The International Journal of Biostatistics

Volume 8, *Issue* 1 2012 *Article* 27

Bayesian inference for a nonlinear mixed-effects Tobit model with multivariate skew-t distributions: application to AIDS studies

Getachew Dagne, *University of South Florida*
Yangxin Huang, *University of South Florida*

Recommended Citation:
Dagne, Getachew and Huang, Yangxin (2012) "Bayesian inference for a nonlinear mixed-effects Tobit model with multivariate skew-t distributions: application to AIDS studies," *The International Journal of Biostatistics*: Vol. 8: Iss. 1, Article 27.
DOI: 10.1515/1557-4679.1387

Bayesian inference for a nonlinear mixed-effects Tobit model with multivariate skew-t distributions: application to AIDS studies

Getachew Dagne and Yangxin Huang

Abstract

Censored data are characteristics of many bioassays in HIV/AIDS studies where assays may not be sensitive enough to determine gradations in viral load determination among those below a detectable threshold. Not accounting for such left-censoring appropriately can lead to biased parameter estimates in most data analysis. To properly adjust for left-censoring, this paper presents an extension of the Tobit model for fitting nonlinear dynamic mixed-effects models with skew distributions. Such extensions allow one to specify the conditional distributions for viral load response to account for left-censoring, skewness and heaviness in the tails of the distributions of the response variable. A Bayesian modeling approach via Markov Chain Monte Carlo (MCMC) algorithm is used to estimate model parameters. The proposed methods are illustrated using real data from an HIV/AIDS study.

KEYWORDS: Bayesian method, nonlinear HIV dynamics, Tobit model, skew-normal distribution

Author Notes: The authors gratefully acknowledge the editor and an anonymous referee for their insightful comments and constructive suggestions that led to a marked improvement of the article. This research was partially supported by NIMH grant R01MH040859-22 to G. Dagne, and NIAID grant AI080338 to Y. Huang.

1 Introduction

In AIDS studies, researchers have recently shown great interest in modeling viral load (plasma HIV-1 RNA copies) data after initiation of a potent antiretroviral (ARV) treatment (Paxton et al., 1997; Coombs et al., 1998). Viral load is a measure of the amount of actively replicating virus and is used as a marker of disease progression among HIV-infected patients. Viral load measurements are often subject to left censoring due to a lower limit of quantification. A limit of detection (LOD) depends upon the assay used, ranging from 500 copies/ml for the first assays available in the mid-nineties to 50 copies/ml for today's ultra sensitive assay (Schockmel et al., 1997). Despite the improvement in assay sensitivity recently, left censoring of viral load data still remains a critical issue, and the methods proposed in the literature for addressing this issue use either the observed LOD or some arbitrary value, such as LOD/2 (Hornung and Reed, 1990; Huang and Dagne, 2010). Those approaches usually lead to biased predictions that are systematically higher than predictions based on the true unknown values of LOD (Paxton et al., 1997; Hughes, 1999; Jacqmin-Gadda et al., 2000; Thi*é*baut et al., 2006). Thus, the objective of this paper is to correctly model left-censored data and explore the use of flexible skew-elliptical distributions to properly account for skewness and heaviness in tails of an asymmetrical distribution of viral loads.

Our approach is to treat all left-censored observations as missing values of a latent variable, instead of replacing them by an arbitrary value as is usually done in the literature (Reilly et al., 2004; Scirica, 2007; Huang and Dagne, 2010), and those above LOD as observed values for making inference about HIV viral dynamics. For modeling left-censored data, the standard linear Tobit model often uses the maximum likelihood estimation (MLE) method (Lynn, 2001; Thi*é*baut et al., 2006; Sattar et al., 2011) to estimate the parameters of interest. However, its drawback is that it gives inconsistent results when the normality assumption is violated (Arabmazar and Schmidt, 1982; Lorimer and Kiermeier, 2007).

To address these issues in the context of left-censoring and skewness, we propose to use skew-elliptical distributions (Sahu et al, 2003; Genton, 2004) instead of the normal distribution which is currently used. Skew-elliptical distributions in which multivariate skew-normal (SN) and skew-t (ST) distributions are special cases are appropriate for analyzing skewed data such as those presented in Figure 1. Figure 1(a) displays the histogram of repeated viral load measurements (in natural log scale) for 44 subjects enrolled in the AIDS clinical trial study–A5055 (Acosta, 2004). It seems that for this data set,

which is analyzed in this paper, the viral load responses are highly skewed even after log-transformation. To properly analyze such skewed and left-censored data, we develop a nonlinear dynamic mixed-effects model using skew-elliptical distributions under a Bayesian approach. A Bayesian method via MCMC also allows us to treat the left-censored observations as *missing values* and to predict them using a predictive posterior distribution.

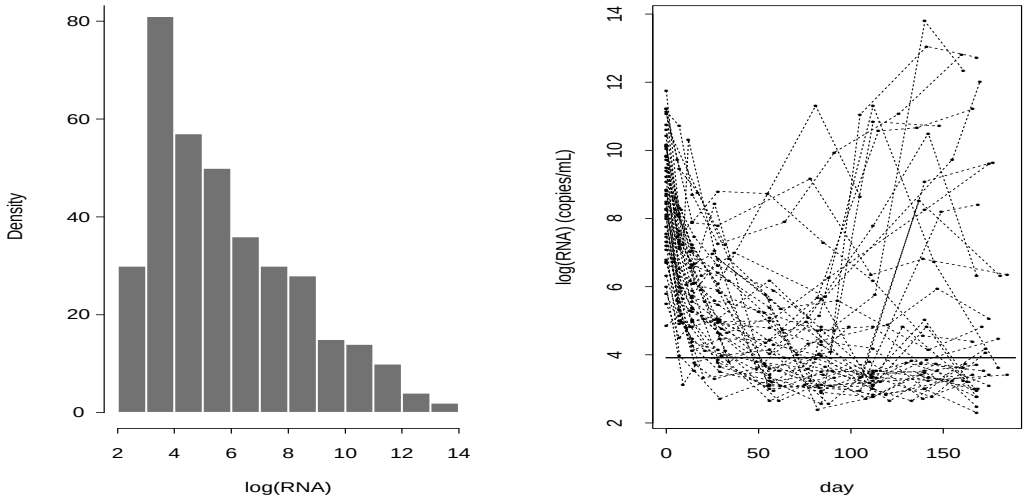

Figure 1: (a) The histogram of viral load; (b) spaghetti plot of viral load (in *log* scale) measured from 44 patients in an AIDS clinical trial study. The horizontal line represents the value of LOD at $log(50)$

A flexible hierarchical representation of our proposed skew-elliptical models makes it easier to implement the MCMC algorithm using a freely available WinBUGS software (Lunn, 2000) and has a computational effort similar to the one necessary to fit the normal version of the models. In particular, the multivariate skew-normal Tobit model and multivariate skew-t model are fit to AIDS data to compare their performance with that of the standard normal Tobit model. The rest of the paper is organized as follows. In Section 2, we develop nonlinear mixed-effects (NLME) Tobit models with multivariate skew-elliptical distributions in general forms. In Section 3, we present the Bayesian inferential procedures. The proposed methodologies are illustrated using the AIDS data set and the results are reported in Section 4. Finally, the paper concludes with discussions in Section 5.

2 Skew-elliptical nonlinear mixed-effects models

2.1 Motivating data

Our research was motivated by the A5055 study considered in (Acosta et al., 2004; Huang et al, 2006). In this study, 44 HIV-1 infected patients were treated with a potent ARV regimen. RNA viral load was measured in copies/mL at study days 0, 7, 14, 28, 56, 84, 112, 140 and 168 of follow-up. Covariates such as CD4 cell counts were also measured throughout the study. Among the 44 eligible patients, the number of viral load measurements for each patient varies from 3 to 9 measurements, with an average of 8.11 and a standard deviation of 1.40. In this study, the viral load detectable limit is 50 copies/mL, and there are 54 out of 357 (about 15 percent) of all viral load measurements that are below the LOD. The HIV-1 RNA measures below this limit are not considered reliable, therefore we impute them based on the Tobit model discussed in the next Section.

