UNIVERSIDADE ESTADUAL DE CAMPINAS

Instituto de Matemática, Estatística e Computação Científica

MAYARA CAROLINE MAIOLI

Univariate and bivariate regression models based on centered skew scale mixture of normal distributions

Modelos de regressão univariados e bivariados baseados nas distribuições de mistura de escala normal assimétrica sob a parametrização centrada

Campinas

2018

Mayara Caroline Maioli

Univariate and bivariate regression models based on centered skew scale mixture of normal distributions

Modelos de regressão univariados e bivariados baseados nas distribuições de mistura de escala normal assimétrica sob a parametrização centrada

Dissertação apresentada ao Instituto de Matemática, Estatística e Computação Científica da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestra em Estatística.

Dissertation presented to the Institute of Mathematics, Statistics and Scientific Computing of the University of Campinas in partial fulfillment of the requirements for the degree of Master in Statistics.

Supervisor: Caio Lucidius Naberezny Azevedo

Este exemplar corresponde à versão final da Dissertação defendida pela aluna Mayara Caroline Maioli e orientada pelo Prof. Dr. Caio Lucidius Naberezny Azevedo.

Campinas

2018

Agência(s) de fomento e nº(s) de processo(s): FAPESP, 2015/25867-2; CNPq, 131979/2016-7

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Matemática, Estatística e Computação Científica
Ana Regina Machado - CRB 8/5467

M285u

Maioli, Mayara Caroline, 1993-
Univariate and bivariate regression models based on centered skew scale mixture of normal distributions / Mayara Caroline Maioli. – Campinas, SP : [s.n.], 2018.

Orientador: Caio Lucidius Naberezny Azevedo.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Matemática, Estatística e Computação Científica.

1. Inferência bayesiana. 2. Distribuição normal assimétrica. 3. Modelos de regressão (Estatística). 4. Modelos lineares (Estatística). I. Azevedo, Caio Lucidius Naberezny, 1979-. II. Universidade Estadual de Campinas. Instituto de Matemática, Estatística e Computação Científica. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: Modelos de regressão univariados e bivariados baseados nas distribuições de mistura de escala normal assimétrica sob a parametrização centrada
Palavras-chave em inglês:
Bayesian inference
Skew-normal distributions
Regression models (Statistics)
Linear models (Statistics)
Área de concentração: Estatística
Titulação: Mestra em Estatística
Banca examinadora:
Caio Lucidius Naberezny Azevedo [Orientador]
Filidor Edilfonso Vilca Labra
Carlos Alberto Ribeiro Diniz
Data de defesa: 08-03-2018
Programa de Pós-Graduação: Estatística

Dissertação de Mestrado defendida em 08 de março de 2018 e aprovada pela banca examinadora composta pelos Profs. Drs.

Prof(a). Dr(a). CAIO LUCIDIUS NABEREZNY AZEVEDO

Prof(a). Dr(a). FILIDOR EDILFONSO VILCA LABRA

Prof(a). Dr(a). CARLOS ALBERTO RIBEIRO DINIZ

As respectivas assinaturas dos membros encontram-se na Ata de defesa

This dissertation is dedicated to God and my family.

Acknowledgements

First of all, I would like to thank God, for providing me this opportunity and by giving me strength and courage to do this work. Without Him, I would not be able to overcome the challenges faced during the completion of this work.

I would like to thank to FAPESP-Brazil (Grant 2015/25867-2) for the financial support between May 2016 and February 2018 and CNPq-Brazil (Grant 131979/2016-7) for the financial support between March 2016 and April 2016.

I would like to express my deepest gratitude to my advisor, Professor Caio Lucidius Naberezny Azevedo, for his friendship, patience, crucial advices and suggestions.

I would also like to express my deepest gratitude to my parents, Katia and João, my future husband Daniel and my brother Danilo, who have always been by my side, giving me all the emotional support to finish this project, helping me to face the difficult moments and sharing with me the joys. Thank you for believing in me and giving me the strength to move on.

I would also like to thank my colleagues in the master's degree, for the moments and shared knowledge.

And last but not least, I would like to thank all my graduating friends: Danielle, Fabiana, Fernanda M., Fernanda E., Livia, Patricia, Tamilyn, Thais and Thalita, who have always supported me and encouraged me to follow my dreams.

Resumo

Situações em que a variável resposta é contínua ou binária são bastante comuns em diversas áreas do conhecimento. Apesar de existirem diversos modelos para essas situações, em muitos casos, características como assimetria e caudas pesadas, não são contempladas adequadamente. Além disso, conjuntos de respostas bivariadas, contendo uma variável contínua e uma discreta, são comuns em muitos problemas reais, as quais também podem apresentar assimetria e caudas pesadas. A abordagem mais comum, no caso bivariado, é modelar cada variável separadamente, ignorando a potencial correlação entre elas, ou decompor a distribuição conjunta na distribuição marginal para a variável binária e na distribuição condicional para a variável contínua, dada a variável binária. A decomposição na distribuição marginal da variável contínua e na distribuição condicional da variável binária, dada a variável contínua, também é possível. Neste projeto desenvolvemos: uma classe de modelos de regressão linear baseada nas distribuições de mistura de escala normal assimétrica sob a parametrização centrada (MENAC), uma classe de modelos de regressão para dados binários com função de ligação associada a alguma distribuição MENAC, e uma classe de modelos de regressão misto para dados bivariados contínuo e binário, em que tanto a resposta contínua, quanto a função de ligação para a resposta binária pertencem a classe MENAC. Para introduzir a estrutura de dependência entre as duas variáveis resposta, consideramos uma estrutura de efeitos aleatórios comuns, cujas distribuições também pertencem a classe MENAC. Desenvolvemos procedimentos de estimação sob o paradigma bayesiano, assim como ferramentas de diagnóstico, contemplando análise residual e medidas de influência, bem como medidas de comparação de modelos. Realizamos estudos de simulação, considerando diferentes cenários de interesse, com o intuito de avaliar o desempenho das estimativas e das medidas de diagnóstico. As metodologias propostas foram ilustradas tanto com dados provenientes de estudos de simulação, quanto com conjuntos de dados reais.

Palavras-chave: Inferência bayesiana; distribuições de mistura de escala normal assimétrica; dados bivariados contínuos e binários; modelo de regressão linear; modelo de regressão binário.

Abstract

Situations where the response variable is either continuous or binary are quite common in several fields of knowledge. Although there are several models for these situations, in many cases, characteristics such as asymmetry and heavy tails, are not properly treated. In addition, bivariate responses, containing one continuous and one discrete variable, are common in many real problems, which may also exhibit asymmetry and heavy tails. The most common approach in the bivariate case is to model each variable separately, ignoring the potential correlation between them, or to decompose the joint distribution into the marginal distribution of the binary variable and the conditional distribution of the continuous variable, given the binary variable. The decomposition into the marginal distribution of the continuous variable and the conditional distribution of the binary variable, given the continuous variable, it is also possible. In this project we developed: a class of linear regression models based on the skew scale mixture of normal distributions under the centered parameterization (SSMNC), a class of regression models for binary data with link function associated with some SSMNC distribution, and a class of mixed regression models for bivariate continuous and binary data, in which both the continuous response and the link function for the binary response, belong to the SSMNC class. To introduce the dependency structure between the two response variables, we consider a common random effects structure, whose distributions also belong to the SSMNC class. We developed estimation procedures under the Bayesian paradigm, also, diagnostic tools, including residual analysis and influence measures, as well as model comparison measures. We performed simulation studies, considering different scenarios of interest, in order to evaluate the performance of estimates and diagnostic measures. The proposed methodologies were illustrated with both data from simulation studies and with real data sets.

Keywords: Bayesian inference; skew scale mixture of normal distributions; bivariate continuous and binary data; linear regression model; binary regression model

List of Figures

List of Tables

List of abbreviations and acronyms

SMSN	Scale Mixture of Skew Normal.
SMN	Scale Mixture of Normal.
SSMNC	Scale Mixture of Skew Normal distributions under the centered parameterization.
SN	Skew Normal.
ST	Skew-t.
SSL	Skew slash.
STgen	Skew generalized t
MCMC	Markov chain Monte Carlo
cdf	cumulative distribution function.
RMSE	Square root of the mean square error.

Contents

Introduction

Situations where the response variable is either continuous or binary are quite common in several fields of knowledge, specially in social and health science. In general, in these studies, multiple responses are collected in order to characterize or evaluate their relationships with some covariates of interest. For example, to evaluate the efficacy of an experimental treatment on vision for macular degeneration, a study was performed in patients with age-related macular degeneration (see (GUYER et al., 1997)). For each patient it was evaluated their patient's visual acuity in the beginning and after one year of study. This acuity is measured by counting how many letters of a standardized vision chart are corrected read. These charts display line letters of decreasing size that the patient must read from the top (large letters) to bottom (small letters). In this study, two outcomes were obtained in order to evaluate the efficacy of the treatment: the binary outcome was defined as the loss of at least three lines of vision at one year compared with their baseline performance and the continuous outcome are defined as the difference between patient's visual acuity from one year and the beginning of the study.

The usual modeling strategy for this type of data is to perform separate analysis for each response variable. As notes in (TEIXEIRA-PINTO; NORMAND, 2009) this strategy is less efficient, since it ignores the extra information contained in the correlation among the outcomes. In the bivariate context the analysis can be made using the factorization method, discussed in (COX; WERMUTH, 1992), which consists to write the likelihood function as the product of the marginal distribution of one of the outcomes and the conditional distribution of second outcome given the first one. (FITZMAURICE; LAIRD, 1995) and (CATALANO; RYAN, 1992) extend this approach to situations with clustered data, in which the method proposed in (FITZMAURICE; LAIRD, 1995) is based on the general location model of (OLKIN; TATE, 1961).

All models cited earlier are based on the normality assumption for the continuous response and assume symmetrical link functions, with a normal type kurtosis, to the binary outcome. However, some data sets may not satisfy these assumptions. As an alternative to the normal distribution, (AZZALINI, 1985) introduced the skew-normal distribution. It has attracted a lot of attention in the literature, due to its mathematical tractability and for having the normal distribution as a special case, as noted in (GUPTA; NGUYEN; SANQUI, 2004). In these distributions the asymmetry is modeled by a single parameter. For this reason, it is an alternative to normal distribution when the data presents non-symmetric behavior, but the kurtosis is not so different from the normal one. However, as noted by (ARELLANO-VALLE; AZZALINI, 2008), the skew normal distribution under the direct parameterization, defined by the location parameter $\alpha \in \mathbb{R}$, scale parameter $\beta \in \mathbb{R}^+$

and skewness parameter $\lambda \in \mathbb{R}$, has some problems in terms of parameter estimation, at least near $\lambda = 0$, since the log-likelihood presents a non-quadratic shape. Even under the Bayesian paradigm, this fact can lead to some problems. (PEWSEY, 2000) addressed various issues related to direct parameterization and explained why it should not be used for estimation procedures. (AZZALINI, 1985) noticed that when $\lambda = 0$ the Fisher Information is singular and (PEWSEY, 2000) linked this singularity to the parameter redundancy of the parameterization for the normal case. The problems cited earlier are solved by using the centered parameterization of the skew-normal distribution, defined with the parameters: $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}^+$ and $\gamma \in (-0.99527, .99527)$, that are, respectively, the mean, variance and Person's skewness coefficient of the skew normal distribution.

On the other hand, a very useful class of models, that can handle asymmetry and heavy tails, is the scale mixtures of skew-normal distribution (SMSN) (see (FERREIRA; BOLFARINE; LACHOS, 2011)), that is a extension of the scale mixture of normal distributions, see (ANDREWS; MALLOWS, 1974). This class developed in (BRANCO; DEY, 2001), includes the normal and skew normal distribution as special cases as well as several asymmetric distributions, as the skew-t, skew slash, skew generalized t and skew contaminated normal and their corresponding symmetric cases.

For binary data, the use of probit and logistic models are not adequate when we have evidence that the probability of success increases at a different rate than decreases. (CZADO; SANTNER, 1992) showed, through a simulation study, using a data generated by a skewed link function, that the link misspecification can yield a substantial bias in the estimates of the regression coefficients. Such problem is circumvented by the use of asymmetric link functions, that can be obtained, for example, through the cumulative distribution function of an asymmetric distribution.

Many skewed binary regressions have been proposed in the literature. (STUKEL, 1988), (CZADO; SANTNER, 1992), and (GUERRERO; JOHNSON, 1982) introduced asymmetry replacing the linear predictor by a nonlinear function of the linear predictor and a parameter that controls the asymmetry. Another approach is to replace the linear predictor by a polynomial function, see for example (COLLETT, 2002). Finally, the third option is to consider the cumulative distribution function of an asymmetric distribution. The most popular example of this method is the complementary log-log link function, that is constructed from the cdf of the Gumbel distribution. (CHEN; DEY; SHAO, 1999) proposed an asymmetric probit link, considering a class of mixture of normal distributions. (BAZáN; BOLFARINE; BRANCO, 2010) presented a unified approach for two skew probit links. In (BAZáN; ROMEO; RODRIGUES, 2014), it was introduced two new asymmetric links, one based on the cdf of the power-normal distribution and the other based on the cdf (cumulative distribution function) of the reciprocal power-normal distribution. (NAGLER, 1994) introduced the asymmetrical link by using the Burr-10 distribution ((BURR, 1942)).

In addition, since probit and logistic regression estimates are not robust in the presence of outliers, (LIU, 2005) proposed a new binary model, named robit regression, in which the normal distribution in probit regression is replaced by a Student-t distribution with known or unknown degrees of freedom. Both logistic and probit models can be approximated by the robit regression, as showed in (LIU, 2005). Instead using the Student-t distribution, (KIM; CHEN; DEY, 2008) introduced a class of skewed generalized t-link models, that accommodate heavy tail and asymmetric link functions.

Motivated by all these issues, we developed a univariate regression model for the continuous response and a binary regression model using the skew scale mixture of normal distribution under the centered parameterization. Also we proposed a bivariate regression model that accommodates skewness and heavy tails for the continuous response as well for the binary data in which the probability of success increases at a different rate than decreases. The organization of the dissertation is as follows:

Chapter 2: We provide some background about the skew normal distribution and the skew scale mixture of normal distributions. We discuss the problems arising from the direct parameterization of the skew-normal distribution introduced in (AZZALINI, 1985). Then we introduce the skew scale mixture of normal distributions under the centered parameterization and the linear regression model based on this distribution class was developed, that accommodate asymmetry and heavy tails error distributions. Bayesian inference was made, and we presented residual and influence analysis and model comparisons techniques. The model was applied in simulated and real datasets. For the simulation study we discussed about the quality of the estimates for $\boldsymbol{\nu}$ and γ and discussed about prior choice for these parameters.

Chapter 3: We developed a binary regression model for with heavy tail and asymmetrical link functions that accommodate binary data in which the probability of success increases at a different rate than decreases and are robust in the presence of outliers. We developed a Bayesian estimation procedure and we described a simple way for checking goodness of fit. We performed simulation studies in order the evaluate the parameter recovery, residual and influence analysis and the effect of incorrect specification of the link function. We discussed about the quality of the estimates for $\boldsymbol{\nu}$ and γ and discussed about prior choice for these parameters. Finally, we applied the developed model to real data to illustrate the use of the model.

Chapter 4: We developed a regression model for bivariate continuous and binary responses assuming the possibility of heavy tails and asymmetry. We developed a Bayesian estimation procedure and appropriate residuals based on latent variables. The model was applied in simulated and real datasets.

Chapter 5: We present final remarks and further researches related to this dissertation.

1 Linear regression model based on skew scale mixture of normal distributions based on the centered parameterization

1.1 Introduction

The normal distribution has been used for many years on diverse fields of knowledge. Despite its simplicity and popularity, it is well known that several phenomena cannot be appropriately modeled by this distribution, due the presence at least one of the following characteristics: asymmetry, heavy tails and multi-modality. For example, as noted by (ARELLANO-VALLE; GENTON; LOSCHI, 2009), the empirical distribution of data sets often exhibits skewness and tails that are lighter or heavier than the normal distribution.

The skew-normal distribution, proposed by (AZZALINI, 1985), has attracted a lot of attention in the literature, due to its mathematical tractability and for having the normal distribution as a special case, as noted in (GUPTA; NGUYEN; SANQUI, 2004). For more details about the skew-normal distribution see (AZZALINI; CAPITANIO, 2013), (GENTON, 2004). In this distributions the asymmetry is modeled by a single parameter. For this reason, it is an alternative to normal distribution when the data set presents non-symmetric behavior.

On the other hand, a very useful class of models that can handle asymmetry and heavy tails is the Scale Mixtures of Skew-Normal distribution (SMSN) (see (FERREIRA; BOLFARINE; LACHOS, 2011)), that is a extension of the scale mixture of normal distributions, see (ANDREWS; MALLOWS, 1974). This class, as described in (FERREIRA; BOLFARINE; LACHOS, 2011) includes the normal and skew normal distribution as special cases as well as several asymmetric distributions, as the skew-t, skew-slash, skew generalized t and skew contaminated normal and their corresponding symmetric cases.

In this work we introduce the scale mixture of skew-normal distributions under the centered parameterization as an alternative to the parameterization given in (FERREIRA; BOLFARINE; LACHOS, 2011), which circumvents some problems related to the use of the skew normal distribution under direct parameterization, including its respective extensions, as the SMSN family.

1.2 Skew Scale Mixtures of Normal Distribution

1.2.1 Skew-Normal distribution

We say that Y has skew-normal distribution with location parameter $\alpha \in \mathbb{R}$, scale parameter $\beta \in \mathbb{R}^+$ and skewness parameter $\lambda \in \mathbb{R}$, denoted by $Y \sim SN(\alpha, \beta^2, \lambda)$ if its probability density function (p.d.f) is

$$f(y|\alpha,\beta,\lambda) = 2\beta^{-1}\phi\left(\frac{y-\alpha}{\beta}\right)\Phi\left(\lambda\left(\frac{y-\alpha}{\beta}\right)\right)I_{(-\infty,\infty)}(y) \tag{1.1}$$

where $\phi(.)$ and $\Phi(.)$ denote the standard normal density and distribution function, respectively. It is straightforward to see that when $\lambda = 0$, the skew-normal reduces to normal distribution. Using the results from (AZZALINI, 1985), (PEWSEY, 2000) and (AZZALINI, 2005), the mean, variance, Pearson's index of skewness (γ_1) and kurtosis (γ_2) are given by, respectively

$$\begin{aligned} &E(Y) = \alpha + \beta b\delta \quad Var(Y) = \beta^2(1-b^2\delta^2) \\ &\gamma_1 = \frac{E((Y-E(Y))^3)}{Var(Y)^{3/2}} = \frac{4-\pi}{2}\frac{(b\delta)^3}{(1-b^2\delta^2)^{3/2}} \quad \gamma_2 = 2(\pi-3)\frac{(b\delta)^4}{(1-b^2\delta^2)^2} \end{aligned} \tag{1.2}$$

where $b = \sqrt{\dfrac{2}{\pi}}$ and $\delta = \dfrac{\lambda}{\sqrt{1+\lambda^2}}$.

As noted by (ARELLANO-VALLE; AZZALINI, 2008), the direct parameterization of the skew-normal distribution has some problems in terms of parameter estimation, at least near $\lambda = 0$, since the log-likelihood presents a non-quadratic shape. Even under the Bayesian paradigm, this fact can lead to some problems. (PEWSEY, 2000) addressed various issues related to direct parameterization and explained why it should not be used for estimation procedures. (AZZALINI, 1985) noticed that when $\lambda = 0$ the Fisher Information is singular and (PEWSEY, 2000) linked this singularity to the parameter redundancy of the parameterization for the normal case. More details of these discussions are found in (GENTON, 2004).

In order to show the behavior of the likelihood under the centered and direct parameterization, we calculated the profiled log-likelihood as described in (SANTOS, 2012) and (AZZALINI; CAPITANIO, 2013). First, 200 samples were generated by the skew normal distribution with $\alpha = 0$, $\beta = 1$ and $\lambda = 4$, and twice the profiled relative log-likelihood was calculated for both centered and direct parameterization. In Figure 1 it is possible to see that under the direct parameterization the log-likelihood exhibits a non-quadratic shape. However, under the centered parameterization, the log-likelihood presents a concave shape.

As illustrated in Figure 1, the problems cited earlier are solved by using the centered parameterization of the skew-normal distribution. Indeed, (AZZALINI, 1985)

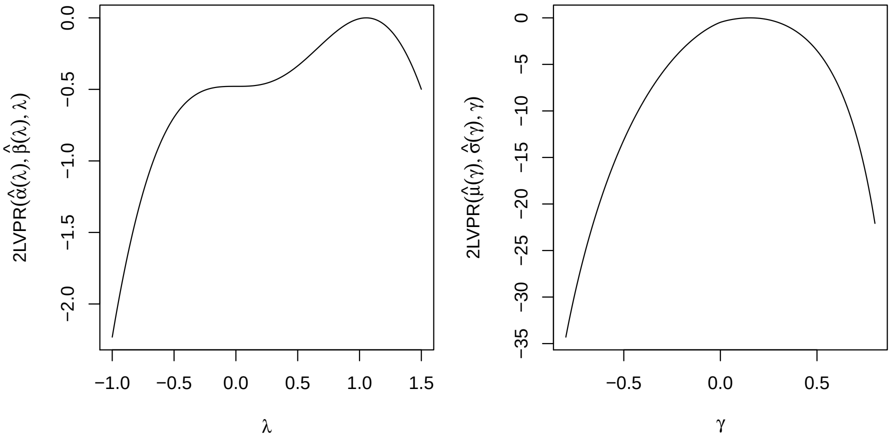


Figure 1 – Twice profiled relative log-likelihood for λ in the direct parameterization (left panel) and for γ in the centered parameterization (right panel)

proposed an alternative parameterization for $Y \sim SN(\alpha, \beta^2, \lambda)$, which is defined by

$$Y = \mu + \sigma Z_0 \tag{1.3}$$

where $Z_0 = \dfrac{Z - \mu_z}{\sigma_z}$ with $Z \sim SN(0, 1, \lambda)$ and $\mu_z = b\delta$ and $\sigma_z = \sqrt{1 - b^2\delta^2}$.

The alternative parameterization is then formed by the centered parameters $\mu \in \mathbb{R}$, $\sigma \in \mathbb{R}^+$ and $\gamma \in (-0.99527, .99527)$, whose explicit expression are

$$\mu = E(Y) = \alpha + \beta\mu_z \quad \sigma^2 = \beta^2(1 - \mu_z^2) \quad \gamma = \frac{4 - \pi}{2}\frac{(b\delta)^3}{(1 - b^2\delta^2)^{3/2}} \tag{1.4}$$

where γ denotes the Person's skewness coefficient. The centered parameterization of the skew-normal distribution will be denoted by $Y \sim SN_c(\mu, \sigma^2, \gamma)$, where μ and σ^2 are the mean and variance of Y, respectively.

Using the Jacobian transformation, the density of (1.3), after some algebra, is given by

$$f(y|\mu, \sigma^2, \gamma) = 2\omega^{-1}\phi(\omega^{-1}(y - \xi))\Phi\left(\lambda\left(\frac{y - \xi}{\omega}\right)\right) \tag{1.5}$$

where

$$\begin{aligned} s &= \left(\frac{2}{4 - \pi}\right)^{1/3} \quad \xi = \mu - \sigma\gamma^{1/3}s \\ \omega &= \sigma\sqrt{1 + s^2\gamma^{2/3}} \quad \lambda = \frac{s\gamma^{1/3}}{\sqrt{b^2 + s^2\gamma^{2/3}(b^2 - 1)}} \end{aligned} \tag{1.6}$$

(HENZE, 1986) introduced a useful stochastic representation of the skew-normal, which is given by

$$Y \stackrel{d}{=} \alpha + \beta(\delta H + \sqrt{1 - \delta^2}T), \tag{1.7}$$

where $\stackrel{d}{=}$ means "distributed as" and $H \sim HN(0,1) \perp T \sim N(0,1)$, where $HN(.)$ denotes the half-normal distribution. Therefore, using (1.7) and the skew-normal under the centered parameterization as described in (1.3), we have that the stochastic representation of the skew-normal under the centered parameterization $Y \sim SN_c(\mu, \sigma^2, \gamma)$ is

$$Y \stackrel{d}{=} \xi + \omega(\delta H + \sqrt{1-\delta^2} T) \tag{1.8}$$

where ξ and ω are defined in (1.6).

Figure 2 shows the density of the skew normal distribution under the centered parameterization, with $\mu = 5$ and $\sigma^2 = 4$, for some values of γ. Negative and positive asymmetry are observed, respectively, as γ assumes negative and positive values, whereas the normal model is obtained when $\gamma = 0$.

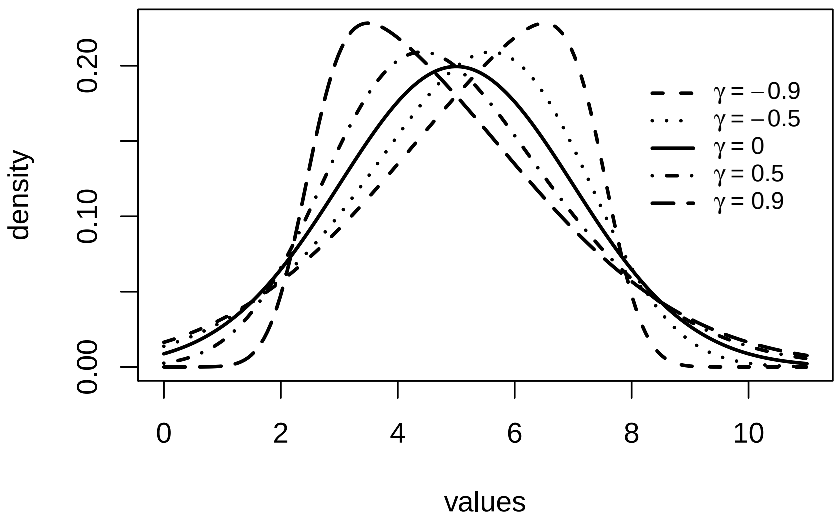


Figure 2 – Density of the skew normal distribution under the centered parameterization for some values of γ

1.2.2 Scale Mixture of Skew-Normal Distribution under the centered parameterization

The scale mixtures of normal distributions (SMN) were first introduced by (ANDREWS; MALLOWS, 1974) and it is often used to model symmetrical data ((FERREIRA; BOLFARINE; LACHOS, 2011)). A wide class of unimodal and symmetrical distributions can be written as a scale mixture of normals, as Student-t, contaminated normal, slash, among others ((WEST, 1987)). In (BRANCO; DEY, 2001) the authors proposed a general class of multivariate skew-elliptical distribution that included the scale mixture of normal distributions as special case. (FERREIRA; BOLFARINE; LACHOS, 2011) introduced an easy representation of the scale mixture of skew-normal distribution class and presented some of its probability and inferential properties, considering the EM algorithm for parameter estimation.

Following (FERREIRA; BOLFARINE; LACHOS, 2011), a random variable Y has a scale mixture of normal (SMN) distribution if it can be written as

$$Y \stackrel{d}{=} \mu + k(U)^{1/2}Z, \tag{1.9}$$

where $k(.)$ is a strictly positive function and $Z \sim N(0, \sigma^2)$ is independent of the scale random variable U with cumulative distribution function (cdf) $G(.|\boldsymbol{\nu})$. Then, a random variable Y follows a Scale Mixtures of Normal Distribution with mean $\mu \in \mathbb{R}$, scale $\sigma \in \mathbb{R}^+$ if its p.d.f can be written as

$$f(y|\mu, \sigma^2, \boldsymbol{\nu}) = \int_0^\infty \phi(y|\mu, k(u)\sigma^2)dH(u|\boldsymbol{\nu}) \tag{1.10}$$

As discussed in the previous section, the use of the direct parameterization of skew-normal distribution can lead to some inferential problems. Then, since the scale mixture of the skew-normal distribution is defined through the direct parameterization of the skew normal distribution, we now define the scale mixtures of skew-normal distribution under the centered parameterization.

Definition 1.2.1. A random variable Y follows a scale mixtures of skew normal distribution under the centered parameterization, if Y can be stochastically represented by

$$Y \stackrel{d}{=} \mu + k(U)^{1/2}Z, \tag{1.11}$$

where $Z \sim SN_c(0, \sigma^2, \gamma)$ and U is a scale random variable with c.d.f $G(.|\boldsymbol{\nu})$.

We use the notation $Y \sim SMSN_c(\mu, \sigma^2, \gamma, G, \boldsymbol{\nu})$ for a random variable represented as in Definition 1.2.1. From Definition 1.2.1, it follows that $E(Y) = \mu$, since $E(Z) = 0$ and $Var(Y) = \sigma^2 E(k(U))$. It also can be noted that when $\gamma = 0$ we get the corresponding scale mixtures of normal distribution family, introduced by (ANDREWS; MALLOWS, 1974), since when the skewness parameter is equal to zero, $Z \sim N(0, \sigma^2)$.

From Definition 1.2.1, the hierarchical representation of Y is given by

$$\begin{aligned} Y|U = u &\sim SN_c(\mu, \sigma^2 k(u), \gamma) \\ U &\sim G(.|\boldsymbol{\nu}) \end{aligned} \tag{1.12}$$

Using Henze's representation of the skew-normal distribution under the centered parameterization as described in 1.8, we also have that Y can be hierarchical represented as

$$\begin{aligned} Y|U=u,H=h &\sim N(\xi_u+\omega_u\delta h,\omega_u^2(1-\delta^2)) \\ H &\sim HN(0,1) \\ U &\sim G(.|\boldsymbol{\nu}) \end{aligned} \tag{1.13}$$

where

$$\begin{aligned} s &= \left(\frac{2}{4-\pi}\right)^{1/3} \quad \xi_u=\mu-\sigma\sqrt{k(u)}\gamma^{1/3}s \\ \omega_u &= \sigma\sqrt{k(u)}\sqrt{1+s^2\gamma^{2/3}} \quad \lambda=\frac{s\gamma^{1/3}}{\sqrt{b^2+s^2\gamma^{2/3}(b^2-1)}} \quad \delta=\frac{\lambda}{\sqrt{1+\lambda^2}} \end{aligned} \tag{1.14}$$

1.2.3 Examples of Scale Mixture of Skew-Normal Distribution under the centered parameterization

In this section we will present some members of the SMSN family under the centered parameterization. For this work, we will restrict this family considering $k(u)=\frac{1}{u}$. Then, it follows that $Var(Y)=\sigma^2E(U^{-1})$. Before introducing some members of this family, let us first define

$$\begin{aligned} d &= \frac{(y-\mu)}{\sigma\omega_1} \quad \omega_1=\sqrt{1+s^2\gamma^{\frac{2}{3}}} \quad \xi_1=-\gamma^{1/3}s \\ \lambda &= \frac{s\gamma^{1/3}}{\sqrt{b^2+s^2\gamma^{2/3}(b^2-1)}} \quad s=\left(\frac{2}{4-\pi}\right)^{1/3} \quad b=\sqrt{\frac{2}{\pi}} \end{aligned} \tag{1.15}$$

1.2.3.1 Skew-t distribution under the centered parameterization:

Considering $U\sim \text{gamma}(\nu/2,\nu/2)$, where

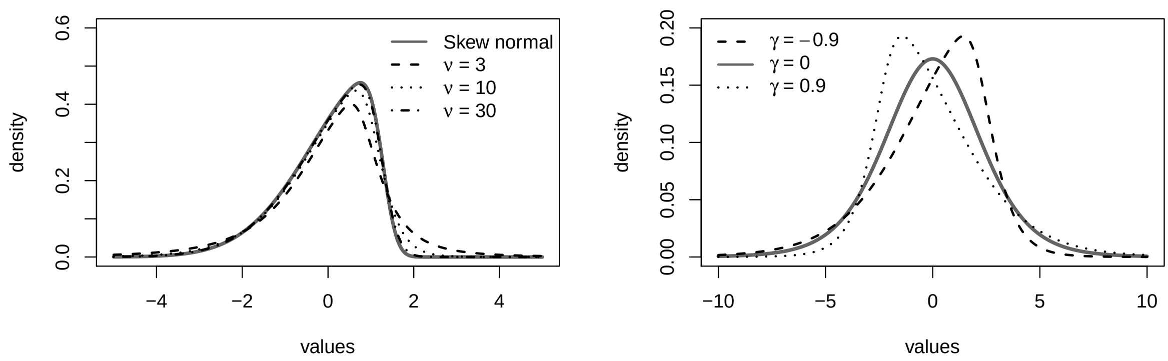
$$h(u|\nu)=\frac{\frac{\nu}{2}^{\frac{\nu}{2}}}{\Gamma(\frac{\nu}{2})}u^{\frac{\nu}{2}-1}e^{-\frac{\nu}{2}u}I_{(0,\infty)}(u)$$

we have the skew-t distribution under the centered parameterization, denoted by $ST_c(\mu,\sigma^2,\gamma,\nu)$, that is

$$f(y|\mu,\sigma^2,\gamma,\nu)=\frac{2\frac{\nu}{2}^{\frac{\nu}{2}}}{\sigma\omega_1\sqrt{2\pi}\Gamma(\frac{\nu}{2})}e^{\frac{-\xi_1^2}{2\omega_1^2}}\int_0^\infty u^{\frac{\nu+1}{2}-1}e^{-\frac{1}{2}\left[u(d^2+\nu)-2\sqrt{u}d\frac{\xi_1}{\omega_1}\right]}\Phi\left(\lambda\left(\sqrt{u}d-\frac{\xi_1}{\omega_1}\right)\right)du, \tag{1.16}$$

where μ denotes the mean, γ the skewness parameter, ν the degree of freedom and σ^2 is related to the variance of Y through $Var(Y)=\sigma^2\frac{\nu}{\nu-2}$, since $E(U^{-1})=\frac{\nu}{\nu-2}$.

Figure 3a presents the density of the skew-t distribution for three values of ν ($\nu=3$ is the dashed line, $\nu=10$, dotted line and $\nu=30$, dot-dashed line) and the skew normal (solid line) for $\sigma^2=1$, $\gamma=-0.9$ and $\mu=0$. As ν increases the skew-t approaches to the skew normal distribution. For $\nu=30$ it is possible to see that the skew-t and skew



(a) (b)

Figure 3 – Probability density function of skew-t distribution under the centered parameterization (a) varying the values of ν, with $\gamma = -0.9$, $\sigma^2 = 1$, $\mu = 0$ and (b) varying the parameter γ, with $\nu = 8$, $\sigma^2 = 5$, $\mu = 0$.

normal curves almost overlap. In Figure 3b, the density of skew-t distribution is drawn for $\gamma = 0$ (solid line), $\gamma = 0.9$ (dashed line) and $\gamma = -0.9$ (dotted line). It is possible to see right/left asymmetry as γ assumes positive/negative values, respectively.

