Rel. Bias	RMSE	Est. (SE)	Rel. Bias	RMSE
$E[\beta_{10}+b_{i1}]$	12	12.008(0.167)	0.001	0.163	11.911(0.170)	-0.007	0.188	15.208(0.215)	0.269	3.215	15.371(0.208)	0.28	3.376
$E[\beta_{10}+b_{i2}]$	12	12.071(0.397)	0.005	0.393	11.883(0.387)	-0.009	0.397	-	-	-	-	-	-
β_{12}	-3	-2.706(0.096)	-0.097	0.308	-2.720(0.149)	-0.093	0.315	-1.038(0.101)	-0.653	1.964	-1.039(0.147)	-0.653	1.965
β_{13}	-2	-1.965(0.326)	-0.017	0.32	-1.813(0.338)	-0.093	0.38	-1.866(0.393)	-0.066	0.412	-1.979(0.401)	-0.01	0.396
σ_e^2	1	1.023(0.327)	0.023	0.319	0.914(0.436)	-0.085	0.436	7.844(0.461)	6.844	6.859	7.990(0.617)	6.99	7.017
$Var[b_{i1}]$	2	1.774(0.586)	-0.113	0.614	1.967(0.677)	-0.016	0.664	0.087(0.033)	-0.956	1.912	0.104(0.048)	-0.947	1.896
$cov[b_{i1}, b_{i2}]$	0	-0.001(0.002)	-	0.002	-0.001(0.002)	-	0.002	0.076(0.185)	-	0.198	0.052(0.249)	-	0.251
$Var[b_{i2}]$	2	1.861(0.293)	-0.069	0.316	1.858(0.256)	-0.071	0.288	1.921(0.321)	-0.039	0.327	2.112(0.317)	0.056	0.332
δ_e	4	4.006(0.180)	0.002	0.175	4.046(0.206)	0.011	0.207	-	-	-	-	-	-
δ_{b_1}	-	2.147(0.375)	-	-	2.262(0.407)	-	-	-	-	-	-	-	-
δ_{b_2}	-	2.226(0.181)	-	-	2.226(0.157)	-	-	-	-	-	-	-	-

estimated without bias and values of RMSE are small. But under normal distribution assumption some parameters such as intercept and variance of random effects are estimated with some biases and the values of RMSE for these parameters are relatively large.

The model for other simulation study is the same as models (3)-(5), but $b_{1i}, b_{2i} \sim Gamma(2,1)$ and b_{1i} and b_{2i} are uncorrelated. The results of this simulation study are presented in Table **3**. Similar to previous table, we report $E\left[\beta_{10}+b_{1i}\right]$, $E\left[\beta_{10}+b_{2i}\right]$, $Var\left[b_{1i}\right]$, $cov\left[b_{1i},b_{2i}\right]$ and $Var\left[b_{2i}\right]$. These results give a better evidence for good performance of skew-normal model, than the previous simulation study. As can be seen in Table **3**, the skew-normal model has not any bias estimate, but the parameter estimates of normal model have some biases.

The relative bias of $cov\left[b_{1i},b_{2i}\right]$ is not reported since the real value of this parameter is zero.

4.3. The Skew Random Effects and Errors

In this section, again, we generate 500 samples with sample size $n=200$ with two rates of missingness, 15% and 30%. We simulate data from model (3)-(5), where $b_{1i}, b_{2i} \sim Gamma(2,1)$, b_{1i} and b_{2i} are uncorrelated and $\varepsilon_{ij} \sim SN(0,\sigma_e^2,\delta_e)$, where $\sigma_e^2=1$ and $\delta_e=4$. In all the simulations β_{21}, β_{22} and β_{23} are selected such that the desired rates of missingness are obtained. Results of this simulation study are given in Table **4**.

This table shows that the performance of the skew-normal model is good but some biases can be found under normal distribution assumption.

4.4. Performance of Nonignorablity of the Model

In this section, we generate using as models (3)-(5) with sample size $n=300$ with two rates of missingness, 15% and 30%. The model is analyzed using an ignorable mechanism. We use available cases without any missingness mechanism. The results of this simulation study are presented in Table **5**. These results show that when data are generated under a non-ignorable missingness mechanism, using a ignorable mechanism leads to some high biased parameter estimates. The higher is the rate of missingness, the higher are the biases.

4.5. Performance of the Proposed Model in Normal Model

In this section, the performance of the proposed method is investigated, when the outcomes are actually normal. The data are simulated from models (3)-(5), but under normal distributional assumption.

We have analyzed the data set using the proposed method and using the pure normal model. The results of this simulation study are presented in Table **6**.

The results show that when the outcomes are actually normal, the skewness parameter is estimated near zero i.e. as a non-significant parameter. The results of the proposed skew-normal model are very close to those of shared random effects model under normal distribution assumption.