2.2 Model specification

In AIDS studies, either viral load or CD4 count or both may be treated as outcome variables. However, CD4 count is more often used as an outcome variable for long follow-up trials or advanced patient populations. But for trials (e.g., A5055) which have short follow-up periods, viral load is often used as an outcome variable of interest, and CD4 count is considered as a covariate to help predict viral load in the HIV dynamic models considered here. The viral load is measured by the numbers of HIV-1 RNA copies per mL in plasma, and it is subject to left censoring due to limitation of the assay. An approach we present in this paper treats censored values as latent (unobserved) continuous observations that have been left-censored. This idea was popularized by Tobit (Tobin, 1958) and the resulting model is commonly referred to as the Tobit model. Analytically, letting $y_{ij}^*(t)$ denote the latent HIV viral load (on log scale) that would be measured if the assay did not have a lower detectible limit τ_{ij} for patient i $(i = 1, ..., m)$ at occasion j $(j = 1, ..., n_i)$, the Tobit model can be formulated as:

$$y_{ij}(t) = \begin{cases} y_{ij}^*(t), & \text{if } y_{ij}^*(t) > \tau_{ij} \\ missing, & \text{if } y_{ij}^*(t) \leq \tau_{ij}, \end{cases} \tag{1}$$

where τ_{ij} is a non-stochastic LOD, which in our example is equivalent to log(50). Note that the value of $y_{ij}(t)$ is missing when $y^*_{ij}(t)$ is less than or equal to τ_{ij}, and the missingness process, as a reviewer pointed out, is informative since the probability of missingness depends on an unobserved value where the process is known.

Let $\boldsymbol{y}_i = (y_{i1}(t), \cdots, y_{in_i}(t))^T$, and $\boldsymbol{X}_i$ be an $n_i \times q$ time-invariant and time-varying covariates (e.g., CD4 value) measured from the ith subject. For modeling the response variable $\boldsymbol{y}_i$, a viral dynamic model can be derived from a dynamic compartmental analysis (Huang et al., 2006; Ho et al., 1995; Wu and Ding, 1999), which is represented by the following NLME Tobit model.

$$\boldsymbol{y}_i \;\; = \;\; h(\boldsymbol{t}_i, \boldsymbol{X}_i, \boldsymbol{\beta}, \boldsymbol{b}_i) + \boldsymbol{e}_i, \tag{2}$$

where $h(\cdot)$ is a known nonlinear dynamic function (a specific example is given in equation (8) of Section 4 below), $\boldsymbol{t}_i = (t_{i1}, \cdots, t_{in_i})^T$ is time points at which measurements are taken, $\boldsymbol{\beta}$ is a $q \times 1$ vector of population level parameters associated with fixed-effects, and $\boldsymbol{b}_i$ is an $s-$dimensional random effects for subject i, which is distributed as normal with mean zero and variance-covariance matrix $\boldsymbol{\Sigma}_b$. The within-subject random error from (2) is $\boldsymbol{e}_i = (e_{i1}, \ldots, e_{in_i})^T$, and we assume that it follows a multivariate skew-elliptical distribution (Sahu et al., 2003; Azzalini and Dalla-Valle, 1996; Arellano-Valle and Genton, 2005) in order to incorporate skewness which is very often inherent in virologic responses (Genton, 2004). Thus, conditional on $\boldsymbol{b}_i$, the response variable $\boldsymbol{y}_i$ is distributed as

$$\boldsymbol{y}_i|\boldsymbol{b}_i; \boldsymbol{\beta}, \boldsymbol{\Sigma}_{n_i}, \boldsymbol{\Delta}_{n_i}, g^{n_i} \sim SE_{n_i}\left(h(\boldsymbol{t}_i, \boldsymbol{X}_i, \boldsymbol{\beta}, \boldsymbol{b}_i), \boldsymbol{\Sigma}_{n_i}, \boldsymbol{\Delta}_{n_i}, g^{n_i}\right), \tag{3}$$

where $SE_{n_i}(\mu_i, \boldsymbol{\Sigma}_{n_i}, \boldsymbol{\Delta}_{n_i}, g^{n_i})$ refers to a n_i-dimensional skew-elliptical distribution with n_i-dimensional location vector μ_i, n_i-dimensional scale matrix $\boldsymbol{\Sigma}_{n_i}$, n_i-dimensional skewness matrix $\boldsymbol{\Delta}_{n_i}$, and a density generator function g^{n_i}. The two commonly used density generator functions are $g^{n_i}(u) = (2\pi)^{-n_i/2} exp(-u/2)$ and $g^{n_i}(u) = \frac{\Gamma((\nu+n_i)/2)}{\Gamma(\nu/2)(\nu\pi)^{n_i/2}}$
$\times (1 + u/\nu)^{-(\nu+n_i)/2}$ for $u > 0$, where $\nu > 0$ is the degrees of freedom parameter. These functions lead to the multivariate SN and ST models, respectively.

For purpose of accounting both skewness and heaviness in tails, we focus our attention to the multivariate ST distribution for the conditional model $\boldsymbol{y}_i|\boldsymbol{\beta}, \boldsymbol{b}_i, \boldsymbol{\Sigma}_{n_i}, \boldsymbol{\Delta}_{n_i} \sim ST_\nu\left(h(\boldsymbol{t}_i, \boldsymbol{X}_i, \boldsymbol{\beta}, \boldsymbol{b}_i), \boldsymbol{\Sigma}_{n_i}, \boldsymbol{\Delta}_{n_i}\right)$, where the conditional mul-

tivariate ST density function is given by

$$\begin{aligned} f(\boldsymbol{y}_i|\boldsymbol{b}_i) &= \eta_{n_i}|Q_{n_i}|^{-1/2}\left[1+\frac{(\boldsymbol{y}_i-h(\boldsymbol{t}_i,\boldsymbol{X}_i,\boldsymbol{\beta},\boldsymbol{b}_i))^T Q_{n_i}^{-1}(\boldsymbol{y}_i-h(\boldsymbol{t}_i,\boldsymbol{X}_i,\boldsymbol{\beta},\boldsymbol{b}_i))}{\nu}\right]^{-\frac{\nu+n_i}{2}} \\ &\quad \times \int_0^\infty t_\nu(\boldsymbol{z}|\boldsymbol{\Delta}_{n_i}Q_{n_i}^{-1}(\boldsymbol{y}_i-h(\boldsymbol{t}_i,\boldsymbol{X}_i,\boldsymbol{\beta},\boldsymbol{b}_i)),D_i^*)d\boldsymbol{z} \end{aligned} \tag{4}$$

where $\eta_{n_i}=\frac{2^{n_i}\Gamma((\nu+n_i)/2)}{\Gamma(\nu/2)(\nu\pi)^{n_i/2}}$, $t_\nu(\cdot|\boldsymbol{\zeta},\boldsymbol{\Omega})$ is a n_i-variate t-distribution with ν degrees of freedom, location $\boldsymbol{\zeta}$, and scale $\boldsymbol{\Omega}$, and $Q_{n_i}=\boldsymbol{\Sigma}_{n_i}+\boldsymbol{\Delta}_{n_i}^2$, $D_i=[\boldsymbol{y}_i-h(\boldsymbol{t}_i,\boldsymbol{X}_i,\boldsymbol{\beta},\boldsymbol{b}_i)]^T Q_{n_i}^{-1}[\boldsymbol{y}_i-h(\boldsymbol{t}_i,\boldsymbol{X}_i,\boldsymbol{\beta},\boldsymbol{b}_i)]$, $D_i^*=\frac{\nu+D_i}{\nu+n_i}(I-\boldsymbol{\Delta}_{n_i}Q_{n_i}^{-1}\boldsymbol{\Delta}_{n_i})$, $\boldsymbol{\Delta}_{n_i}=diag(\delta_1,\delta_2,\cdots,\delta_{n_i})$ is a diagonal matrix with skewness vector $\boldsymbol{\delta}_{n_i}=(\delta_1,\cdots,\delta_{n_i})^T$ and its elements determine the level of skewness. For example, if $\delta_j>0$, then the density is skewed to the right in the jth dimension, while it is skewed to the left when $\delta_j<0$. However, if $\delta_j=0$ for all j we get a symmetric multivariate t distribution. For a full discussion of the properties of this distribution, see Sahu et al. (2003), Liu and Dey (2004) and the Appendix. When $\nu\to\infty$ in (4), the multivariate ST distribution approaches the multivariate SN distribution and reduces to the multivariate normal distribution when $\boldsymbol{\Delta}_{n_i}=\mathbf{0}$. Furthermore, for a finite ν and for $n_i=1$ for all i, it becomes the univariate skew-t distribution (Azzalini and Capitanio, 2003). The parameters of model (4) are estimated using a Bayesian approach which is discussed below.