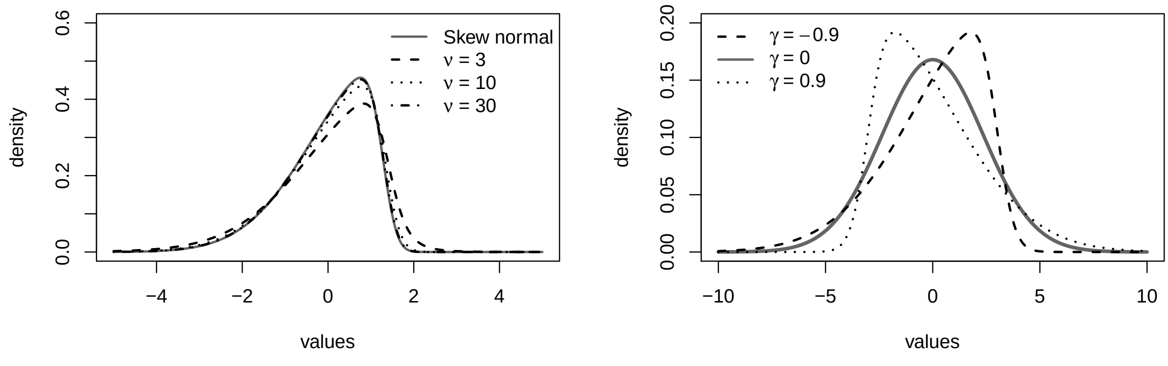
1.2.3.2 Skew-slash distribution under the centered parameterization:

If we consider $U \sim beta(\nu, 1)$ then

$$f(y|\mu, \sigma^2, \gamma, \nu) = \frac{2\nu}{\sigma\omega_1\sqrt{2\pi}} e^{\frac{-\xi_1^2}{2\omega_1^2}} \int_0^1 u^{\nu+\frac{1}{2}-1} e^{-\frac{1}{2}\left[ud^2 - 2\sqrt{u}d\frac{\xi_1}{\omega_1}\right]} \Phi\left(\lambda\left(\sqrt{u}d - \frac{\xi_1}{\omega_1}\right)\right) du \quad (1.17)$$

where μ is the mean, γ the skewness parameter, ν the degree of freedom and σ^2 is related to the variance of Y through $Var(Y) = \sigma^2 \dfrac{\nu}{\nu - 1}$. This distribution is denoted by $SSL_c(\mu, \sigma^2, \gamma, \nu)$.

Again, from Figures 4a and 4b it is possible to see a similar behavior as observed in the skew-t distribution.



(a) (b)

Figure 4 – Probability density function of skew-slash distribution under the centered parameterization (a) varying the values of ν, with $\gamma = -0.9$, $\sigma^2 = 1$, $\mu = 0$ and (b) varying the parameter γ, with $\nu = 8$, $\sigma^2 = 5$, $\mu = 0$.

1.2.3.3 Skew-contaminated normal distribution under the centered parameterization:

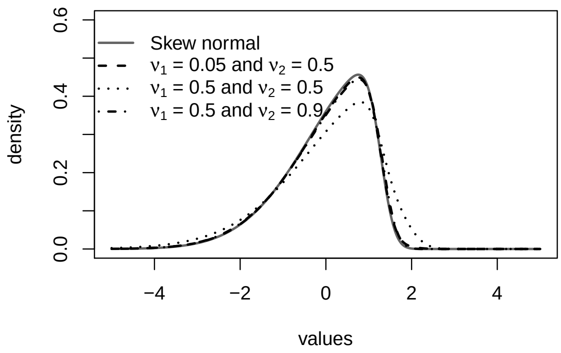
Considering U a discrete random variable assuming only two values, with the following probability function

$$h(u|\boldsymbol{\nu}) = \nu_1 I(u = \nu_2) + (1 - \nu_1) I(u = 1)$$

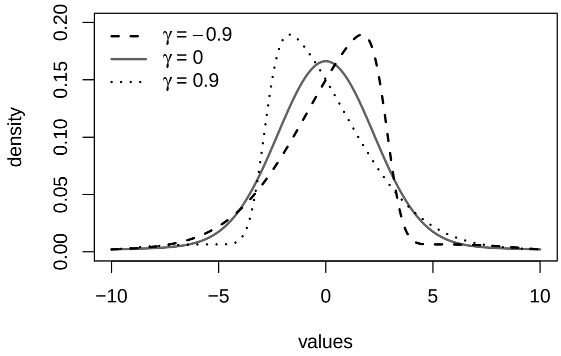
and $E(U^{-1}) = \dfrac{\nu_1 + \nu_2(1-\nu_1)}{\nu_2}$. Then the respective density is given by

$$\begin{aligned} f(y|\mu, \sigma^2, \gamma, \nu_1, \nu_2) &= 2\left\{\nu_1 \frac{\sqrt{\nu_2}}{\sigma\omega_1\sqrt{2\pi}} e^{-\frac{1}{2}\left(\sqrt{\nu_2}d - \frac{\xi_1}{\omega_1}\right)^2} \Phi\left(\lambda\left(\sqrt{\nu_2}d - \frac{\xi_1}{\omega_1}\right)\right)\right. \\ &\quad \left. + (1-\nu_1)\frac{1}{\sigma\omega_1\sqrt{2\pi}} e^{-\frac{1}{2}\left(d - \frac{\xi_1}{\omega_1}\right)^2} \Phi\left(\lambda\left(d - \frac{\xi_1}{\omega_1}\right)\right)\right\} \end{aligned} \tag{1.18}$$

denoted by $SCN_c(\mu, \sigma^2, \gamma, \boldsymbol{\nu})$, where μ is the mean, γ is the skewness parameter. According to (GARAY; LACHOS; ABANTO-VALLE, 2011) the parameters ν_1 and ν_2 can be interpreted as the proportion of outliers and a scale factor, respectively. For this distribution, the variance of Y is equal to $\sigma^2 \dfrac{\nu_1 + \nu_2(1-\nu_1)}{\nu_2}$. From Figure 5a it is possible to see that as ν_2 approaches 1 and/or ν_1 approaches 0, the skew contaminated normal density tends to the skew normal distribution.



(a)



(b)

Figure 5 – Probability density function of skew contaminated normal distribution under the centered parameterization (a) varying the values of ν_1 and ν_2, with $\sigma^2 = 1$, $\gamma = -0.9$ and $\mu = 0$ and (b) varying the parameter γ, with $\sigma^2 = 5$, $\mu = 0$, $\nu_1 = .1$ and $\nu_2 = 0.1$.

1.2.3.4 Skew generalized t distribution under the centered parameterization:

Skew generalized t distribution is obtained by considering $U \sim \text{gamma}(\nu_1/2, \nu_2/2)$ as the scale random variable, where

$$h(u|\boldsymbol{\nu}) = \frac{\frac{\nu_2}{2}^{\frac{\nu_1}{2}}}{\Gamma(\frac{\nu_1}{2})} u^{\frac{\nu_1}{2}-1} e^{-\frac{\nu_2}{2}u} I_{(0,\infty)}(u).$$

However, before provide further details, we will discuss about a model identification problem. Without loss of generality, consider the case when $\gamma = 0$ and $\mu = 0$, that is, $Z \sim N(0, \sigma^2)$. From the stochastic representation, we have that

$$f(y|\mu, \sigma^2, \gamma, \nu_1, \nu_2) = \int_0^\infty \phi(y|0, \frac{\sigma^2}{u}) h(u|\boldsymbol{\nu}) du$$

which leads, after some algebra, to

$$f(y|\mu, \sigma^2, \gamma, \nu_1, \nu_2) = \frac{\Gamma(\frac{\nu_1+1}{2})}{\sqrt{\nu_2 \sigma^2 \pi} \Gamma(\frac{\nu_1}{2})} \left(1 + \frac{y^2}{\nu_2 \sigma^2}\right)^{\frac{\nu_1+1}{2}} \tag{1.19}$$

From equation (1.19) it is evident that different values of σ^2 and ν_2 can produce the same value for the likelihood. Problems with identifiability can also be noted in skew generalized t distribution under the centered parameterization. To illustrate this, consider the plot in Figure 6. It is visible that the two curves overlap each other. So, to avoid this problem, we will fix $\sigma = 1$. In this way, we will have the skew-t with $\sigma^2 = 1$ and the skew normal with $\sigma^2 = 1$, as special cases of the skew generalized t distribution.

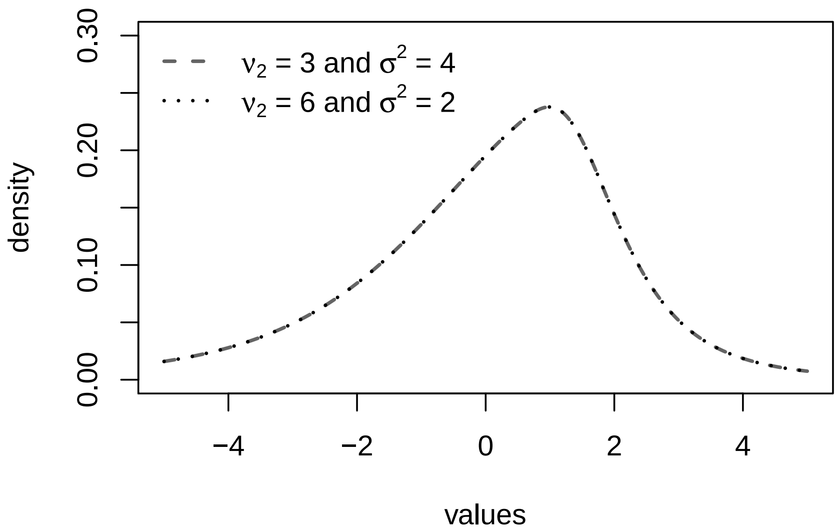


Figure 6 – Probability density function of skew generalized t distribution under the centered parameterization for two parameter sets of ν_2 and σ^2.

The density of the skew generalized t distribution under the centered parameterization, denoted by

$SGT_c(\mu, \gamma, \nu_1, \nu_2)$, is given by

$$f(y|\mu, \gamma, \nu_1, \nu_2) = \frac{2 \frac{\nu_2}{2}^{\frac{\nu_1}{2}}}{\omega_1 \sqrt{2\pi}\Gamma(\frac{\nu_1}{2})} e^{\frac{-\xi_1^2}{2\omega_1^2}} \int_0^\infty u^{\frac{\nu_1+1}{2}-1} e^{-\frac{1}{2}\left[u\left(\frac{(y-\mu)^2}{\omega_1^2}+\nu_2\right)-2\sqrt{u}(y-\mu)\frac{\xi_1}{\omega_1^2}\right]} \Phi\left(\lambda\left(\sqrt{u}\frac{(y-\mu)}{\omega_1} - \frac{\xi_1}{\omega_1}\right)\right) du, \tag{1.20}$$

where μ is the mean, γ is the skewness parameter, and ν_1 and ν_2 control de variance of Y, since $Var(Y) = \dfrac{\nu_2}{\nu_1 - 2}$.

Figure 7 presents density of skew generalized t for three combinations of ν_1 and ν_2 values. The solid line represents the density of the skew normal distribution, dashed line for skew generalized t with $\nu_1 = 30$ and $\nu_2 = 5$, dotted line when $\nu_1 = 30$ and $\nu_2 = 5$, and dot-dashed line when $\nu_1 = 30$ and $\nu_2 = 30$. We can see that the generalized t distribution approaches to the skew normal only when both ν_1 and ν_2 tend to infinite, simultaneously. In Figure 8, the density of skew generalized t distribution is drawn to $\gamma = 0$ (solid line), $\gamma = 0.9$ (dashed line) and $\gamma = -0.9$ (dotted line).

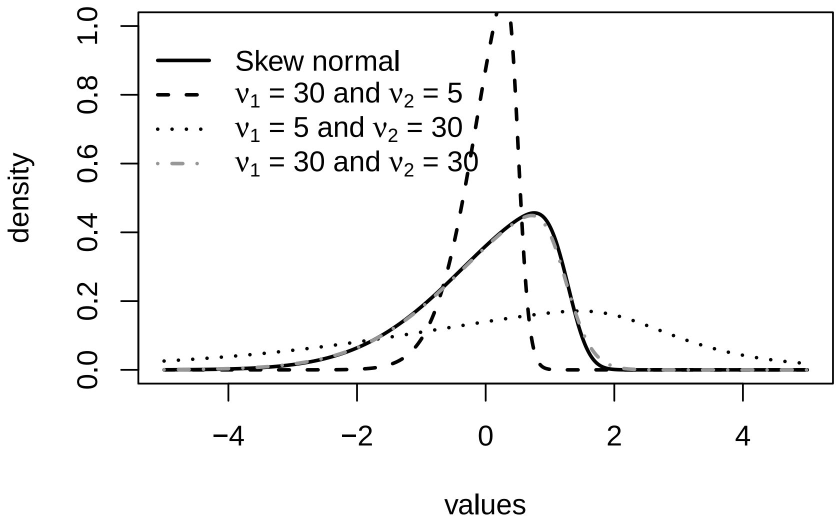


Figure 7 – Probability density function of skew generalized t distribution under the centered parameterization for some values of ν_1 and ν_2, $\sigma^2 = 1$, $\mu = 0$ and $\gamma = -0.9$.

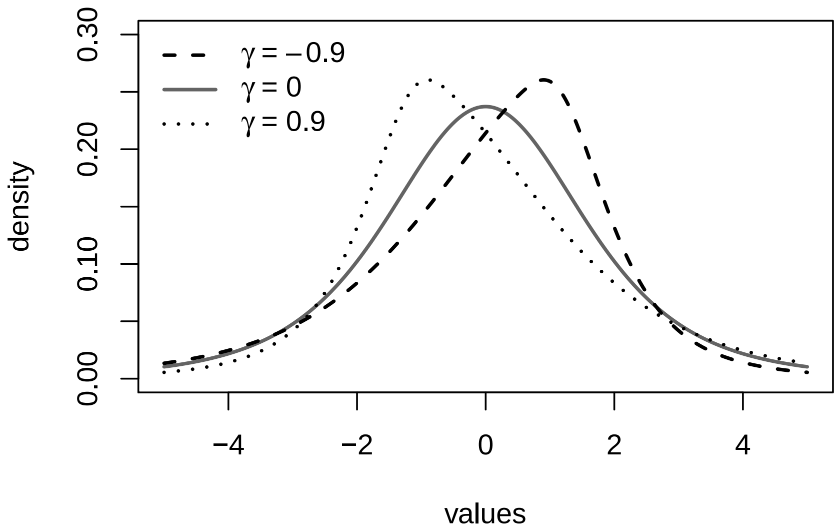


Figure 8 – Probability density function of skew generalized t distribution under the centered parameterization for some values of γ, $\sigma^2 = 5$, $\mu = 0$, $\nu_1 = 8$ and $\nu_2 = 8$.

1.3 Profiled log-likelihood for SMSN class of distributions

By introducing the skew normal distribution, we pointed out some inferential problems related to its direct parameterization. Since the Scale Mixture of Skew Normal distribution depends on the direct parameterization ((FERREIRA; BOLFARINE; LACHOS, 2011)), this family may heritage the same problems and, then the respective centered parameterization of this family can circumvents these problems. In this section, our objective is to show that under the direct parameterization for the SMSN distributions, log-likelihoods for λ presents problems around . In addition, we will show that under the centered parameterization, this problem is no longer observed. Based on the results from some simulation studies that the authors performed previously, we are also interest in analyze the log-likelihood for ν, since depending on its behavior, the estimation procedure can be a complicated task. As noted in (VILCA; AZEVEDO; BALAKRISHNAN, 2017), for the sinh-contaminated normal distribution developed in this work, large posterior standard deviation and large length of credibility intervals for ν_1 and ν_2 may be explained by the ill-behavior of the profiled log-likelihoods for these parameters.

Consider a random sample $\boldsymbol{y} = (y_1, y_1, \ldots, y_n)^t$ from the SMSN family under the direct and centered parameterizations and $\boldsymbol{\theta} = (\mu, \sigma, \lambda, \boldsymbol{\nu})$ the parameter vector. We shall denote the log-likelihood under the centered parameterization by $l(\boldsymbol{\theta}|y)_{CP}$. Under direct parameterization we shall use the density functions as described in (FERREIRA; BOLFARINE; LACHOS, 2011) and denote the respective log-likelihoods by $l(\boldsymbol{\theta}|y)_{DP}$. Log-likelihoods for the direct and centered parameterizations are presented in Appendix A.

1.3.1 Profiled log-likelihood for parameter λ

Following the ideas of (CHAVES, 2015) and (SANTOS, 2012), the profiled log-likelihood for λ is calculated by obtaining the maximum-likelihood estimator for the parameters $\mu, \sigma^2, \boldsymbol{\nu}$, for a specif value of λ and then plugging these estimates in the log-likelihood. for each λ defined in its parametric space we repeat the former steps.

We denote the profiled log-likelihood of the direct parameterization as $l_{DP}(\hat{\mu}(\lambda), \hat{\boldsymbol{\nu}}(\lambda), \lambda)$ for the skew generalized t distribution or $l_{DP}(\hat{\mu}(\lambda), \hat{\sigma^2}(\lambda), \hat{\boldsymbol{\nu}}(\lambda), \lambda)$ for the others distributions, where $\hat{\mu}(\lambda), \hat{\sigma^2}(\lambda), \hat{\boldsymbol{\nu}}(\lambda)$ stand for values of $\mu, \sigma^2, \boldsymbol{\nu}$ that maximizes l_{DP} for a specif value of λ. To make interpretations easier, we have calculated the relative profiled log-likelihood, as suggested by (ARELLANO-VALLE; AZZALINI, 2008) obtained by subtracting $l_{DP}(\hat{\mu}(\lambda), \hat{\sigma^2}(\lambda), \hat{\boldsymbol{\nu}}(\lambda), \lambda)$ of $l_{DP}(\hat{\mu}(\lambda), \hat{\sigma^2}(\lambda), \hat{\boldsymbol{\nu}}(\lambda), \hat{\lambda})$.

Similarly to the direct parameterization, the relative profiled log-likelihood under the centered parameterization is obtained by subtracting $l_{CP}(\hat{\mu}(\gamma), \hat{\sigma^2}(\gamma), \hat{\boldsymbol{\nu}}(\gamma), \gamma)$ of $l_{CP}(\hat{\mu}(\gamma), \hat{\sigma^2}(\gamma), \hat{\boldsymbol{\nu}}(\gamma), \hat{\gamma})$.

We generated a random sample of size 100 from a Scale Mixture of Normal distribution under the centered parameterization with $\mu = 0$, $\sigma^2 = 1$, $\gamma = 0.7$ and $\nu = 4$ for skew slash, $\nu_1 = 0.4$ and $\nu_2 = 0.6$ for the skew contaminated normal, $\nu = 5$ for skew-t and $\nu_1 = 15$ $\nu_2 = 5$ for the skew generalized-t

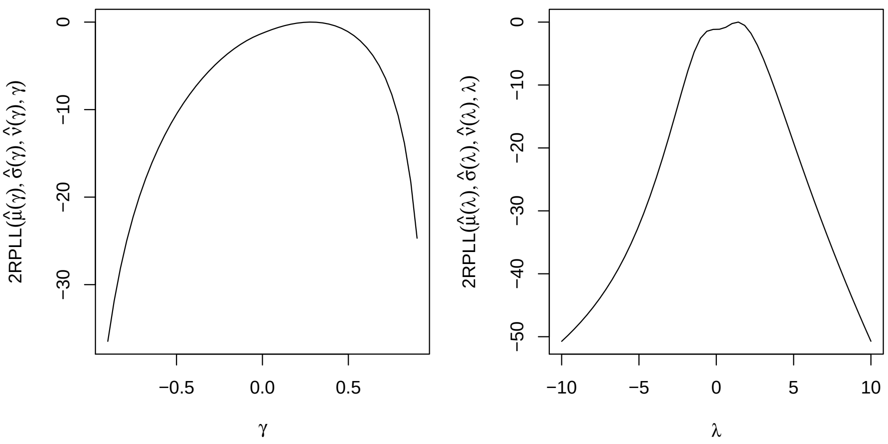


Figure 9 – Profile twice the relative log-likelihood for γ in the centered parameterization (left panel) and for λ in the direct parameterization (right panel) for the skew contaminated model

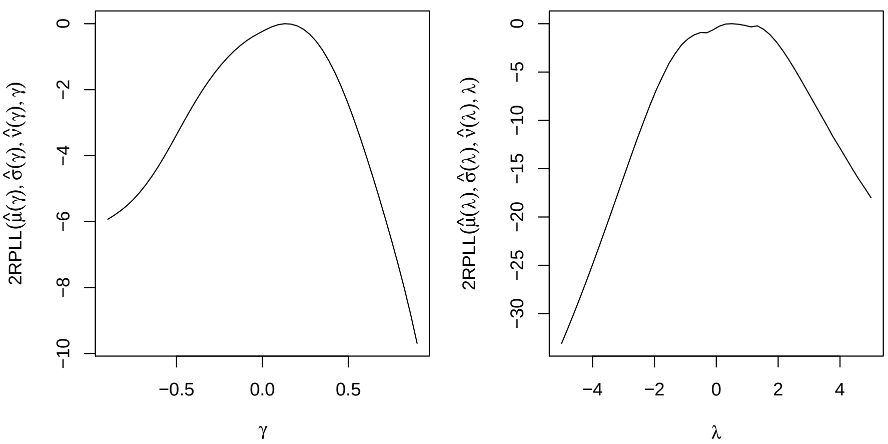


Figure 10 – Profile twice the relative log-likelihood for γ in the centered parameterization (left panel) and for λ in the direct parameterization (right panel) for the skew slash model

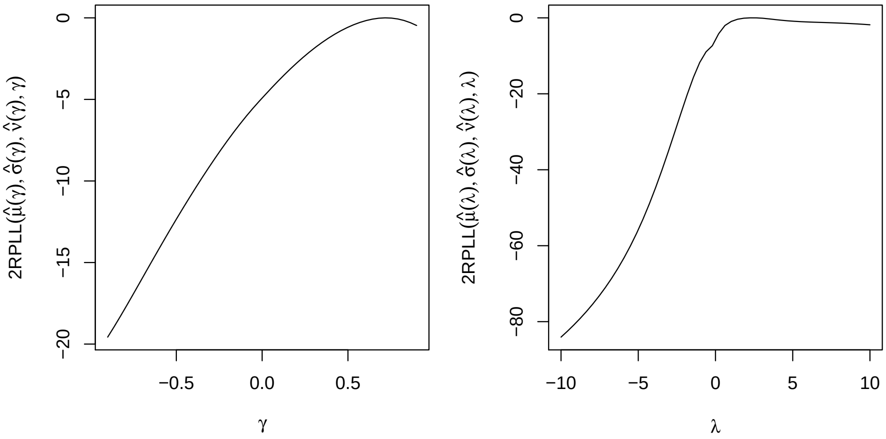


Figure 11 – Profile twice the relative log-likelihood for γ in the centered parameterization (left panel) and for λ in the direct parameterization (right panel) for the skew-t model

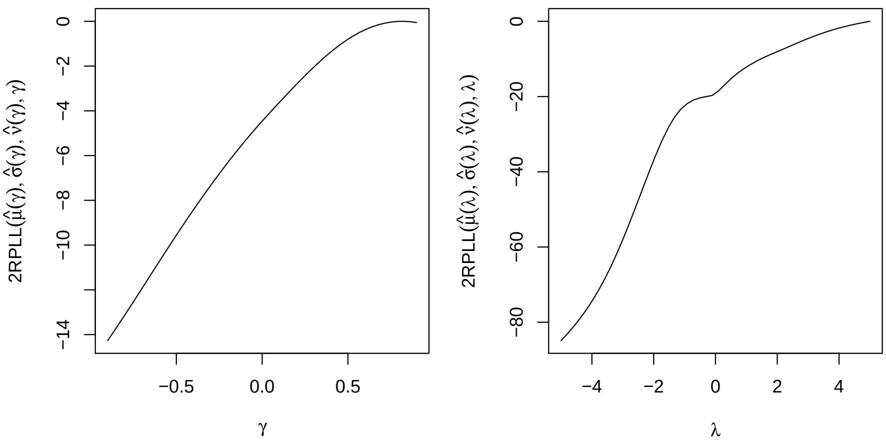
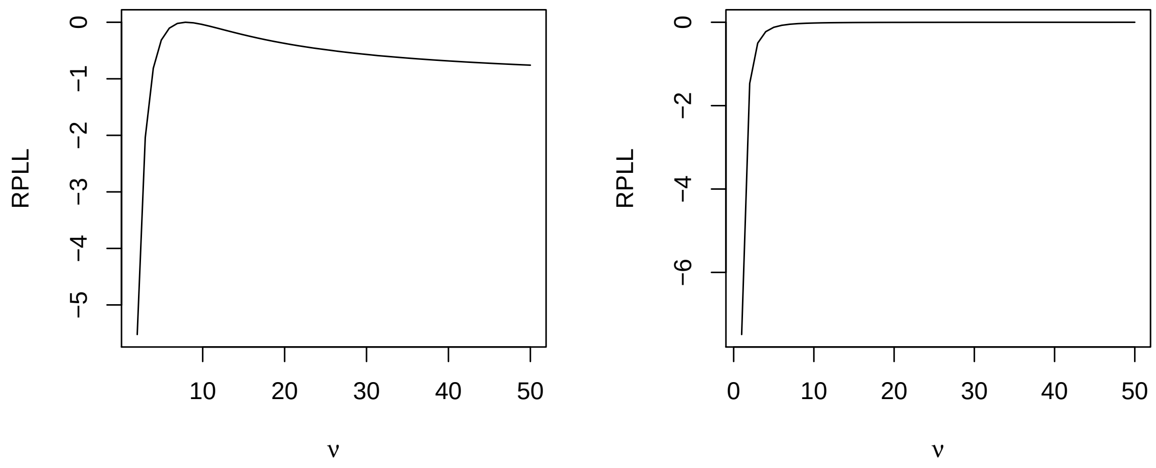


Figure 12 – Profile twice the relative log-likelihood for γ in the centered parameterization (left panel) and for λ in the direct parameterization (right panel) for the skew generalized t model

From Figures 9 - 12, it is possible to see that under the centered parameterization twice the relative log-likelihood presents a concave behavior while in direct parameterization the relative log-likelihood presents a non-quadratic shape around point zero. In figure 11, for values of $\lambda > 0$ the function is almost linear. These figures illustrate that under the centered parameterization we have better behavior of log-likelihood for the parameter γ than for the parameter λ in direct parameterization.

1.3.2 Profiled log-likelihood for parameter ν

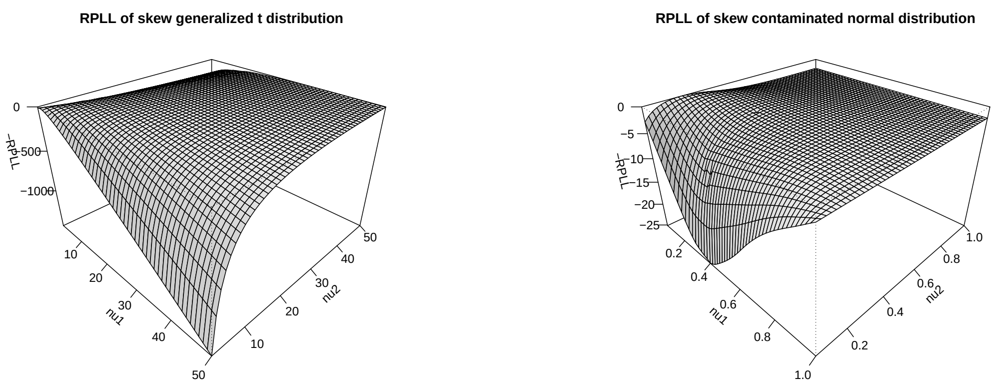
The profiled log-likelihood for parameter $\boldsymbol{\nu}$ was calculated in the same way as described previously. The only difference from the former procedure is that now we replaced the maximum-likelihood estimator of parameters μ,σ^2 and γ in the log-likelihood function, calculated by fixing values of $\boldsymbol{\nu}$.



(a) (b)

Figure 13 – Profiled relative log-likelihood for ν for the: skew-t (a) and skew slash (b)distributions.

From Figures 13 and 14, it is possible to observe that as ν increases, the profiled relative log-likelihood tends to be a straight line. That is, for larger values of ν, the estimates tends to range within such interval with almost the same probability. Similar behavior can be noted for the skew generalized t and skew contaminated normal distributions, as we can see from Figures 14a and 14b. As long as at least ν_1 or ν_2 increases, flatter is the surface. Analysis of these figures indicates that the profiled log-likelihoods are ill-behaved, which can result in large posterior standard deviations and large credibility intervals for $\boldsymbol{\nu}$.



(a) (b)

Figure 14 – Profiled relative log-likelihood for ν_1 and ν_2 for the: skew generalized t (a) and skew contaminated normal (b) distributions.

1.4 Regression Model

Let $\boldsymbol{X} = (1, \boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_{p-1})^t$ a $p \times n$ design matrix of fixed covariates, $\boldsymbol{Y} = (Y_1, \ldots, Y_n)^t$ a $n \times 1$ vector of response variables and $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_{p-1})^t$ a $p \times 1$ vector of regression coefficients. The regression model is given by

$$Y_i = {X_i}^t \boldsymbol{\beta} + \varepsilon_i \quad i = 1, \ldots, n \tag{1.21}$$

where ε_i are independent random variables identically distributed as a member of SMSN family. In this work, we say that $\varepsilon_i \overset{iid}{\sim} SSL_c(0,1,\gamma,\nu)$ or $\varepsilon_i \overset{iid}{\sim} ST_c(0,1,\gamma,\nu)$ or $\varepsilon_i \overset{iid}{\sim} SCN_c(0,1,\gamma,\nu_1,\nu_2)$ or $\varepsilon_i \overset{iid}{\sim} SGT_c(0,1,\gamma,\nu_1,\nu_2)$, where SSL_c, ST_c, SCN_c and SGT_c denote, respectively, the skew slash, skew-t, skew contaminated and skew generalized t distributions under the centered parameterization. From the Scale Mixture of Skew-Normal distributions under the centered parameterization it follows that $E(Y_i) = \mu_i = \boldsymbol{X}_i^t\boldsymbol{\beta}$, where $\boldsymbol{X}_i$ is the i-th row of matrix $\boldsymbol{X}$. From the hierarchical representation described in (1.12), we have that

$$\begin{aligned} Y_i|U_i = u_i &\sim SN_c\left(\boldsymbol{X}_i^t\boldsymbol{\beta}, \frac{\sigma^2}{u_i}, \gamma\right) \\ U_i &\sim G(.|\boldsymbol{\nu}) \end{aligned}$$

We also can represent Y as

$$Y_i|U_i = u_i \overset{d}{=} \boldsymbol{X}_i^t\boldsymbol{\beta} + \frac{\sigma}{\sqrt{u_i}}\left(\frac{V_i - \mu_v}{\sigma_v}\right)$$

where $V_i \overset{iid}{\sim} SN(0,1,\lambda)$, μ_v and σ_v are the mean and the variance of V_i, respectively. Using Henze's stochastic representation for V_i, then

$$Y_i|U_i = u_i \overset{d}{=} \boldsymbol{X}_i^t\boldsymbol{\beta} - \frac{\sigma}{\sqrt{u_i}}\frac{\mu_v}{\sigma_v} + \frac{\sigma}{\sigma_v\sqrt{u_i}}(\delta H + \sqrt{1-\delta^2}Z) \tag{1.22}$$

Since $\mu_v = \delta b$ and $\sigma_v = \sqrt{1 - b^2\delta^2}$ then (1.22) becomes

$$\begin{aligned} Y_i|U_i = u_i &\overset{d}{=} \boldsymbol{X}_i^t\boldsymbol{\beta} - \frac{\sigma\delta b}{\sqrt{u_i}\sqrt{1-b^2\delta^2}} + \frac{\sigma\delta}{\sqrt{u_i}\sqrt{1-b^2\delta^2}}H_i + \frac{\sigma\sqrt{1-\delta^2}}{\sqrt{u_i}\sqrt{1-b^2\delta^2}}Z_i \\ &\boldsymbol{X}_i^t\boldsymbol{\beta} + \frac{\sigma\delta}{\sqrt{u_i}\sqrt{1-b^2\delta^2}}(H_i - b) + \frac{\sigma\sqrt{1-\delta^2}}{\sqrt{u_i}\sqrt{1-b^2\delta^2}}Z_i \end{aligned} \tag{1.23}$$

Setting $\Delta = \dfrac{\sigma\delta}{\sqrt{1-b^2\delta^2}}$ and $\tau = \dfrac{\sigma^2(1-\delta^2)}{1-b^2\delta^2}$, we have an one by one transformation such that it is possible to recover σ and δ through: $\delta = \dfrac{\Delta}{\sqrt{\tau + \Delta^2}}$ and $\sigma^2 = \tau + \Delta^2(1-b^2)$. Then

$$Y_i|U_i = u_i \overset{d}{=} \boldsymbol{X}_i^t\boldsymbol{\beta} + \frac{\Delta}{\sqrt{u_i}}(h_i - b) + \frac{\sqrt{\tau}}{\sqrt{u_i}}Z_i \tag{1.24}$$

Using (1.24), the hierarchical representation of Y becomes

$$\begin{aligned} Y_i|U_i = u_i, H_i = h_i &\sim N(\boldsymbol{X}_i^t\boldsymbol{\beta} + \frac{\Delta}{\sqrt{u_i}}(h_i - b), \frac{\tau}{u_i}) \\ H_i &\sim HN(0,1) \\ U_i &\sim G(.|\boldsymbol{\nu}) \end{aligned} \tag{1.25}$$

1.4.1 Bayesian Inference

To use the Bayesian paradigm, it is essential to obtain the joint posterior distribution. However, since the necessary integrals are not easy to calculate, it is not possible to obtain such distribution, analytically. However, it is possible to obtain numerical approximation for the marginal posterior distributions of interest by using MCMC algorithms, see (GEMAN; GEMAN, 1984) and (HASTINGS, 1970).