5. REAL DATA ANALYSIS: HIV DATA SET

We consider a longitudinal study on 467 HIV infected patients who had failed or were intolerant of zidovudine (AZT) therapy. The data had analyzed before by [31-32]. The aim of the study was to compare the efficacy and safety of two alternative antiretroviral

Table 5: Results of simulation study for 500 samples generated by skew-normal within-subject errors and analyzed under an ignorable mechanism, d_{11}, d_{12} and d_{22} are distinct elements of the matrix D. (Est.: posterior mean, SE: standard error, Rel. Bias: relative bias and RMSE: root of mean square error)

Rate of missingness		15%			30%		
parameters	real	Est.(SE)	Rel.Bias	RMSE	Est.(SE)	Rel.Bias	RMSE
β_{11}	10	10.090(0.188)	0.009	0.205	10.485(0.593)	0.048	0.765
β_{12}	-3	-2.591(0.108)	-0.136	0.422	-2.038(0.321)	-0.32	1.013
β_{13}	-2	-1.952(0.292)	-0.023	0.291	-1.972(0.506)	-0.014	0.506
d_{11}	1	2.294(0.231)	1.294	1.313	3.425(0.374)	2.425	2.454
d_{12}	0.5	-0.008(0.058)	-1.016	0.511	-0.104(0.128)	-1.209	0.618
d_{22}	1	1.136(0.063)	0.136	0.149	2.509(0.212)	1.508	1.523
δ_e	4	4.073(0.148)	0.018	0.163	3.516(0.655)	-0.121	0.813
σ_e^2	1	0.319(0.171)	-0.681	0.701	0.793(0.681)	-0.207	0.71

Table 6: Results of simulation study for 500 samples generated by normal within-subject errors and normal random effects and analyzed under two distribution assumptions with 15% and 30% rates of missingness, also d_{11}, d_{12} and d_{22} are distinct elements of the matrix D. (Est.: posterior mean, SE: standard error, Rel. Bias: relative bias and RMSE: root of mean square error)

Model		Skew-normal Scenario						Normal scenario					
Rate of missingness		15%			30%			15%			30%		
parameters	real	Est. (SE)	Rel. Bias	RMSE	Est. (SE)	Rel. Bias	RMSE	Est. (SE)	Rel. Bias	RMSE	Est. (SE)	Rel. Bias	RMSE
β_{10}	10	9.965(0.508)	-0.003	0.507	9.999(0.428)	0	0.421	10.021(0.118)	0.002	0.12	10.007(0.066)	0.001	0.065
β_{11}	-3	-2.980(0.079)	-0.006	0.081	-2.991(0.0.046)	-0.002	0.046	-2.974(0.079)	-0.008	0.082	-3.001(0.045)	0.001	0.044
β_{12}	-2	-1.964(0.182)	-0.017	0.185	-2.001(0.100)	0.001	0.098	-2.022(0.180)	0.011	0.181	-1.991(0.123)	-0.004	0.121
d_{11}	1	0.671(0.174)	-0.328	0.371	0.902(0.126)	-0.097	0.157	0.693(0.169)	-0.306	0.349	0.861(0.074)	-0.139	0.157
d_{12}	0.5	0.437(0.096)	-0.125	0.114	0.462(0.058)	-0.074	0.068	0.436(0.083)	-0.126	0.103	0.457(0.048)	-0.085	0.063
d_{22}	1	0.875(0.094)	-0.124	0.156	0.912(0.055)	-0.087	0.103	0.865(0.085)	-0.134	0.158	0.913(0.057)	-0.086	0.103
δ_0	0	0.062(0.613)	-	0.613	0.014(0.538)	-	0.529	-	-	-	-	-	-
σ_e^2	1	0.884(0.134)	-0.115	0.176	0.871(0.091)	-0.012	0.156	1.072(0.079)	0.072	0.106	1.014(0.046)	0.014	0.048