3 Bayesian modeling approach

In this section, we now describe a Bayesian estimation procedure for the model (3) with the ST distribution. Bayesian analysis rests upon computing the posterior probability distribution for model parameters. The posterior probability distribution is the conditional probability distribution of the unknown parameters, given the observed data and weighted by the prior information. To obtain such posterior distributions for the model parameters, first, we present the likelihood function for the observed data under the ST NLME Tobit model. Before doing this, we denote that the observed dependent variable $y_{ij}=y_{ij}^*$ if $c_{ij}=0$, and y_{ij} is left censored if $c_{ij}=1$, where c_{ij} is a censoring indicator, and the latent variable y_{ij}^* was discussed in Section 2. Let $f(\cdot|\cdot)$, $F(\cdot|\cdot)$ denote a probability density function (pdf) and cumulative density function (cdf) of ST, respectively. Conditional on the random variables and some unknown parameters, a detectable measurement y_{ij} contributes $f(y_{ij}|\boldsymbol{b}_i)$, whereas a non-detectable measurement contributes $F(\tau_{ij}|\boldsymbol{b}_i)\equiv Pr(y_{ij}<\tau_{ij}|\boldsymbol{b}_i)$ in the likelihood. We assume that $\boldsymbol{\Sigma}_{n_i}=\sigma^2\boldsymbol{I}_{n_i}$ and letting $\boldsymbol{\Theta}=(\boldsymbol{\beta},\sigma^2,\boldsymbol{\delta}_{n_i},\boldsymbol{\Sigma}_b,\nu)$

be the collection of unknown parameters, the joint distribution (likelihood function) of the observed data $\boldsymbol{y} = (\boldsymbol{y}_1^T, \cdots, \boldsymbol{y}_m^T)^T$ is

$$\begin{aligned} f(\boldsymbol{y}|\boldsymbol{\Theta}) &= \prod_{i=1}^{m} \int \left[\prod_{j=1}^{n_i} (f(y_{ij}|t_{ij}, \boldsymbol{X}_i, \boldsymbol{b}_i, \boldsymbol{\Theta}))^{1-c_{ij}} (F(\tau_{ij}|t_{ij}, \boldsymbol{X}_i, \boldsymbol{b}_i, \boldsymbol{\Theta}))^{c_{ij}} \right] \\ &\quad \times f(\boldsymbol{b}_i|\boldsymbol{\Sigma}_b) d\boldsymbol{b}_i. \end{aligned} \tag{5}$$

Directly using the likelihood in (5) is computationally challenging since it requires the calculation of complex integrals. An alternative is to use the MCMC procedure for estimating the model parameters by exploiting the stochastic representation of the ST distribution given in Sahu et al. (2003). In order to specify the model (3) for MCMC computation it can be shown that, by introducing a random variable vector $\boldsymbol{w}_i = (w_{i1}, \ldots, w_{in_i})^T$, the distribution of $\boldsymbol{y}_i$ conditional on random-effects $\boldsymbol{b}_i$ and $\boldsymbol{w}_i$ can be hierarchically formulated as follows.

$$\begin{aligned} \boldsymbol{y}_i|\boldsymbol{b}_i, \boldsymbol{w}_i; \boldsymbol{\beta}, \sigma^2, \boldsymbol{\delta}_{n_i}, \nu &\sim t_\nu \left(h(\boldsymbol{t}_i, \boldsymbol{X}_i, \boldsymbol{\beta}, \boldsymbol{b}_i) + \boldsymbol{\Delta}_{n_i}[\boldsymbol{w}_i - J(\nu)\mathbf{1}_{n_i}], \sigma^2 \boldsymbol{I}_{n_i} \right), \\ \boldsymbol{w}_i|\sigma^2 &\sim t_\nu(\mathbf{0}, \sigma^2 \boldsymbol{I}_{n_i}) I(\boldsymbol{w}_i > 0),\ \boldsymbol{b}_i|\boldsymbol{\Sigma}_b \sim N\left(\mathbf{0}, \boldsymbol{\Sigma}_b\right), \\ \boldsymbol{\beta} &\sim N_q(\boldsymbol{\beta}_0, \boldsymbol{\Lambda}),\ \nu \sim \exp(\lambda) I(\nu > 3), \\ \sigma^2 &\sim IG(\omega_1, \omega_2),\ \boldsymbol{\Sigma}_b \sim IW(\boldsymbol{\Omega}_b, \zeta_b),\ \boldsymbol{\delta}_{n_i} \sim N(\mathbf{0}, \boldsymbol{\Psi}), \end{aligned} \tag{6}$$

where $J(\nu) = (\nu/\pi)^{1/2}\Gamma((\nu-1)/2)/\Gamma(\nu/2)$ and $\mathbf{1}_{n_i}$ is a vector of unity of size n_i, $\boldsymbol{w}_i$ is a latent variable whose distribution is a truncated multivariate t-distribution, $I(\cdot)$ is an indicator function, $t_\nu(\boldsymbol{\mu}, \boldsymbol{A})$ denote the multivariate $t-$distribution with location parameter $\boldsymbol{\mu}$, scale parameter $\boldsymbol{A}$ and degrees of freedom ν. The degrees of freedom parameter ν is assumed to have an exponential distribution with a range of support greater than 3 for accommodating an ST distribution with parameters of location, scale and skewness. The variance parameters σ^2 and $\boldsymbol{\Sigma}_b$ are assumed to have inverse gamma distribution (IG) and inverse Wishart (IW) distribution, respectively, in order to guarantee proper posteriors. The hyper-parameter matrices $\boldsymbol{\Lambda}$, $\boldsymbol{\Omega}_b$ and $\boldsymbol{\Psi}$ are assumed to be diagonal for ease of implementation.

Next, we write down the joint posterior distribution for the unknown model parameters in (6) based on the observed data $\boldsymbol{y}$ as follow.

$$
\begin{aligned}
f(\boldsymbol{\Theta}|\boldsymbol{y},\boldsymbol{X}) \;\propto\; & \prod_{i=1}^{m}\{\int[\prod_{j=1}^{n_i} t_\nu(h(t_{ij},\boldsymbol{X}_i,\boldsymbol{\beta},\boldsymbol{b}_i)+\delta_{ij}[w_{ij}-J(\nu)],\sigma^2)^{1-c_{ij}} \\
& \times F_\nu(\tau_{ij}|h(t_{ij},\boldsymbol{X}_i,\boldsymbol{\beta},\boldsymbol{b}_i)+\delta_{ij}[w_{ij}-J(\nu)],\sigma^2)^{c_{ij}} \\
& \times t_\nu(w_{ij}|\sigma^2)I(w_{ij}>0)]|\boldsymbol{\Sigma}_b|^{-1/2}\exp(-0.5\boldsymbol{b}_i^T\boldsymbol{\Sigma}_b^{-1}\boldsymbol{b}_i) \\
& \times|\boldsymbol{\Psi}|^{-1/2}\exp(-0.5\boldsymbol{\delta}_{n_i}^T\boldsymbol{\Psi}^{-1}\boldsymbol{\delta}_{n_i})d\boldsymbol{b}_i\}(\sigma^2)^{-(\omega_1+1)}\exp(-\omega_2/\sigma^2) \\
& \times|\boldsymbol{\Lambda}|^{-1/2}\exp(-0.5(\boldsymbol{\beta}-\boldsymbol{\beta}_0)^T\boldsymbol{\Lambda}^{-1}(\boldsymbol{\beta}-\boldsymbol{\beta}_0)) \\
& \times|\boldsymbol{\Sigma}_b|^{-(\eta_b+s+1)/2}\exp(-0.5tr(\Omega_b\boldsymbol{\Sigma}_b^{-1}))\lambda\exp(-\lambda\nu)I(\nu>3),
\end{aligned}
\tag{7}
$$

where η_b is the degree of freedom of inverse Wishart distribution. The integrals in (7) are of high dimensions and also do not have closed forms. Thus, it is quite complicated to obtain marginal posterior distributions for the parameters of interest, $\boldsymbol{\Theta}$. As an alternative, MCMC procedures can be used to simulate direct draws from the full conditional distributions iteratively until convergence is achieved using the Gibbs sampler along with the Metropolis-Hastings algorithm. A single long chain (Geyer, 1992; Raftery and Lewis, 1992) is used for the proposed models. Geyer (1992) argues that using a single longer chain is better than using a number of smaller chains with different initial values. We follow this strategy in our empirical analysis after initial sensitivity analysis.