To obtain the posterior distribution we need to consider the complete likelihood

$$\begin{aligned} L_c(\boldsymbol{\theta}|y,u,h) &\propto \prod_{i=1}^{n} \phi\left(y_i|\mu_i, \tau u_i^{-1}\right) f(h_i) h(u_i|\boldsymbol{\nu}) \\ &\propto \prod_{i=1}^{n} \frac{\sqrt{u_i}}{\sqrt{\tau}} \exp\left\{-\frac{u_i}{2\tau}\left(y_i - \mu_i\right)^2\right\} \exp\left\{-\frac{h_i^2}{2}\right\} h(u_i|\boldsymbol{\nu}) \\ &\propto \frac{\prod_{i=1}^{n}\sqrt{u_i}}{\tau^{n/2}} \exp\left\{-\frac{1}{2\tau}\sum_{i=1}^{n} u_i\left(y_i-\mu_i\right)^2\right\} \exp\left\{-\frac{\sum_{i=1}^{n} h_i^2}{2}\right\} \prod_{i=1}^{n} h(u_i|\boldsymbol{\nu}) \end{aligned}$$

where $\mu_i = \boldsymbol{X}_i^t\boldsymbol{\beta} + \frac{\Delta}{\sqrt{u_i}}(h_i - b)$ and $\boldsymbol{\theta} = (\boldsymbol{\beta}, \Delta, \tau, \boldsymbol{\nu})$. Since we set $\sigma^2 = 1$ for the skew generalized t model, the MCMC algorithm will be slightly different from the other models. For this model, we have $\boldsymbol{\theta} = (\boldsymbol{\beta}, \delta, \boldsymbol{\nu})$, therefore, Δ and τ are functions of only ν. Then, for this model, is preferable to sample directly from δ, instead of sampling for Δ and τ.

We need to consider a prior distribution for $\boldsymbol{\theta}$. We will assume an independence structure, that is $\pi(\boldsymbol{\theta}) = \pi(\boldsymbol{\beta})\pi(\Delta)\pi(\tau)\pi(\boldsymbol{\nu})$ for the skew-t, skew slash and skew contaminated normal models and $\pi(\boldsymbol{\theta}) = \pi(\boldsymbol{\beta})\pi(\delta)\pi(\boldsymbol{\nu})$, for the skew generalized t model. Furthermore, we will assume conditional conjugate prior distributions, as in (GELMAN, 2006), for $\boldsymbol{\beta}$, τ^{-1}, Δ and $\delta \in U(-1,1)$. On the other hand, for $\boldsymbol{\nu}$, the choice of the prior distribution will depend on the model.

1.4.2 Full conditional distributions

In order to implement the MCMC algorithm, we have to simulate iteratively from the full conditionals described bellow.

Denoting by $\boldsymbol{\theta}_{-\boldsymbol{\theta}_i}$ the parameter vector $\boldsymbol{\theta}$ without the component $\boldsymbol{\theta}_i$, the full conditional distributions are

For β:

$$\pi(\boldsymbol{\beta}|\boldsymbol{\theta}_{-\beta},\boldsymbol{y},\boldsymbol{u},\boldsymbol{h}) \propto \exp\left\{-\frac{1}{2}\left(\boldsymbol{\beta}^t\Sigma_*^{-1}\boldsymbol{\beta}-2\mu_*^t\Sigma_*^{-1}\boldsymbol{\beta}\right)\right\} I_{\mathbb{R}^p}(\boldsymbol{\beta})$$

which can be recognized as the kernel of p-variate normal distribution with variance $\boldsymbol{\Sigma_*} = \left(\frac{\sum_{i=1}^n u_i x_i x_i^t}{\tau}+\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{-1}\right)^{-1}$ and mean $\boldsymbol{\mu_*} = \left(\sum_{i=1}^n \frac{u_i}{\tau}\left(y_i-\frac{\Delta}{\sqrt{u_i}}(h_i-b)\right)\boldsymbol{x_i^t}+\boldsymbol{\mu_\beta^t}\boldsymbol{\Sigma_\beta^{-1}}\right)\boldsymbol{\Sigma_*}$.

For h_i:

$$f(\boldsymbol{\theta},u_i,h_i|y_i) \propto \exp\left\{-\frac{1}{2}\left(\frac{\Delta^2+\tau}{\tau}\right)\left[h_i^2-2h_i\left(\frac{\Delta^2 b+\Delta\sqrt{u_i}(y_i-\boldsymbol{X}_i^t\boldsymbol{\beta})}{\Delta^2+\tau}\right)\right]\right\} I_{(0,\infty)}(h_i)$$

which can be recognized as the kernel of a truncated normal distribution, so

$$h_i|\boldsymbol{\theta},u_i,y_i \sim TN\left(\frac{\Delta^2 b+\Delta\sqrt{u_i}(y_i-\boldsymbol{X}_i^t\boldsymbol{\beta})}{\Delta^2+\tau},\frac{\tau}{\Delta^2+\tau}\right) I(0,\infty)$$

For u_i:

- Skew slash:

  $$f(\boldsymbol{\theta},u_i,h_i|y_i) \propto u_i^{\nu+1/2-1}\exp\left\{-\frac{u_i}{2\tau}\left[(y_i-\boldsymbol{X}_i^t\boldsymbol{\beta})^2-2\frac{\Delta}{\sqrt{u_i}}(h_i-b)(y_i-\boldsymbol{X}_i^t\boldsymbol{\beta})\right]\right\} I_{(0,1)}(u_i)$$

- Skew-t:

  $$f(\boldsymbol{\theta},u_i,h_i|y_i) \propto u_i^{\frac{\nu+1}{2}-1}\exp\left\{-\frac{u_i}{2}\left[\frac{(y_i-\boldsymbol{X}_i^t\boldsymbol{\beta})^2}{\tau}+\nu\right]+\frac{\Delta\sqrt{u_i}}{\tau}(h_i-b)(y_i-\boldsymbol{X}_i^t\boldsymbol{\beta})\right\} I_{(0,\infty)}(u_i)$$

- Skew generalized t:

  $$f(\boldsymbol{\theta},u_i,h_i|y_i) \propto u_i^{\frac{\nu_1+1}{2}-1}\exp\left\{-\frac{u_i}{2}\left[\frac{(y_i-\boldsymbol{X}_i^t\boldsymbol{\beta})^2}{\tau}+\nu_2\right]+\frac{\Delta\sqrt{u_i}}{\tau}(h_i-b)(y_i-\boldsymbol{X}_i^t\boldsymbol{\beta})\right\} I_{(0,\infty)}(u_i)$$

- Skew-contaminated normal: the discrete conditional distribution of u_i assumes ν_2 with probability $\frac{p_i}{p_i+q_i}$ and 1 with probability $\frac{q_i}{p_i+q_i}$ where

  $$p_i = \nu_1\sqrt{\nu_2}\exp\left\{-\frac{\nu_2}{2\tau}\left[(y_i-\boldsymbol{X}_i^t\boldsymbol{\beta})^2-2\frac{\Delta}{\sqrt{\nu_2}}(h_i-b)(y_i-\boldsymbol{X}_i^t\boldsymbol{\beta})\right]\right\}$$
  $$q_i = (1-\nu_1)\exp\left\{-\frac{1}{2\tau}\left[(y_i-\boldsymbol{X}_i^t\boldsymbol{\beta})^2-2\Delta(h_i-b)(y_i-\boldsymbol{X}_i^t\boldsymbol{\beta})\right]\right\}$$

For ν:

- Skew slash: Considering a gamma distribution left truncated at 1 as prior with mean $\frac{\alpha_1}{\alpha_2}$ and variance $\frac{\alpha_1}{\alpha_2^2}$, it follows that

$$\pi(\nu|\boldsymbol{\theta}_{-\nu},\boldsymbol{y},\boldsymbol{u},\boldsymbol{h}) \propto \nu^{n+\alpha_1-1}\exp\left\{-\nu(\alpha_2-\sum_{i=1}^{n} ln(u_i))\right\} I_{(1,\infty)}(\nu)$$

that is, $\nu|\boldsymbol{\theta}_{-\nu},\boldsymbol{y},\boldsymbol{u},\boldsymbol{h} \sim TG(n+\alpha_1,\alpha_2-\sum_{i=1}^{n} ln(u_i))I(1,\infty)$, where TG denotes the Truncated Gamma distribution.

- Skew-t: For Skew-t, we have adopted a very useful hierarchical prior distribution as noted in (CABRAL; LACHOS; MADRUGA, 2012), which consists in $\nu|\lambda \sim exp(\lambda)I(\nu)_{(2,\infty)}$ and $\lambda \sim U(\rho_0,\rho_1)$ where $0<\rho_0<\rho_1$ are known. The exponential distribution is left truncated at 2 to insure finite variance. Then

$$\pi(\nu|\boldsymbol{\theta}_{-\nu},\lambda,\boldsymbol{y},\boldsymbol{u},\boldsymbol{h}) \propto \frac{\frac{\nu}{2}^{\frac{n\nu}{2}}}{\Gamma(\nu/2)^n}\left(\prod_{i=1}^{n} u_i\right)^{\nu/2-1}\exp\left\{-\nu\left(\frac{\sum_{i=1}^{n} u_i}{2}+\lambda\right)\right\} I_{(2,\infty)}(\nu)$$

$$\pi(\lambda|\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{u},\boldsymbol{h}) \propto \lambda\exp-\lambda(\nu-2)I_{(\rho_0,\rho_1)}(\lambda)$$

that is, $\lambda|\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{u},\boldsymbol{h} \sim TG(2,\nu-2)I(\rho_0,\rho_1)$.

- Skew generalized t: Assuming $\nu_1|\lambda_1 \sim exp(\lambda_1)I(\nu_1)_{(2,\infty)}$ and $\lambda_1 \sim U(\rho_0,\rho_1)$ where $0<\rho_0<\rho_1$ are known and $\nu_2|\lambda_2 \sim exp(\lambda_2)$ and $\lambda_2 \sim U(\psi_0,\psi_1)$ where $0<\psi_0<\psi_1$ are known, we have

$$\pi(\nu_1|\boldsymbol{\theta}_{-\nu_1},\lambda_1,\boldsymbol{y},\boldsymbol{u},\boldsymbol{h}) \propto \frac{\nu_2/2^{n\nu_1/2}}{\Gamma(\nu_1/2)^n}\left(\prod_{i=1}^{n} u_i\right)^{\nu_1/2-1}\exp\left\{-\lambda_1(\nu_1-2)\right\} I_{(2,\infty)}(\nu_1)$$

$$\pi(\lambda_1|\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{u},\boldsymbol{h}) \propto \lambda_1\exp\left\{-\lambda_1(\nu_1-2)\right\} I_{(\rho_0,\rho_1)}(\lambda_1)$$

that is, $\lambda_1|\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{u},\boldsymbol{h} \sim TG(2,\nu_1-2)I(\rho_0,\rho_1)$.

Also, we have that

$$\pi(\nu_2|\boldsymbol{\theta}_{-\nu_2},\lambda_2,\boldsymbol{y},\boldsymbol{u},\boldsymbol{h}) \propto \nu_2/2^{n\nu_1/2}\exp\left\{-\nu_2\left(\frac{\sum_{i=1}^{n} u_i}{2}+\lambda_2\right)\right\} I_{(0,\infty)}(\nu_2)$$

and

$$\pi(\lambda_2|\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{u},\boldsymbol{h}) \propto \lambda_2\exp\{-\lambda_2(\nu_2)\}I_{(\psi_0,\psi_1)}(\lambda_2)$$

that is, $\nu_2|\boldsymbol{\theta}_{-\nu_2},\lambda_2,\boldsymbol{y},\boldsymbol{u},\boldsymbol{h} \sim \text{gamma}(\frac{n\nu_1}{2}+1,\frac{\sum_{i=1}^{n} u_i}{2}+\lambda_2)$

and $\lambda_2|\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{u},\boldsymbol{h} \sim TG(2,\nu_2-2)I(\xi_0,\xi_1)$

- Skew-contaminated normal: Observe that distribution of U can be written as

$$h(u|\boldsymbol{\nu}) = \nu_1^{\frac{1-u}{1-\nu_2}}(1-\nu_1)^{\frac{u-\nu_2}{1-\nu_2}} I_{\{\nu_2,1\}}(u)$$

Setting as prior distributions $\nu_1 \sim beta(\alpha_1, \beta_1)$, $\nu_2 \sim beta(\alpha_2, \beta_2)$, it follows that the conditional distributions of ν_1 and ν_2 are

$$\pi(\nu_1|\boldsymbol{\theta}_{-\nu_1}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{h}) \propto \nu_1^{\frac{n-\sum_{i=1}^n u_i}{1-\nu_2}+\alpha_1-1}(1-\nu_1)^{\frac{\sum_{i=1}^n u_i - n\nu_2}{1-\nu_2}+\beta_1-1} I_{(0,1)}(\nu_1)$$

which can be recognized as the kernel of a beta distribution. So, $\nu_1|\boldsymbol{\theta}_{-\nu_1}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{h} \sim beta\left(\frac{n-\sum_{i=1}^n u_i}{1-\nu_2}+\alpha_1, \frac{\sum_{i=1}^n u_i - n\nu_2}{1-\nu_2}+\beta_1\right)$ And

$$\pi(\nu_2|\boldsymbol{\theta}_{-\nu_2}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{h}) \propto \nu_1^{\frac{n-\sum_{i=1}^n u_i}{1-\nu_2}}(1-\nu_1)^{\frac{\sum_{i=1}^n u_i - n\nu_2}{1-\nu_2}} \nu_2^{\alpha_2-1}(1-\nu 2)^{(\beta_2-1)} I_{(0,1)}(\nu_2)$$

Finally, for the skew generalized t distribution the conditional distribution of δ is

For δ:

$$\pi(\delta|\boldsymbol{\theta}_{-\delta}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{h}) \propto \left(\frac{\sqrt{1-b^2\delta^2}}{\sqrt{1-\delta^2}}\right)^n \exp\left\{-\frac{1-b^2\delta^2}{2(1-\delta^2)}\sum_{i=1}^n u_i\left(y_i - \left(\boldsymbol{X}_i^t\boldsymbol{\beta} + \frac{\delta}{\sqrt{u_i}\sqrt{1-b^2\delta^2}}(h_i-b)\right)\right)\right\} I_{(-1,1)}(\delta)$$

and for the other distributions:

For Δ:

$$\pi(\Delta|\boldsymbol{\theta}_{-\Delta}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{h}) \propto \exp\left\{-\frac{1}{2}\left(\frac{\sigma_\Delta^2\sum_{i=1}^n(h_i-b)^2+\tau}{\tau\sigma_\Delta^2}\right)\left[\Delta^2 - 2\Delta\frac{\sigma_\Delta^2\sum_{i=1}^n(h_i-b)\sqrt{u_i}(y_i-\boldsymbol{X}_i^t\boldsymbol{\beta})+\mu_\Delta\tau}{\sigma_\Delta^2\sum_{i=1}^n(h_i-b)^2+\tau}\right]\right\} I_{\mathbb{R}}(\Delta)$$

which can be recognized as the kernel of normal distribution with mean $\frac{\sigma_\Delta^2\sum_{i=1}^n(h_i-b)\sqrt{u_i}(y_i-\boldsymbol{X}_i^t\boldsymbol{\beta})+\mu_\Delta\tau}{\sigma_\Delta^2\sum_{i=1}^n(h_i-b)^2+\tau}$ and variance $\frac{\tau\sigma_\Delta^2}{\sigma_\Delta^2\sum_{i=1}^n(h_i-b)^2+\tau}$

For τ^{-1}:

$$\pi(\tau^{-1}|\boldsymbol{\theta}_{-\tau^{-1}}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{h}) \propto (\tau^{-1})^{n/2+c-1}\exp\left\{-\tau^{-1}\left\{d+\sum_{i=1}^n\frac{u_i}{2}\left(y_i-\left(\boldsymbol{X}_i^t\boldsymbol{\beta}+\frac{\Delta}{\sqrt{u_i}}(h_i-b)\right)\right)^2\right\}\right\} I_{(0,\infty)}$$

that can be recognized as the kernel of a gamma distribution. So, $\tau^{-1}|\boldsymbol{\theta}_{-\tau^{-1}}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{h} \sim$ gamma $\left(n/2+c, d+\sum_{i=1}^n\frac{u_i}{2}\left(y_i-\left(\boldsymbol{X}_i^t\boldsymbol{\beta}+\frac{\Delta}{\sqrt{u_i}}(h_i-b)\right)\right)^2\right)$

1.4.3 Model fit assessment and model comparison

1.4.4 Model comparison

In Bayesian inference, there are several methodologies to compare a set of competing models for a given data set. According to (SPIEGELHALTER et al., 2002) under the Bayesian context, the selection criteria are obtained through the parameters posterior distribution. Moreover, when MCMC algorithms are used to approximate the posterior distribution, such criteria can be easily calculated. In this work we will consider expected the Akaike information criterion (EAIC), expected Bayesian information criterion (EBIC), deviance information criterion (DIC) and log pseudo-marginal likelihood (LPML) criterion. The LPML criterion is defined based on the conditional predictive ordinate (CPO), which is based on the cross validation criterion. Let $\boldsymbol{\theta} = (\boldsymbol{\beta}, \sigma^2, \gamma, \boldsymbol{\nu})$ and $D(\boldsymbol{\theta}) = -2l(\boldsymbol{\theta}|\boldsymbol{y})$, where $l(\boldsymbol{\theta}|\boldsymbol{y})$ is the log-likelihood given by equations (A.2), (A.4), (A.6) and (A.8), according to the fitted model. Also, let $\boldsymbol{\theta}^{(m)}$, $m = 1, \ldots, M$ a valid MCMC sample (after discharging the burn-in and the spacing between the values), $\bar{\boldsymbol{\theta}}$ the vector with the posterior expectation, and $\overline{D(\boldsymbol{\theta})} = (1/M) \sum_{m=1}^{M} D(\boldsymbol{\theta}^{(m)})$. According to (GARAY; LACHOS; ABANTO-VALLE, 2011), using the MCMC sample, the approximation for the CPO_i for the i-th observation, is defined as

$$\widehat{CPO_i} = \left\{ \frac{1}{M} \sum_{m=1}^{M} \frac{1}{L(\boldsymbol{\theta}^{(m)}|y_i)} \right\}^{-1} \tag{1.26}$$

A summary statistic of the CPO is the LPML criterion, which is defined as

$$LPML = \sum_{i=1}^{n} ln(\widehat{CPO_i})$$

For calculating DIC, EAIC and EBIC criteria, we need to define $D(\bar{\boldsymbol{\theta}}) = -2l(\bar{\boldsymbol{\theta}}|\boldsymbol{y})$, therefore $EAIC = D(\bar{\boldsymbol{\theta}}) + 2k$ and $EBIC = D(\bar{\boldsymbol{\theta}}) + klog(n)$ and DIC is defined as $DIC = D(\bar{\boldsymbol{\theta}}) + 2p_D$ where $p_D = \overline{D(\boldsymbol{\theta})} - D(\bar{\boldsymbol{\theta}})$. Smaller values of DIC, EAIC and EBIC indicate the best model, occurring the opposite with the LPLM. More discussions about these criteria can be found in (ANDO, 2007).

1.4.5 Influential observations

As noted in (CHO et al., 2009), influential observations in a data set can have a strong impact in statistical inference and the related conclusions. Computation of divergence measures between posterior distributions with and without a given subset of the data is a useful way of quantifying influence, and the most popular Bayesian case deletion influence diagnostics is based on the Kullback-Leibler divergence, which is a type of q-divergence measure ((GARAY et al., 2015)). Let $K(P, P_{(-i)})$ denote the KL divergence between two densities P and $P_{(-i)}$ for $\boldsymbol{\theta}$, which is defined as

$$K(P, P_{(-i)}) = \int \pi(\boldsymbol{\theta}|\boldsymbol{y}) log \left(\frac{\pi(\boldsymbol{\theta}|\boldsymbol{y})}{\pi(\boldsymbol{\theta}|\boldsymbol{y}_{(-i)})} \right) d\boldsymbol{\theta} \tag{1.27}$$

where P denotes the posterior distribution of $\boldsymbol{\theta}$ based on the full data $\boldsymbol{y}$, and $P_{(-i)}$ denotes the posterior distribution obtained from the data $\boldsymbol{y}$ without the i-th observation. The $K(P, P_{(-i)})$ measures the effect of deleting the i-th observation from the full data on the joint posterior distribution of $\boldsymbol{\theta}$. From expression (1.27) it is possible to rewrite the KL divergence as a posterior expectation, that is

$$K(P, P_{(-i)}) = E_{\boldsymbol{\theta}|\boldsymbol{y}} \left(log \left(\frac{\pi(\boldsymbol{\theta}|\boldsymbol{y})}{\pi(\boldsymbol{\theta}|\boldsymbol{y}_{(-i)})} \right) \right) \tag{1.28}$$

From (1.28), the computation of this measure can be approximated by using the MCMC posterior samples. Also, (CHO et al., 2009) developed a simplified expressions for computing the KL divergence, through the CPO statistic and the log-likelihood as

$$K(P, P_{(-i)}) = -log(\widehat{CPO}_i) + \frac{1}{M} \sum_{m=1}^{M} l(\boldsymbol{\theta}^{(m)}|y_i) \quad i = 1, \dots, n \tag{1.29}$$

where $\widehat{CPO}_i$ is expressed as in (1.26). As usual, we need to establish a cut-off point, in order to determine whether an observation is influent or not. As pointed by (CHO et al., 2009), the calibration of KL divergence can be done by solving for p_i the equation

$$K(P, P_{(-i)}) = K(Ber(1/2), Ber(p_i)) = -\frac{1}{2} log(4p_i(1 - p_i)) \tag{1.30}$$

where $Ber(p_i)$ is the Bernoulli distribution with success probability p_i. The equality $K(P, P_{(-i)}) = K(Ber(1/2), Ber(p_i))$ we have that describing outcomes using $\pi(\boldsymbol{\theta}|\boldsymbol{y})$ instead of $\pi(\boldsymbol{\theta}|\boldsymbol{y}_{(-i)})$ is compatible with describing an unobserved event as having probability p_i when correct probability is .5 ((CHO et al., 2009)). Solving 1.30, the calibration of the KL divergence is

$$p_i = .5[1 + \sqrt{1 - exp\{-2K(P, P_{(-i)})\}}]$$

This equation implies that $0.5 \leq p_i \leq 1$. For p_i much greater than .5 implies that the i-th observation is influential. In this work we are going to consider a influential observation if $p_i \geq .8$, as used in (CHAVES, 2015) and (GARAY et al., 2015). So, for KL divergence measure greater than $K(Ber(1/2), Ber(0.8)) \approx .2231436$, the observation is considered influential.

1.4.6 Residual analysis

Residual analysis is an important tool for model fit assessment, including detection of departing from model assumptions as well as the presence of outliers. For this section we consider a residual analysis based on Bayesian estimate for each unknown parameter:

$$R_i = \frac{Y_i - \boldsymbol{X}_i^t\hat{\boldsymbol{\beta}}}{\hat{\sigma}}, \quad i = 1, \ldots, n \tag{1.31}$$

We expected that the residuals in equation (1.31) approximately follows, under the good fit of the model, a $ST_c(0,1,\gamma,\nu)$, $SSL_c(0,1,\gamma,\nu)$, $SCN_c(0,1,\gamma,\nu_1,\nu_2)$ or $SGT_c(0,1,\gamma,\nu_1,\nu_2)$ distribution, according to the respective adopted distribution, with $\boldsymbol{\nu}$ and γ equal to the Bayesian estimates. For checking goodness of fit, we can build envelope plot, using the above mentioned distributions to simulate the envelopes, For outliers detection, we can graph the residual versus the index of the observations and the residuals against the fitted values.

1.5 Simulation study

We performed simulation studies in order to evaluate the performance of the model and the estimation method proposed in this work. All these models were implemented in JAGS ((PLUMMER, 2003)) through the interface provided by the rjags package ((PLUMMER, 2016)) available in R program ((R Development Core Team, 2008)). The codes are available from the authors upon request.

1.5.1 Parameter recovery

We considered different scenarios based on the crossing of the levels of some factors of interest. For the four SMSN distribution examples exposed in this work, we simulate from samples of size n=50, 250, 500 and 1000, varying values of parameter $\boldsymbol{\nu}$ and R=50 replicas were made. A sample of the regression model were simulated considering

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad i = 1, \ldots, n$$

where $\beta_0 = 1$, $\beta_1 = 2$, ε_i belongs to the SMSN family with $\sigma^2 = 1$ and $\gamma \in \{-0.9, 0, .9\}$, which allows the model to have strong negative, null and strong positive asymmetry, respectively. Also, we set $\nu \in \{3, 10, 30\}$ for the skew-t and skew-slash distribution; $\boldsymbol{\nu} = (\nu_1, \nu_2) = (0.1, .1), (0.9, .9), (0.9, .1)$ and $(0.5, .5)$ for the skew contaminated normal distribution and $\boldsymbol{\nu} = (\nu_1, \nu_2) = (15, 5), (5, 15)$ and $(30, 30)$ for the skew generalized-t distribution. These values for $\boldsymbol{\nu}$ were chosen in order to have distributions with heavy tails and tails close to the skew normal distribution. The covariates were simulated from N(0,1) and centered in their respective means. To eliminate the effect of the initial values and to avoid correlations problems, we run a MCMC chain of size 120,000 with a burn-in of 20,000 and thin 100, so we retain a valid MCMC chain of size 1000. The values of the Gelman-Rubin statistics and the analyses of traceplots, Geweke and autocorrelation plots indicated that the MCMC algorithm converged and the autocorrelation were negligible.

To compare the performance of the estimation methods we considered some appropriate statistics. Let ϑ be an element of $\boldsymbol{\theta} = (\boldsymbol{\beta}, \gamma, \sigma^2, \boldsymbol{\nu})$ and $\hat{\vartheta}_r$ the respective posterior estimate from the r-th replica. These statistics are: mean of the estimates of parameter ϑ (Est) $\bar{\hat{\vartheta}} = \frac{\sum_{r=1}^{R} \hat{\vartheta}_r}{R}$, variance of the estimates $Var_\vartheta = \frac{\sum_{r=1}^{R}(\hat{\vartheta}_r - \bar{\hat{\vartheta}})^2}{R-1}$, bias of the estimates $\bar{\hat{\vartheta}} - \vartheta$, relative bias $\frac{|Bias_\vartheta|}{\vartheta}$, square root of the mean square error $RMSE_\vartheta = \sqrt{Bias_\vartheta^2 + Var_\vartheta}$, the length of the credibility interval and the Coverage Ratio of the 95% credibility interval of parameter ϑ. To obtain it, we have calculated the numbers of intervals which contain the true value of the parameter, and then divide it by the total of replicas.

Posterior mean, median and mode were calculated for each parameter ϑ, in each replica. For skew-t, skew slash and skew generalized t distributions, we considered the posterior mean as punctual estimate for ν whereas for the skew contaminated normal model we considered the mode, since it presented to be closer to the true values of ν_1 and ν_2, when compared with the others.

We adopted weakly informative priors for all parameters, that is: $\Delta \sim N(0, 100)$, $\tau \sim \text{gamma}(0.01, .01)$, $\beta_0 \sim N(0, 1000)$ and $\beta_1 \sim N(0, 1000)$. For the skew-t model we set $\nu \sim exp(\boldsymbol{\theta})T(2,)$ and $\boldsymbol{\theta} \sim unif(0.02, .49)$; for the skew contaminated model we used $\nu_1 \sim beta(1, 1)$ and $\nu_2 \sim beta(1, 1)$ and for the skew generalized t model we set $\nu_1 \sim exp(\boldsymbol{\theta}_1)T(2,)$ and $\boldsymbol{\theta}_1 \sim unif(0.02, .49)$ and $\nu_2 \sim exp(\boldsymbol{\theta}_2)$ and $\boldsymbol{\theta}_2 \sim unif(0.02, .49)$.

For the skew slash model, we performed prior sensitivity study, considering several priors suggested in preview works, as well as others that we considered as possible useful, which are $\nu \sim \text{gamma}(1, .1)$, $\nu \sim \text{gamma}(1.5, .05)$, $\nu \sim \text{gamma}(\boldsymbol{\theta}, \boldsymbol{\theta})$ and $\boldsymbol{\theta} \sim unif(0.002, .2)$, $\nu \sim \text{gamma}(\boldsymbol{\theta}, \boldsymbol{\theta})$ and $\boldsymbol{\theta} \sim unif(0.002, .05)$, $\nu \sim \text{gamma}(\boldsymbol{\theta}, \boldsymbol{\theta})$ and $\boldsymbol{\theta} \sim unif(0.02, .5)$, $\nu \sim exp(\boldsymbol{\theta})$ and $\boldsymbol{\theta} \sim unif(0.02, .5)$, $\nu \sim \text{gamma}(1, .005)$, the independent Jeffrey's prior, and $\nu \sim BSSNC(0.5, 10)$, $\nu \sim BSSNC(1, 10)$, $\nu \sim BSSNC(1, 20)$, as in (CHAVES, 2015) and they correspond to the Birnbaum–Saunders Skew Normal distribution under the centered parameterization. All these priors are truncated at $(1, \infty)$, in order to have finite variance. The results revealed that for $\nu = 3$, the estimates under all priors were quite close to the true value. For $\nu = 10$, depending on the prior choice, estimates were either close to 10 or 30. The same was observed for $\nu = 30$. So, we decided to implement the skew model under two priors: $\nu \sim \text{gamma}(1, .1)$, in which the estimates are centered in 10, with relative small variance, and $\nu \sim \text{gamma}(1.5, .05)$, which is centered around 30 with large variance, and give us estimates around 30. After the implementation we choose the best model using the EAIC, EBIC and DIC criterion. In order to facilitate the reading of the results, we present here only some of the simulated scenarios. The other tables are arranged in Appendix B.

The results for the skew-t model are showed in Tables 1, and the remaining tables are presented in Appendix B. Under $\nu = 3$, for all sample sizes, the estimates of all parameters were very close to the true values, and the length of the credibility intervals decreases as the sample size increases. For γ, from the simulations with $\gamma = -0.9, 0$ and .9, only for n=50 the estimate were not so accurate. On the other hand, under $\nu = 10$ and $\nu = 30$, the estimates for ν became more accurate as the sample size increased. Also, the estimates for $\boldsymbol{\beta}, \gamma$ and σ^2 were close to true values.

For the skew slash model, as mentioned before, we used the EAIC, EBIC and DIC criteria to select between the adjust using the priors $\nu \sim$ gamma$(1, .1)$ and $\nu \sim$ gamma$(1.5, .05)$ that we shall denoted by models 1 and 2, respectively. In general, when $\nu = 10$, model 1 was chosen almost all the time, occurring the opposite when $\nu = 30$. As can be seen from tables 2, 3 and the others arranged in Appendix B, only for sample sizes higher than 500, we have more accurate estimates for parameter ν, when the true value is either 10 or 30. For $\nu = 3$, the estimates are close to the real value, for $n > 250$. For all other parameters, the results were quite good, regardless the scenario. In general, the variance, bias, relative bias, RMSE and the length of the credibility intervals become smaller as the sample size increases.

For the skew contaminated normal distribution, in general, as the sample size increases the estimates become less unbiased and more precise. We can notice that, for some parameters ($\beta_0, \beta_1, \sigma^2$ and ν_1), the credibility intervals do not contain the true values, which can be explained by their short length, however, the true value is close to the upper or lower bound. When the true values of ν_1 and ν_2 are .9 and .1, respectively, the estimates of σ^2 were not so good, specially for small sample sizes. However, as the sample size increases, the estimate become closer to the real value. In order to confirm that, we ran another scenario, with sample size of 2,000, whose results are shown in Table 4. In this case, we can see that the estimate of σ^2 is very close to the true value.

For the skew generalized t distribution, when the true values of $\boldsymbol{\nu} = (\nu_1, \nu_2)$ are (15,5) and (5,15), the estimates of all parameters become more precise, as the sample size increases. Also the length of the credibility intervals decreases. When $\boldsymbol{\nu} = (30, 30)$, although the length of the credibility interval for ν_1 and ν_2 increased, the estimates became more precise. Comparing all the results obtained in the simulations, it is observed that the estimates, in general, are more accurate as the sample size increases. For small sample sizes (< 250), estimates of $\boldsymbol{\beta}$, σ^2 and γ approach to the true values, while that for ν it is close only when the true model has heavier tails than the skew normal model.