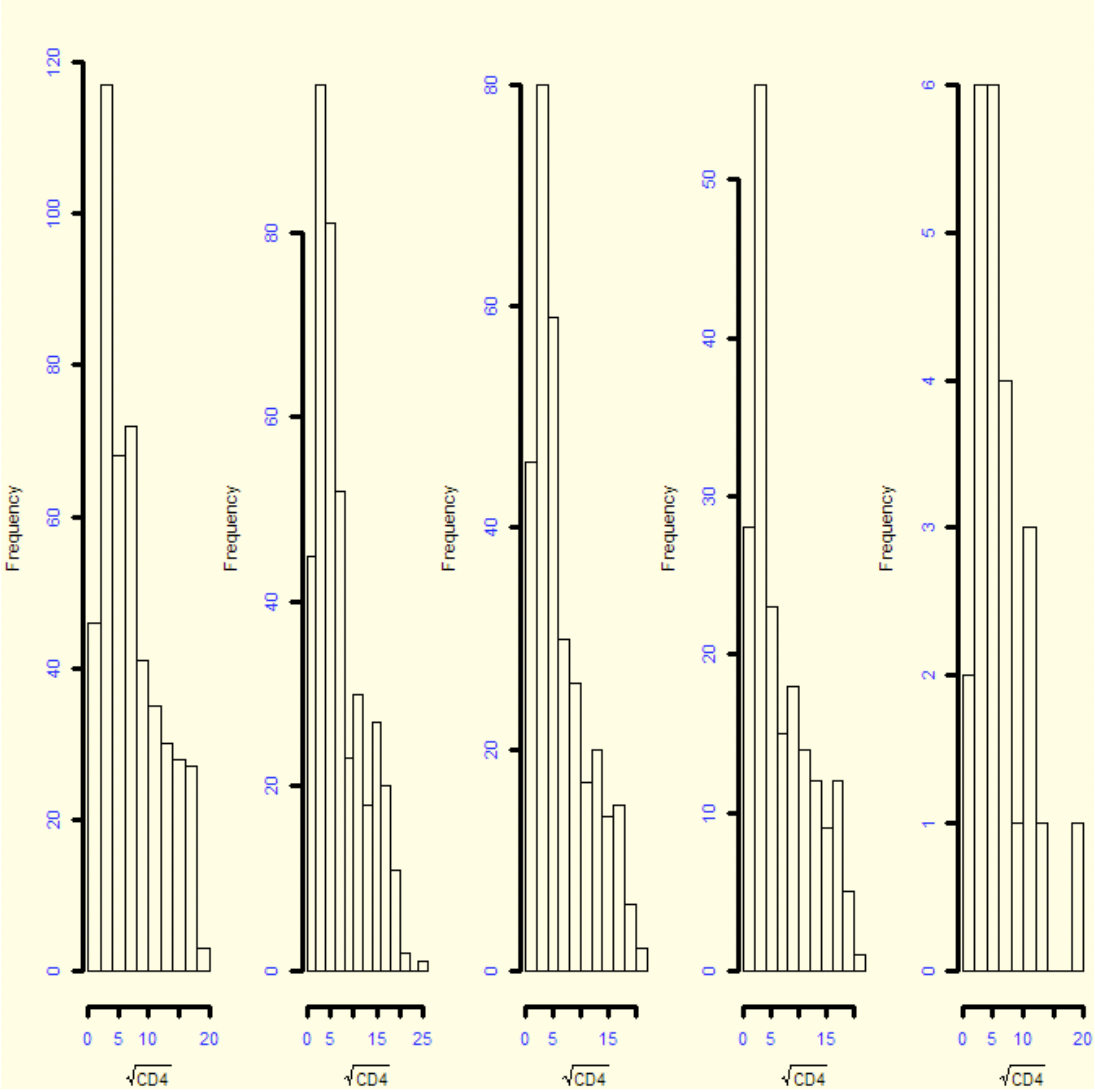

Figure 1: Histogram of AIDS data set in each calender time.

drugs, namely didanosine (ddI) and zalcitabine (ddC). Patients were randomly assigned to receive either ddI or ddC, and CD4 cell counts were recorded at study entry, where randomization took place, as well as 2, 6, 12, and 18 months thereafter. Because of right skewness in the CD4, for all of the analysis, we will work with the squared root of the CD4 cell values. Figure **1** presents histogram of $\sqrt{CD4}$ for each calender time, this figure shows that $\sqrt{CD4}$ are severely skew.

Figure **2** shows $\sqrt{CD4}$ trajectories for all individuals (panel a) and fifty randomly selected individuals from the study (panel b). These figures show a sharply increasing degree of missing data over time due to deaths, dropouts, and missed clinic visits. More details about this data set can be found in [33]. The red lines are mean profile for all observed individuals in each panel.

We have used the structure of the hierarchical model (2) for modeling the data set, that is:

$$Y_{ij}|b_i,\beta,\sigma_e^2,\delta_e,u_{eij} \overset{ind.}{\sim} N(\mu_{1ij}+\delta_e u_{eij},\sigma_e^2)$$

$$U_{eij} \overset{ind.}{\sim} N(0,1)I_{\{u_{eij}>0\}}$$

$$b_i|\sigma_b^2,\delta_b,u_{bi} \overset{ind.}{\sim} N_2(\delta_b u_{bi},\sigma_b^2)$$

$$U_{bi} \overset{ind.}{\sim} N(0,1)I_{\{u_{bi}>0\}}$$

$$R_{ij}|b_i,\beta_2,\gamma \overset{ind.}{\sim} Ber\left(\frac{\exp(\mu_{2ij})}{1+\exp(\mu_{2ij})}\right)$$