The commonly used criteria for model selection like BIC and AIC are not appropriate for the hierarchical models (in the presence of random effects), which complicate the counting of the true number of free parameters. To overcome such a hurdle, Spiegelhalter et al. (2002) proposed a Bayesian model comparison criterion, called Deviance Information Criterion (DIC). It can easily be calculated from MCMC samples, and a smaller value of DIC indicates a better fit.

4 Analysis of AIDS data using NLME Tobit models

We illustrate the proposed methods by applying them to the A5055 data set (Acosta et al., 2004) described in Section 2.1. As is evident from Figure 1(b), the inter-patient variations in viral load appear to be large and these variations appear to change over time. Previous studies suggest that the inter-patient variation in viral load may be partially explained by time-varying CD4 cell counts (Huang et al., 2006). Thus, we use CD4 as a covariate in the models to be fitted next. A natural log-transformation is also used in the analysis of viral

load data in order to stabilize the variation of measurement error and speed up estimation algorithm. In addition, to avoid very small (large) estimates which may be unstable, we rescale the original time t (in days) so that the time scale is between 0 and 1.

4.1 Analytical framework

For modeling the viral load, viral dynamic models can be formulated through a system of ordinary differential equations (Huang et al., 2006; Wu and Ding, 1999; Wu and Ding, 1998), especially for two infected cell compartments. It has been thought that they produce a biphasic viral decay (Wu and Ding, 1999; Perelson et al., 1997) in which an effective parametric model may be formulated to estimate viral dynamic parameters. This model plays an important role in modeling HIV dynamics and is defined as

$$y_{ij}(t) \quad = \quad h(t_{ij}, \boldsymbol{X}_i, \boldsymbol{\beta}, b_i) + e_{ij} = \log(V(t_{ij}, \boldsymbol{X}_i, \boldsymbol{\beta}, b_i)) + e_{ij}, \tag{8}$$

where $y_{ij}(t)$ is the natural logarithmic transformation of the observed total viral load measurement for the ith patient ($i = 1, \cdots, 44$) at the jth time point ($j = 1, \cdots, n_i$), and the total viral load is expressed as

$$V(t_{ij}, \boldsymbol{X}_i, \boldsymbol{\beta}, b_i)) = \exp[\alpha_{1i} - \lambda_{1i} t_{ij}] + \exp[\alpha_{2i} - \lambda_{2ij} t_{ij}], \tag{9}$$

where $\exp(\alpha_{1i}) + \exp(\alpha_{2i})$ is the baseline viral load at time $t = 0$ for patient i, λ_{1i} is the first-phase viral decay rate which may represent the minimum turnover rate of productively infected cells (Perelson et al., 1997) and λ_{2ij} is the second-phase viral decay rate which may represent the minimum turnover rate of latently or long-lived infected cells (Perelson et al., 1997). It is of particular interest to estimate the viral decay rates λ_{1i} and λ_{2ij} because they quantify the antiviral effect and hence can be used to assess the efficacy of the antiviral treatments (Ding and Wu, 1999).

Looking at the profiles of the total viral load in logarithmic scale for the 44 patients in Figure 1(b), one can see that the rate change in viral load appears to vary substantially across patients, reflecting both biological variation and systematic associations between and within subjects. Because the inter-subject variations seem substantial, we include subject-specific parameters in model (9). These subject-specific dynamic parameters are then modeled as functions of covariates with fixed-effects and random-effects, describing the association between changes in parameters and changes in covariate values (Liu and Wu, 2007; Wu, 2002). Among potential covariates, it has been suggested that variation in the dynamic parameters may be partially associated

with CD4 cell count (Wu, 2002). Thus, we include CD4 which is associated with the second-phase viral decay rate for capturing viral rebound. The subject-specific parameters in (9) are then expressed as

$$\begin{array}{rclcl} \alpha_{1i} & = & \beta_1 + b_{1i}, & \lambda_{1i} = \beta_2 + b_{2i}, \\ \alpha_{2i} & = & \beta_3 + b_{3i}, & \lambda_{2ij} = \beta_4 + \beta_5 t_{ij} + \beta_6 CD4_{ij} + b_{4i}, \end{array} \tag{10}$$

where $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_6)^T$ are population-level parameters, and $\boldsymbol{b}_i = (b_{1i}, \ldots, b_{4i})^T$ are individual-level random-effects which are normally distributed with mean zero and variance $\boldsymbol{\Sigma}_b$.

As shown in Figure 1(a), the histogram of the viral load in logarithms scale clearly indicates its asymmetric nature and it seems logical to fit skew-elliptical NLME model to the data, which also incorporates left-censoring. Accordingly, we consider the following NLME Tobit models with ST and SN distributions which are special cases of the skew-elliptical distribution as described in detail in Section 2.

- **Model I:** A nonlinear mixed-effects model with independent multivariate normal distributions of random errors;
- **Model II:** A nonlinear mixed-effects model with independent multivariate skew-normal distributions of random errors;
- **Model III:** A nonlinear mixed-effects model with independent multivariate skew-t distributions of random errors.

In order to carry out the Bayesian analysis for these models, we need to prescribe the prior distributions for the parameters. Prior distributions for the parameters involved are the same for the three models for comparison purposes. In particular, (i) fixed-effects are taken to be independent normal distribution $N(0, 100)$ for each component of the population parameter vectors $\boldsymbol{\beta}$. (ii) For the precision parameter σ^2, we assume an inverse gamma prior distribution, $IG(0.01, 0.01)$ so that the distribution has mean 1 and variance 100. (iii) The prior for the variance-covariance matrix of the random-effects $\boldsymbol{\Sigma}_b$ is taken to be inverse Wishart distributions $IW(\boldsymbol{\Omega}_b, \zeta_b)$ with covariance matrix $\boldsymbol{\Omega}_b = diag(0.01, 0.01, 0.01, 0.01)$ and degrees of freedom $\zeta_b = 5$. (iv) For the skewness parameter δ , we choose independent normal distribution $N(0, 100)$, where we assume that $\boldsymbol{\delta}_{n_i} = \delta \mathbf{1}_{n_i}$ to indicate that we are interested in skewness of overall viral load data.

Based on the likelihood function and the prior distributions specified above, the MCMC sampler was implemented using WinBUGS, and the program codes are available from the first author upon request. Convergence of

the MCMC implementation was assessed using standard tools (such as trace plots, ACF plots) within WinBUGS, and it was achieved after 100,000 iterations. After initial 100,000 burn-in iterations, every 40th MCMC sample thereafter is retained from the next 400,000 iterations, obtaining 10,000 samples for subsequent posterior inference of the unknown parameters. The computational burden is mild considering the highly non-linear nature of the models fitted. To fit Model III (skew-t) it took about 4 hours on a Window PC with Intel Core 2 Quad CPU 2.66GHz and RAM of 8.0 GB.

4.2 Results

Table 1 reports the values of DIC, which help us determine how assumptions of skew-elliptical distributions may contribute to virologic response in contrast to that of the normal distribution. First, we see that Model I with normality assumption does not fit the data well since its DIC value (862.781) is the biggest. Next, Model II with the SN distribution has a smaller DIC value than that of Model I, showing that accounting for skewness gives a better fit. Finally, Model III with the ST distribution has the smallest DIC value (484.653) which gives a relatively better fit to the data where skewness and heaviness at the tails may exist than either Model II or Model I .

Table 1: Comparison Indices among Model I (Normal), Model II (SN) and Model III (ST)

Model	Deviance	p_D	DIC
I	791.471	71.310	862.781
II	561.221	51.986	613.206
III	404.323	80.330	484.653

To assess the goodness-of-fit of the three models, the plots of residuals against fitted values (left panel), fitted values versus observed values (middle panel) and Q-Q plots (right panel) are presented in Figure 2. The residuals for Models II and III tend to have narrow ranges around zero than those of Model I, implying a better fit. Looking at the plots of the observed values versus the fitted values for the three models in the second column of Figure 2, it seems that Model II and Model III provide better fit to the observed data compared with Model I where the random error is assumed to be normal. In these plots, the horizontal line refers to the lower detection limit, which is log(50).

The observed values below this limit are unreliable and we do not know the "true" values corresponding to them either. The best thing we can do is to predict them based on the proposed Tobit models using posterior predictive distributions of the viral load. The Q-Q plots in the right panel suggest that both Model II (SN) and Model III (ST) give a better goodness-of-fit to the data than that of Model I (normal). Furthermore, Model III does a relatively better job in accounting heaviness in the tails in addition to skewness since the extreme values at right side of the plot are closer to the straight line than those of Model II. In the lower tail, however, the extreme values are not observed values but rather predicted values from the posterior predictive distributions of missing values below detection limit. The Q-Q plots at the lower tails for both Models II and III seem to provide similar results because of a lesser effect of kurtosis in these predicted viral loads.