Table 1 – Results of the simulation study for the skew-t model with $\nu = 3$ and $\gamma = -0.9$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
	β_0	1.0000	1.1741	.0001	.1741	.1741	.1741	1.0000	.7704
	β_1	2.0000	1.9855	.0001	-0.0145	.0073	.0145	1.0000	.9597
50	γ	-0.9000	-0.6873	.0004	.2127	.2363	.2127	1.0000	1.4047
	σ^2	1.0000	.9382	.0002	-0.0618	.0618	.0618	1.0000	1.3460
	ν	3.0000	2.9472	.0017	-0.0528	.0176	.0528	1.0000	3.3977
	β_0	1.0000	1.0463	$<$ .0001	.0463	.0463	.0463	1.0000	.3856
	β_1	2.0000	2.0490	$<$ .0001	.0490	.0245	.0490	1.0000	.3060
250	γ	-0.9000	-0.8340	.0003	.0660	.0733	.0660	1.0000	.7583
	σ^2	1.0000	1.3646	.0001	.3646	.3646	.3646	.8400	.8354
	ν	3.0000	3.4223	.0008	.4223	.1408	.4223	1.0000	2.6695
	β_0	1.0000	.9407	$<$ .0001	-0.0593	.0593	.0593	1.0000	.2373
	β_1	2.0000	2.0571	$<$ .0001	.0571	.0286	.0571	1.0000	.1925
500	γ	-0.9000	-0.8588	.0001	.0412	.0457	.0412	1.0000	.4304
	σ^2	1.0000	.9649	$<$ .0001	-0.0351	.0351	.0351	1.0000	.4397
	ν	3.0000	2.9570	.0004	-0.0430	.0143	.0430	1.0000	1.5980
	β_0	1.0000	1.0240	$<$ .0001	.0240	.0240	.0240	1.0000	.1671
	β_1	2.0000	2.0616	$<$ .0001	.0616	.0308	.0616	.9800	.1377
1000	γ	-0.9000	-0.9531	$<$ .0001	-0.0531	.0590	.0531	1.0000	.1311
	σ^2	1.0000	1.0448	$<$ .0001	.0448	.0448	.0448	1.0000	.3402
	ν	3.0000	3.2101	.0002	.2101	.0700	.2101	1.0000	1.2472

Table 2 – Results of the simulation study for the skew slash model with $\nu = 10$ and $\gamma = 0.9$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
	β_0	1.0000	1.0680	.0001	.0680	.0680	.0680	1.0000	.7648
	β_1	2.0000	2.1067	$<$ .0001	.1067	.0533	.1067	1.0000	.5885
50	γ	.9000	.9051	$<$ .0001	.0051	.0057	.0051	1.0000	.3290
	σ^2	1.0000	1.5340	.0004	.5340	.5340	.5340	1.0000	1.7415
	ν	10.0000	11.2042	.1552	1.2042	.1204	1.2141	1.0000	35.8608
	β_0	1.0000	1.0802	$<$ .0001	.0802	.0802	.0802	1.0000	.2681
	β_1	2.0000	1.9643	$<$ .0001	-0.0357	.0178	.0357	1.0000	.1810
250	γ	.9000	.9417	$<$ .0001	.0417	.0464	.0417	1.0000	.1377
	σ^2	1.0000	1.0737	.0002	.0737	.0737	.0737	1.0000	.5336
	ν	10.0000	10.7952	6.3059	.7952	.0795	6.3559	1.0000	31.9375
	β_0	1.0000	1.0055	$<$ .0001	.0055	.0055	.0055	1.0000	.1764
	β_1	2.0000	1.9622	$<$ .0001	-0.0378	.0189	.0378	1.0000	.1234
500	γ	.9000	.9200	$<$ .0001	.0200	.0223	.0200	1.0000	.1390
	σ^2	1.0000	.8875	.0001	-0.1125	.1125	.1125	1.0000	.3980
	ν	10.0000	9.5510	1.0960	-0.4490	.0449	1.1844	1.0000	29.8340
	β_0	1.0000	1.0901	$<$ .0001	.0901	.0901	.0901	.0000	.1321
	β_1	2.0000	1.9749	$<$ .0001	-0.0251	.0126	.0251	1.0000	.0928
1000	γ	.9000	.9015	$<$ .0001	.0015	.0017	.0015	1.0000	.1238
	σ^2	1.0000	1.0272	.0007	.0272	.0272	.0272	1.0000	.3589
	ν	10.0000	12.0765	24.4792	2.0765	.2076	24.5671	1.0000	38.4406

Table 3 – Results of the simulation study for the skew slash model with $\nu = 30$ and $\gamma = 0.9$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
	β_0	1.0000	.9926	$<.0001$	-0.0074	.0074	.0074	1.0000	.5396
	β_1	2.0000	1.7755	$<.0001$	-0.2245	.1123	.2245	.0000	.3893
50	γ	.9000	.9102	.0001	.0102	.0113	.0102	1.0000	.8599
	σ^2	1.0000	.9363	.0001	-0.0637	.0637	.0637	1.0000	.8418
	ν	30.0000	34.5909	.9577	4.5909	.1530	4.6897	1.0000	91.6302
	β_0	1.0000	.8902	$<.0001$	-0.1098	.1098	.1098	1.0000	.2406
	β_1	2.0000	2.0422	$<.0001$	.0422	.0211	.0422	1.0000	.1572
250	γ	.9000	.9047	$<.0001$	.0047	.0052	.0047	1.0000	.1789
	σ^2	1.0000	.9085	.0007	-0.0915	.0915	.0915	1.0000	.4298
	ν	30.0000	21.6728	86.1437	-8.3272	.2776	86.5453	1.0000	59.5103
	β_0	1.0000	1.0504	$<.0001$	.0504	.0504	.0504	1.0000	.1768
	β_1	2.0000	1.9264	$<.0001$	-0.0736	.0368	.0736	.0000	.1253
500	γ	.9000	.8958	$<.0001$	-0.0042	.0047	.0042	1.0000	.1272
	σ^2	1.0000	1.0184	$<.0001$	.0184	.0184	.0184	1.0000	.3229
	ν	30.0000	33.7918	8.1521	3.7918	.1264	8.9908	1.0000	89.6331
	β_0	1.0000	.9752	$<.0001$	-0.0248	.0248	.0248	1.0000	.1223
	β_1	2.0000	1.9851	$<.0001$	-0.0149	.0074	.0149	1.0000	.0866
1000	γ	.9000	.8960	$<.0001$	-0.0040	.0045	.0040	1.0000	.0998
	σ^2	1.0000	.9503	.0002	-0.0497	.0497	.0497	1.0000	.2450
	ν	30.0000	29.8191	58.0334	-0.1809	.0060	58.0337	1.0000	77.7648

Table 4 – Results of the simulation study for the skew contaminated normal model with $\nu_1 = 0.9$ and $\nu_2 = 0.1$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	-0.0791	.0008	-1.0791	1.0791	1.0791	.0000	2.1279
	β_1	2.0000	1.3118	.0007	-0.6882	.3441	.6882	1.0000	2.2450
	γ	-0.9000	-0.9850	$<$ .0001	-0.0850	.0944	.0850	1.0000	.3305
	σ^2	1.0000	11.7727	.0932	10.7727	10.7727	10.7731	.0000	17.1898
	ν_1	0.9000	.6451	.0002	-0.2549	.2833	.2549	1.0000	.9542
	ν_2	0.1000	.7579	.0001	.6579	6.5789	.6579	1.0000	.8119
250	β_0	1.0000	1.2141	$<$ .0001	.2141	.2141	.2141	1.0000	.6752
	β_1	2.0000	2.2152	$<$ .0001	.2152	.1076	.2152	1.0000	.5672
	γ	-0.9000	-0.8087	$<$ .0001	.0913	.1014	.0913	1.0000	.3758
	σ^2	1.0000	6.8983	.0383	5.8983	5.8983	5.8985	.2000	6.2386
	ν_1	0.9000	.2000	.0008	-0.7000	.7778	.7000	1.0000	.9471
	ν_2	0.1000	.5763	.0003	.4763	4.7632	.4763	1.0000	.7788
500	β_0	1.0000	.9943	$<$ .0001	-0.0057	.0057	.0057	1.0000	.5606
	β_1	2.0000	1.9129	$<$ .0001	-0.0871	.0435	.0871	1.0000	.4449
	γ	-0.9000	-0.8681	$<$ .0001	.0319	.0354	.0319	1.0000	.2347
	σ^2	1.0000	8.3673	.1441	7.3673	7.3673	7.3687	.1000	9.7310
	ν_1	0.9000	.8311	.0008	-0.0689	.0765	.0689	1.0000	.9263
	ν_2	0.1000	.5439	.0010	.4439	4.4389	.4439	1.0000	.8644
1000	β_0	1.0000	1.2459	$<$ .0001	.2459	.2459	.2459	.0000	.3412
	β_1	2.0000	2.0330	$<$ .0001	.0330	.0165	.0330	1.0000	.2544
	γ	-0.9000	-0.8962	$<$ .0001	.0038	.0043	.0038	1.0000	.1391
	σ^2	1.0000	2.2930	1.1360	1.2930	1.2930	1.7212	.9000	6.3600
	ν_1	0.9000	.8848	.0042	-0.0152	.0169	.0158	1.0000	.7238
	ν_2	0.1000	.2640	.0140	.1640	1.6402	.1646	1.0000	.7394
2000	β_0	1.0000	1.0612	.0005	.0612	.0612	.0612	1.0000	.2423
	β_1	2.0000	1.9588	$<$ .0001	-0.0412	.0206	.0412	1.0000	.1879
	γ	-0.9000	-0.8891	$<$ .0001	.0109	.0122	.0109	1.0000	.0761
	σ^2	1.0000	1.1679	.4704	.1679	.1679	.4995	1.0000	1.7399
	ν_1	0.9000	.8552	.0020	-0.0448	.0497	.0448	.9800	.1611
	ν_2	0.1000	.1190	.0049	.0190	.1897	.0196	.9800	.1848

Table 5 – Results of the simulation study for the skew generalized t model with $\nu_1 = 15$ and $\nu_2 = 5$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	1.0552	$<$ .0001	.0552	.0552	.0552	1.0000	.3375
	β_1	2.0000	1.9662	$<$ .0001	-0.0338	.0169	.0338	1.0000	.2808
	γ	.0000	-0.0989	.0001	-0.0989	Inf	.0989	1.0000	1.1529
	ν_1	15.0000	13.2828	.5377	-1.7172	.1145	1.7994	1.0000	31.1524
	ν_2	5.0000	4.6102	.0821	-0.3898	.0780	.3983	1.0000	11.4945
250	β_0	1.0000	.9788	$<$ .0001	-0.0212	.0212	.0212	1.0000	.1521
	β_1	2.0000	1.9288	$<$ .0001	-0.0712	.0356	.0712	.2200	.1372
	γ	.0000	-0.0372	$<$ .0001	-0.0372	Inf	.0372	1.0000	.6340
	ν_1	15.0000	13.0119	1.8672	-1.9881	.1325	2.7274	1.0000	26.2708
	ν_2	5.0000	4.4867	.2864	-0.5133	.1027	.5878	1.0000	10.2284
500	β_0	1.0000	1.0208	$<$ .0001	.0208	.0208	.0208	1.0000	.1058
	β_1	2.0000	1.9531	$<$ .0001	-0.0469	.0235	.0469	.9800	.1029
	γ	.0000	-0.0023	$<$ .0001	-0.0023	Inf	.0023	1.0000	.4319
	ν_1	15.0000	14.2629	3.8366	-0.7371	.0491	3.9067	1.0000	26.8264
	ν_2	5.0000	4.7916	.5539	-0.2084	.0417	.5918	1.0000	10.3291
1000	β_0	1.0000	.9912	$<$ .0001	-0.0088	.0088	.0088	1.0000	.0712
	β_1	2.0000	1.9719	$<$ .0001	-0.0281	.0141	.0281	1.0000	.0691
	γ	.0000	.0783	.0001	.0783	Inf	.0783	1.0000	.4237
	ν_1	15.0000	15.9252	3.9802	.9252	.0617	4.0863	1.0000	23.7698
	ν_2	5.0000	5.0506	.5453	.0506	.0101	.5476	1.0000	8.9080

1.5.2 Model selection

To analyze the performance of the selection criteria presented in Section 1.4.3, we have conducted a simulation study considering different scenarios based on the crossing of the levels of some factors of interest. For the four SMSN distribution examples exposed in this work, we simulate from samples of size n=50, 250, 500 and 1000. We set $\boldsymbol{\beta} = (1, 2)$, $\sigma^2 = 1$, $\gamma = -0.9$ and $\nu = 3$ for the skew-t and skew slash model, $\boldsymbol{\nu} = (0.1, .15)$ for the skew contaminated model and $\boldsymbol{\nu} = (3, 1)$ for the skew generalized t model. In this part, our goal is to check if as long as sample size increases, the selection criteria are able to select the correct model. We generated samples for each of the above distributions and fitted for each sample all four models (SSL_c, ST_c, SCN_c and SGT_c), then we calculated the EBIC, EAIC, DIC and LPML criteria. The adopted priors were the same as described previously in Section 1.5.1. The number of times that the skew slash model was selected is the sum of times that the slash 1 model (SSL1) or slash 2 model (SSL2) was chosen by the selection criteria.

In general, the true underlying model was selected in all almost of the replicas by EBIC, EAIC, DIC and LPML criteria for all models, when the sample size is higher than 500. For the skew-t model, when n=1,000, the percentual that the true model was chosen was smaller than for n = 500. This was, probably, due to the fact that the estimates of ν_1 and ν_2 were very close to each other and to $\nu = 3$. Since we set $\sigma^2 = 1$ for the skew-t model, in this case the skew-t model is a special case of the skew generalized t model with $\nu_1 = \nu_2 = 3$. As shown in Table 60, all criteria, for the skew-t and skew generalized t models, are practically the same, for n=1,000. For the skew slash model, under n=1,000, the correct model was chosen in almost all replicas by EAIC, EBIC and LPML criteria, although the DIC criterion has not chosen the model correctly. Finally, for the skew generalized t model, the skew-t model was preferred than the correct model for almost all sample sizes, except for n=1,000. This was, probably, due to the fact that the estimates of ν_1 and ν_2 were not so accurate as for n=1,000.

Table 6 – Percentage of times that the correct model was selected, in a total of 50 replicas

		criteria			
model	n	EAIC	EBIC	DIC	LPML
ST_c	50	100	98	14	14
	250	0	0	0	2
	500	100	100	100	100
	1000	52	52	62	62
SSL_c	50	0	0	26	0
	250	2	2	34	0
	500	100	100	98	80
	1000	100	100	28	98
SCN_c	50	0	0	8	84
	250	0	0	8	0
	500	100	100	100	100
	1000	100	100	100	100
SGT_c	50	0	0	0	0
	250	0	0	0	0
	500	0	0	0	0
	1000	100	100	98	96

1.5.3 Residual analysis

In this section we analyzed the behavior of the residuals, presented in Section 1.4.6, under some conditions of interest. We have conducted a simulation study, considering a sample size of 1000. We simulated data sets, for each regression model, considering $\beta_0 = 1$, β_1, $\sigma^2 = 1$ and $\gamma = -0.9$. Also, we considered: $\nu = 3$ for skew-t and skew slash distributions; $\boldsymbol{\nu} = (3, 1)$ for skew generalized t and $\boldsymbol{\nu} = (0.1, .15)$ for skew contaminated normal distribution.

For each simulated data, we fitted the skew normal, skew-t, skew slash, skew generalized t and skew contaminated normal models using the priors described in Section 1.5.1. We built suitable quantile-quantile plots for all models, where the confidence bands were made considering the underlying distribution, with $\mu = 0$ and $\sigma^2 = 1$.

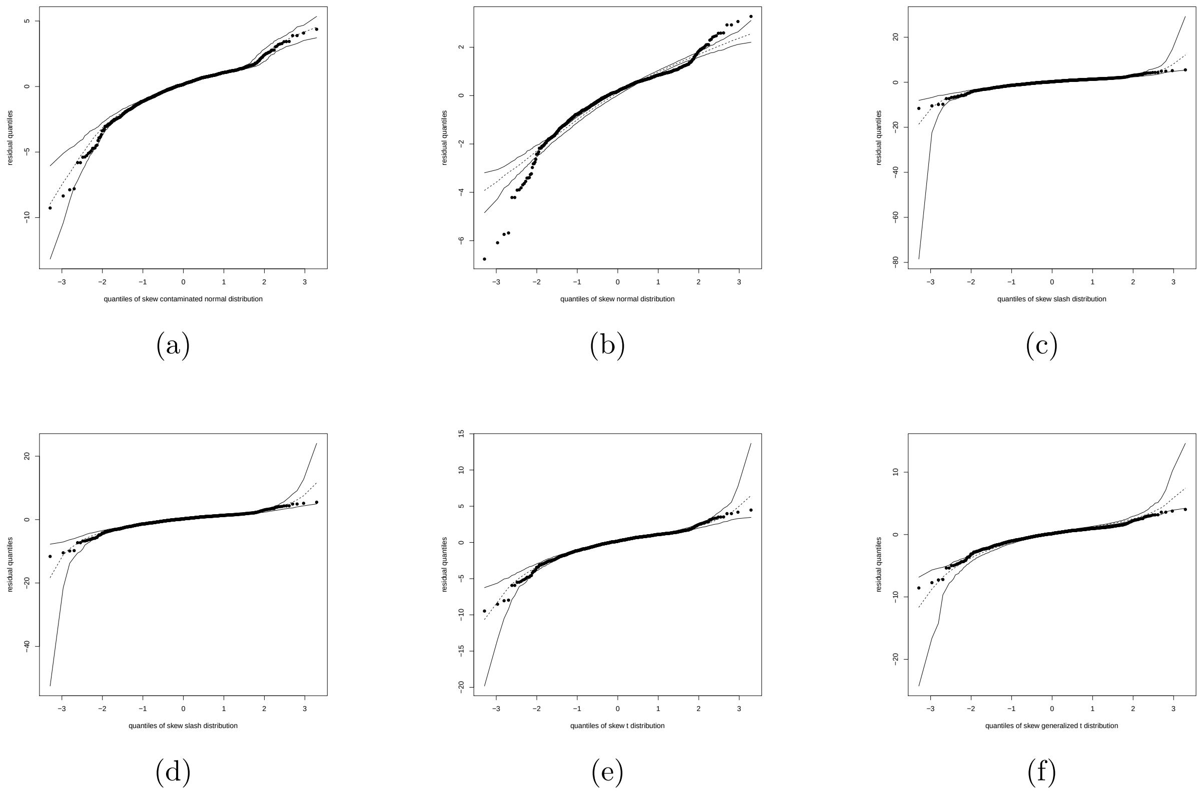
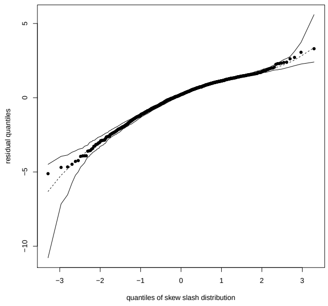


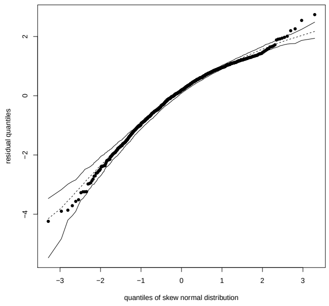
Figure 15 – QQ plots using the data set generated by the skew contaminated normal distribution and adjusting by: skew contaminated normal (a), skew normal (b), skew slash 1 (c), skew slash 2 (d), skew-t (e) and skew generalized t (f)

For all simulated data, we can see from Figures 15 - 18 that when data exhibits heavy tails, residuals obtained from the skew normal fit indicated that this model did not fit well to the data, since the residuals lying outside the confidence bands. For all models, when we adjusted the correct model to the data, there is no residual lying outside the confidence bands, indicating that the model is well adjusted.

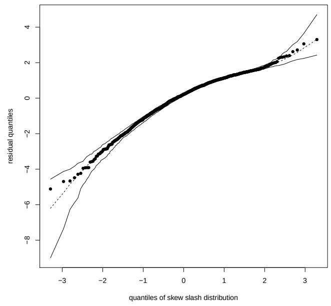
Comparing the adjust by another members of SMSN class, we noted that when observations are generated by skew slash and skew contaminated normal distributions and we adjusted with the skew generalized t, the fit using this distribution is not as good as the rest. However, when the data is generated by the skew-t or the skew generalized t model, the residuals obtained from the skew-t and skew generalized t models are well adjusted.



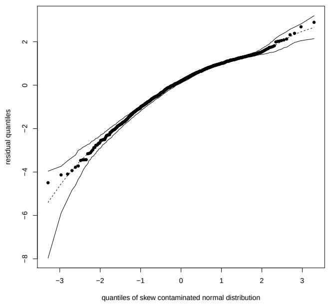
(a)



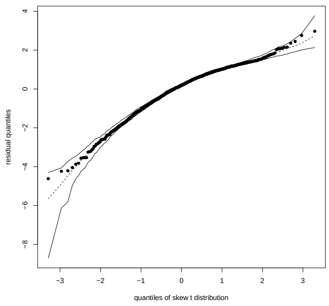
(b)



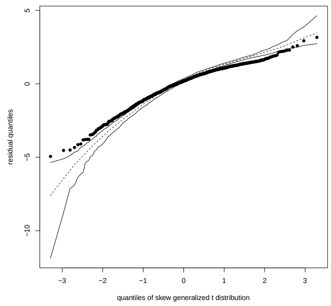
(c)



(d)

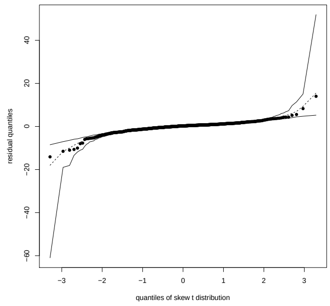


(e)

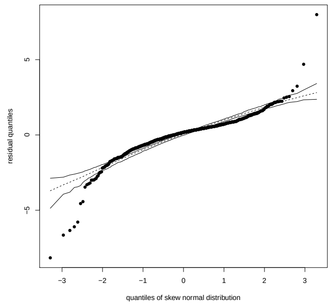


(f)

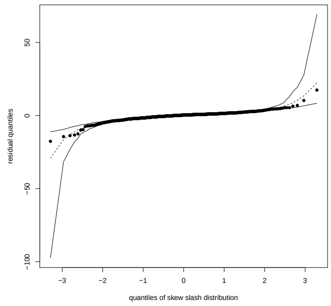
Figure 16 – QQ plots using the data set generated by the skew slash distribution and adjusting by: skew slash 1 (a), skew normal (b), skew slash 2 (c), skew contaminated normal (d), skew-t (e) and skew generalized t (f)



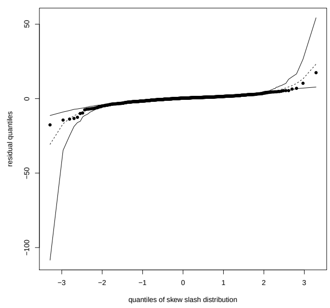
(a)



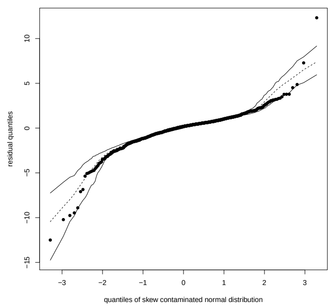
(b)



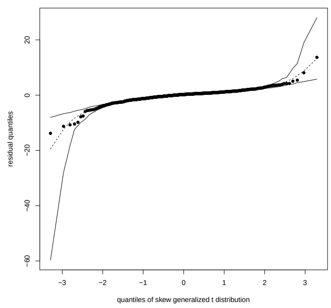
(c)



(d)



(e)



(f)

Figure 17 – QQ plots using the data set generated by the skew-t distribution and adjusting by: skew-t (a), skew normal (b), skew slash 1 (c), skew slash 2 (d), skew contaminated normal (e) and skew generalized t (f)

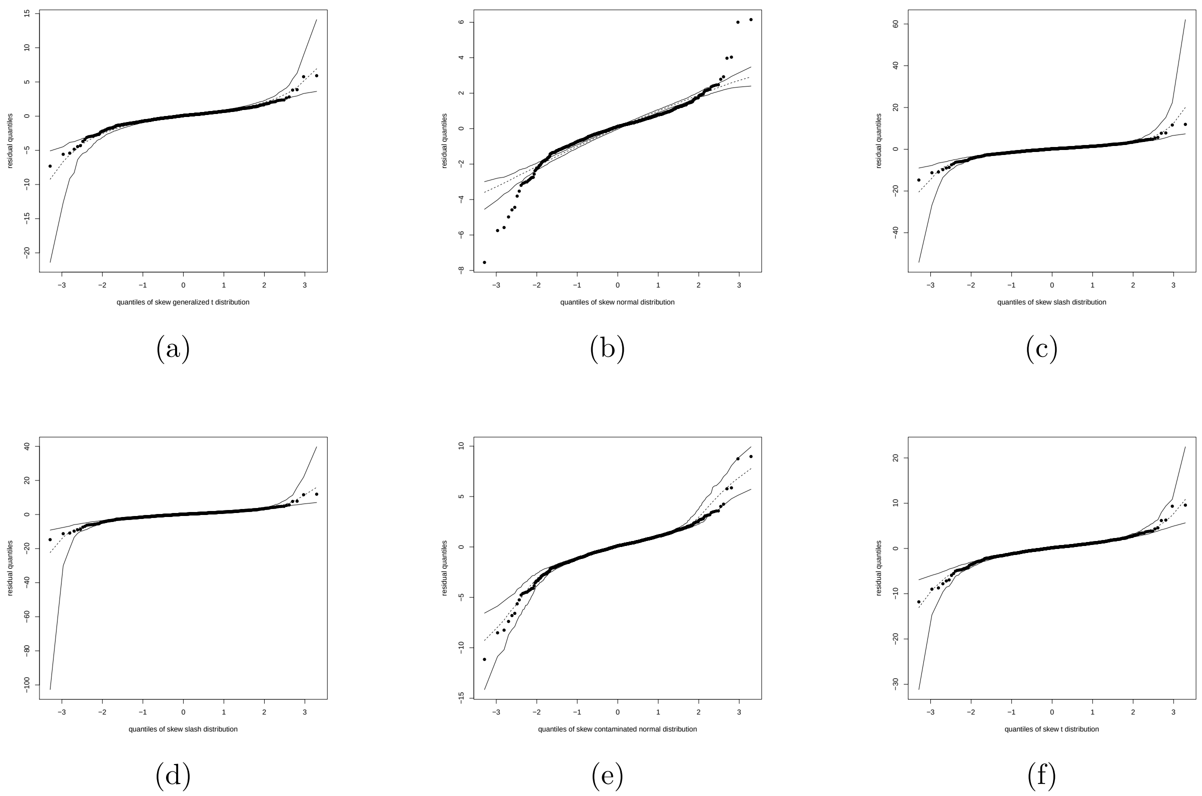


Figure 18 – QQ plots using the data set generated by the skew generalized t distribution and adjusting by: skew generalized t (a), skew normal (b), skew slash 1 (c), skew slash 2 (d), skew contaminated normal (e) and skew-t (f)

1.5.4 Influence analysis

To analyze the behavior of the influence diagnostic analysis technique presented in this work using the K-L divergence measure, We have conducted a simulation study, considering a sample size of 1000. We simulated data sets, for each regression model, considering $\beta_0 = 1$, β_1, $\sigma^2 = 1$ and $\gamma = -0.9$. Also, we considered: $\nu = 3$ for skew-t and skew slash distributions; $\boldsymbol{\nu} = (3, 1)$ for skew generalized t and $\boldsymbol{\nu} = (0.1, .15)$ for skew contaminated normal distribution.

For each simulated data, we fitted the skew normal, skew-t, skew slash, skew generalized t and skew contaminated normal models using the priors described in Section 1.5.1. For each fitted model, plots with K-L divergence are made. As described in section 1.4.5, an influential observation is considered if $p_i \geq .8$, that is, if $K(P, P_{(-i)}) \geq .2231436$.

In all cases, we can see from figures 19 - 22 that the skew normal model indicates possible influential observations that are no influential, for the other models, indicating that our models accommodate, properly, all observations, differently from the normal model. As we can see, there is one observation that is higher than all other in the skew-t model, but is still smaller than the cut-off point. However, when the skew normal model is adjusted, that same observation appears as influential. It can also be noted that when the data are simulated from skew-t, skew slash and skew generalized t model and we fitted the skew contaminated model, at least one observation is considered as potentially influential. This does not happen when the data is simulated from the skew contaminated normal distribution. This indicates that the skew contaminated model does not accommodate so well the extreme observations, compared with the other models.

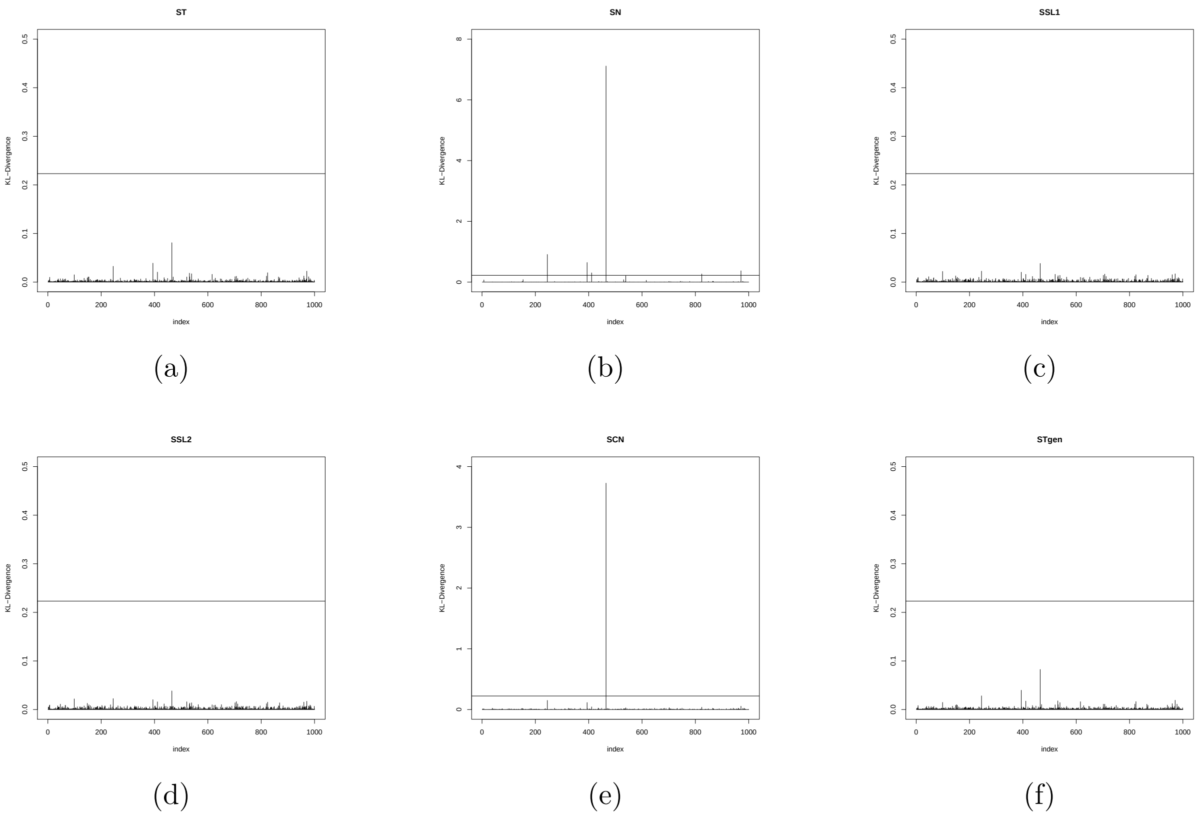


Figure 19 – Index plots of $K(P, P_{(i)})$ for the data set generated by the skew-t distribution adjusting by: skew-t (a), skew normal (b), skew slash 1 (c), skew slash 2 (d), skew contaminated normal (e) and skew generalized t (f)

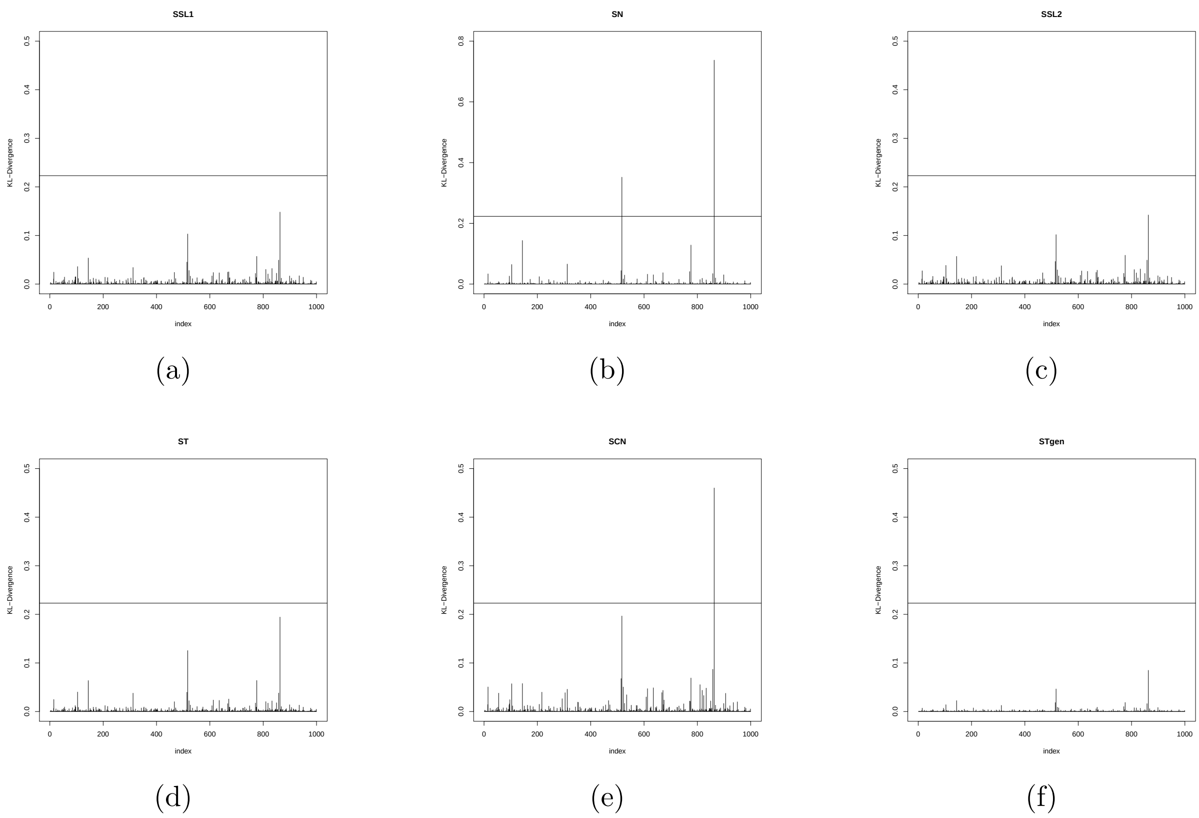


Figure 20 – Index plots of $K(P, P_{(i)})$ for the data set generated by the skew slash distribution adjusting by: skew slash 1 (a), skew normal (b), skew slash 2 (a), skew-t (d), skew contaminated normal (e) and skew generalized t (f)