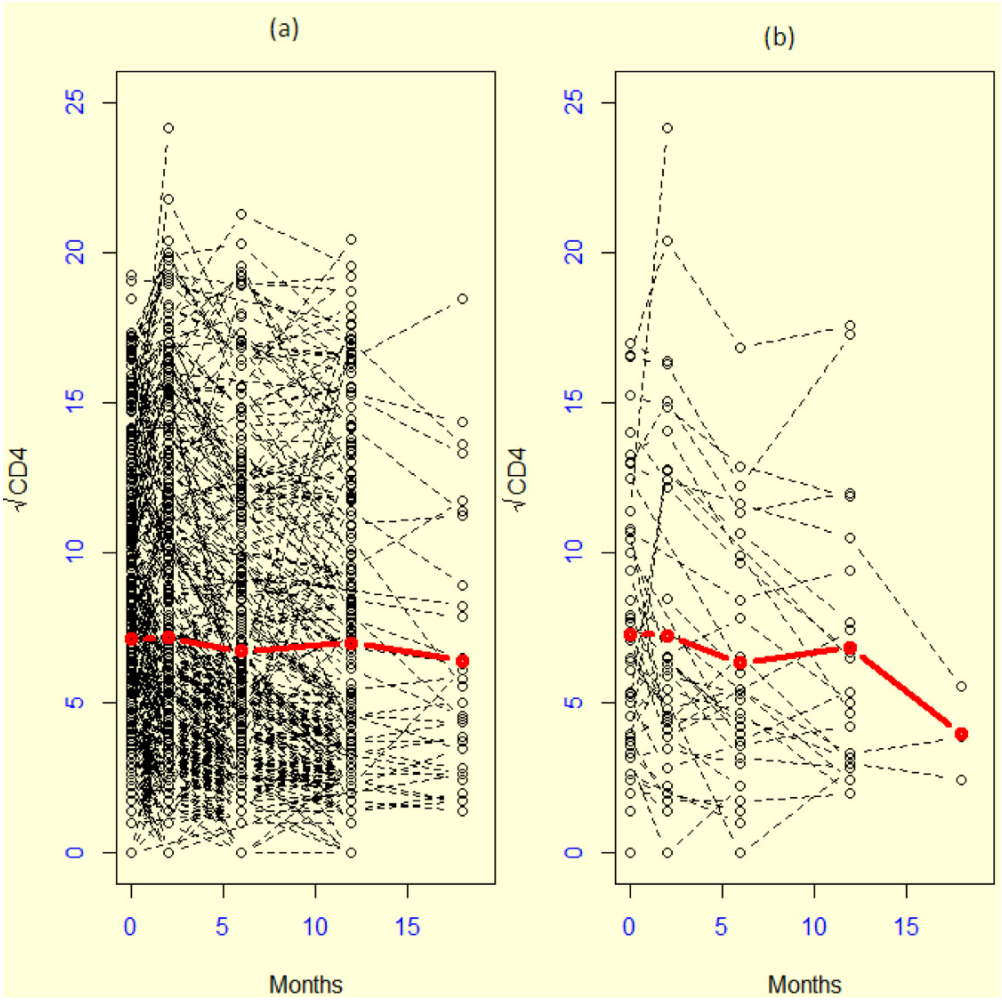

Figure 2: Profiles of CD4 measurements over time. (**a**) All individuals. (**b**) Fifty randomly selected individuals. The red lines are mean profile for all observed individuals in each panel.

where,

$$\mu_{1ij} = \beta_{10} + \beta_{11}t_{ij} + \beta_{12}t_{ij}Drug_i + \beta_{13}Sex_i$$

$$+\beta_{14}PrevOI_i + \beta_{15}Stratum_i + b_{1i} + b_{2i}t_{ij}$$

and,

$$\mu_{2ij} = \beta_{20} + \beta_{21}b_{1i} + \beta_{22}b_{2i}.$$

Here, Sex_i is a gender indicator (0=female, 1=male), also other three explanatory variables are $Drug_i$ (0=ddC, 1= ddI), $PrevOI_i$, previous opportunistic infection, (0=AIDS diagnosis, 1=no AIDS diagnosis), and $Stratum_i$ (0=AZT failure, 1=AZT intolerance). In our approach, we generate two parallel MCMC chains with different starting values for $300{,}000$ iterations each. Then, we discarded the first $20{,}000$ iterations as pre-convergence burn-in and the retained iterations are used for the posterior analysis. Then, we have checked convergence of parameter estimates using Gelman and Rubin diagnostic test [34] for all models. In all models priors are $\beta_k \sim N_{p_k}(0,10^3)$, $k=1,2, p_1=6, p_2=3$, and $D \sim IW(100I_2,2)$, $\sigma_e, \sigma_b \sim IG(0.1,0.1)$ and $\delta_b, \delta_e \sim N(0,0.01)$. The results of various models can be found in Table **6**, where different distributional assumptions are specified in the header. In this table, four statistical models, differing in the error and random effects distribution assumptions, are compared. These models are:

A model with multivariate skew-normal distributional assumption for both random effects and errors.