Table 2 gives posterior means (PM), standard deviations (SD), and the 95 percent credible intervals (in terms of the 2.5 and 97.5 percentiles) of the parameters based on Models I, II and III. The findings in Table 2, particularly for the fixed effects $(\beta_2, \beta_4, \beta_5, \beta_6)$ which are parameters of the first-phase decay rate λ_1 and the second-phase decay rate λ_2 in the exponential HIV viral dynamics, show that the estimates are not all significantly different from zero. Nevertheless, the estimate of β_6 which is the coefficient of CD4 covariate is positive and significant since the 95% credible intervals do not contain zero. This means that CD4 has a significantly positive effect on the second-phase viral decay rate, suggesting that the CD4 covariate may be an important predictor of the second-phase viral decay rate during the HIV-1 RNA process. An increase in CD4 cell count may be associated with a faster viral decay in the late stage, because of the fact that higher CD4 cell counts suggest a higher turnover rate of lymphocyte cells.

Even though the posterior means of β_6 are slightly different in values across the three models considered, there are marked differences in posterior means of the precision parameter σ^2 of the viral load: 1.949 for Model I, 0.4739 for Model II, and 0.4112 for Model III. The corresponding estimates of the variances of the error terms for skew-normal and skew-t are computed using the variance formulas given in Appendix A (equations (11-12)), and they are 1.486 and 1.304, respectively. These results show that skew-t has smaller error variability than skew-normal for our data. This fact is also confirmed by the smaller value of DIC from Table 1 favoring skew-t, which shows a relatively better fit in picking up the skewness and kurtosis simultaneously. Thus, Model III, which is based on the ST distribution for the random error term, seems a preferred model for use in modeling HIV-1 RNA viral load when there are censoring and skewness in the data.

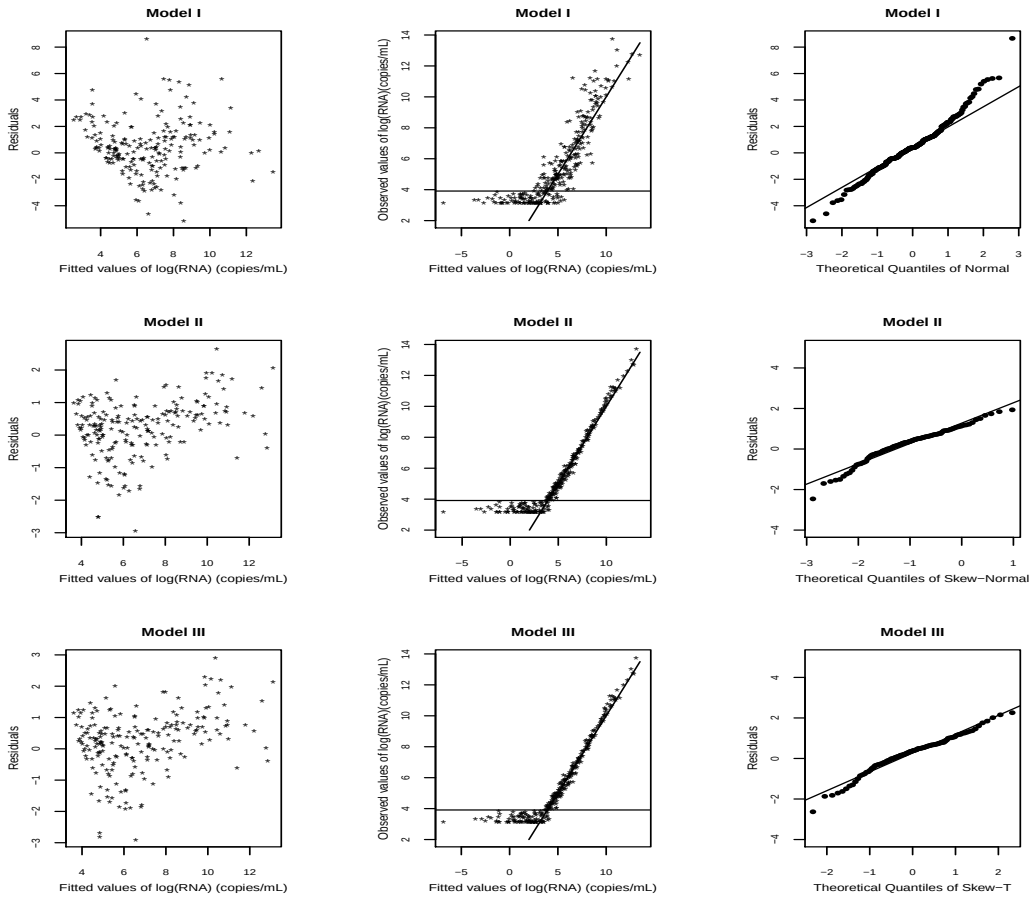

Figure 2: Plots of goodness-of-fit statistics for three models: (i) The first column has the residual plots against the fitted values, (ii) the second has plots of the fitted against observed log(RNA) values, and (iii) the last column has the Q-Q plots of residuals.

Table 2: Estimated posterior means (PM) of fixed-effects, precision, skewness, and degree of freedom parameters, standard deviation (SD) and 95% credible intervals with lower limit (L_{CI}) and upper limit (U_{CI}) based on Models I-III.

Model		β_1	β_2	β_3	β_4	β_5	β_6	σ	δ	ν
I	PM	8.253	25.380	2.592	-4.073	2.122	2.873	1.949	–	–
	L_{CI}	7.692	16.450	-1.355	-14.440	-4.292	1.597	1.498	–	–
	U_{CI}	8.811	34.94	5.049	3.897	9.622	4.285	2.521	–	–
	SD	0.286	4.735	1.679	4.896	3.692	0.691	0.263	–	–
II	PM	8.267	24.590	2.751	-2.837	0.749	3.029	0.474	1.864	–
	L_{CI}	7.679	15.780	-0.314	-11.27	-4.954	1.722	0.016	0.486	–
	U_{CI}	8.836	33.900	5.167	4.462	6.742	4.488	1.689	2.417	–
	SD	0.295	4.592	1.415	4.026	3.099	0.714	0.431	0.477	–
III	PM	8.312	25.900	2.977	-2.425	0.460	2.885	0.411	1.748	139.1
	L_{CI}	7.679	16.240	-1.083	-12.45	-4.843	1.585	0.014	-0.181	4.191
	U_{CI}	8.920	35.290	5.387	5.147	6.899	4.391	1.524	2.406	634.5
	SD	0.3154	4.776	1.633	4.469	3.21	0.727	0.391	0.569	178.0

The posterior mean of the skewness parameter (δ) of Model II (SN) is 1.864 with 95% credible intervals $(0.4857, 2.417)$, suggesting that there is a positive and significant skewness in the data; this confirms the fact that the distribution of the original data is skewed even after taking log-transformation. Thus, incorporating skewness parameter in the modeling of the data is recommended. Furthermore, when heaviness in the tails is taken into account using the ST distribution, the estimate of the skewness parameter is 1.748 which is positive indicating positive skewness but not statistically strong since the 95% credible interval contains zero. This may be due to the fact that an additional parameter ν for heaviness in the tails was estimated lessening the effect of skewness.

We now focus on the lower end of the distribution of the viral load where there is left-censoring. As it was mentioned in the introduction section, the current assay techniques for quantifying HIV-RNA viral load may not give accurate readings below a LOD, which in our data is 50 copies/mL. In our analysis, we treated those inaccurate observed viral loads as missing values and predict them using the proposed NLME Tobit models. Note that the main advantage of our proposed Tobit models is their ability to predict the true viral loads below LOD based on a latent variable approach with different specifications of error distributions. The results of the fits of these models for values below LOD are depicted in Figure 3, where the histograms show the distribution of the observed but inaccurate values (upper left) LOD and the predicted values (on log-scale) under normal, SN, and ST distributions (Figures 3(b-d)). The dotted line shows the LOD value at log(50). It can be seen from the histograms that most observed values are piled up in the

lower end of the range in the first histogram (upper left) due to left-censoring, whereas for the NLME Tobit models under all the three models (normal, SN and ST), the predicted values of the unobserved viral load below LOD are spread out as expected (see Figures 3(b-d)). However, for the case of NLME Tobit model with normal error some predicted values exceeded the LOD, suggesting bad fits. The NLME Tobit model with SN distribution (Model II) shows an improvement over the NLME Tobit model with normal distribution (Model I) by giving few predicted values greater than LOD. Likewise, NLME Tobit model with ST distribution (Model III) provides better predictions than that of Model I (see Figure 3(d)). When we compare Model II and Model III in terms of their performance in predicting viral loads below LOD, we see that Model III gives more plausible values in the sense that the distribution of the predicted viral loads is portrayed as proportionally increasing towards the LOD which is log(50). This distribution is relatively smooth and closely fits the lower part of the whole distribution of the predicted viral load values under Model III as expected implying that Model III with ST distribution is the best model. This finding also confirms the conclusion made using DIC.