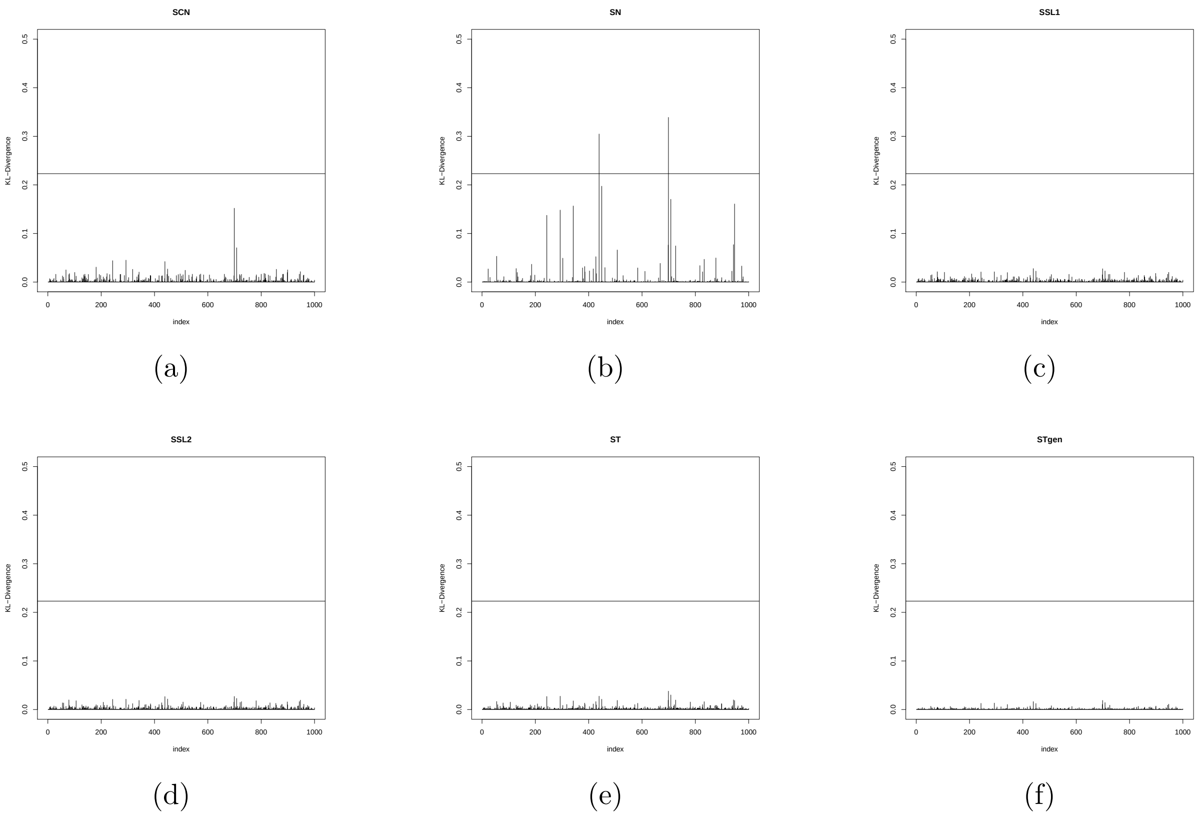


Figure 21 – Index plots of $K(P, P_{(i)})$ for the data set generated by the skew contaminated normal distribution adjusting by: skew-t (a), skew normal (b), skew slash 1 (c), skew slash 2 (d), skew-t (e) and skew generalized t (f)

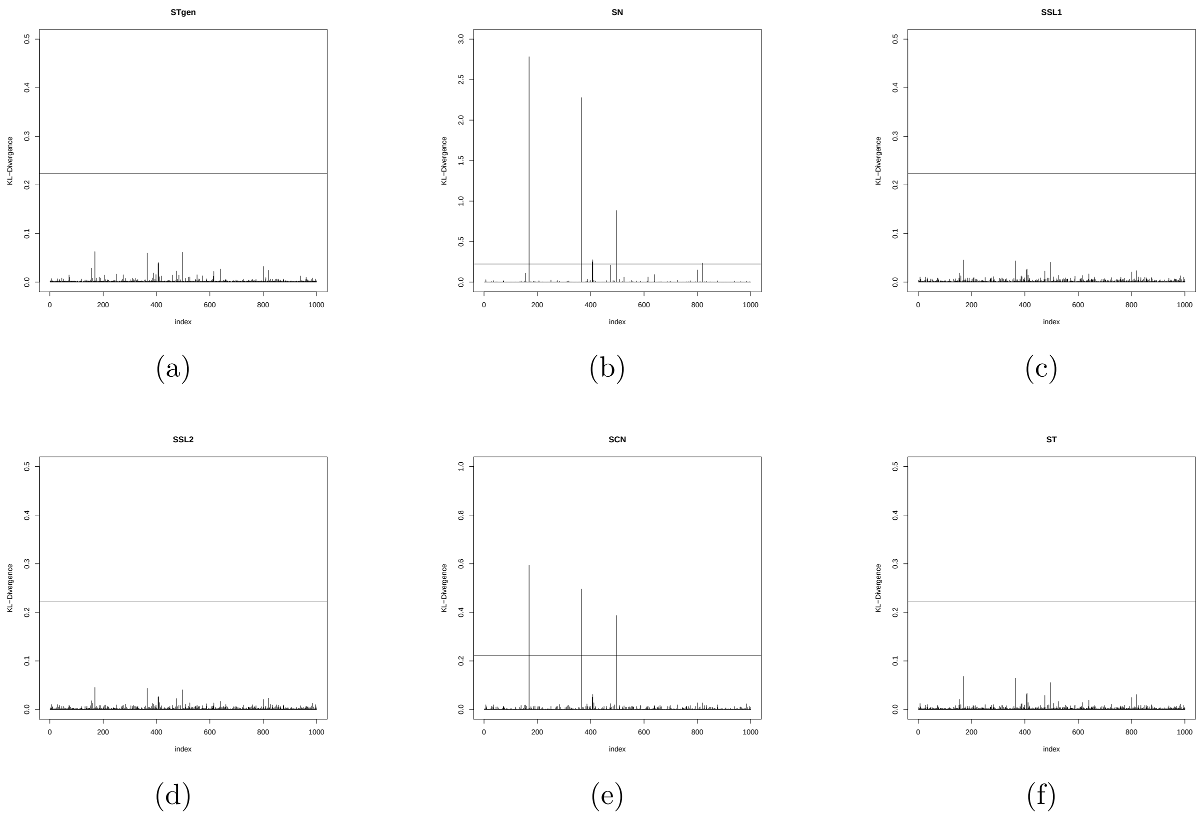


Figure 22 – Index plots of $K(P, P_{(i)})$ for the data set generated by the skew generalized t distribution adjusting by: skew-t (a), skew normal (b), skew slash 1 (c), skew slash 2 (d), skew contaminated normal (e) and skew-t (f)

1.6 Application

The application is made on Australian Athletes dataset described in (WEISBERG, 2005). This data consist on sample of 202 elite athletes who were in training at the Australian Institute of Sport. We consider the lean body mass (LBM) as our response and the height in cm (Ht), weight in kg (Wt), and the sex (0 = male and 1 = female) as our covariates for the regression model. We consider a regression model of the form $Y_i = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon_i$, for $i = 1, 2, \ldots, 202$, where $x_1 = 0$ if male and 1 if female, x_2 and x_3 are, respectively, the covariates Ht and Wt centered in their respective mean. We fitted six models, assuming that: $\varepsilon \overset{iid}{\sim} ST_c(0, \sigma^2, \gamma, \nu)$, or , $\varepsilon \overset{iid}{\sim} SCN_c(0, \sigma^2, \gamma, \nu_1, \nu_2)$, or $\varepsilon \overset{iid}{\sim} SGT_c(0, 1, \gamma, \nu_1, \nu_2)$, or $\varepsilon \overset{iid}{\sim} SN_c(0, \sigma^2, \gamma)$, or $\varepsilon \overset{iid}{\sim} SSL_c(0, \sigma^2, \gamma, \nu)$ assuming the prior for ν, $\nu \sim \text{gamma}(1, .1)T(1, \infty)$ and $\varepsilon \overset{iid}{\sim} SSL_c(0, \sigma^2, \gamma, \nu)$ assuming the prior for ν, $\nu \sim \text{gamma}(1.5, .05)T(1, \infty)$, that we denote, respectively, by ST, SCN, SGT, SN, SSL1 and SSL2. The values for the MCMC algorithm were the same used in the simulation study. For the slash distribution, we fitted two models, with different priors for ν, as described in Section 1.5.1. The priors for all other parameters were chosen according to the simulation study available in Section 1.5.1. Table 7 presents the statistics for model comparison. The skew t model (ST) was selected by EAIC, EBIC and LPML. The slash model using as prior for ν, $\nu \sim gamma(1.5, .05)T(1,)$, was select by DIC criterion. For the SCN model, it is not clear if the skew normal model is preferable to the skew contaminated model, since the width of the credibility intervals are large. Analyzing the posterior distribution of ν_1, presented in Figure 23, we can noted that it is concentrated toward .2. For ν_2, the histogram shows it is unlikely ν_2 is large. In fact, the probability of $\nu_1 < 0.05$ is .075 and $\nu_2 > 0.9$ is equal to .008, which indicates that skew contaminated normal model is preferable to skew normal. For skew-t and skew slash (SSL1) models, the credibility intervals do not include values of $\nu > 30$. From the posterior distribution of ν_1 for the skew generalized t model, we have the indicative that this model is preferred to the skew normal model. For the skew slash model (SSL2) we noted the credibility interval for ν indicated that the skew normal model may be preferred to the skew slash model. In this way, we have calculated from the posterior distributions of ν, the probability of $\nu > 30$, which was equal to .13. This indicate that the skew slash model is also preferable to the skew normal model. From Figure 24, QQ plot with envelopes for all fitted models are shown. It is possible to see that for skew slash, skew contaminated normal and skew normal models, there are some points lying outside the confidence bands, which do not happen with skew-t and skew generalized t models. The analysis of influential observations, presented in Figure 25, indicated that there are two influential observations for the skew-t model, three for the skew slash, skew contaminated normal and skew normal models and only one for the skew generalized t model. Under this criterion, we observe that both skew-t and skew generalized t models outperform the others models. The residual analysis also indicated that skew-t and skew generalized t models provide the best fit. Since the EAIC, EBIC and LPML criteria selected the skew-t as the best model, and the residual and influential observations analyses indicated that the skew-t model fit the data very well, this model was preferred to the others.

From Table 8, it is possible to see that all regression parameters are different from zero. Also, the model indicated that the height and weight have a positive influence in the lean body mass, and females presented lean body mass inferior to males.

Table 7 – **AIS dataset**: Statistics for model comparison

	Model					
criterion	ST	SSL1	SSL2	SCN	SGT	SN
EAIC	**976.83**	977.54	979.07	979.40	982.85	982.27
EBIC	**999.99**	1000.69	1002.23	1005.87	1006.00	1002.12
DIC	970.06	971.02	**956.54**	970.09	975.47	976.33
LPML	**-485.60**	-486.49	-488.14	-486.90	-488.38	-489.94

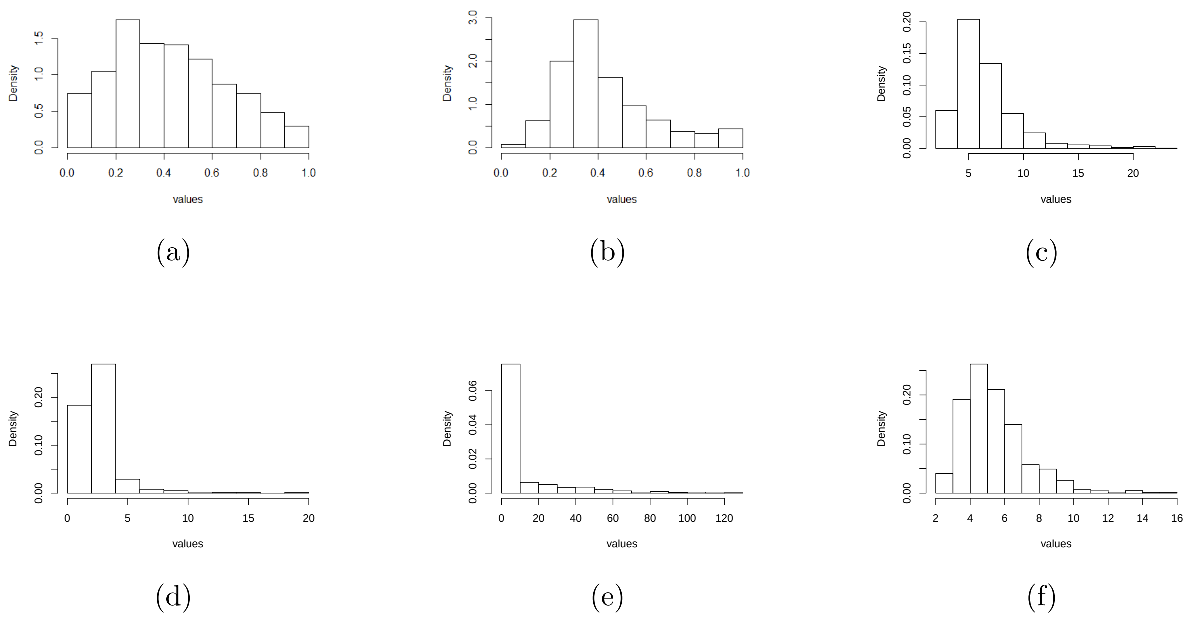


Figure 23 – Posterior distribution of ν_1 (a) and ν_2 (b) for the skew contaminated model and posterior distributions of ν for the skew-t (c) and skew slash (SSL1) (d) and skew slash (SSL2)(e) models and posterior distributions of ν_1 for the skew generalized t model (f).

Table 8 – **AIS dataset**: Posterior parameter estimates for the skew-t, skew slash, skew contaminated normal and skew generalized t models.

Model	parameter	Est	SD	CI (95%)
ST	β_0	68.7646	.3279	(68.1614, 69.4599)
	β_1	-7.9095	.5459	(-9.0050, -6.9025)
	β_2	.1007	.0322	(0.0384, .1619)
	β_3	.6680	.0257	(0.6210, .7176)
	γ	-0.7700	.2090	(-0.9861, -0.2589)
	ν	6.5260	2.8705	(3.1711, 14.2663)
	σ^2	5.3632	.8981	(3.7166, 7.1824)
SSL1	β_0	68.7350	.3652	(67.9851, 69.4309)
	β_1	-7.8757	.6090	(-8.9991,-6.6541)
	β_2	.0993	.0318	(0.0393, .1616)
	β_3	.6699	.0265	(0.6156, .7201)
	γ	-0.7237	.2197	(-0.9870, -0.2352)
	ν	2.6267	1.6607	(1.2747, 7.2775)
	σ^2	4.4623	1.0251	(2.6407, 6.8239)
SSL2	β_0	68.8016	.3512	(68.1157, 69.4860)
	β_1	-8.0060	.5888	(-9.0944,-6.7838)
	β_2	.0944	.0337	(0.0270, .1589)
	β_3	.6713	.0272	(0.6196, .7252)
	γ	-0.6137	.2282	(-0.9612, -0.0509)
	ν	11.6930	19.9007	(1.3640, 76.5994)
	σ^2	5.2014	1.5549	(2.8939, 8.4160)
SCN	β_0	68.9962	.3507	(68.1643, 69.5254)
	β_1	-8.1766	.5831	(-9.1280, -6.8339)
	β_2	.0990	.0335	(0.0263, .1554)
	β_3	.6635	.0278	(0.6190, .7291)
	γ	-0.6586	.1904	(-0.8891, -0.1038)
	ν_1	.1055	.2047	(0.0316, .8096)
	ν_2	.2452	.1567	(0.1124, .7496)
	σ^2	5.3521	1.3146	(2.3737, 7.7965)
SGT	β_0	69.0891	.3227	(68.4256, 69.6798)
	β_1	-8.3022	.4987	(-9.2603, -7.2947)
	β_2	.0989	.0301	(0.0409, .1573)
	β_3	.6588	.0236	(0.6144, .7058)
	γ	-0.3742	.3413	(-0.9337, .0790)
	ν_1	5.3696	1.8766	(2.8445, 9.9040)
	ν_2	32.5412	17.3355	(10.7662, 77.0973)
SN	β_0	68.9325	.3477	(68.2271, 69.5839)
	β_1	-8.2888	.5628	(-9.3732, -7.1902)
	β_2	.0754	.0320	(0.0158, .1380)
	β_3	.6829	.0263	(0.6312, .7353)
	γ	-0.4792	.1485	(-0.7316, -0.1725)
	σ^2	7.5311	.8356	(6.0504, 9.3382)

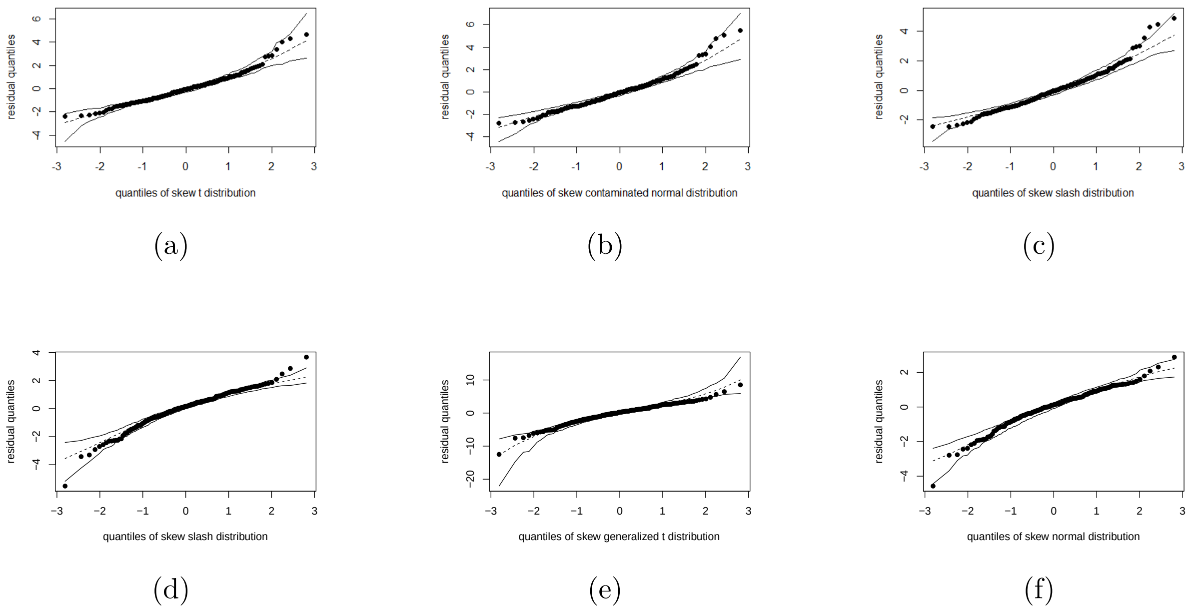


Figure 24 – QQ plots with envelopes for the ais dataset using the models: skew-t (a), skew contaminated normal (b), skew slash 1 (c), skew slash 2 (d), skew generalized t (e) and skew normal (f).

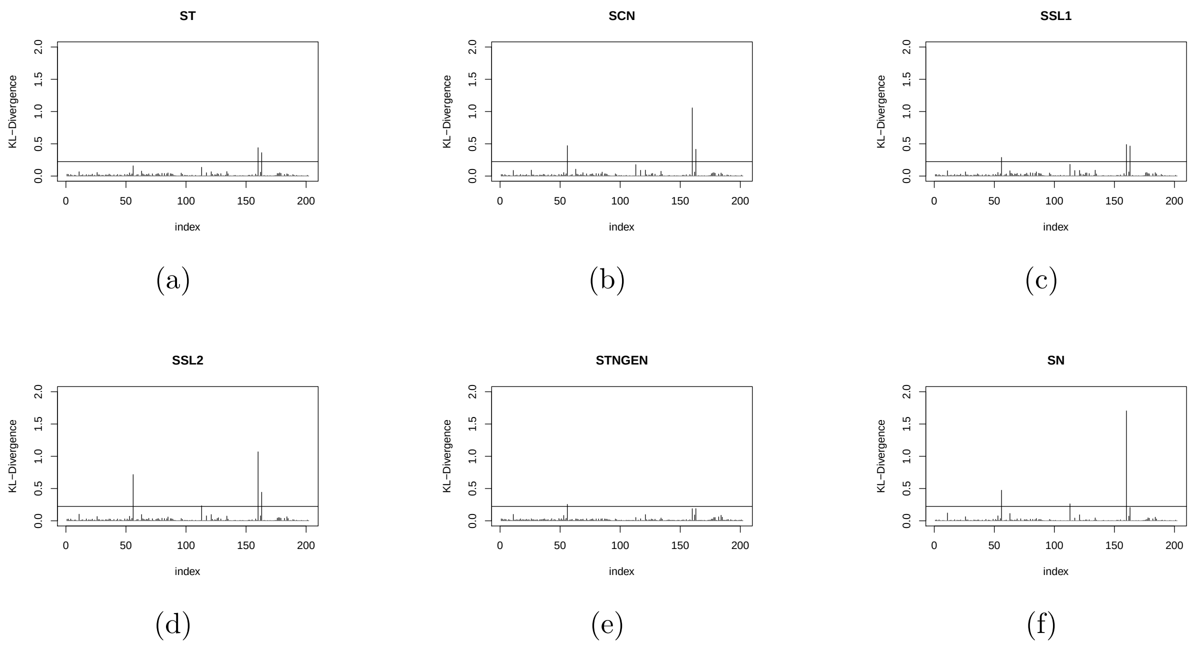


Figure 25 – Index plots of $K(P, P_{(i)})$ for the ais dataset: skew-t (a), skew contaminated normal (b), skew slash 1 (c), skew slash 2 (d), skew generalized t (e) and skew normal (f) models.

1.7 Conclusions

In this chapter we developed a scale mixture of skew-normal distribution under the centered parameterization class of probability distributions as an alternative to the parameterization used in (FERREIRA; BOLFARINE; LACHOS, 2011). It was decided to use a new parameterization for this class for several reasons, among them, the simplicity of parameter interpretation compared to the parameterization used in (FERREIRA; BOLFARINE; LACHOS, 2011). Another motivation was the issues related to the estimation process of parameter λ in the direct parameterization. We have showed, through profiled log-likelihood, that the SMSN class under direct parameterization can heritage the problem caused by the non quadratic likelihood shape of the direct parameterization.

A class of linear regression models based on the SMSN family under the centered parameterization was introduced, and we developed the Bayesian estimation approach. Also, we described model comparison criteria, and we developed analysis of influential observations and residual analysis. Simulation studies were performed in order to evaluate the parameter recovery under different scenarios. We concluded that for values of $\boldsymbol{\nu}$ that generate distributions with heavy tails the estimates are very accurate. On the other hand, for values of $\boldsymbol{\nu}$ close to the skew normal (or the symmetric) model, the estimates tend to be biased and the credibility intervals to be large. However, as the sample size increases, the estimates are improved. An application of the proposed model in a real dataset was performed in order to show that heavy tails models (special cases of the developed class of linear regression model) provide better fits than the skew normal linear regression.

2 Binary regression model with skew scale mixture of normal link function based on the centered parameterization

2.1 Introduction

Binary regression models are adequate to analyze data when the response variable assumes only two values. In these models the expected probability of success of a binary response is estimated based on one or more covariates through the specification of link function, that linearizes the relationship between the success probability and the covariates. As characterized in (CHEN; DEY; SHAO, 1999), the degree of asymmetry of the link function can be measured by the rate at which the probability of a response approaches 0 or 1. According to (CHEN; DEY; SHAO, 1999), a link function is symmetric if the approximation rate at 0 is the same as the approximation rate at 1. The most popular members of the binary regression models, the probit and logit, are examples of symmetric link functions. In the same way, a link function is positively asymmetric if the approximation rate at 1 is faster than the approximation rate at 0, and negatively asymmetric, otherwise.

From this definition of (CHEN; DEY; SHAO, 1999), we can say that the use of probit and logistic models are not adequate when we have evidence that the probability of success increases at a different rate than decreases. (CZADO; SANTNER, 1992) showed, through a simulation study, using a data generated by a skewed link function, that the link misspecification can yield a substantial bias in the estimates of the regression coefficients. Such problem is circumvented by the use of asymmetric link functions, that can be obtained, for example, through the cumulative distribution function of asymmetric distributions.

Many skewed binary regressions have been proposed in the literature. (STUKEL, 1988), (CZADO; SANTNER, 1992), and (GUERRERO; JOHNSON, 1982) introduced asymmetry replacing the linear predictor by a nonlinear function of the linear predictor and a parameter that controls the asymmetry. Another approach is to replace the linear predictor by a polynomial function, see for example (COLLETT, 2002). Finally, the third option is to consider the cumulative distribution function of an asymmetric distribution. The most popular example of this method is the complementary log-log link function, that is constructed from the cdf of the Gumbel distribution. (CHEN; DEY; SHAO, 1999) proposed an asymmetric probit link, considering a class of mixture of normal distributions. (BAZáN; BOLFARINE; BRANCO, 2010) presented a unified approach for two skew probit links. In (BAZáN; ROMEO; RODRIGUES, 2014), it was introduced two new asymmetric links, one based on the cdf of the power-normal distribution and another based on the cdf of the reciprocal power-normal distribution. (NAGLER, 1994) introduced the asymmetrical link by using the Burr-10 distribution ((BURR, 1942)).

Since probit and logistic regression estimates are not robust in the presence of outliers, (LIU, 2005) proposed a new binary model, named robit regression, in which the normal distribution in probit regression is replaced by a t-distribution with known or unknown degrees of freedom. Both the logistic model and the probit model can be approximated by the robit regression, as showed in (LIU, 2005). Instead using the t-distribution, (KIM; CHEN; DEY, 2008) introduced a class of skewed generalized t-link models, that accommodate heavy tail and asymmetric link functions.

In this work we developed a wide class of link functions for binary regression models that accommodate asymmetrical and heavy tails link functions, and includes the skewed generalized t-link, probit, skew probit, skew slash, skew contaminated normal, skew cauchy and probit models, among other, as special cases. This class is constructed based on the SMSN family under the centered parameterization introduced in Section 1.2.2.

2.2 Binary regression model

Let $\boldsymbol{X} = (1, \boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_{p-1})^t$ be a $p \times n$ known design matrix of fixed covariates, $\boldsymbol{Y} = (Y_1, \ldots, Y_n)^t$ be a $n \times 1$ vector of dichotomous response variables, such that $y_i = 1$ with probability p_i and $y_i = 0$ with probability $1-p_i$, and $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_{p-1})^t$ be a $p \times 1$ vector of regression coefficients. The binary regression model additionally assumes that

$$p_i = F(\eta_i) = F({X_i}^t\boldsymbol{\beta}), \quad i = 1, \ldots, n, \tag{2.1}$$

where $\eta_i = {X_i}^t\boldsymbol{\beta}$, $F(\cdot)$ denotes the cumulative distribution function and F^{-1} is a link function that linearizes the relationship between the success probability and the covariates. In this work, we assume that

$$p_i = F({X_i}^t\boldsymbol{\beta}|\gamma, \boldsymbol{\nu}), \quad i = 1, \ldots, n, \tag{2.2}$$

where $F(\cdot|\gamma, \boldsymbol{\nu})$ is the cdf of the skew-t, skew slash, skew generalized t or the skew contaminated normal distribution under the centered parametrization, as presented in 1.2.3, γ is the skewness parameter and $\boldsymbol{\nu} = \nu$ is the degree of freedom for the skew-t and skew slash distributions; for the skew contaminated normal distribution, $\boldsymbol{\nu} = (\nu_1, \nu_2)$ are, respectively, the proportion of outliers and scale factor, and for the skew generalized t distribution, $\boldsymbol{\nu} = (\nu_1, \nu_2)$ controls the tail and variance of the distribution.

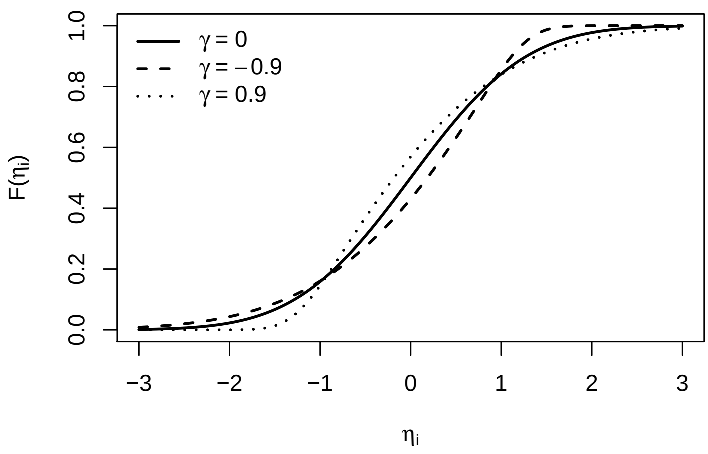
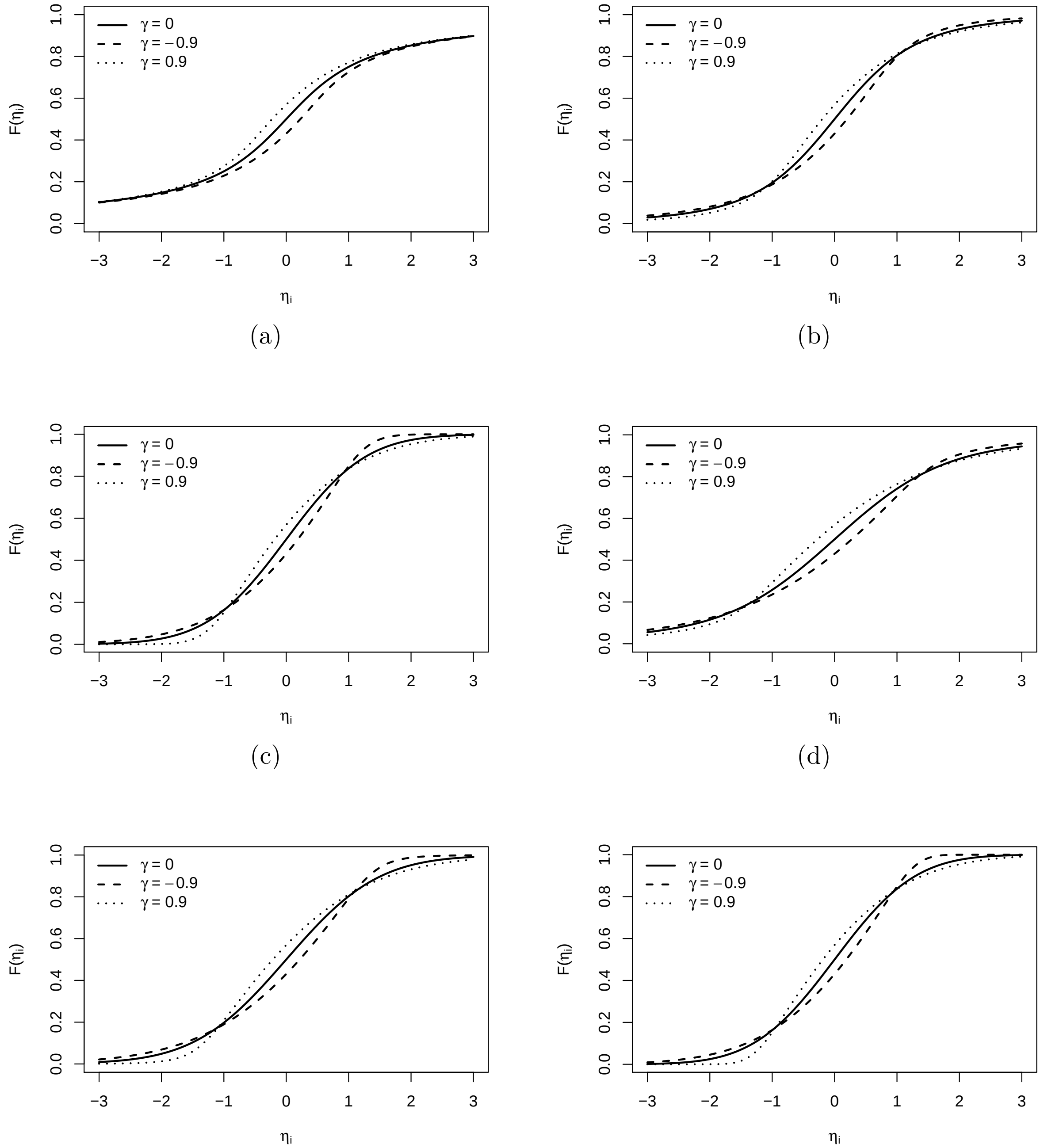


Figure 26 – Probability of success as a function of η_i for the skew normal distribution

The use of this distribution class in the binary regression model allows us a great flexibility in the choice of the link function, since this class includes heavy tails, symmetric and asymmetric distributions. From Figures 27 and 28 we can see the effect of heavy tails on the cumulative distribution function by observing that the probability of success grows slowly when compared to the normal skew cdf, as presented in Figure 26, for small values of ν for the skew-t, skew slash distributions, as well for small values of ν_1 in the skew generalized t distribution and when ν_1 approaches to 1 and ν_2 to 0, in the skew contaminated normal distribution. Panel (c) of the Figures 27 and 28 shows the cdf when these distributions approach to the skew normal case. From these figures, we can also see that when $\delta = -0.9$ ($\delta = .9$) the probability p_i approaches 1 (0) at a faster rate than it approaches 0 (1). When $\delta = 0$ the probability approaches 1 or 0 at the same rate.

These figures suggest that the use of heavy tails distribution is appropriate in the cases where extreme values of the linear predictor are expected. Also, the use of the asymmetric link is appropriate when we expected that the rate approaches 1 is different to the rate approaches . In addition, heavy tails links help to control the rate of convergence to 0 and 1, providing more flexibility in the modeling of the influence of the covariates in the response variable.



(a) (b)

(c) (d)

(e) (f)

Figure 27 – Probability of success as a function of η_i for various ν for the skew-t distribution with (a) $\nu = 1$, (b) $\nu = 3$ and (c) $\nu = 30$, and for the skew slash distribution with (d) $\nu = 1$, (e) $\nu = 3$ and (f) $\nu = 30$.