A model with multivariate skew-normal distributional assumption for random effects and normal distributional assumption for errors.

A model with multivariate normal distributional assumption for random effects and skew-normal distributional assumption for errors.

Table 7: Bayesian parameter estimates (posterior mean and standard deviation, S.D.) and 95% HPD for analysing the HIV data set by some longitudinal model, notation $M_r - M_e$ is used to indicate the implemented model where M_r and M_e are the distributions of the shared random effect and longitudinal residuals, respectively. N and SN are abbreviations for the normal and the skew-normal distributions, respectively.

Model	N-N		SN-SN		N-SN		SN-N	
Variable	Est. (S.D)	95% HPD	Est. (S.D)	95% HPD	Est. (S.D)	95% HPD	Est. (S.D)	95% HPD
Intercept	8.123(0.349)	(7.440,8.806)	2.523(0.772)	(1.348,3.938)	3.227(0.387)	(2.431,3.945)	3.798(0.667)	(2.507,5.124)
Time	-0.237(0.025)	(-0.287,-0.190)	-0.236(0.063)	(-0.361,-0.112)	-0.397(0.046)	(-0.488,-0.309)	-0.223(0.062)	(-0.349,-0.105)
Time *Drug	0.023(0.030)	(-0.036,0.083)	0.021(0.030)	(-0.037,0.081)	0.025(0.053)	(-0.075,0.133)	0.024(0.031)	(-0.036,0.084)
Gender	-0.348(0.322)	(-0.985,0.276)	-0.070(0.244)	(-0.537,0.422)	-0.086(0.292)	(-0.670,0.479)	-0.140(0.288)	(-0.711,0.419)
Prevol	-2.055(0.241)	(-2.522,-1.586)	-1.097(0.208)	(-1.504,-0.715)	-1.511(0.225)	(-1.942,1.057)	-1.586(0.232)	(-2.045,-1.135)
Stratum	-0.128(0.232)	(-0.579,0.327)	-0.146(0.175)	(-0.494,0.193)	-0.156(0.203)	(-0.563,0.235)	-0.156(0.206)	(-0.559,0.247)
β_{21}	0.012(0.055)	(-0.097,0.119)	-0.401(0.176)	(-0.737,-0.049)	-0.029(0.078)	(-0.184,0.123)	-0.465(0.183)	(-0.819,-0.097)
β_{22}	0.077(0.015)	(0.046,0.106)	0.081(0.013)	(0.054,0.109)	0.0181(0.055)	(0.078,0.292)	0.102(0.023)	(0.056,0.146)
β_{23}	2.472(0.543)	(1.500,3.461)	2.481(0.478)	(1.522,3.396)	1.845(0.175)	(1.518,2.199)	2.154(0.574)	(0.989,3.247)
δ_{b_1}	-	-	6.523(0.579)	(5.987,7.153)	-	-	6.016(0.384)	(5.240,6.763)
δ_{b_2}	-	-	-0.001(0.074)	(-0.151,0.140)	-	-	-0.004(0.072)	(-0.146,0.137)
δ_e	-	-	-0.023(0.688)	(-1.405,1.274)	5.841(0.318)	(5.186,6.437)	-	-
σ_{11}	16.152(1.196)	(13.800,18.470)	0.461(0.925)	(0.019,1.128)	3.179(0.680)	(1.972,4.563)	4.287(0.963)	(2.620,6.220)
σ_{21}	-0.088(0.068)	(-0.225,0.043)	-0.019(0.026)	(-0.081,0.023)	-0.034(0.071)	(-0.175,0.104)	-0.067(0.057)	(-0.181,0.044)
σ_{22}	-0.088(0.068)	(0.033,0.060)	0.044(0.007)	(0.032,0.059)	0.431(0.034)	(0.364,0.500)	0.044(0.006)	(0.031,0.058)
σ_e	3.015(0.162)	(2.700,3.336)	2.994(0.161)	(2.693,3.321)	2.935(0.160)	(2.612,3.243)	2.961(0.158)	(2.648,3.267)
DIC	8613.911		8579.51		8543.592		8551.643	

A model with multivariate normal distributional assumptions for both random effects and errors.

According to DIC criterion $N - SN$ model has the smallest DIC and $N - N$ has the largest amount of DIC.