5 Discussion and conclusion

In this paper, we have considered a Bayesian approach for estimating parameters in asymmetric NLME Tobit models and compared them with the symmetric NLME Tobit models in the presence of left censoring due to a lower limit detection and skewness. Among the models considered, Model III with ST distribution for error term was found to be the best. This model has two phases for describing the HIV dynamic process as given in (8). The first-phase decay rate, which is assumed to be time invariant, is estimated as $\hat{\lambda}_1 = 25.9$, while the second-phase decay rate, which is assumed to be time-varying, is estimated as $\hat{\lambda}_2 = -2.425 + 0.460t + 2.885\text{CD4}$ on population level. These first and second phase viral decay rates represent the minimum turnover rate of productively infected cells and that of latently or long-lived infected cells, respectively. For the second-phase decay rate $\hat{\lambda}_2$, the coefficient of CD4 is positive and significantly different from zero (see Table 2). This suggests that CD4 count is a clinically important predictor of the second-phase viral decay rate during the treatment process. More rapid increase in CD4 cell count may be associated with faster viral decay in the late stage. This may be explained by the fact that higher CD4 cell count suggest a higher turnover rate of lymphocyte cells, which may cause a positive correlation between viral decay and

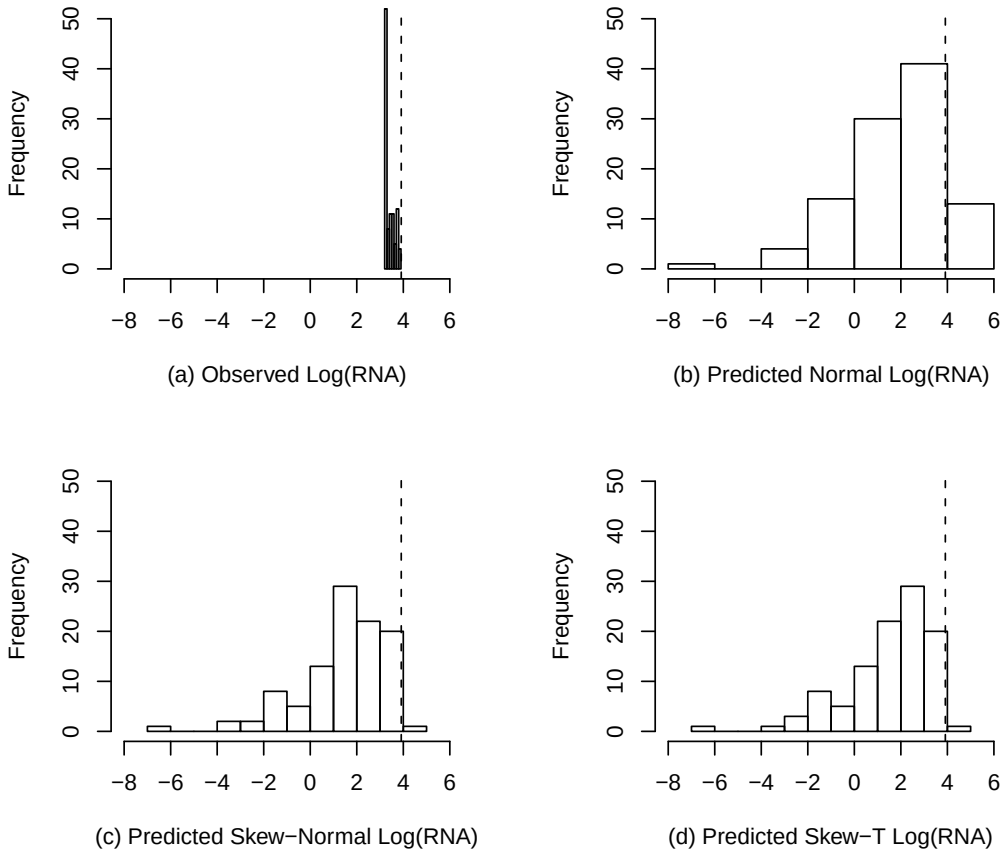

Figure 3: Histograms of (a) inaccurate raw data below LOD (dotted line), and predicted viral loads under (b) normal model, (c) SN model, and (d) ST model.

the CD4 cell count. We did not find the coefficient of time to be significant for the second-phase viral decay though it shows a tendency for rebound.

We also found that Model III provides good predictions for unobserved viral loads due to left-censoring (see Figure 3), which may be used for assessing treatment efficacy and minimum viral levels after treatment. In this prediction, we have assumed that the viral dynamic model continues to hold for left-censored viral loads. This assumption may be reasonable because the bi-exponential model considered here was derived based on reasonable biological arguments rather than on empirical arguments (Wu and Ding, 1999). It is to be noted that a Bayesian estimation procedure based on MCMC facilitates the prediction as it gives the entire posterior predictive distribution of the viral loads after accounting for uncertainties in censoring and CD4 counts.

There has been a limited research in Bayesian extensions of the Tobit model (Hamilton, 1999; Wei, 1999). For example, Hamilton (1999) implemented an extension of the Tobit model by allowing the error terms to follow a t-distribution, rather than a normal distribution. In a similar manner, our Model III allows the error term to follow a multivariate ST distribution. However, our model is novel in that it allows for non-symmetry (skewness). An important advantage of this modeling alternative is that our model can be easily fitted using freely available WinBUGS software and has a computational cost similar to the standard normal model. This makes our approach quite powerful and accessible to practitioners and applied statisticians.

In order to examine the sensitivity of parameter estimates to the prior distributions and initial values, we also conducted a sensitivity analysis using different values of hyperparameters of prior distributions and different initial values (data not shown). The results of the sensitivity analysis showed that the estimated dynamic parameters were not sensitive to changes of both priors and initial values. Thus, the final results are reasonable and robust, and the conclusions of our analysis remain unchanged.

There are certain limitations to our study, though. The current study is not intended to be an exhaustive study of the HIV dynamic models. We could have fitted more elaborate nonlinear dynamic models with a larger number of determinants of HIV viral loads. However, the purpose of this paper is to explore the use of flexible skew-elliptical distributions and Bayesian methods for extending the Tobit model to account for left-censoring and skewness, thus allowing more realistic models to be constructed. Therefore, we chose a small number of covariates, particularly CD4, that would be related to viral load, a priori. However, it would be straightforward to extend the proposed methods for incorporating several covariates including, as the editor pointed out, lagged values of CD4 and a further study may be warranted to assess their impact.

Another limitation of our proposed models is that they are based on skew-elliptical (skew-normal and skew-t) distributions for the response variable, and these distributions may not be robust enough with respect to normality for handling possible outliers or sub-populations. It is plausible to extend the skew-elliptical distribution to account atypical data by using mixtures of skew-elliptical distributions similar to that of Moulton et al. (2002), which used a mixture of Gaussian and Bernoulli distribution to account for an excess of observations below detection limit. Another area where both skew-normal and skew-t can be extended to make them robust is to use other more flexible distributions such as nonparametric distributions, which follow a Dirichlet process prior for model error and/or random effects. Recently, there appear few papers focusing on left-censored data (e.g., Vock et al., 2011; Sattar et al., 2011), which relaxed the normality assumption for random effects by using a semi-nonparametric distribution. These, however, are beyond the focus of this article, but a further study may be warranted and we are actively investigating these interesting issues now.

In conclusion, we have examined the use of flexible Bayesian methods for analyzing HIV viral load data, which allow one to incorporate left-censoring and skewness in the observed data. We believe that these proposed methods may have an important impact on HIV/AIDS research because, in the presence of left-censoring, skewness and heaviness in the tails of response variables, appropriate inference for HIV dynamics is important for making reliable conclusions and appropriate clinical decisions.

