University of Connecticut

Bolfarine's Contributions in Asymmetric Models

Victor H. Lachos
University of Connecticut, USA

June 2021

ISI Mahalanobis International Award Session
63rd ISI World Statistics Congress 2021 (ISI WSC 2021)

Introduction

- Prof. H. Bolfarine is one of my most significant collaborators, so far we have 21 joint papers that have appeared in refereed journals.
- Some of these publications are product of co-advising PhD students.
- Joint Collaboration started in 2002, as he was my PhD advisor.
- I have divided his main contribution related to asymmetric models in 4 areas, to the best of my knowledge.
- Linear mixed models (topic of my PhD thesis), measurement error models, finite mixtures and binary regression (skew-probit).

The skew-normal LMM

In general, a linear mixed effects model is defined as

$$\mathsf{Y}_i = \mathsf{X}_i\boldsymbol{\beta} + \mathsf{Z}_i\mathsf{b}_i + \boldsymbol{\epsilon}_i, \quad i = 1, \ldots, n\,, \tag{1}$$

where

- Usual assumptions: $\mathsf{b}_i \overset{\text{iid}}{\sim} N_q(\mathbf{0}, \mathsf{D}) \perp \boldsymbol{\epsilon}_i \overset{\text{ind}}{\sim} N_{n_i}(\mathbf{0}, \boldsymbol{\Sigma}_i)$.
- There is an increasing interest to consider more flexible distributions.
- The $\mathrm{SN}_p(\boldsymbol{\mu}, \boldsymbol{\Sigma}, \boldsymbol{\lambda})$ distribution (Azzalini and Valle, 1996) is defined as:

$$f(\mathsf{y}) = 2\phi_p(\mathsf{y}; \boldsymbol{\mu}, \boldsymbol{\Sigma})\Phi(\boldsymbol{\lambda}^\top \boldsymbol{\Sigma}^{-1/2}(\mathsf{y} - \boldsymbol{\mu})), \quad \mathsf{y} \in \mathbb{R}^p,$$

The skew-normal linear mixed model (SN-LMM)

Arellano-Valle, Bolfarine and Lachos (2005), define the SN-LMM as:

$$\begin{pmatrix} \mathbf{b}_i \\ \boldsymbol{\epsilon}_i \end{pmatrix} \overset{\text{ind}}{\sim} \text{SN}_{q+n_i}\left(\begin{pmatrix} 0 \\ \mathbf{0} \end{pmatrix}, \begin{pmatrix} \mathbf{D} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_i \end{pmatrix}, \begin{pmatrix} \boldsymbol{\lambda} \\ \mathbf{0} \end{pmatrix}\right), i = 1, \ldots, n.$$

- An EM-type algorithm is proposed for ML estimation.
- Interesting properties are developed. For instance, the marginal distribution of the response is still skew-normal family.
- So far, this paper has around 230 citations.
- In a subsequent paper, Arellano-Valle, Bolfarine and Lachos (2007), developed a Bayesian approach for the SN-LMM.

Some publications related to the SN-LMM

- Lachos, Bolfarine, Arellano-Valle and Montenegro (2007). Likelihood-based inference for multivariate skew-normal regression models (Communication in Statistics TM)

- Ferreira, Bolfarine and Lachos (2021). Linear mixed models based on skew scale mixtures of normal distribution (Communications in Statistics – SC).

- Schumacher, Lachos and Matos (2021). skewlmm: Scale Mixture of Skew-Normal Linear Mixed Models (R package)

Skew-normal measurement error models

Arellano-Valle, Ozan, Bolfarine and Lachos (2005) [Journal of Multivariate Analysys] proposed the SN-MEM which is defined by

$$\begin{aligned} Y_i &= \alpha + \beta x_i + e_i, \\ X_i &= x_i + \delta_i, \end{aligned}$$

where

$$\begin{pmatrix} e_i \\ \delta_i \\ x_i \end{pmatrix} \overset{\text{ind}}{\sim} \text{SN}_3 \left(\begin{pmatrix} 0 \\ 0 \\ \mu_x \end{pmatrix}, \begin{pmatrix} \sigma_e^2 & 0 & 0 \\ 0 & \sigma_\delta^2 & 0 \\ 0 & 0 & \sigma_x^2 \end{pmatrix}, \begin{pmatrix} \lambda_e \\ \lambda_\delta \\ \lambda_x \end{pmatrix} \right),$$

- An EM-type algorithm is proposed for ML estimation.
- Bayesian inference is also discussed for the family of SN-MEM.
- Interesting properties are developed. For instance, the marginal distribution of (X,Y) belongs to a skew-normal family.
- So far, this paper has around 63 citations.

Some publications related to the SN-MEM

- Lachos and Bolfarine (2006). Skew binary regression with measurement errors (Statistics).

- Lachos, Labra, Bolfarine and Ghosh (2010). Multivariate measurement error models based on scale mixtures of the skew–normal distribution (Statistics)

- Arellano-Valle, Azzalini, Ferreira and Santoro (2020). A two-piece normal measurement error model (Computational Statistics and Data Analysis)

Finite mixtures of skew distributions

A *finite mixture of SN distributions (FM-SN)* is defined by its pdf as

$$g(\mathbf{y}|\boldsymbol{\Theta}) = \sum_{j=1}^{G} p_j \mathrm{SN}_q(\mathbf{y}|\boldsymbol{\mu}_j, \boldsymbol{\Sigma}_j, \boldsymbol{\Delta}_j),$$

where $p_j \geq 0$ are such that $\sum_j^G p_j = 1$.

- Cabral, Bolfarine and Pereira (2008). Bayesian density estimation using skew student-t-normal mixtures (CSDA)
- I co-advised a PhD Student Luis Benites Sanchez with the thesis entitled "Finite Mixtures of Regression Models" (2014-2018)

Binary regression with skew-link

This idea consider the CDF of the skew-normal distribution instead the probit link in binary regression. Some important publications in this area are:

- Lachos and Bolfarine (2006). Skew binary regression with measurement errors (Statistics).
- Bazan, Bolfarine and Branco (2006). A skew item response model (Bayesian Analysis)

The later introduce a new skew-probit link for item response theory (IRT) by considering the CDF of the skew-normal distribution. This paper has 120 citation according to Google Scholar.

Thank you Prof. Heleno Bolfarine for the enormous contribution to my academic career and inspiring young researchers