In the best fitted model, time and Prevol are two significant factors in the study, also skewness parameter of error is highly significant. An important result of this analysis is the estimated values of β_{21} and β_{22}, which show NRD mechanism for response variables. Also estimated variance of shared random effect is another important notice of this table. The table shows also that other estimated values of regression coefficients such as types of drug (ddI and ddC) and gender are not significant.

6. CONCLUSION

Use of appropriate and general distributional assumptions in analysing data sets, specially in the presence of missing values, is important. In this paper, we have developed a shared parameter model under skew-normal distribution assumption and Bayesian approach for analysing data with nonignorable missingness. Some simulation studies show that our model has an adequately well performance. The model is also robust with respect to some modifications of distributional assumptions. The proposed method is illustrated using a real AIDS data set. For implementation of the proposed model we have used WinBUGS package. Our model is quite general and can be applied to other structures of mixed effects model. The proposed model can be also developed to non-linear or semiparametric mixed effect models.

APPENDIX

Let $\Xi = (Y^{obs}, R, X_1, X_2)$ and $\vartheta = (\beta_1, \beta_2, \gamma, D, \sigma_e^2, \delta_e, \delta_b, b, U_e, U_b, Y^{mis})$, and let ϖ be one of the component of it, we define $\vartheta_{(-\varpi)}$ for the above-mentioned vector when ϖ is omitted from it. Then under the hierarchical model (2), the full conditional distributions are given by:

$$\beta_1 | \vartheta_{(-\beta_1)}, \Xi \sim N(\Omega_{\beta_1}^{-1} \eta_{\beta_1}, \Omega_{\beta_1}^{-1}),$$

where, $\Omega_{\beta_1}=\frac{1}{\sigma_e^2}\sum_{i=1}^{n}x_{1i}^{'}x_{1i}+\Sigma_{\beta_1}^{-1}$ and

$$\eta_{\beta_1}=\frac{1}{\sigma_e^2}\sum_{i=1}^{n}x_{1i}^{'}(y_i-z_i b_i-\delta_e u_{ei})+\Sigma_{\beta_1}^{-1}\ \mu_{\beta_1}.$$

$$\pi\left(\beta_2|\vartheta_{(-\beta_2)},\Xi\right)\propto\prod_{i=1}^{n}\prod_{k=1}^{m}f(r_{ik}|b_i,\beta_2,\gamma)\times\phi(\beta_2|\mu_{\beta_2},\Sigma_{\beta_2}),$$

$$U_{b_i}|\vartheta_{(-U_{b_i})},\Xi\sim N(\Omega_{U_{b_i}}^{-1}\Delta_b D^{-1}b_i,\Omega_{U_{b_i}}^{-1})I_{\{u_{bi}>0\}},$$

such that, $\Omega_{U_{b_i}}=(I+\Delta_b^{'}D^{-1}\Delta_b)$.

$$\pi\left(\gamma|\vartheta_{(-\gamma)},\Xi\right)\propto\prod_{i=1}^{n}\prod_{k=1}^{m_i}f(r_{ik}|b_i,\beta_2,\gamma)\times\pi(\gamma),$$

$$\sigma_e^2|\vartheta_{(-\sigma_e^2)},\Xi\sim I\Gamma(\alpha^*_{\sigma_e^2},\alpha_e+np+1)$$

where $\alpha^*_{\sigma_e^2}=\frac{1}{2}\sum_{i=1}^{n}(y_i-\mu_{\sigma_e^2})'(y_i-\mu_{\sigma_e^2})+\frac{1}{\tau_e}$ and $\mu_{\sigma_e^2}=x_{1i}\beta_1+z_i b_i+\delta_e u_{ei}$.

$$\pi(D|\vartheta_{(-D)},\Xi)\propto\prod_{i=1}^{n}\exp\left\{-\frac{1}{2}(b_i-\Delta_b u_{bi})'D^{-1}(b_i-\Delta_b u_{bi})\right\}$$

$$\times|D|^{-\frac{\tau_D+p+1}{2}}\exp\left\{tr\left(\Psi_D D^{-1}\right)\right\},$$

$$\delta_e|\vartheta_{(-\delta_e)},\Xi\sim N(\Omega_{\delta_e}^{-1}\eta_{\delta_e},\Omega_{\delta_e}^{-1}),$$

where, $\eta_{\delta_e}=\sum_{i=1}^{n}\frac{u_{ei}^{'}}{\sigma_e^2}(y_i-X_{1i}\beta_1-Z_i b_i)+\frac{\mu_{\delta_e}}{\kappa_{\delta_e}}$,

$\Omega_{\delta_e}=\frac{\sum_{i=1}^{n}u_{ei}^{'}u_{ei}}{\sigma_e^2}+\frac{1}{\kappa_{\delta_e}}$. and

$$\pi(\delta_b|\vartheta_{(-\delta_b)},\Xi)\propto\prod_{i=1}^{n}f(b_i|D,\Delta_b,u_{bi})\times\pi(\delta_b),$$

also, $b_i|\vartheta_{(-b_i)},\Xi\sim N_q(\psi_b\left(\frac{Z_i^{'}(y_i-X_{1i}\beta_1)}{\sigma_e^2}+D^{-1}\Delta_b u_{b_i}\right),\psi_b)$

where, $\psi_b=\left[\frac{Z_i^{'}Z_i}{\sigma_e^2}+D^{-1}\right]^{-1}$.