Appendix: Multivariate skew distributions

A new class of multivariate skew elliptical distributions, which includes the multivariate skew–normal(SN), skew-t (ST) and normal (N) distributions as special cases, was introduced in the literature (Sahu et al., 2003; Genton, 2004; Azzalini and Capitanio, 2003) to provide flexibility in capturing asymmetrical behavior of response variables. The structure of the skew-elliptical distribution was given in Section 2. The multivariate skew-normal(SN) and skew-t (ST) distributions are described in details below. Assume an n-dimensional random vector $\boldsymbol{Y}$ follows an n variate SN or ST distribution with location vector $\boldsymbol{\mu}$, $n \times n$ positive (diagonal) dispersion matrix $\boldsymbol{\Sigma}$ and $n \times n$ skewness matrix $\boldsymbol{\Delta} = diag(\delta_1, \delta_2, \ldots, \delta_n)$ with skewness parameter vector $\boldsymbol{\delta} = (\delta_1, \delta_2, \ldots, \delta_n)^T$ and the degrees of freedom ν.

A.1 *Skew-t distribution*

An n-dimensional random vector $\boldsymbol{Y}$ follows an n-variate ST distribution if its probability density function (pdf) is given by

$$f(\boldsymbol{y}|\boldsymbol{\mu},\boldsymbol{\Sigma},\boldsymbol{\Delta},\nu) = 2^n t_\nu(\boldsymbol{y}|\boldsymbol{\mu},\boldsymbol{Q})P(\boldsymbol{V}>\boldsymbol{0}), \tag{A.1}$$

where $\boldsymbol{Q} = \boldsymbol{\Sigma} + \boldsymbol{\Delta}^2$, we denote the n-variate t distribution with parameters $\boldsymbol{\mu}$, $\boldsymbol{Q}$ and degree of freedom ν by $t_\nu(\boldsymbol{\mu},\boldsymbol{Q})$ and the corresponding pdf by $t_\nu(\boldsymbol{y}|\boldsymbol{\mu},\boldsymbol{Q})$ henceforth, $\boldsymbol{V}$ follows the t distribution t_ν. We denote this distribution by $ST_\nu(\boldsymbol{\mu},\boldsymbol{\Sigma},\boldsymbol{\Delta})$. In particular, when $\boldsymbol{\Sigma} = \sigma^2\boldsymbol{I}_n$ and $\boldsymbol{\Delta} = \delta\boldsymbol{I}_n$, the equation (A.1) simplifies to

$$f(\boldsymbol{y}|\boldsymbol{\mu},\sigma^2,\delta,\nu) = 2^n(\sigma^2+\delta^2)^{-n/2}\frac{\Gamma((\nu+n)/2)}{\Gamma(\nu/2)(\nu\pi)^{n/2}}\left\{1+\frac{(\boldsymbol{y}-\boldsymbol{\mu})^T(\boldsymbol{y}-\boldsymbol{\mu})}{\nu(\sigma^2+\delta^2)}\right\}^{-(\nu+n)/2}$$
$$\times F_\nu\left[\left\{\frac{\nu+(\sigma^2+\delta^2)^{-1}(\boldsymbol{y}-\boldsymbol{\mu})^T(\boldsymbol{y}-\boldsymbol{\mu})}{\nu+n}\right\}^{-1/2}\frac{\delta(\boldsymbol{y}-\boldsymbol{\mu})}{\sigma\sqrt{\sigma^2+\delta^2}}\right],$$

where $F_\nu(\cdot)$ denotes the cumulative distribution function (cdf) of $t_\nu(\boldsymbol{0},\boldsymbol{I}_n)$.

The mean and covariance matrix of the ST distribution $ST_\nu(\boldsymbol{\mu},\sigma^2\boldsymbol{I}_n,\boldsymbol{\Delta})$ are given by $E(\boldsymbol{Y}) = \boldsymbol{\mu}+(\nu/\pi)^{1/2}\frac{\Gamma((\nu-1)/2)}{\Gamma(\nu/2)}\boldsymbol{\delta}$ and $cov(\boldsymbol{Y}) = \left[\sigma^2\boldsymbol{I}_n+\boldsymbol{\Delta}^2\right]\frac{\nu}{\nu-2} - \frac{\nu}{\pi}\left[\frac{\Gamma\{(\nu-1)/2\}}{\Gamma(\nu/2)}\right]^2\boldsymbol{\Delta}^2$.

It is noted that when $\boldsymbol{\delta} = \boldsymbol{0}$, the ST distribution reduces to usual t distribution. In order to better understand the shape of a ST distribution, plots of an univariate ST density as a function of the skewness parameter with $\delta = -3, 0, 3$ are shown in Figure 4(a).

A.2 *Skew-normal distribution*

An n-dimensional random vector $\boldsymbol{Y}$ follows an n-variate SN distribution, if its pdf is given by

$$f(\boldsymbol{y}|\boldsymbol{\mu},\boldsymbol{\Sigma},\boldsymbol{\Delta}) = 2^n|\boldsymbol{Q}|^{-1/2}\phi_n\{\boldsymbol{Q}^{-1/2}(\boldsymbol{y}-\boldsymbol{\mu})\}P(\boldsymbol{V}>\boldsymbol{0}), \tag{A.2}$$

where $\boldsymbol{V}\sim N\{\boldsymbol{\Delta}\boldsymbol{Q}^{-1}(\boldsymbol{y}-\boldsymbol{\mu}), \boldsymbol{I}_n-\boldsymbol{\Delta}\boldsymbol{Q}^{-1}\boldsymbol{\Delta}\}$, and $\phi_n(\cdot)$ is the pdf of $N(\boldsymbol{0},\boldsymbol{I}_n)$. we denote the above distribution by $SN(\boldsymbol{\mu},\boldsymbol{\Sigma},\boldsymbol{\Delta})$. An appealing feature of equation (A.2) is that it gives independent marginal when $\boldsymbol{\Sigma} = diag(\sigma_1^2,\sigma_2^2,\ldots,\sigma_n^2)$. The pdf (A.2) thus reduces to

$$f(\boldsymbol{y}|\boldsymbol{\mu},\boldsymbol{\Sigma},\boldsymbol{\Delta}) = \prod_{i=1}^n\left[\frac{2}{\sqrt{\sigma_i^2+\delta_i^2}}\phi\left\{\frac{y_i-\mu_i}{\sqrt{\sigma_i^2+\delta_i^2}}\right\}\Phi\left\{\frac{\delta_i}{\sigma_i}\frac{y_i-\mu_i}{\sqrt{\sigma_i^2+\delta_i^2}}\right\}\right],$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of the standard normal distribution, respectively.

The mean and covariance matrix are given by $E(\boldsymbol{Y}) = \boldsymbol{\mu}+\sqrt{2/\pi}\boldsymbol{\delta}$ and $cov(\boldsymbol{Y}) = \boldsymbol{\Sigma}+(1-2/\pi)\boldsymbol{\Delta}^2$. It is noted that when $\boldsymbol{\delta} = \boldsymbol{0}$, the SN distribution reduces to usual normal distribution. In addition, the SN distribution is a special case of the ST distribution. That is, the ST distribution reduces to the SN distribution when the degree of freedom $\nu\to\infty$. In order to better

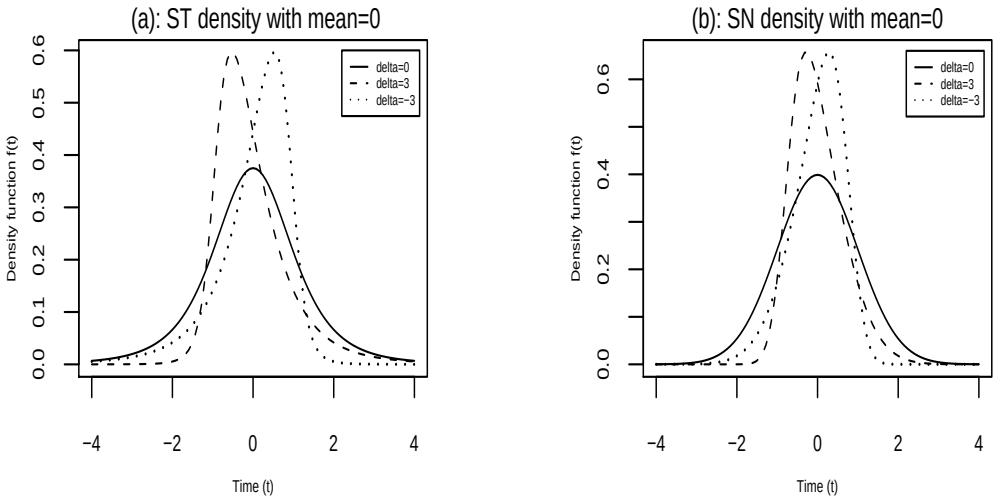

Figure 4: The univariate skew-t (df $\nu = 4$) and skew-normal density functions with precision $\sigma^2 = 1$ and skewness parameter $\delta = 0, -3$ and 3, respectively.

understand the shape of an SN distribution, plots of an univariate SN density as a function of the skewness parameter with $\delta = -3, 0$, and 3 are shown in Figure 4(b).