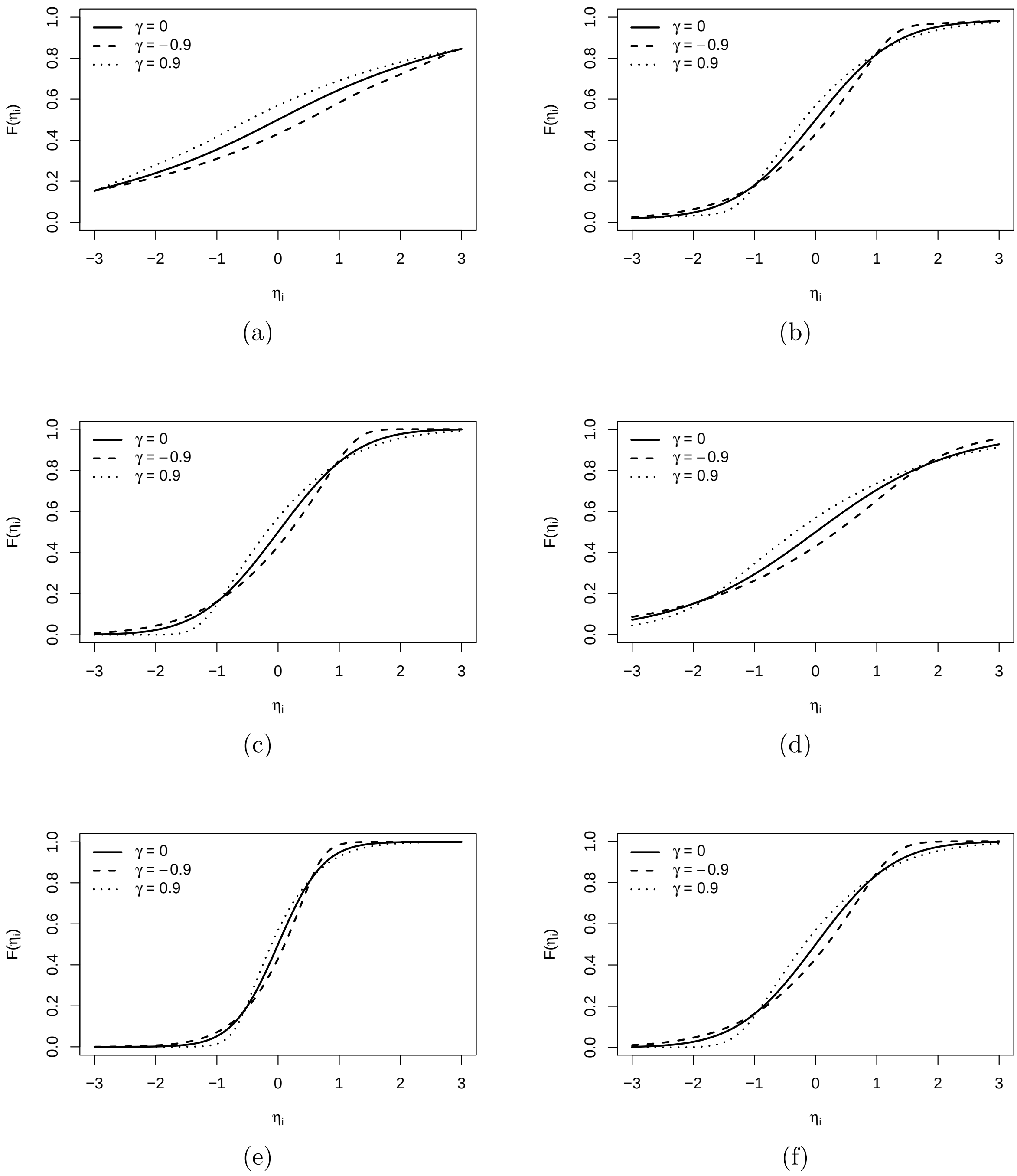


Figure 28 – Probability of success as a function of η_i for various $\boldsymbol{\nu}$ for the skew contaminated normal distribution with (a) $\boldsymbol{\nu} = (.9, 0.1)$, (b) $\boldsymbol{\nu} = (.1, 0.1)$ and (c) $\boldsymbol{\nu} = (.1, 0.9)$, and for the skew generalized t distribution with (d) $\boldsymbol{\nu} = (5, 15)$, (e) $\boldsymbol{\nu} = (15, 5)$ and (f) $\boldsymbol{\nu} = (30, 30)$.

To perform Bayesian inference, an approach based on data augmentation, as considered in (ALBERT; CHIB, 1993) and (BAZáN; BOLFARINE; BRANCO, 2010), will be used. The main advantage of using this approach is the ability to introduce a hierarchical structure, which simplifies the Bayesian estimation process. Before introducing this alternative representation for the binary model, we consider $Z_i \sim SMSN_c(0, 1, -\gamma, G, \boldsymbol{\nu})$ which implies that

$$Z_i = X_i{}^t\boldsymbol{\beta} + U_i{}^{-1/2}\varepsilon_i \quad i = 1, \ldots, n, \tag{2.3}$$

where $\varepsilon_i \overset{iid}{\sim} SN_c(0, 1, -\gamma)$ and $U_i \overset{iid}{\sim} H(\cdot|\boldsymbol{\nu})$ and $I(\cdot)$ denotes the indicator function. From the stochastic representation of the skew-normal presented in 1.7, the skew-normal under the centered parametrization (ε_i) can be written as

$$\varepsilon_i = -\Delta(H_i - b) + \sqrt{\tau}T_i \quad i = 1, \ldots, n, \tag{2.4}$$

where $\lambda = \dfrac{s\gamma^{1/3}}{\sqrt{b^2 + s^2\gamma^{2/3}(b^2-1)}}$, $\delta = \dfrac{\lambda}{\sqrt{1+\lambda^2}}$, $\Delta = \dfrac{\delta}{\sqrt{1-b^2\delta^2}}$, $\tau = \dfrac{1-\delta^2}{1-b^2\delta^2}$, $b = \sqrt{\dfrac{2}{\pi}}$, $s = \left(\dfrac{2}{4-\pi}\right)^{1/3}$ and $H_i \overset{iid}{\sim} HN(0, 1) \perp T_i \overset{iid}{\sim} N(0, 1)$.

Using 2.3 and 2.4 we have the following proposition:

Proposition 2.2.1. The binary model $Y_i \sim Ber(p_i)$ and $p_i = F(X_i{}^t\boldsymbol{\beta}|\gamma, \boldsymbol{\nu})$ is equivalent to consider

$$y_i = I(Z_i > 0) = \begin{cases} 1 & \text{if } Z_i > 0 \\ 0 & \text{if } Z_i \leq 0 \end{cases} \quad i = 1, \ldots, n, \tag{2.5}$$

with

$$\begin{aligned} &Z_i = X_i{}^t\boldsymbol{\beta} + U_i{}^{-1/2}(\Delta(b - H_i) + \sqrt{\tau}\varepsilon_i) \\ &\varepsilon_i \sim N(0, 1) \quad H_i \sim HN(0, 1) \quad U_i \sim G(.|\boldsymbol{\nu}) \end{aligned} \tag{2.6}$$

Following 2.5 and 2.6, the hierarchical formulation of the model is given as follow:

$$\begin{aligned} Z_i|U_i = u_i, H_i = h_i, y_i &\sim N\left(X_i{}^t\boldsymbol{\beta} + u_i{}^{-1/2}\Delta(b - h_i), \frac{\tau}{u_i}\right) I(z_i, y_i) \\ H_i &\sim HN(0, 1) \\ U_i &\sim G(.|\boldsymbol{\nu}), \end{aligned} \tag{2.7}$$

where $I(z_i, y_i) = I(z_i > 0)I(y_1 = 1) + I(z_i \leq 0)I(y_i = 0)$.

2.3 Bayesian Inference

To use the Bayesian paradigm, it is essential to obtain the joint posterior distribution. However, since the necessary integrals are not easy to calculate, it is not possible to obtain such distribution, analytically. However, it is possible to obtain numerical approximation for the marginal posterior distributions of interest by using MCMC algorithms, see (GEMAN; GEMAN, 1984) and (HASTINGS, 1970).

To obtain the posterior distribution we need first to consider the complete likelihood

$$\begin{aligned} L_c(\boldsymbol{\theta}|y,z,u,h) &\propto \prod_{i=1}^{n} \phi\left(z_i|\mu_i, \tau u_i^{-1}\right) I(z_i,y_i) f(h_i) h(u_i|\boldsymbol{\nu}) \\ &\propto \prod_{i=1}^{n} \frac{\sqrt{u_i}}{\sqrt{\tau}} \exp\left\{-\frac{u_i}{2\tau}\left(z_i-\mu_i\right)^2\right\} I(z_i,y_i) \exp\left\{-\frac{h_i^2}{2}\right\} h(u_i|\boldsymbol{\nu}) \\ &\propto \frac{\prod_{i=1}^{n}\sqrt{u_i}}{\tau^{n/2}} \exp\left\{-\frac{1}{2\tau}\sum_{i=1}^{n} u_i\left(z_i-\mu_i\right)^2\right\} I(z_i,y_i) \exp\left\{-\frac{\sum_{i=1}^{n} h_i^2}{2}\right\} \prod_{i=1}^{n} h(u_i|\boldsymbol{\nu}), \end{aligned}$$

where $\mu_i = \boldsymbol{X}_i^t\boldsymbol{\beta} + \frac{\Delta}{\sqrt{u_i}}(b-h_i)$ and $\boldsymbol{\theta} = (\boldsymbol{\beta}, \gamma, \boldsymbol{\nu})$. We need to consider a prior distribution for $\boldsymbol{\theta}$ such that $\pi(\boldsymbol{\theta}) = \pi(\boldsymbol{\beta})\pi(\gamma)\pi(\boldsymbol{\nu})$. Furthermore, we will assume conditional conjugate prior distributions, as in (GELMAN, 2006), for $\boldsymbol{\beta}$. For γ, we assume the same prior used in (AZEVEDO; BOLFARINE; ANDRADE, 2012), that is $\pi(\gamma) \propto (1+\gamma)^{\alpha_{\gamma_1}-1}(1-\gamma)^{\alpha_{\gamma_2}-1} I(\gamma \in A_\gamma)$, where $A_\gamma = (-0.99527, 0.99527)$. For $\boldsymbol{\nu}$, the choice of the prior distribution will depend on the model.

2.3.1 Full conditional distributions

In order to implement the MCMC algorithm, we have to simulate iteratively from the full conditionals described bellow.

Denoting by $\boldsymbol{\theta}_{-\theta_i}$ the parameter vector $\boldsymbol{\theta}$ without the component θ_i, the full conditional distributions are

For β:

$$\pi(\boldsymbol{\beta}|\boldsymbol{\theta}_{-\beta}, \boldsymbol{y}, \boldsymbol{z}, \boldsymbol{u}, \boldsymbol{h}) \propto \exp\left\{-\frac{1}{2}\left(\boldsymbol{\beta}^t\Sigma_*^{-1}\boldsymbol{\beta} - 2\mu_*^t\Sigma_*^{-1}\boldsymbol{\beta}\right)\right\} I_{\mathbb{R}^p}(\boldsymbol{\beta}),$$

which can be recognized as the kernel of p-variate normal distribution with variance $\boldsymbol{\Sigma_*} = \left(\frac{\sum_{i=1}^{n} u_i x_i x_i^t}{\tau} + \boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{-1}\right)^{-1}$ and mean $\boldsymbol{\mu_*} = \left(\sum_{i=1}^{n} \frac{u_i}{\tau}\left(z_i - \frac{\Delta}{\sqrt{u_i}}(b-h_i)\right)\boldsymbol{x}_i^t + \boldsymbol{\mu}_{\boldsymbol{\beta}}^t\boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{-1}\right)\boldsymbol{\Sigma_*}$.

For z_i:

$$f(z_i|\boldsymbol{\theta}, y_i, h_i, u_i) \propto \phi\left(z_i|\mu_i, \tau u_i^{-1}\right) I(z_i, y_i).$$

Then,

$$z_i|\boldsymbol{\theta}, u_i, h_i, y_i = 1 \sim TN\left(\mu_i, \tau u_i^{-1}\right) I(0, \infty)$$
$$z_i|\boldsymbol{\theta}, u_i, h_i, y_i = 0 \sim TN\left(\mu_i, \tau u_i^{-1}\right) I(-\infty, 0)$$

For h_i:

$$f(h_i|\boldsymbol{\theta}, y_i, z_i, u_i) \propto \exp\left\{-\frac{1}{2}\left(\frac{\Delta^2+\tau}{\tau}\right)\left[h_i^2 - 2h_i\left(\frac{\Delta^2 b - \Delta\sqrt{u_i}(z_i - \boldsymbol{X}_i^t\boldsymbol{\beta})}{\Delta^2+\tau}\right)\right]\right\} I_{(0,\infty)}(h_i),$$

which can be recognized as the kernel of a truncated normal distribution, so

$$h_i|\boldsymbol{\theta}, u_i, y_i \sim TN\left(\frac{\Delta^2 b - \Delta\sqrt{u_i}(z_i - \boldsymbol{X}_i^t\boldsymbol{\beta})}{\Delta^2+\tau}, \frac{\tau}{\Delta^2+\tau}\right) I(0, \infty).$$

For u_i:

- Skew slash:

$$f(u_i|\boldsymbol{\theta}, y_i, h_i, z_i) \propto u_i^{\nu+1/2-1} \exp\left\{-\frac{u_i}{2\tau}\left[(z_i - \boldsymbol{X}_i^t\boldsymbol{\beta})^2 - 2\frac{\Delta}{\sqrt{u_i}}(b-h_i)(z_i - \boldsymbol{X}_i^t\boldsymbol{\beta})\right]\right\} I_{(0,1)}(u_i).$$

- Skew-t:

$$f(u_i|\boldsymbol{\theta}, y_i, h_i, z_i) \propto u_i^{\frac{\nu+1}{2}-1} \exp\left\{-\frac{u_i}{2}\left[\frac{(z_i - \boldsymbol{X}_i^t\boldsymbol{\beta})^2}{\tau} + \nu\right] + \frac{\Delta\sqrt{u_i}}{\tau}(b-h_i)(z_i - \boldsymbol{X}_i^t\boldsymbol{\beta})\right\}$$
$$\times I_{(0,\infty)}(u_i).$$

- Skew generalized t:

$$f(u_i|\boldsymbol{\theta}, y_i, h_i, z_i) \propto u_i^{\frac{\nu_1+1}{2}-1} \exp\left\{-\frac{u_i}{2}\left[\frac{(z_i - \boldsymbol{X}_i^t\boldsymbol{\beta})^2}{\tau} + \nu_2\right] + \frac{\Delta\sqrt{u_i}}{\tau}(b-h_i)(z_i - \boldsymbol{X}_i^t\boldsymbol{\beta})\right\}$$
$$\times I_{(0,\infty)}(u_i).$$

- Skew-contaminated normal: the discrete conditional distribution of u_i assumes ν_2 with probability $\dfrac{p_i}{p_i+q_i}$ and 1 with probability $\dfrac{q_i}{p_i+q_i}$ where

$$p_i = \nu_1\sqrt{\nu_2}\exp\left\{-\frac{\nu_2}{2\tau}\left[(z_i - \boldsymbol{X}_i^t\boldsymbol{\beta})^2 - 2\frac{\Delta}{\sqrt{\nu_2}}(b-h_i)(z_i - \boldsymbol{X}_i^t\boldsymbol{\beta})\right]\right\}$$
$$q_i = (1-\nu_1)\exp\left\{-\frac{1}{2\tau}\left[(z_i - \boldsymbol{X}_i^t\boldsymbol{\beta})^2 - 2\Delta(b-h_i)(z_i - \boldsymbol{X}_i^t\boldsymbol{\beta})\right]\right\}.$$

For ν:

- Skew slash: Considering a gamma distribution as prior, with mean $\frac{\alpha_1}{\alpha_2}$ and variance $\frac{\alpha_1}{\alpha_2^2}$, it follows that

$$\pi(\nu|\boldsymbol{\theta}_{-\nu}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{h}) \propto \nu^{n+\alpha_1-1} \exp\left\{-\nu(\alpha_2 - \sum_{i=1}^{n} ln(u_i))\right\} I_{(0,\infty)}(\nu),$$

  that is, $\nu|\boldsymbol{\theta}_{-\nu}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{h} \sim \text{gamma}(n + \alpha_1, \alpha_2 - \sum_{i=1}^{n} ln(u_i))$.

- Skew-t: We have adopted a very useful hierarchical prior distribution as noted in (CABRAL; LACHOS; MADRUGA, 2012), which consists on $\nu|\lambda \sim exp(\lambda)$ and $\lambda \sim U(\rho_0, \rho_1)$ where $0 < \rho_0 < \rho_1$ are known. Then

$$\pi(\nu|\boldsymbol{\theta}_{-\nu}, \lambda, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{h}) \propto \frac{\frac{\nu}{2}^{\frac{n\nu}{2}}}{\Gamma(\nu/2)^n} \left(\prod_{i=1}^{n} u_i\right)^{\nu/2-1} \exp\left\{-\nu\left(\frac{\sum_{i=1}^{n} u_i}{2} + \lambda\right)\right\} I_{(0,\infty)}(\nu)$$

$$\pi(\lambda|\boldsymbol{\theta}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{h}) \propto \lambda \exp -\lambda(\nu) I_{(\rho_0,\rho_1)}(\lambda),$$

  that is, $\lambda|\boldsymbol{\theta}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{h} \sim TG(2, \nu)I(\rho_0, \rho_1)$.

- Skew generalized t: Assuming $\nu_1|\lambda_1 \sim exp(\lambda_1)$ and $\lambda_1 \sim U(\rho_0, \rho_1)$ where $0 < \rho_0 < \rho_1$ are known and $\nu_2|\lambda_2 \sim exp(\lambda_2)$ and $\lambda_2 \sim U(\psi_0, \psi_1)$ where $0 < \psi_0 < \psi_1$ are known, we have

$$\pi(\nu_1|\boldsymbol{\theta}_{-\nu_1}, \lambda_1, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{h}) \propto \frac{\nu_2/2^{n\nu_1/2}}{\Gamma(\nu_1/2)^n} \left(\prod_{i=1}^{n} u_i\right)^{\nu_1/2-1} \exp\left\{-\lambda_1(\nu_1 - 2)\right\} I_{(0,\infty)}(\nu_1)$$

$$\pi(\lambda_1|\boldsymbol{\theta}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{h}) \propto \lambda_1 \exp\left\{-\lambda_1(\nu_1)\right\} I_{(\rho_0,\rho_1)}(\lambda_1),$$

  that is, $\lambda_1|\boldsymbol{\theta}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{h} \sim TG(2, \nu_1)I(\rho_0, \rho_1)$.

  Also, we have that

$$\pi(\nu_2|\boldsymbol{\theta}_{-\nu_2}, \lambda_2, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{h}) \propto \nu_2/2^{n\nu_1/2} \exp\left\{-\nu_2\left(\frac{\sum_{i=1}^{n} u_i}{2} + \lambda_2\right)\right\} I_{(0,\infty)}(\nu_2)$$

  and

$$\pi(\lambda_2|\boldsymbol{\theta}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{h}) \propto \lambda_2 \exp\{-\lambda_2(\nu_2)\} I_{(\psi_0,\psi_1)}(\lambda_2),$$

  that is, $\nu_2|\boldsymbol{\theta}_{-\nu_2}, \lambda_2, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{h} \sim \text{gamma}(\frac{n\nu_1}{2} + 1, \frac{\sum_{i=1}^{n} u_i}{2} + \lambda_2)$
  and $\lambda_2|\boldsymbol{\theta}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{h} \sim TG(2, \nu_2)I(\xi_0, \xi_1)$

- Skew-contaminated normal: Observe that the distribution of U can be written as

$$h(u|\boldsymbol{\nu}) = \nu_1^{\frac{1-u}{1-\nu_2}}(1-\nu_1)^{\frac{u-\nu_2}{1-\nu_2}} I_{\{\nu_2,1\}}(u).$$

Setting as prior distributions $\nu_1 \sim beta(\alpha_1, \beta_1)$, $\nu_2 \sim beta(\alpha_2, \beta_2)$, it follows that the conditional distributions of ν_1 and ν_2 are

$$\pi(\nu_1|\boldsymbol{\theta}_{-\nu_1}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{h}) \propto \nu_1^{\frac{n-\sum_{i=1}^n u_i}{1-\nu_2}+\alpha_1-1}(1-\nu_1)^{\frac{\sum_{i=1}^n u_i - n\nu_2}{1-\nu_2}+\beta_1-1} I_{(0,1)}(\nu_1),$$

which can be recognized as the kernel of a beta distribution. So, $\nu_1|\boldsymbol{\theta}_{-\nu_1}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{h} \sim beta\left(\frac{n-\sum_{i=1}^n u_i}{1-\nu_2}+\alpha_1, \frac{\sum_{i=1}^n u_i - n\nu_2}{1-\nu_2}+\beta_1\right)$ and

$$\pi(\nu_2|\boldsymbol{\theta}_{-\nu_2}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{h}) \propto \nu_1^{\frac{n-\sum_{i=1}^n u_i}{1-\nu_2}}(1-\nu_1)^{\frac{\sum_{i=1}^n u_i - n\nu_2}{1-\nu_2}} \nu_2^{\alpha_2-1}(1-\nu2)^{(\beta_2-1)} I_{(0,1)}(\nu_2).$$

For γ:

$$\pi(\gamma|\boldsymbol{\theta}_{-\gamma}, \boldsymbol{y}, \boldsymbol{u}, \boldsymbol{z}, \boldsymbol{h}) \propto \tau^{-n/2} \exp\left\{-\frac{1}{2\tau}\sum_{i=1}^n u_i \left(z_i - \boldsymbol{X}_i^t\boldsymbol{\beta} - \frac{\Delta}{\sqrt{u_i}}(b-h_i)\right)^2\right\}(1+\gamma)^{\alpha_{\gamma_1}-1}(1-\gamma)^{\alpha_{\gamma_2}-1}$$
$$I(\gamma \in A_\gamma).$$

2.3.2 Residual analysis

For the binary regression models, the ordinary residual can be defined as $Y_i - \hat{p}_i$, where $\hat{p}_i$ is the i-th fitted observation, based on an appropriate estimate of $\boldsymbol{\beta}$ (related to a consistent estimator). We can also define other types of residual, as the Pearson, deviance, among others ((MCCULLAGH; NELDER, 1989)). However, due to the discrete nature of the binary response variable, these residuals have unknown distribution, which can affect the interpretation and outlying detection. (ALBERT; CHIB, 1995) proposed a continuous Bayesian residual based on latent variable. Let $(\boldsymbol{\beta}^{(m)}, \gamma^{(m)}, \boldsymbol{\nu}^{(m)})$, $m = 1, \ldots, M$ a valid MCMC sample, another possibility is considered the residual $r_i^{(m)} = Y_i - p_i^{(m)}$ where $p_i^{(m)} = F(\boldsymbol{X}_i^t\boldsymbol{\beta}^{(m)}|\gamma^{(m)}, \boldsymbol{\nu}^{(m)})$. In the Bayesian context, this residual has continuous distribution, with support on the interval $(Y_i - 1, Y_i)$ ((FARIAS; BRANCO, 2012)). Then, an observation will be outlying for $y_i = 0$ if the posterior distribution of r_i is concentrated towards -1. On the other hand, for $y_i = 1$, the observation will be outlying if the distribution of r_i is concentrated towards 1.

Similarly to the latent Bayesian residual for the skew probit regression developed in (FARIAS; BRANCO, 2012), we can define the latent residual for the binary regression model with link function based on the SMSN family under the centered parameterization from the stochastic representation given in 2.2.1. From that, we can define the residual

$$\epsilon_i = \frac{Z_i - \boldsymbol{X}_i^t \boldsymbol{\beta} - U_i^{-1/2}\Delta(b - H_i)}{\sqrt{\tau}}, \tag{2.8}$$

where Z_i, U_i and H_i are the latent variables. It follows that, conditioned in $\boldsymbol{\beta}, \gamma$, the residual 2.8 is normally distributed a priori. A way to check outliers is to analyze the posterior distribution of the residual given in 2.8.

For model checking, we can use the Deviance residual, defined as

$$d_i = sign(y_i - \hat{p}_i)\sqrt{-2\left(y_i log(\hat{p}_i) + (1 - y_i) log(1 - \hat{p}_i)\right)}. \tag{2.9}$$

A way to check lack of fit is to build envelope plot for the deviance residuals, under a Bayesian perspective. The considered methodology for constructing the simulated envelopes is described in Appendix E.

2.4 Simulation study

We performed simulation studies in order to evaluate the performance of the model, estimation method and behavior of the residuals and influence diagnostic analysis proposed in this work. All these models were implemented in JAGS ((PLUMMER, 2003)) through the interface provided by the *rjags* package ((PLUMMER, 2016)) available in R program ((R Development Core Team, 2008)). The codes are available from the authors upon request.

We adopted weakly informative priors for all parameters, that is: $\beta_0 \sim N(0, 1000)$ and $\beta_1 \sim N(0, 1000)$. For the γ_b, γ_c and γ_w we used the prior described in Section 2.3 with $\alpha_{\gamma_1} = \alpha_{\gamma_2} = 0.5$. For the skew-t model we set $\nu \sim exp(\theta)$ and $\theta \sim unif(.05, 0.5)$; for the skew slash model we adopt $\nu \sim gamma(1, 0.05)$; for the skew contaminated model we used $\nu_1 \sim beta(1, 1)$ and $\nu_2 \sim beta(1, 1)$ and for the skew generalized t model we set $\nu_1 \sim gamma(1, 0.05)$ and $\nu_2 \sim exp(\theta_2)$ and $\theta_2 \sim unif(0.05, 0.5)$. All these priors for $\boldsymbol{\nu}$ were chosen based on previous sensitivity study done by the authors, taking into account different values of hyperparameters. The results presented here, using the selected priors, were those that presented more accurate estimates among the tested values.

To eliminate the effect of the initial values and to avoid correlations problems, we run a MCMC chain of size 600,000 with a burn-in of 100,000 and a thin of 1,000, retaining a valid MCMC chain of size 500. The values of the Gelman-Rubin statistics and the analyses of traceplots, Geweke and autocorrelation plots indicated that the MCMC algorithm converged and the autocorrelations were negligible, with such spacement.

2.4.1 Simulation study I

The objective of this simulation study is to measure the impact of the sample size on the parameter recovery. We considered different scenarios based on the crossing of the levels of some factors of interest. For the four regression models explored in this work, we simulate from samples of size n=50, 250, 500 and 1000, varying the value of $\boldsymbol{\nu}$, considering a single replica. That is, we generated replicas from

$$\begin{aligned} Y_i &= I(Z_i > 0) \\ Z_i &= \beta_0 + \beta_1 x_i + \varepsilon_i \quad i = 1, \ldots, n \end{aligned}$$

where $\beta_0 = 1$, $\beta_1 = 2$, ε_i belongs to the SMSN family with $\sigma^2 = 1$ and $\gamma \in \{-0.9, 0, 0.9\}$, which allows the model to have strong negative, null and strong positive asymmetry, respectively. Also, we set $\nu \in \{3, 10, 50\}$ for the skew-t and skew-slash distribution; $\boldsymbol{\nu} = (\nu_1, \nu_2) = (.1, 0.1), (.1, 0.9)$, and $(.9, 0.1)$ for the skew contaminated normal distribution and $\boldsymbol{\nu} = (\nu_1, \nu_2) = (15, 5), (5, 15)$ and $(50, 50)$ for the skew generalized-t distribution. These values for $\boldsymbol{\nu}$ were chosen in order to have distributions with heavy tails and tails close to the skew normal distribution. The covariate was simulated from N(0,1) and centered in its respective mean.

Based on small studies performed by the author about the choice of prior distribution for γ we noted that γ estimates tend to have credibility intervals that cover practically all parameter space, and the estimates can be very dependent on the prior choice. In the work of (AZEVEDO; BOLFARINE; ANDRADE, 2012) this problem is discussed in the Item Response Model context. Also based in some previous studies, we decided use the beta modified prior described in (AZEVEDO; BOLFARINE; ANDRADE, 2012). The authors intend to carry out a more detailed study on the accuracy of the estimates and propose alternatives to improve them.

The results for the skew-t and skew slash models are showed in Tables 9 and 10, and the remaining tables are presented in Appendix F. From these two tables we can notice that when the true value of ν is small, the estimates are accurate, but the skew-t model presents better results than the skew slash model, when we compare the standard deviations obtained from these two models. When $\nu = 10$, for both models, we observe overestimation of this parameter, since the estimates tend to be around 20. The opposite occurs when $\nu = 50$, however, for this scenario the estimates are close to 30, which indicate that the skew-t model approaches to the skew-normal model. The parameter vector $\boldsymbol{\beta}$ were adequately estimated, for sample sizes greater than 50. Comparing the results for γ, we noted that the estimates were accurate only for large sample sizes.

For the skew generalized t model, we see that estimates of ν_1 tend to be around 5 and of ν_2 around 15, even though the respective true do not correspond to these values. The same occurs for the skew contaminated normal model, where the estimates of ν_1 are close to .6 and ν_2 to .2. For this model we can see that the respective credibility intervals cover almost all the parameter space of ν_1 and ν_2. It was also observed that the estimates of $\boldsymbol{\beta}$ were not as accurate as for the skew-t and skew slash models, as can be seen in Table 11 and 12.

Table 9 – Results of the simulation study for the skew-t model with $\nu = 3$.

		Sample size											
		50			250			500			1000		
		Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_0	$\gamma = -0.9$	.396	.431	(-0.376;1.273)	.939	.208	(.567;1.361)	.99	.202	(.637;1.377)	.925	.132	(.66;1.169)
	$\gamma = 0$	.265	.343	(-0.319; .967)	1.1	.265	(.685;1.644)	1.097	.217	(.678;1.491)	.841	.131	(.603;1.089)
	$\gamma = .9$	1.457	.701	(.462;2.921)	.975	.245	(.549;1.439)	1.122	.224	(.747;1.581)	.874	.137	(.611;1.142)
β_1	$\gamma = -0.9$	3.112	1.402	(1.02;5.757)	1.949	.447	(1.232;2.932)	2.107	.38	(1.419;2.871)	2.033	.249	(1.537;2.493)
	$\gamma = 0$	2.183	.848	(.904;3.898)	2.027	.489	(1.265;2.963)	2.108	.349	(1.481;2.751)	1.798	.241	(1.361;2.249)
	$\gamma = .9$	2.475	1.222	(.648;5.014)	1.91	.377	(1.288;2.722)	2.093	.365	(1.506;2.863)	1.764	.241	(1.321;2.214)
γ	$\gamma = -0.9$	.895	.57	(-0.788; .995)	-0.896	.467	(-0.995; .523)	.681	.449	(-0.53; .994)	-0.525	.456	(-0.995; .52)
	$\gamma = 0$	.891	.518	(-0.692; .995)	-0.718	.559	(-0.995; .868)	.89	.442	(-0.383; .995)	.397	.363	(-0.293; .992)
	$\gamma = .9$	.835	.591	(-0.87; .995)	.923	.448	(-0.408; .995)	.949	.257	(.214; .995)	.936	.228	(.273; .995)
ν	$\gamma = -0.9$	16.096	18.039	(2.001;51.907)	12.831	14.404	(2.001;44.905)	9.782	16.197	(2.003;39.819)	3.262	2.202	(2.002;5.942)
	$\gamma = 0$	18.282	17.935	(2.007;54.049)	11.033	14.523	(2.001;38.94)	6.692	10.681	(2;22.238)	4.694	4.745	(2;13.186)
	$\gamma = .9$	14.769	16.9	(2.01;51.577)	14.812	17.32	(2.024;50.557)	7.038	8.373	(2.001;23.623)	5.087	5.147	(2.002;15.37)

Table 10 – Results of the simulation study for the skew slash model with $\nu = 3$.

		Sample size											
		50			250			500			1000		
		Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_0	$\gamma = -0.9$	1.15	.483	(.225;2.126)	.78	.118	(.575;1.019)	.885	.134	(.649;1.11)	.977	.141	(.729;1.239)
	$\gamma = 0$	1.04	.424	(.306;1.973)	.722	.172	(.404;1.068)	.892	.195	(.577;1.292)	.842	.136	(.627;1.157)
	$\gamma = .9$	.762	.346	(.07;1.392)	.723	.163	(.398;1.03)	.884	.16	(.591;1.193)	1.021	.188	(.66;1.358)
β_1	$\gamma = -0.9$	2.363	.698	(.962;3.691)	1.494	.251	(1.039;1.95)	1.999	.301	(1.517;2.578)	1.957	.289	(1.515;2.559)
	$\gamma = 0$	1.913	.659	(.793;3.074)	1.459	.319	(.937;2.113)	1.953	.396	(1.365;2.808)	1.724	.251	(1.374;2.309)
	$\gamma = .9$	1.565	.509	(.766;2.594)	1.521	.285	(1.033;2.115)	1.813	.253	(1.399;2.289)	2.098	.337	(1.44;2.707)
γ	$\gamma = -0.9$	-0.85	.689	(-0.973; .995)	-0.953	.289	(-0.995;-0.174)	-0.71	.258	(-0.963;-0.078)	-0.939	.216	(-0.995;-0.289)
	$\gamma = 0$	-0.85	.606	(-0.995; .872)	-0.345	.477	(-0.995; .573)	.411	.373	(-0.354; .994)	.198	.263	(-0.348; .66)
	$\gamma = .9$	-0.797	.623	(-0.995; .948)	.144	.491	(-0.675; .99)	.96	.189	(.413; .995)	.962	.139	(.566; .995)
ν	$\gamma = -0.9$	20.14	20.677	(1.061;58.042)	21.011	20.703	(1.084;60.079)	18.746	17.714	(1.003;56.469)	9.888	12.958	(1.084;38.736)
	$\gamma = 0$	18.659	19.151	(1.017;56.187)	14.18	17.255	(1.003;50.321)	12.878	17.559	(1.005;47.746)	16.331	17.369	(1.006;53.824)
	$\gamma = .9$	18.268	19.3	(1.03;59.027)	18.028	18.797	(1.013;57.341)	18.396	18.226	(1.028;54.084)	3.382	5.531	(1.002;12.884)

Table 11 – Results of the simulation study for the skew generalized t model with $\nu_1 = 50$ and $\nu_2 = 50$.