Let $x_{1i,mis}$ and $z_{i,mis}$ denote the components of x_{1i} and z_i, respectively. These are associated with the missing components of $y_{i,mis}$. Also, let $Y_{i,mis}$ be a k_i-dimensional vector $k_i<n_i$. The full conditional distribution for $y_{i,mis}$ is given by:

$$Y_{i,mis}|\vartheta_{(-y_{i,mis})},y_{i,obs},t\sim N_{k_i}(x_{1i,mis}\beta_1+z_{i,mis}b_i,\sigma_\varepsilon^2 I_{k_i}).$$

The derivation of the full conditional distributions is straightforward. The MCMC methods such as the Gibbs sampler and the Metropolis-Hastings algorithm can be used to draw samples from the full conditional distributions. The Gibbs sampler works by drawing samples iteratively from conditional posterior distributions.

REFERENCES

[1] Azzalini A. A class of distribution which includes the normal ones. Scand J Statist 1985; 12: 171-78.

[2] Jara A, Quintana F, San Martn E. Linear mixed models with skew-elliptical distributions: A Bayesian approach. Comput Statist Data Anal 2008; 52: 5033-45. http://dx.doi.org/10.1016/j.csda.2008.04.027

[3] De la Cruz R, Branco M. Bayesian analysis for nonlinear regression model under skewed errors, with application in growth curves. Biometr J 2009; 51: 588-609. http://dx.doi.org/10.1002/bimj.200800154

[4] Huang Y, Dagne G. A Bayesian approach to joint mixed-effects models with a skew-normal distribution and measurement errors in covariates. Biometrics 2011; 67: 260-69. http://dx.doi.org/10.1111/j.1541-0420.2010.01425.x

[5] Abanto-Valle CA, Bandyopadhyay D, Lachos VH. Robust Bayesian Analysis of Heavy-tailed Stochastic Volatility Models using Scale Mixtures of Normal Distributions. Computat Statist Data Anal 2010; 54: 2883-98. http://dx.doi.org/10.1016/j.csda.2009.06.011

[6] Abanto-Valle CA, Migon H, Lachos VH. Bayesian Analysis of Heavy-tailed Stochastic Volatility in Mean model using Scale Mixtures of Normal distributions. J Statist Planning Inference 2011; 141: 1875-87. http://dx.doi.org/10.1016/j.jspi.2010.11.039

[7] Zeller CB, Lachos VH, Labra F. Local influence analysis for regression models with skew-normal independent distributions. J Appl Statist 2011; 38: 343-68. http://dx.doi.org/10.1080/02664760903406504

[8] Ferreira CS, Lachos VH, Bolfarine H. Skew scale mixtures of normal distributions: properties and estimation. Statist Methodol 2011; 8: 154-71. http://dx.doi.org/10.1016/j.stamet.2010.09.001

[9] Lachos VH, Garibay V, Ortega E. A nonlinear model with skew-normal errors. Statist Papers 2010; 51: 547-58. http://dx.doi.org/10.1007/s00362-008-0139-y

[10] Huang Y. Segmental modeling of viral load changes for HIV longitudinal data with skewness and detection limits. Statist Med 2013; 32(2): 319-34. http://dx.doi.org/10.1002/sim.5527

[11] Lin I, Ho J, Chen L. Analysis of multivariate skew-normal models with incomplete data. J Multivar Anal 2009; 100: 2337-51. http://dx.doi.org/10.1016/j.jmva.2009.07.005

[12] Baghfalaki T, Ganjali M. An EM estimation approach for analyzing bivariate skew-normal data with non-monotone

missing values. Commun Statist Theory Methods 2011; 40(9): 1671-86. http://dx.doi.org/10.1080/03610921003637454

[13] Baghfalaki T, Ganjali M. An ECM estimation approach for analyzing multivariate skew-normal data with dropout. Commun Statist Simulat Comput 2012; 41: 1970-88. http://dx.doi.org/10.1080/03610918.2011.627099

[14] Ganjali M, Baghfalaki T, Khazaei M. A linear mixed model for analyzing longitudinal skew-normal responses with random dropout. J Korean Statist Soc 2012; 42(2): 149-60. http://dx.doi.org/10.1016/j.jkss.2012.06.004

[15] Huang Y, Dagne G. Bayesian Semiparametric Nonlinear Mixed- Effects Joint Models for Data with Skewness, Missing Responses, and Measurement Errors in Covariates. Biometrics 2012; 68: 943-53. http://dx.doi.org/10.1111/j.1541-0420.2011.01719.x

[16] Little RJ, Rubin D. Statistical analysis with missing data. Second edition. New york, Wiley 2002.