References

Acosta, E.P., Wu, H., Walawander, A., Eron, J., Pettinelli, C., Yu, S., Neath, D., Ferguson, E., Saah, A.J., Kuritzkes, D.R., Gerber, J.G., for the Adult ACTG 5055 Protocol Team. (2004). Comparison of two indinavir/ritonavir regimens in treatment-experienced HIV-infected individuals. *Journal of Acquired Immune Deficiency Syndromes*, **37**, 1358–1366.

Arabmazar, A. and Schmidt, P. (1982). An Investigation of the robustness of the Tobit estimator to non-Normality. *Econometrica*, **50**, 1055–1064.

Arellano-Valle, R.B., and Genton, M.G. (2005). On fundamental skew distributions. *Journal of Multivariate Analysis*, **96**, 93–116.

Azzalini, A., Dalla-Valle, A. (1996). The multivariate skew-normal distribution. *Biometrika*, **83**, 715–726.

Azzalini, A. and Capitanio, A. (2003). Distributions generated by perturbation of symmetry with emphasis on a multivariate skew-t distribution. *Journal of the Royal Statistical Society, B*, **65**, 367–389.

Coombs, R.W., Speck, C.E., Hughes, J.P., Lee, W., Sampoleo, R., Ross, S.O., Dragavon, J., Peterson, G., Hooton, T.M., Collier, A.C., Corey, L., Koutsky, L., Krieger, J.N. (1998). Association between culturable human immunodeficiency virus type-1 (HIV-1) in semen and HIV-1 RNA levels in semen and blood: Evidence for compartmentalization of HIV-1 between semen and blood. *Journal of Infectious Disease*, **177**, 320–330.

Ding, A.A., and Wu, H. (1999). Relationships between antiviral treatment effects and biphasic viral decay rates in modeling HIV dynamics. *Mathematical Biosciences*, **160**, 63–82.

Genton, M. G. (2004). *Skew-Elliptical Distributions and their Applications: A Journey Beyond Normality*. Chapman & Hall / CRC: Boca Raton, FL.

Geyer, C.J. (1992). Practical Markov Chain Monte Carlo (with discussion). *Statistical Science*, **7**, 473–511.

Hamilton, B. (1999). The impact of HMOs on medicare costs: Bayesian MCMC estimation of a robust panel data Tobit model with survival. *Health Economics*, **8**, 403–414.

Ho, D.D., Neumann, A.U., Perelson, A.S., Chen, W., Leonard, J.M., Markowitz, M. (1995). Rapid turnover of plasma virions and CD4 lymphocytes in HIV-1 infection. *Nature*, **373**, 123–126.

Hornung, R.W., and Reed, L.D. (1990). Estimation of average concentration in the presence of nondetectable values. *Applied Occupational and Environmental Hygiene*, **4**, 46–51.

Huang, Y. and Dagne, G. (2010). Skew-normal Bayesian nonlinear mixed-effects models with application to AIDS studies. *Statistics in Medicine*, **29**, 2384–2398.

Huang, Y., Liu, D., and Wu, H. (2006). Hierarchical Bayesian methods for estimation of parameters in a longitudinal HIV dynamic system. *Biometrics*, **62**, 413–423.

Hughes, J.P. (1999). Mixed effects models with censored data with application to HIV RNA levels. *Biometrics*, **55**, 625–629.

Jacqmin-Gadda, H., Thiébaut, R., Chêne, G., and Commenges, D. (2000). Analysis of left-censored longitudinal data with application to viral load in HIV infection. *Biostatistics*, **1**, 355–368.

Liu, L. and Dey, D. (2004). Skew-elliptical distributions. In *Skew-Elliptical Distributions and their Applications: A Journey Beyond Normality*, Genton MG (ed.). Chapman & Hall/CRC: Boca Raton, FL, 43–64.

Liu, W., and Wu, L. (2007). Simultaneous inference for semiparametric nonlinear mixed-effects models with covariate measurement errors and missing responses. *Biometrics*, **63**, 342–350.

Lorimer, M.F. and Kiermeier, A. (2007). Analysing microbiological data: Tobit or not Tobit? *International Journal of Food Microbilogy*, **116**, 313–318.

Lunn, D.J., Thomas, A., Best, N. and Spiegelhalter, D. (2000). WinBUGS – a Bayesian modelling framework: concepts, structure, an extensibility. *Statistics and Computing*, **10**, 325–337.

Lynn,H. S. (2001). Maximum likelihood inference for left-censored HIV RNA data. *Statistics in Medicine*, **20**, 33–45.

Moulton, Curriero, and Barroso (2002). Mixture models for quantitative HIV RNA data. *Statistical Methods in Medical Research*, **11**, 317–325.

Paxton, W. B., Coombs, R. W., McElrath, M. J., Keefer, M.C., Hughes, J., Corey, L. (1997). Longitudinal anlysis of quantitative virologic measures in human immunodeficiency virus-infected subjects with $\geq$ 400 CD4 lymphocytes: Implication for applying measurements to individual patients. *Journal of Infectious Disease*, **175**, 247–254.

Perelson, A.S., Essunger, P., Cao, Y., Vesanen, M., Hurley, A., Saksela, K., Markowitz, M., Ho, D.D. (1997). Decay characteristics of HIV-1-infected compartments during combination therapy. *Nature*, **387**, 188–191.

Raftery, A.E., and Lewis, S. (1992). Comment: One Long Run with Diagnostics: Implementation Strategies for Markov Chain Monte Carlo. *Statistical Science*, **7**,493–497.

Reilly, M.P., Wolfe, M.L., Localio, A.R., Rader, D.J. (2004). Coronary artery calcification and cardiovascular risk factors: impact of the analytic approach. *Atherosclerosis*, **173**, 69–78.

Sahu, S. K., Dey, D.K., and Branco, M.D. (2003). A new class of multivariate skew distributions with applications to Bayesian regression models. *The Canadian Journal of Statistics*, **31**, 129–150.

Schockmel, G.A., Yerly, S., and Perrin, L. (1997). Detection of Low HIV-1 RNA Levels in Plasma. *Journal of Acquired Immune Deficiency Syndromes & Human Retrovirology*, **14**,179–183.

Scirica, C.V., Gold, D.R., Ryan, L., Abulkerim, H., Celedón, J.C., Platts-Mills, T.A., Naccara, L.M., Weiss, S.T., Litonjua, A.A. (2007). Predictors of cord blood IgE levels in children at risk for asthma and atopy. *Journal of Allergy and Clinical Immunology*, **111**, 81–88.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., and Van der Linde, A. (2002). Bayesian measures of model complexity and fit (with Discussion). *Journal of the Royal Statistical Society, Series B*, **64**, 583–639.

Sattar, A., Weissfeld, L. A. and Molenberghs, G. (2011). Analysis of non-ignorable missing and leftcensored longitudinal data using a weighted random effects tobit model. *Statistics in Medicine*, **30**, 3167-3180.

Thiébaut, R., Guedj, J., Jacqmin-Gadda, H., Chêne, G., Trimoulet, P., Neau,

D. and Commenges, D. (2006). Estimation of dynamical model parameters taking into account undetectable marker values. *BMC Medical Research Methodology*, **6**, 38.

Vock, D. M., Davidian, M., Tsiatis, A. A. and Muir, A. J. (2012). Mixed model analysis of censored longitudinal data with flexible random-effects density. *Biostatistics*, **13**, 61–73.

Wei, S. X. (1999). A Bayesian approach to dynamic Tobit models. *Econometric review*, **18**, 417–439.

Wu, H., and Ding, A.A, de Gruttola, V. (1998). Estimation of HIV dynamic parameters. *Statistics in Medicine*, **17**, 2463–2485.

Wu, H., Ding, A.A. (1999). Population HIV-1 dynamics in vivo: applicable models and inferential tools for virological data from AIDS clinical trials. *Biometrics*, **55**, 410–418.

Wu, L. (2002). A joint model for nonlinear mixed-effects models with censoring and covariates measured with error, with application to AIDS studies. *Journal of the American Statistical association*, **97**, 955–964.