		Sample size											
		50			250			500			1000		
		Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_0	$\gamma = -0.9$	1.749	1.302	(.146;4.466)	1.101	1.01	(.216;3.651)	2.507	1.218	(.817;4.835)	.907	.193	(.626;1.28)
	$\gamma = 0$	2.291	1.834	(.476;6.458)	1.43	.934	(.382;3.755)	1.023	.754	(.3;2.674)	1.734	.871	(.448;3.347)
	$\gamma = .9$	1.037	.864	(.101;3.138)	1.04	.815	(.316;2.98)	2.994	1.507	(1.114;6.037)	1.9	.721	(.58;2.861)
β_1	$\gamma = -0.9$	4.478	3.271	(.325;11.516)	1.993	1.784	(.461;6.053)	5.293	2.571	(1.81;10.189)	1.809	.383	(1.347;2.566)
	$\gamma = 0$	3.339	2.61	(.358;9.616)	3.508	2.302	(.934;9.343)	2.052	1.477	(.625;5.055)	3.453	1.74	(.872;6.402)
	$\gamma = .9$	3.207	2.495	(.483;7.938)	1.925	1.441	(.659;5.709)	5.775	2.891	(2.243;12.042)	3.904	1.452	(1.292;5.632)
γ	$\gamma = -0.9$	-0.826	.564	(-0.994; .891)	-0.79	.658	(-0.993; .968)	-0.814	.423	(-0.99; .368)	-0.95	.12	(-0.995;-0.62)
	$\gamma = 0$	.75	.68	(-0.995; .961)	-0.776	.459	(-0.991; .589)	-0.765	.405	(-0.991; .273)	.019	.36	(-0.669; .761)
	$\gamma = .9$	.8	.664	(-0.955; .99)	.937	.366	(-0.115; .995)	.932	.254	(.245; .994)	.96	.118	(.626; .994)
ν_1	$\gamma = -0.9$	11.717	16.805	(2.009;48.632)	20.418	20.04	(2.216;69.431)	9.01	6.205	(2.425;23.158)	21.759	12.645	(5.581;48.166)
	$\gamma = 0$	10.492	10.513	(2.117;34.257)	8.668	5.607	(2.213;19.746)	10.224	6.531	(2.277;23.762)	6.623	4.218	(2.11;15.142)
	$\gamma = .9$	9.688	8.932	(2.02;28.709)	10.06	7.111	(2.016;23.682)	3.969	1.653	(2.022;7.307)	11.96	11.964	(2.026;39.17)
ν_2	$\gamma = -0.9$	11.712	13.983	(.512;38.801)	11.669	12.653	(.569;40.609)	23.707	16.876	(3.332;58.383)	15.558	9.096	(3.703;33.353)
	$\gamma = 0$	15.872	18.956	(.758;52.191)	14.336	15.906	(.884;42.401)	7.971	9.54	(.456;24.662)	14.801	16.218	(1.464;48.792)
	$\gamma = .9$	12.098	13.309	(.888;38.128)	9.264	11.506	(.797;31.168)	16.883	16.325	(1.436;48.65)	24.352	13.434	(6.35;50.644)

Table 12 – Results of the simulation study for the skew contaminated normal model with $\nu_1 = 0.1$ and $\nu_2 = 0.1$.

		Sample size											
		50			250			500			1000		
		Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_0	$\gamma = -0.9$	1.249	1.108	(-0.36;3.719)	1.772	1.393	(.499;5.09)	1.334	.608	(.548;2.344)	1.524	.77	(.792;3.198)
	$\gamma = 0$	3.483	2.472	(.684;8.886)	1.568	1.185	(.393;4.07)	2.322	.897	(.819;4.09)	1.866	1.217	(.61;4.558)
	$\gamma = .9$	1.176	1.368	(-0.45;4.137)	3.361	1.689	(.698;6.36)	2.699	1.614	(.86;6.296)	1.637	.9	(.762;3.578)
β_1	$\gamma = -0.9$	3.888	2.596	(1.172;9.27)	3.791	3.045	(1.107;11.537)	3.12	1.536	(1.296;5.445)	3.039	1.539	(1.622;6.302)
	$\gamma = 0$	5.96	4.036	(1.249;14.453)	3.321	2.227	(1.361;8.547)	4.521	1.811	(1.722;8.527)	4.056	2.637	(1.374;9.938)
	$\gamma = .9$	3.712	3.165	(.342;10.575)	7.011	3.306	(1.741;12.809)	5.06	2.931	(1.673;11.515)	2.793	1.519	(1.36;6.271)
γ	$\gamma = -0.9$	.866	.606	(-0.83; .995)	-0.624	.485	(-0.993; .615)	-0.534	.556	(-0.995; .869)	.327	.384	(-0.692; .823)
	$\gamma = 0$	.854	.581	(-0.821; .995)	.834	.483	(-0.583; .994)	.246	.579	(-0.995; .893)	.597	.341	(-0.236; .987)
	$\gamma = .9$	.876	.515	(-0.629; .995)	.886	.621	(-0.825; .995)	.94	.261	(.235; .995)	.879	.199	(.319; .989)
ν_1	$\gamma = -0.9$	.6	.27	(.078; .975)	.525	.275	(.05; .959)	.334	.187	(.033; .704)	.472	.259	(.05; .94)
	$\gamma = 0$	.565	.263	(.115; .999)	.45	.258	(.068; .957)	.436	.157	(.161; .742)	.587	.25	(.117; .962)
	$\gamma = .9$	.582	.266	(.11; .996)	.588	.212	(.245; .996)	.649	.239	(.192; .997)	.567	.289	(.099; .999)
ν_2	$\gamma = -0.9$	.311	.289	(.002; .911)	.224	.232	(.004; .77)	.08	.077	(.003; .191)	.255	.189	(.018; .662)
	$\gamma = 0$	.229	.256	(.003; .82)	.152	.165	(.003; .517)	.06	.052	(.009; .152)	.191	.202	(.013; .7)
	$\gamma = .9$	.238	.27	(.001; .845)	.107	.157	(.003; .478)	.197	.192	(.012; .612)	.316	.225	(.009; .812)

2.4.2 Simulation study II

In this study we assess the parameter recovery of our model and estimation method, in terms of bias, variance, relative bias and mean square error, when the data are generated by symmetric, asymmetric or heavy-tailed distributions. We simulated three data sets from the model

$$Y_i = I(Z_i > 0) \tag{2.10}$$

$$Z_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad i = 1, \ldots, n \tag{2.11}$$

where $\varepsilon_i \sim ST_c(0, 1, -\gamma, \nu)$ (Scenario 1), $\varepsilon_i \sim SN_c(0, 1, -\gamma)$ (Scenario 2), $\varepsilon_i \sim N(0, 1)$ (Scenario 3), $\boldsymbol{\beta} = (1, 2)$, $\nu = 3$, $\gamma = 0.9$, the covariate x_i was simulated from a $N(0, 1)$ distribution and centered in its respective mean. We considered a sample size of n=500, R=10 replicas were made. For each scenario, the three models (skew-t, skew normal and normal binary models) were fitted and the statistics: mean of the estimate of parameter (Est), standard deviation of the estimates (SD), bias of the estimates (Bias), square root of the mean square error (RMSE), as described in section 1.5.1 were calculated.

Table 13 contains the simulation results for all fitted models for each scenario. From the Table 13, we observe that when the data are generated from a skew link function and the probit regression is fitted, the bias, relative bias and RMSE are greater than for those adjusted with skew and/or heavy tails. Comparing the results of the probit model in the Scenarios 1 and 2, it is possible to observe that the bias of the regression parameters is greater when the data is generated using a heavy tail link function. The results from the skew-t and skew normal models for all scenarios produce estimates close to the real value. Finally, the estimates of the parameter ν is close to the real value when we simulated from the skew-t model. However, for the second scenario, the parameter value tends to be underestimated, but in the third scenario the estimate of ν indicate that the skew-t model approaches to the normal model, as discussed in section 1.2.3, where it was showed that skew-t with $\nu = 30$ and skew normal curves almost overlap (3a).

Table 13 – Study of parameter recovery of the skew-t, skew normal and normal binary models from different scenarios.

		Scenario											
		1				2				3			
Fitted Model	Statistic	$\beta_0 = 1$	$\beta_1 = 2$	$\gamma = 0.9$	$\nu = 3$	$\beta_0 = 1$	$\beta_1 = 2$	$\gamma = 0.9$	$\nu = \infty$	$\beta_0 = 1$	$\beta_1 = 2$	$\gamma = 0$	$\nu = \infty$
	Est	1.104	2.066	.945	4.385	1.153	2.053	.978	14.425	1.057	2.080	.035	26.126
	SD	.009	.018	.027	.810	.009	.013	.008	1.074	.006	.013	.073	.889
ST	Bias	.104	.066	.045	1.385	.153	.053	.078	-	.057	.080	.035	-
	Rel Bias	.104	.033	.050	.462	.153	.026	.087	-	.057	.040	.993	-
	RMSE	.105	.068	.053	1.604	.153	.054	.079	-	.057	.081	.081	-
	Est	1.086	1.965	.979	-	1.085	1.963	.979	-	.998	1.965	.048	-
	SD	.005	.011	.006	-	.004	.009	.008	-	.005	.008	.054	-
SN	Bias	.086	-0.035	.079	-	.085	-0.037	.079	-	-0.002	-0.035	.048	-
	Rel Bias	.086	.017	.088	-	.085	.018	.087	-	.002	.018	-	-
	RMSE	.087	.036	.079	-	.085	.038	.079	-	.005	.036	.072	-
	Est	.813	1.452	-	-	1.010	1.731	-	-	1.001	1.983	-	-
	SD	.002	.003	-	-	.002	.006	-	-	.004	.005	-	-
N	Bias	-0.187	-0.548	-	-	.010	-0.269	-	-	.001	-0.017	-	-
	Rel Bias	.187	.274	-	-	.010	.134	-	-	.001	.008	-	-
	RMSE	.187	.548	-	-	.010	.269	-	-	.004	.017	-	-

2.4.3 Residual analysis

In this section we analyzed the behavior of the residuals, presented in Section 2.3.2, under some conditions of interest. We have conducted a simulation study, considering a sample size of 1000. We simulated data sets, for each binary regression model, considering $\beta_0 = 1$, $\beta_1 = 2$, and $\gamma = -0.9$. Also, we considered: $\nu = 3$ for skew-t and skew slash distributions; $\boldsymbol{\nu} = (1, 2)$ for skew generalized t and $\boldsymbol{\nu} = (0.1, 0.15)$ for skew contaminated normal distribution.

For each simulated data, we fitted the skew normal, skew-t, skew slash, skew generalized t and skew contaminated normal models using the priors described in Section 2.4. We built suitable quantile-quantile plots for all models, where the confidence bands were made considering the methodology described in Section 2.3.2.

For all simulated data, we can see from Figures 29 - 32 that when data exhibits heavy tails, residuals obtained from the skew normal fit indicated that this model did not fit well to the data, since the residuals lying outside the confidence bands. In general, when we adjusted the correct model to the data, there is no residual lying outside the confidence bands, indicating that the model is well adjusted. For the skew generalized t model (Figure 32a) we observe some points lying outside the band. This may have happened due to the estimation problems observed in Section 2.4.1.

Comparing the adjust by another members of SMSN class, we noted that when observations are generated by skew contaminated normal distributions and we adjusted all other heavy tail link function model, the fit using these distributions are not as good as the skew contaminated normal model. For the data generated using the skew-t distribution,

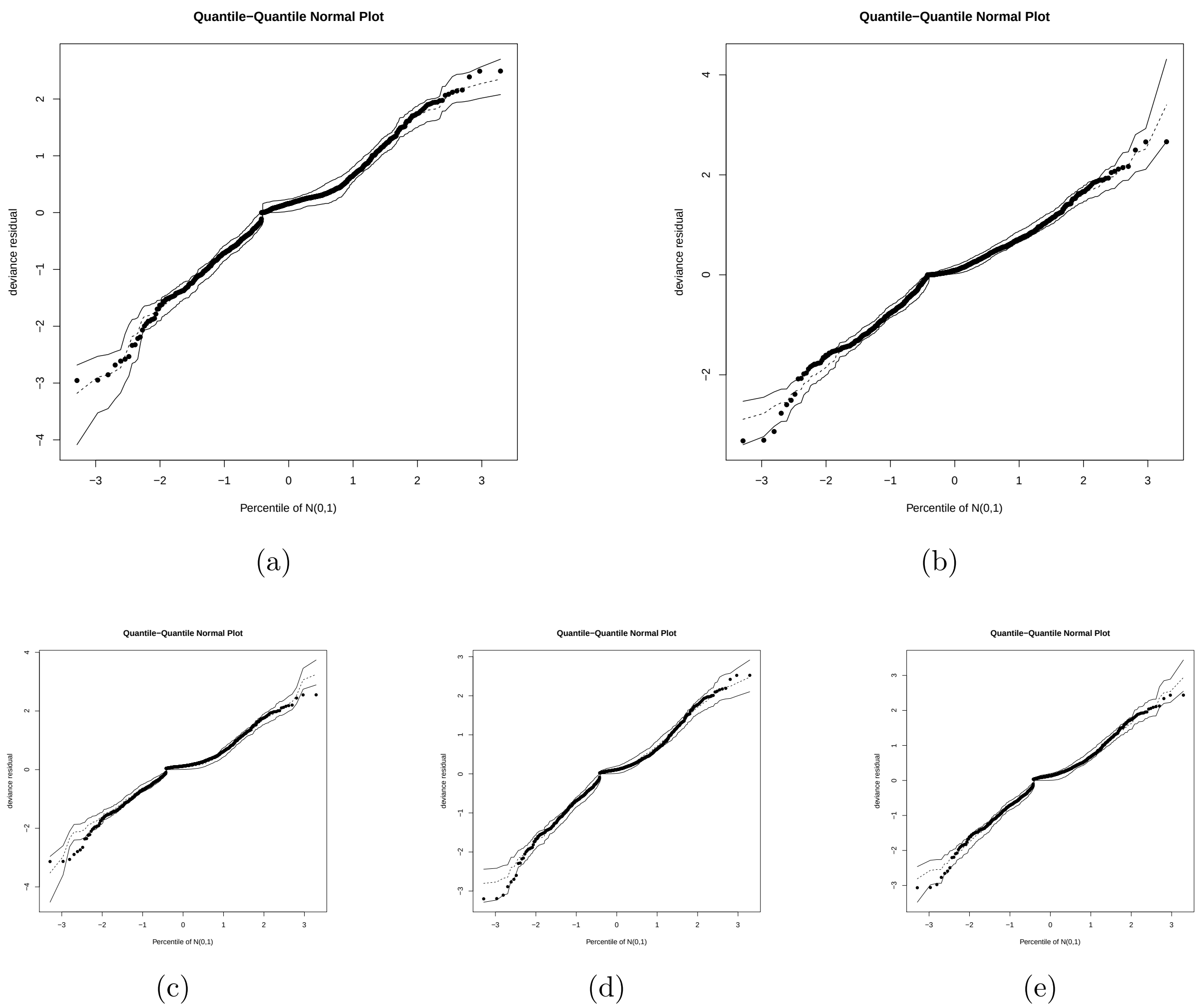
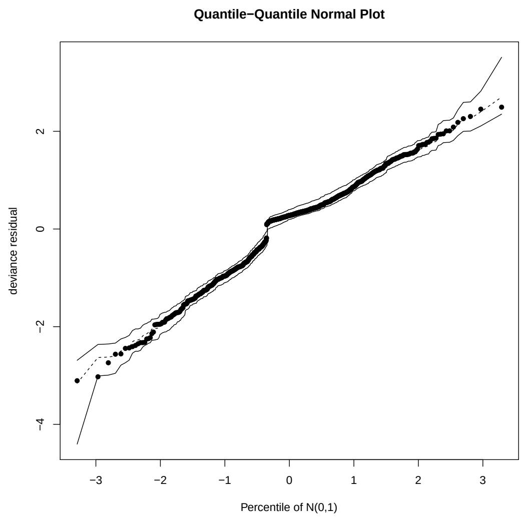
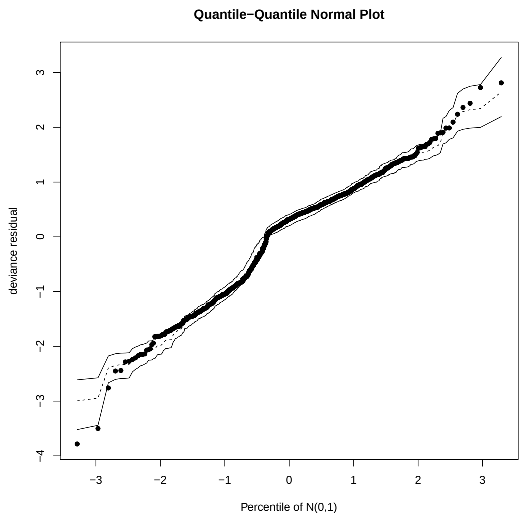


Figure 29 – QQ plots using the data set generated by the skew contaminated normal distribution and adjusting by: skew contaminated normal (a), skew normal (b), skew slash (c), skew-t (d) and skew generalized t (e)

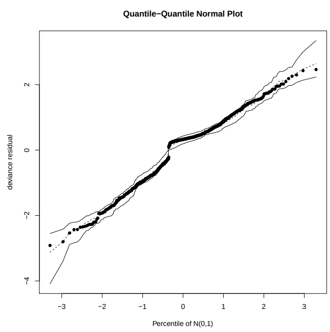
the skew contaminated normal model did not fit well the data. When observations were generated using the skew generalized t model, all models are well adjusted.



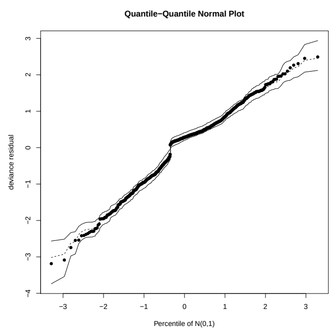
(a)



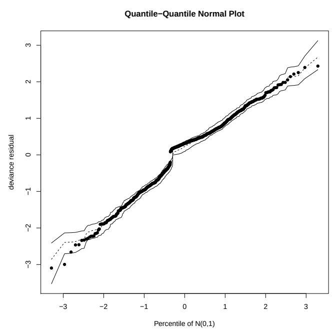
(b)



(c)



(d)



(e)

Figure 30 – QQ plots using the data set generated by the skew slash distribution and adjusting by: skew slash (a), skew normal (b), skew contaminated normal (c), skew-t (d) and skew generalized t (e)

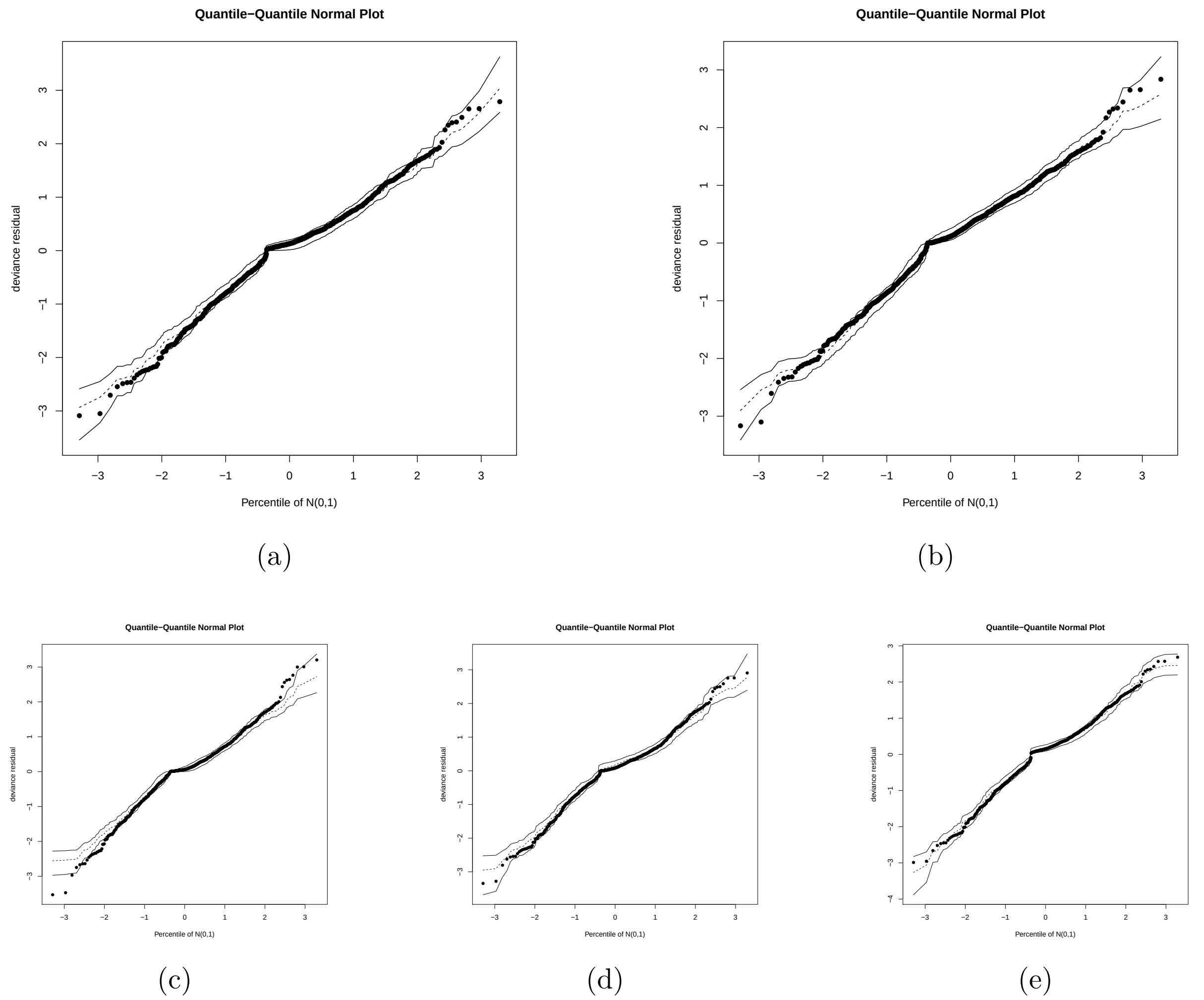
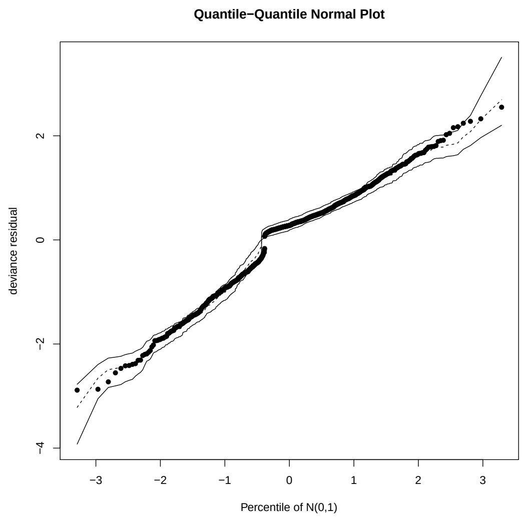
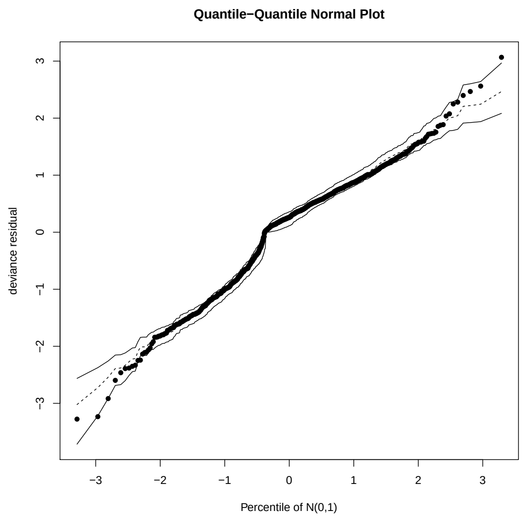


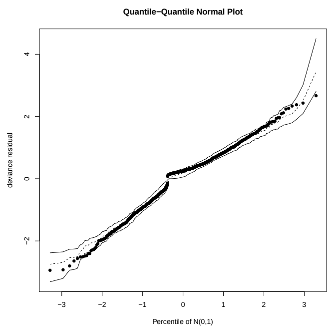
Figure 31 – QQ plots using the data set generated by the skew-t distribution and adjusting by: skew-t (a), skew normal (b), skew slash (c), skew contaminated normal (d) and skew generalized t (e)



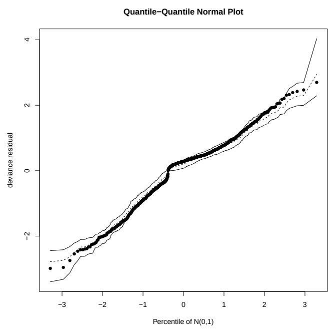
(a)



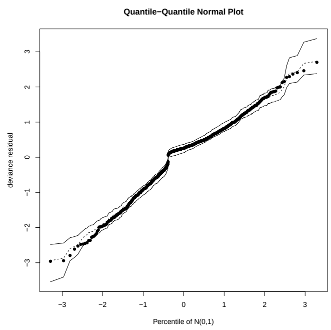
(b)



(c)



(d)



(e)

Figure 32 – QQ plots using the data set generated by the skew generalized t distribution and adjusting by: skew generalized t (a), skew normal (b), skew slash (c),skew contaminated normal (d) and skew-t (e)

2.4.4 Influence analysis

To analyze the behavior of the influence diagnostic analysis technique presented in this work using the K-L divergence measure, We have conducted a simulation study, considering a sample size of 1000. We simulated data sets, for each regression model, considering $\beta_0 = 1$, β_1, $\sigma^2 = 1$ and $\gamma = -0.9$. Also, we considered: $\nu = 3$ for skew-t and skew slash distributions; $\boldsymbol{\nu} = (1, 2)$ for skew generalized t and $\boldsymbol{\nu} = (0.1, 0.15)$ for skew contaminated normal distribution.

For each simulated data, we fitted the skew-normal, skew-t, skew slash, skew generalized t and skew contaminated normal models using the priors described in Section 2.4. For each fitted model, plots with K-L divergence are made. As described in section 1.4.5, an influential observation is considered if $p_i \geq .8$, that is, if $K(P, P_{(-i)}) \geq .2231436$.

In general, we can see from figures 33 - 36 that the skew normal model indicates possible influential observations that are no influential for the other models, indicating that the models considering heavy tails distributions accommodate properly all observations, differently from the skew-normal model. As we can see, the observations that are higher than all others in the skew slash, skew generalized t and skew contaminated normal models, but are still smaller than the cut-off point appear as influential in the skew normal model. For the case when we simulated from the skew-t model, the observations from the skew-normal model are close to the cut-off line (Figure 33b). However, for the skew-t model these points are distant to the cut-off line (Figure 33a). It can also be noted that when the data are simulated from skew-t, skew slash and skew generalized t model and we fitted the skew contaminated model, at least one observation is considered as potentially influential. This does not happen when the data is simulated from the skew contaminated normal distribution. This indicates that the skew contaminated model does not accommodate so well the extreme observations, compared with the other models.

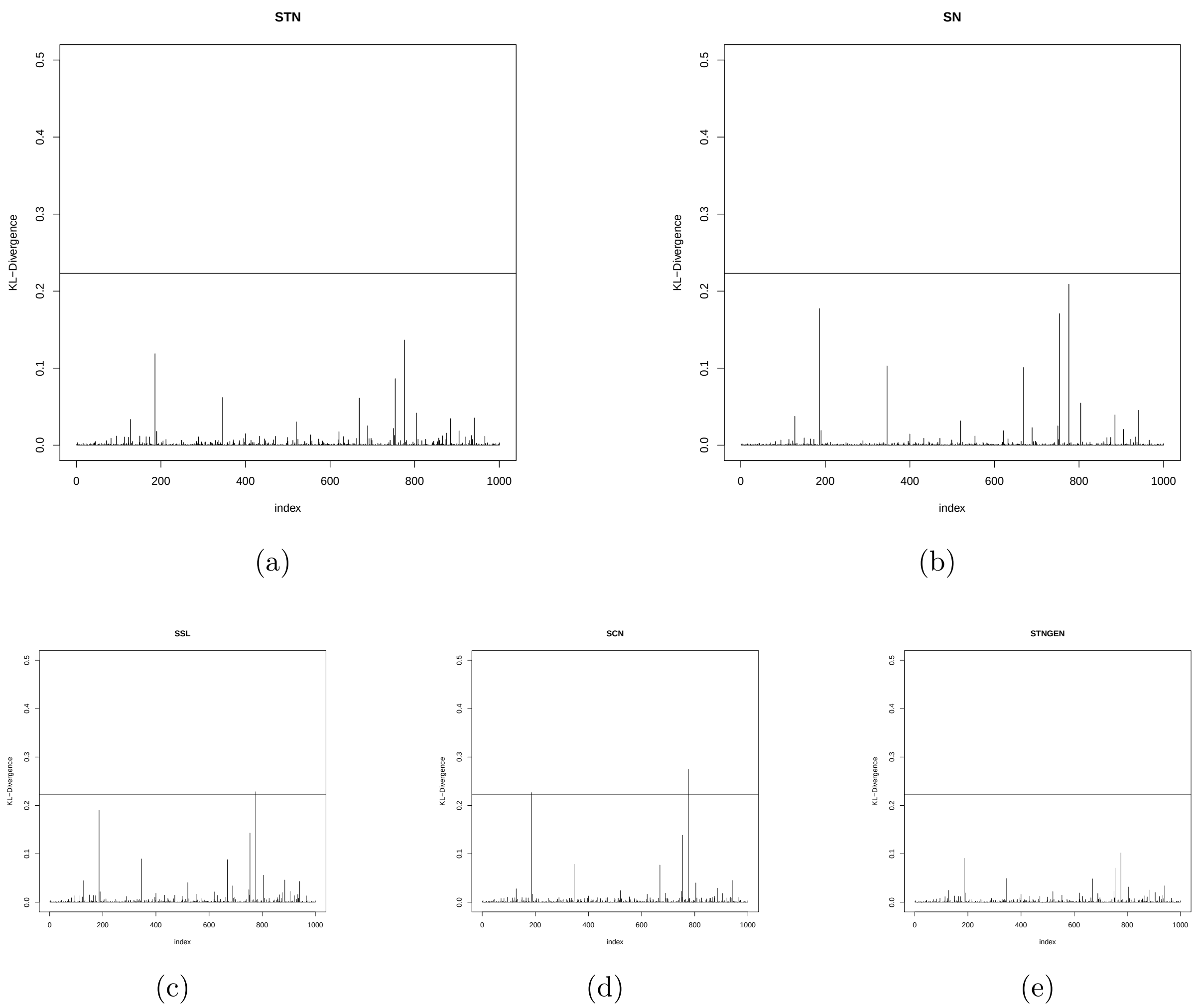


Figure 33 – Index plots of $K(P, P_{(i)})$ for the data set generated by the skew-t distribution adjusting by: skew-t (a), skew normal (b), skew slash (c), skew contaminated normal (d) and skew generalized t (e)

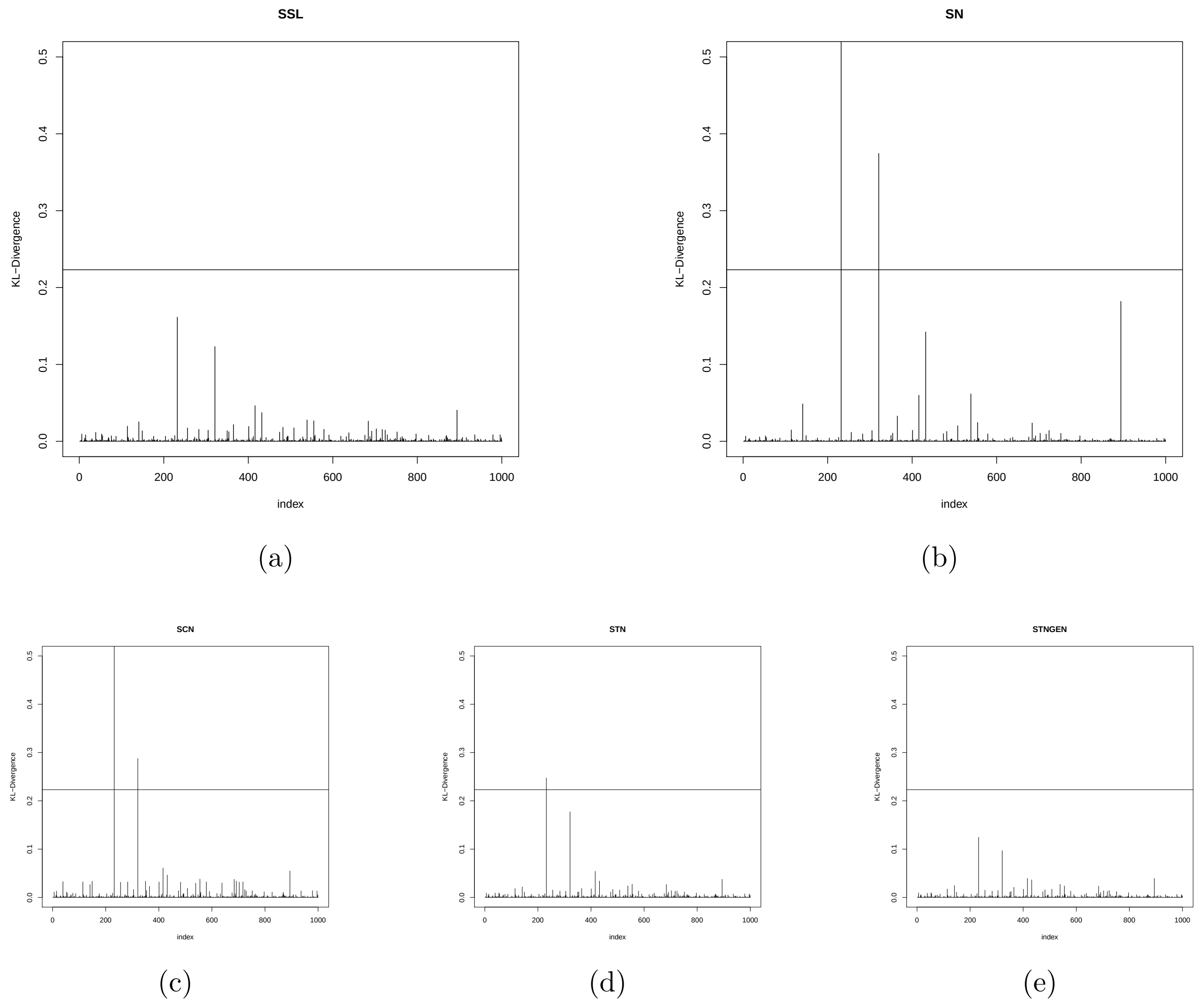
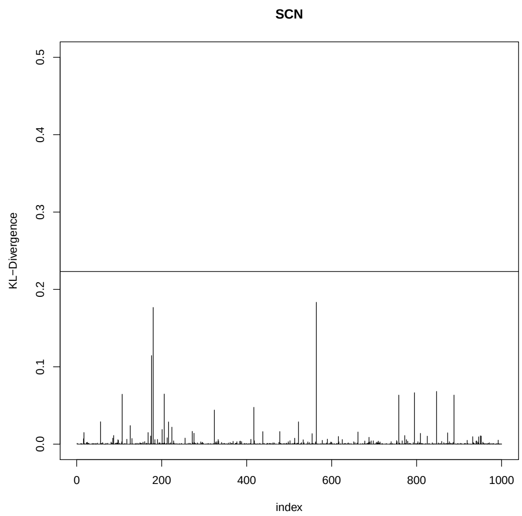
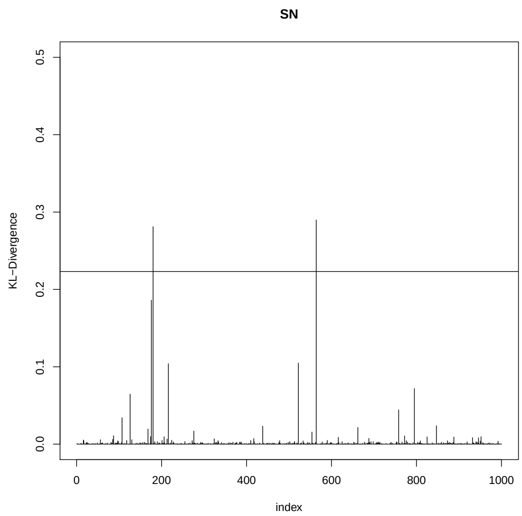