[17] Roy J. Modeling longitudinal data with nonignorable dropouts using a latent dropout class model. Biometrics 2003; 59: 829-36. http://dx.doi.org/10.1111/j.0006-341X.2003.00097.x

[18] Gao S. A shared random effect parameter approach for longitudinal dementia data with nonignorable missing data. Statist Med 2004; 23: 211-19. http://dx.doi.org/10.1002/sim.1710

[19] Albert PS, Follmann DA. A random effects transition model for longitudinal binary data with informative missingness. Statist Neerlandic 2003; 57: 100-11. http://dx.doi.org/10.1111/1467-9574.00223

[20] Albert PS, Follmann DA. Shared-parameter models. In G. Fitzmaurice, M. Davidian, G. Verbeke, and G. Molenberghs (Eds.), Longitudinal data analysis. Boca Raton, FL: Chapman and Hall 2009; pp. 433-452.

[21] Vonesh EF, Greene T, Schluchter MD. Shared parameter models for the joint analysis of longitudinal data with event times. Statist Med 2006; 25: 143-63. http://dx.doi.org/10.1002/sim.2249

[22] Yuan Y, Little RJA. Meta-Analysis of Studies with Missing Data. Biometrics 2009; 65: 487-96. http://dx.doi.org/10.1111/j.1541-0420.2008.01068.x

[23] Azzalini A, Dalla-Valle A. The multivariate skew-normal distribution. Biometrika 1996; 83: 715-26. http://dx.doi.org/10.1093/biomet/83.4.715

[24] Sahu SK, Dey DK, Branco M. A new class of multivariate skew distributions with applications to Bayesian regression models. Can J Statist 2003; 31(2): 129-50. http://dx.doi.org/10.2307/3316064

[25] Lachos VH, Bandyopadhyay D, Dey DK. Linear and nonlinear mixed-effects models for censored HIV viral loads using normal/independent distributions. Biometrics 2011; 67: 1594-604. http://dx.doi.org/10.1111/j.1541-0420.2011.01586.x

[26] Arellano-Valle RB, Bolfarine H, Lachos VH. Bayesian inference for skew-normal linear mixed models. J Appl Statist 2007; 34: 663-82. http://dx.doi.org/10.1080/02664760701236905

[27] Spiegelhalter DJ, Best NG, Carlin BP, Lindevan der A. Bayesian measures of model complexity and fit. J Royal Statist Soc Ser B 2002; 64: 583-16. http://dx.doi.org/10.1111/1467-9868.00353

[28] Wu MC, Carroll RJ. Estimation and comparison of changes in the presence of informative right censoring by modelling the censoring process. Biometrics 1988; 44: 175-88. http://dx.doi.org/10.2307/2531905

[29] Ten Have TR, Kunselman AR, Pulkstenis EP, Landis JR. Mixed effects logistic regression models for longitudinal binary response data with informative drop-out. Biometrics 1998; 54: 367-83. http://dx.doi.org/10.2307/2534023

[30] Albert PS, Follmann DA. Modeling repeated count data subject to informative dropout. Biometrics 2000; 56: 667-77. http://dx.doi.org/10.1111/j.0006-341X.2000.00667.x

[31] Carlin BP, Louis TA. Bayesian Methods for Data Analysis. Boca Raton, FL: Chapman and Hall - CRC Press 2009.

[32] Guo X, Carlin BP. Separate and joint modeling of longitudinal and event time data using standard computer packages. Am Statist 2004; 58: 16-24. http://dx.doi.org/10.1198/0003130042854

[33] Goldman AI, Carlin BP, Crane LR, Launer C, Korvick JA, Deyton L, Abrams DI. Response of CD4+ and Clinical Consequences to Treatment Using ddI or ddC in Patients with Advanced HIV Infection. J Acquired Immune Deficiency Syndromes Human Retrovirol 1996; 11: 161-69. http://dx.doi.org/10.1097/00042560-199602010-00007

[34] Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. Statist Sci 1992; 7: 457-11. http://dx.doi.org/10.1214/ss/1177011136

Received on 30-11-2013 Accepted on 13-02-2014 Published on 30-04-2014

http://dx.doi.org/10.6000/1929-6029.2014.03.02.4