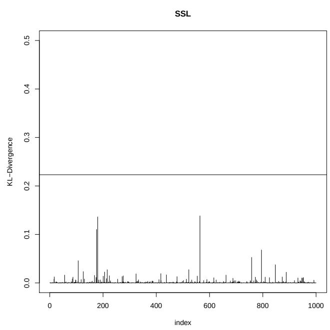
Figure 34 – Index plots of $K(P, P_{(i)})$ for the data set generated by the skew slash distribution adjusting by: skew slash (a), skew normal (b), skew contaminated normal (c), skew-t (d) and skew generalized t (e)



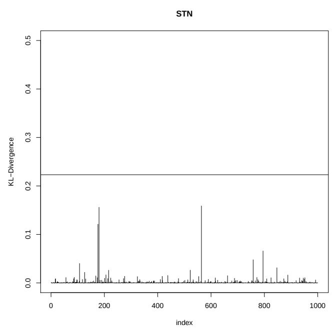
(a)



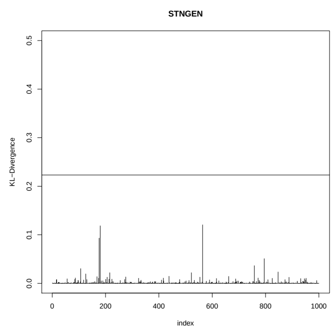
(b)



(c)



(d)



(e)

Figure 35 – Index plots of $K(P, P_{(i)})$ for the data set generated by the skew contaminated normal distribution adjusting by: skew contaminated normal (a), skew normal (b), skew slash (c), skew-t (d) and skew generalized t (e)

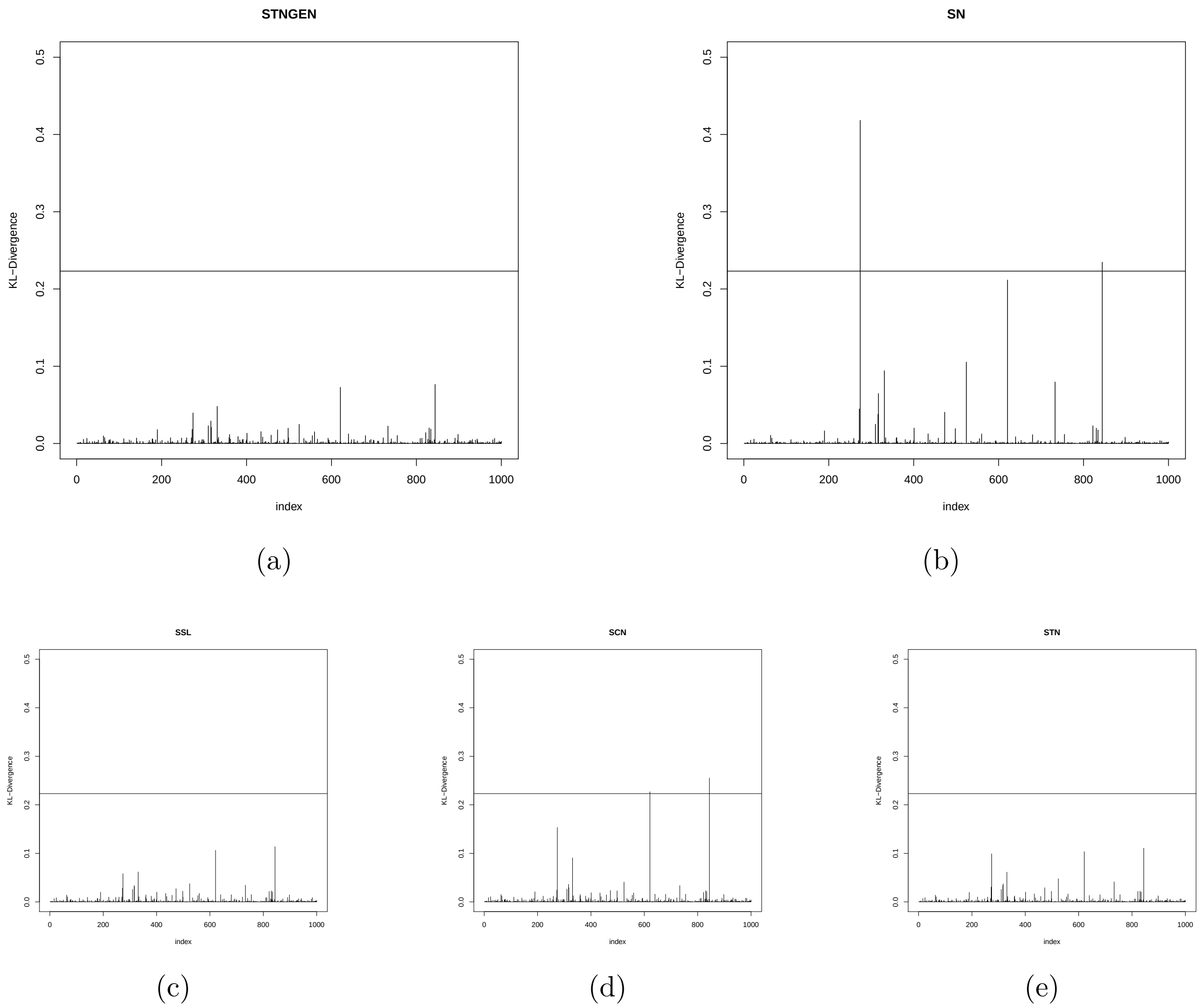


Figure 36 – Index plots of $K(P, P_{(i)})$ for the data set generated by the skew generalized t distribution adjusting by: skew generalized t (a), skew normal (b), skew slash (c), skew contaminated normal (d) and skew-t (e)

2.5 Application

2.5.1 Body Fat Data

We analyze, using the developed models and the probit one, the Body Fat Data collected by (JOHNSON, 1996). This data consist on a sample of 252 body measurements and body fat percentage collected from men who were diagnosed as being normal and abnormal in body fat. The objective of this example is to show that this class of models can provide better fits than the usual binary models. We follow (KIM, 2002) and considered as the binary response Y_i the body fat status, considered a status of normal in body fat if $Y_i = 0$ and abnormal if $Y_i = 1$. A normal status of body fat is considered when the man has less than 15% percentage of body fat and abnormal otherwise. We fitted the model

$$
\begin{aligned}
Y_i &= I(Z_i > 0) \\
Z_i &= \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon_i \quad i = 1, \ldots, 252
\end{aligned}
$$

where x_1 is the age, x_2 is the weight (kg) and x_3 is the height (cm), centered in their respective mean. We fitted six models, assuming that: $\varepsilon \overset{iid}{\sim} ST_c(0, 1, -\gamma, \nu)$, or , $\varepsilon \overset{iid}{\sim} SSL_c(0, 1, -\gamma, \nu)$, or $\varepsilon \overset{iid}{\sim} SCN_c(0, 1, -\gamma, \nu_1, \nu_2)$, or $\varepsilon \overset{iid}{\sim} SGT_c(0, 1, -\gamma, \nu_1, \nu_2)$, or $\varepsilon \overset{iid}{\sim} SN_c(0, 1, -\gamma)$, or $\varepsilon \overset{iid}{\sim} N(0, 1)$, that we denote, respectively, by ST, SSL, SCN, SGT, SN and N. The parameters for the MCMC algorithm and the adopted prior distributions were the same used in the simulation study described in 2.4.1.

Similarly to Section 1.4.4, the model comparison criteria was calculated used the likelihood

$$
L(\boldsymbol{\beta}, \gamma, \boldsymbol{\nu}|\boldsymbol{y}) = \prod_{i=1}^{n} (F({X_i}^t\boldsymbol{\beta}|\gamma, \boldsymbol{\nu}))^{y_i} (1 - F({X_i}^t\boldsymbol{\beta}|\gamma, \boldsymbol{\nu}))^{1-y_i}. \tag{2.12}
$$

Table 14 presents the statistics for model comparison. The skew contaminated normal model (SCN) was selected by DIC and LPML and the skew normal model was select by EAIC and EBIC criteria. From this table we can noted that the skew-t, skew slash, skew contaminated normal and skew normal models outperform the normal model. Since the skew normal model was selected by EAIC and EBIC criteria, we analyzed the posterior distributions of ν_1 and ν_2 from the skew contaminated normal model. From Figure 38, both posterior distributions indicated that the skew contaminated normal and skew model are not equivalent.

In Table 15 we have the estimates of all fitted models. For all models, weight and age have positive influence on the probability of present abnormal body fat while height has a negative influence. We can also note the large credibility intervals for γ in all asymmetric link models. In this way, the posterior distribution of γ is presented in Figure 37 for each fitted model. As we can see, the posterior distributions indicate that

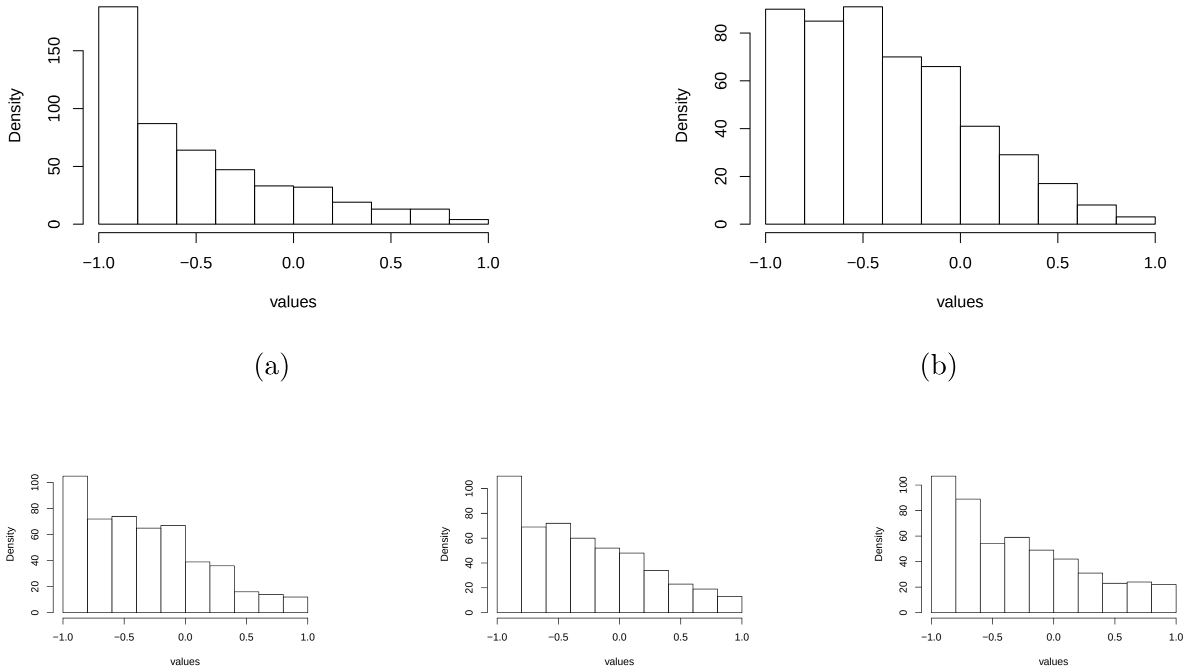


Figure 37 – Posterior distribution of γ for the: skew contaminated normal (a), skew normal (b), skew slash (c), skew-t (d) and skew generalized t (e) models.

the rate that probability goes to 1 is higher than it goes to 0, since γ is concentrated towards negative values. In this way, we conclude that asymmetrical link function is more appropriate for the Body Fat data.

The analysis of influential observations, presented in Figure 40, indicated that there are two influential observations for the skew-normal and skew slash models, only one for the skew contaminated normal model and for the skew-t and skew generalized t there is no influential observation. However, for the probit model, several observations are influential, indicating that this model is not adequate. From Figure 39, QQ plot with envelopes for all fitted models are shown. It is possible to see that for all models there are some points lying outside the confidence bands, possibly caused by the influential observations. Considering the selection criteria, posterior distribution of ν_1 and ν_2 for the skew contaminated normal model, residual and influence analysis, the skew contaminated normal model provides the best fit among all other models.

Table 14 – **Body fat dataset**: Statistics of model comparison

	Model					
criterion	ST	SSL	SCN	SGT	SN	N
EAIC	219.82	219.77	220.64	220.80	217.29	2722.31
EBIC	241.00	240.95	245.34	241.98	234.94	2736.43
DIC	211.52	211.15	204.99	212.45	211.01	5224.97
LPML	-106.28	-106.60	-105.83	-106.77	-106.31	-1357.15

Table 15 – **Body fat dataset**: Bayesian estimates for the skew-t, skew slash, skew contaminated normal, skew generalized t and skew normal models.

	Model					
	ST			**SSL**		
	Est	SD	%95 HPD	Est	SD	%95 HPD
β_0	.973	.203	[.651; 1.41]	1.006	.307	[.649; 1.771]
β_1	.025	.01	[.004; .042]	.025	.013	[.003; .05]
β_2	.061	.014	[.038; .093]	.063	.021	[.036; .105]
β_3	-0.25	.084	[-0.431; -0.108]	-0.253	.109	[-0.466; -0.083]
γ	-0.317	.516	[-0.995; .658]	-0.334	.484	[-0.995; .617]
ν	19.254	19.448	[1.031; 57.703]	17.89	21.862	[.539; 61.855]
	SCN			**SGT**		
	Est	SD	%95 HPD	Est	SD	%95 HPD
β_0	2.645	1.034	[.854; 4.418]	2.238	.859	[.627; 3.834]
β_1	.067	.038	[.009; .139]	.059	.034	[.01; .126]
β_2	.16	.059	[.051; .251]	.142	.054	[.041; .243]
β_3	-0.641	.246	[-0.998; -0.187]	-0.575	.232	[-1.035; -0.19]
γ	-0.513	.467	[-0.995; .469]	-0.296	.546	[-0.995; .783]
ν_1	.73	.189	[.368; 1]	5.376	5.417	[1.236; 14.747]
ν_2	.139	.174	[.013; .553]	21.998	18.655	[1.211; 57.993]
	SN			**N**		
	Est	SD	%95 HPD	Est	SD	%95 HPD
β_0	.845	.12	[.619; 1.097]	.886	.179	[.629; 1.183]
β_1	.022	.008	[.007; .04]	.022	.012	[.004; .039]
β_2	.053	.008	[.038; .068]	.056	.01	[.044; .072]
β_3	-0.211	.052	[-0.309; -0.116]	-0.229	.056	[-0.342; -0.134]
γ	-0.371	.428	[-0.995; .41]	-	-	-

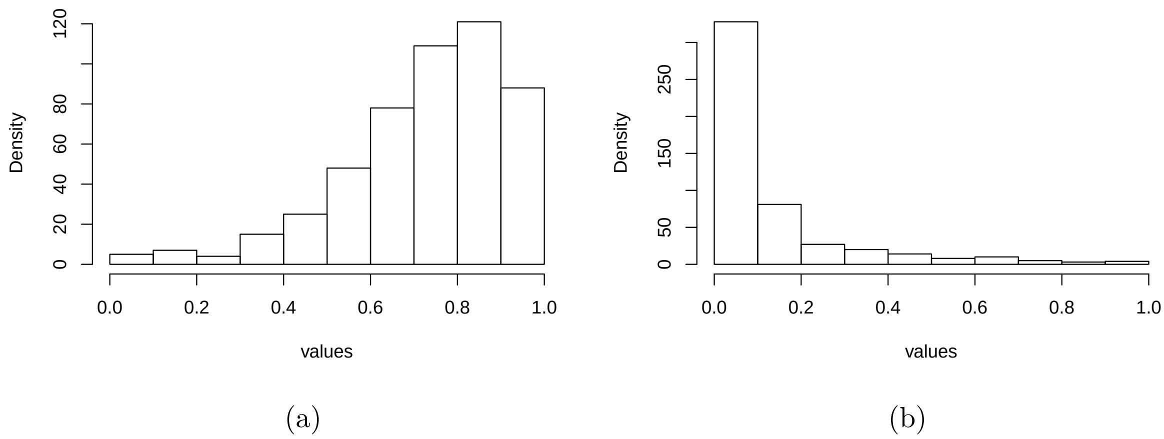
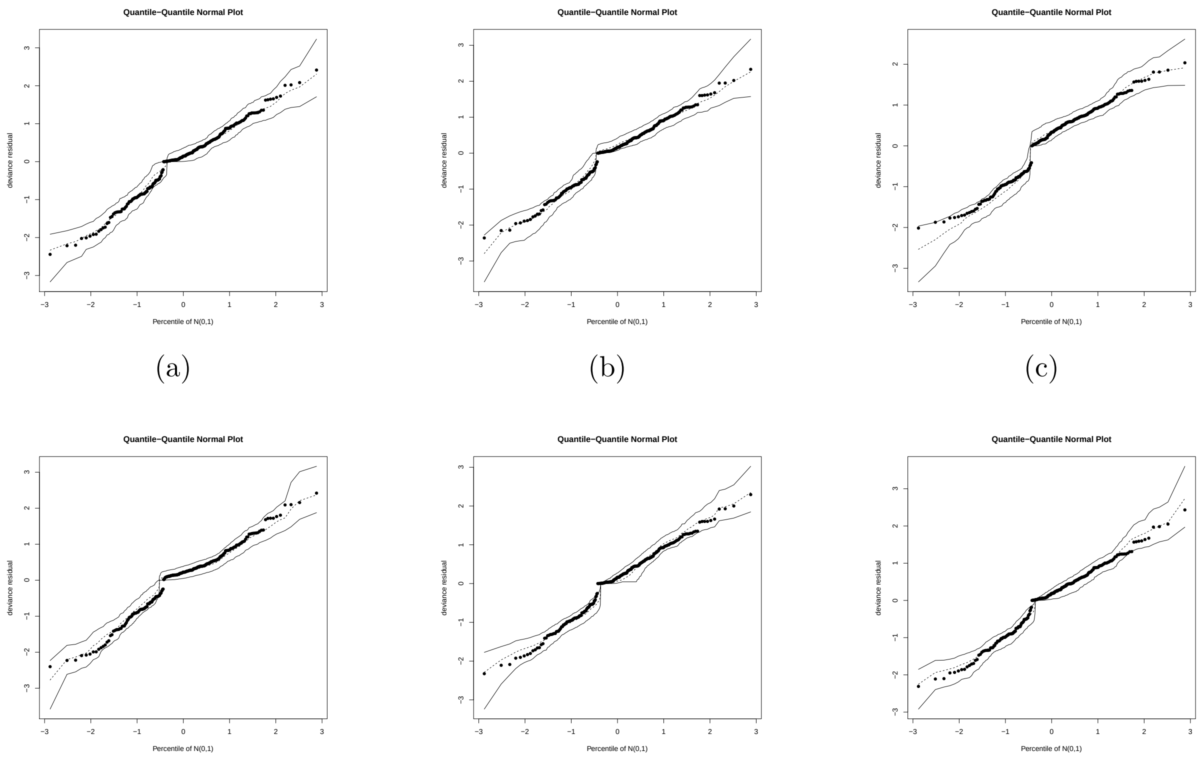


Figure 38 – Posterior distribution of ν_1 (a) and ν_2 (b) for the skew contaminated normal model.



(a) (b) (c) (d) (e) (f)

Figure 39 – QQ plots with envelopes for the Body Fat data using the link functions: skew-t (a), skew slash (b), skew contaminated normal (c), skew generalized t (d), skew normal (e) and normal (f).

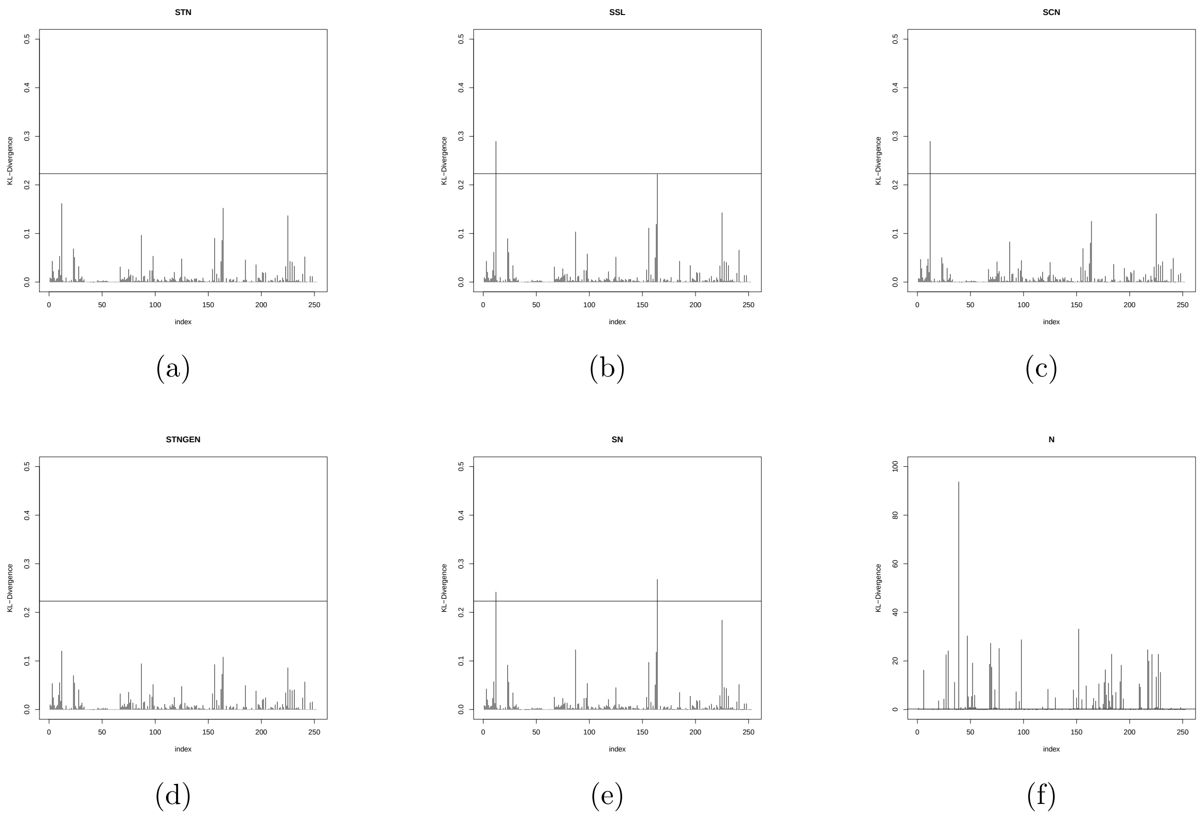


Figure 40 – Index plots of $K(P, P_{(i)})$ for the Body Fat data using the link functions: skew-t (a), skew slash (b), skew contaminated normal (c), skew generalized t (d), skew normal (e) and normal (f).

2.5.2 Beetle Mortality data

We analyze, using the developed models and the probit one, the well-known Beetle Mortality data. This data contains the numbers of beetles killed after 5 hour exposure to carbon disulphide at N = 8 different concentrations. (CZADO, 1994) and (BAZáN; BOLFARINE; BRANCO, 2010) concluded that a asymmetric link is more appropriate and improves the fit comparing with logit and probit models. The objective of this example is to show that this class of models can provide better fits than the usual binary models.

We fitted six models, considering the fixed covariate as the logarithm of the concentration of carbon disulphide and we assumed that the following distributions for the link function: $ST_c(0, 1, -\gamma, \nu)$, $SSL_c(0, 1, -\gamma, \nu)$, $SCN_c(0, 1, -\gamma, \nu_1, \nu_2)$, $SGT_c(0, 1, -\gamma, \nu_1, \nu_2)$, $SN_c(0, 1, -\gamma)$, $N(0, 1)$, that we denote, respectively, by ST, SSL, SCN, SGT, SN and N. The parameters for the MCMC algorithm and the adopted prior distributions were the same used in the simulation study described in 2.4.1.

Table 16 presents the statistics for model comparison. The skew-normal model (SN) was selected by all criteria. In Table 17 we have the estimates of all fitted models and from Figure 41 we can see the estimated link function for each fitted model, compared with the skew-normal and normal models. From these figures, we can see that an asymmetric link model provides better fits than the probit model. The residual analysis for the Beetle Mortality data, available in Figure 42 shows that, with exception of the probit model, all fitted models are well adjusted to the data. From the influence analysis in Figure 43, we noted that the proportion of death beetles for the log concentration of carbon disulphide equal to 1.8113 appears to the an influent observation.

Table 16 – **Beetle Mortality data**: Statistics of model comparison

	Model					
criterion	ST	SSL	SCN	SGT	SN	N
EAIC	376.13	376.57	377.81	375.99	**373.85**	661.16
EBIC b	392.84	393.28	398.69	392.69	**386.38**	669.51
DIC	369.97	371.76	370.38	370.30	**369.75**	942.95
LPML	-185.37	-185.97	-185.31	-185.33	**−185.26**	-328.58

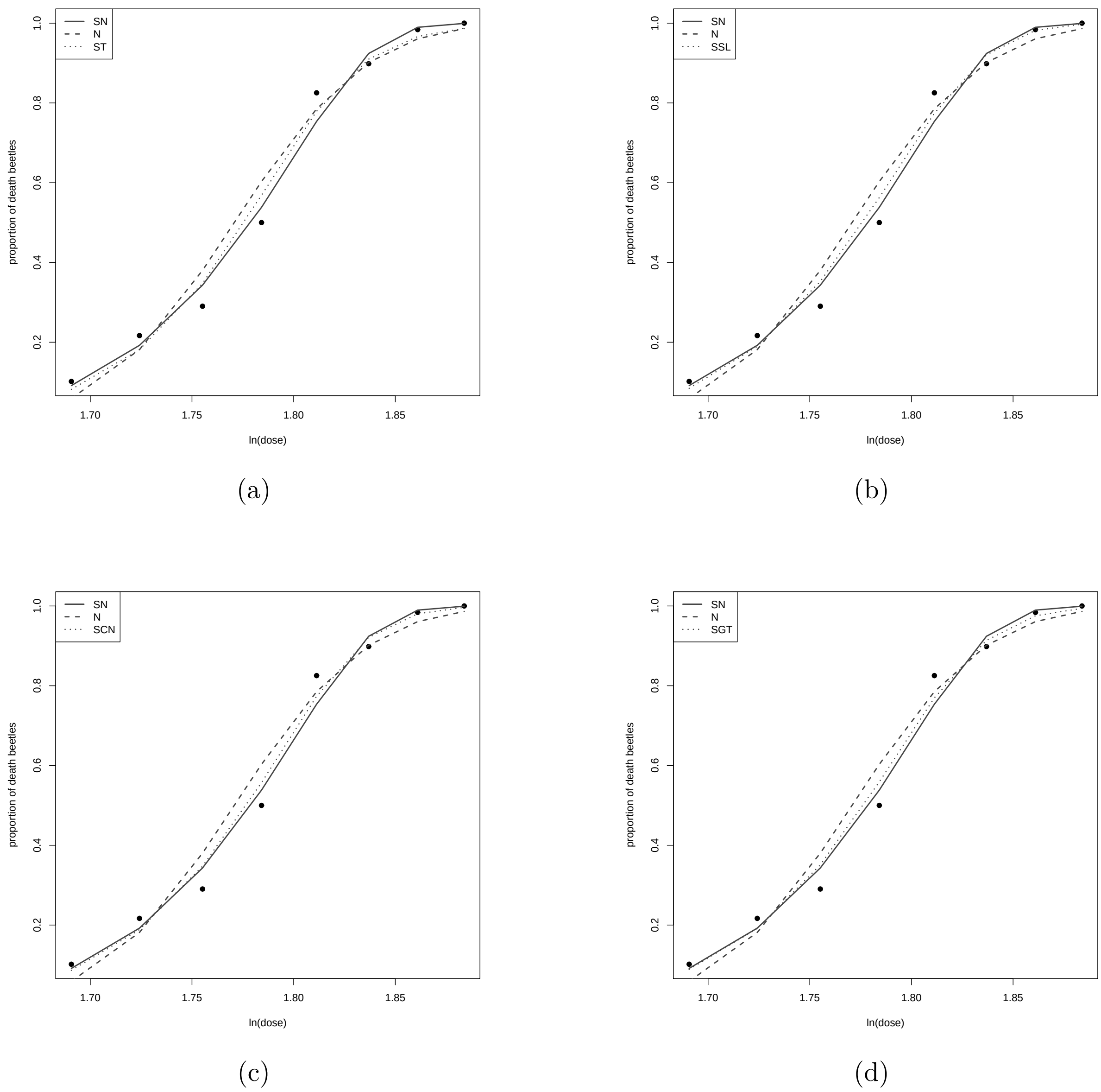


Figure 41 – Fitted probabilities of beetle's death considering the skew normal, normal and the following link functions: skew-t (a), skew slash (b), skew contaminated normal (c), skew generalized t (d).

Table 17 – **Beetle Mortality data**: Bayesian estimates for the skew-t, skew slash, skew contaminated normal, skew generalized t and skew normal models.

	Model					
		ST			**SSL**	
	Est	SD	%95 HPD	Est	SD	%95 HPD
β_0	-35.746	3.786	[-42.623; -29.163]	-33.938	4.224	[-42.004; -24.342]
β_1	20.191	2.125	[16.512; 23.992]	19.179	2.358	[13.909; 23.749]
γ	-0.735	.236	[-0.995; -0.281]	-0.702	.198	[-0.991; -0.306]
ν	20.45	17.264	[1.924; 55.766]	21.342	20.97	[1.581; 66.351]
		SCN			**SGT**	
	Est	SD	%95 HPD	Est	SD	%95 HPD
β_0	-38.104	5.275	[-47.378; -27.209]	-14.153	2.695	[-18.852; -8.695]
β_1	21.532	2.974	[15.447; 26.792]	7.998	1.522	[4.907; 10.639]
γ	-0.747	.188	[-0.995; -0.356]	-0.745	.204	[-0.995; -0.349]
ν_1	.49	.266	[.001; .924]	21.618	13.033	[2.91; 47.468]
ν_2	.571	.23	[.222; .996]	3.88	2.759	[.253; 9.195]
		SN			**N**	
	Est	SD	%95 HPD	Est	SD	%95 HPD
β_0	-32.316	2.799	[-37.795; -27.244]	-34.565	3.074	[-40.87; -29.488]
β_1	18.257	1.562	[15.445; 21.346]	19.519	1.731	[16.662; 23.077]
γ	-0.672	.204	[-0.98; -0.287]	-	-	-

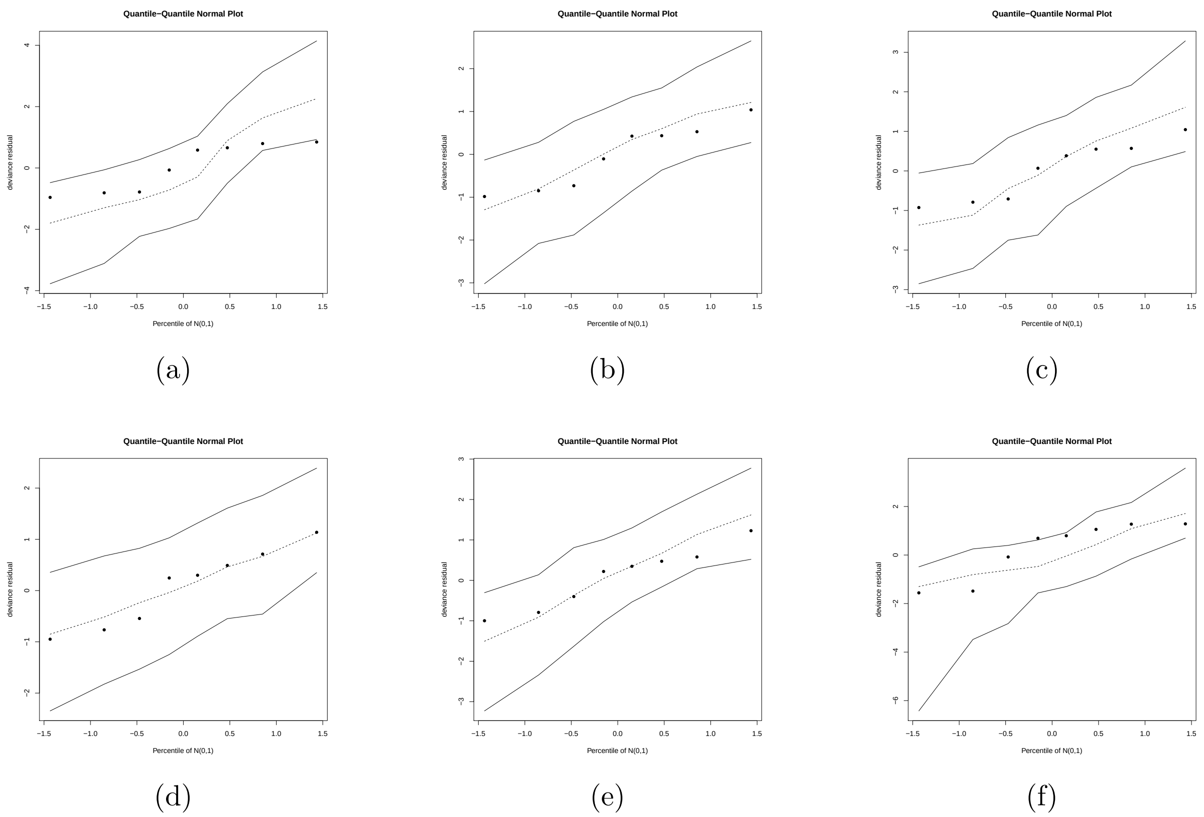


Figure 42 – QQ plots with envelopes for the Beetle Mortality data using the link functions: skew-t (a), skew slash (b), skew contaminated normal (c), skew generalized t (d), skew normal (e) and normal (f).

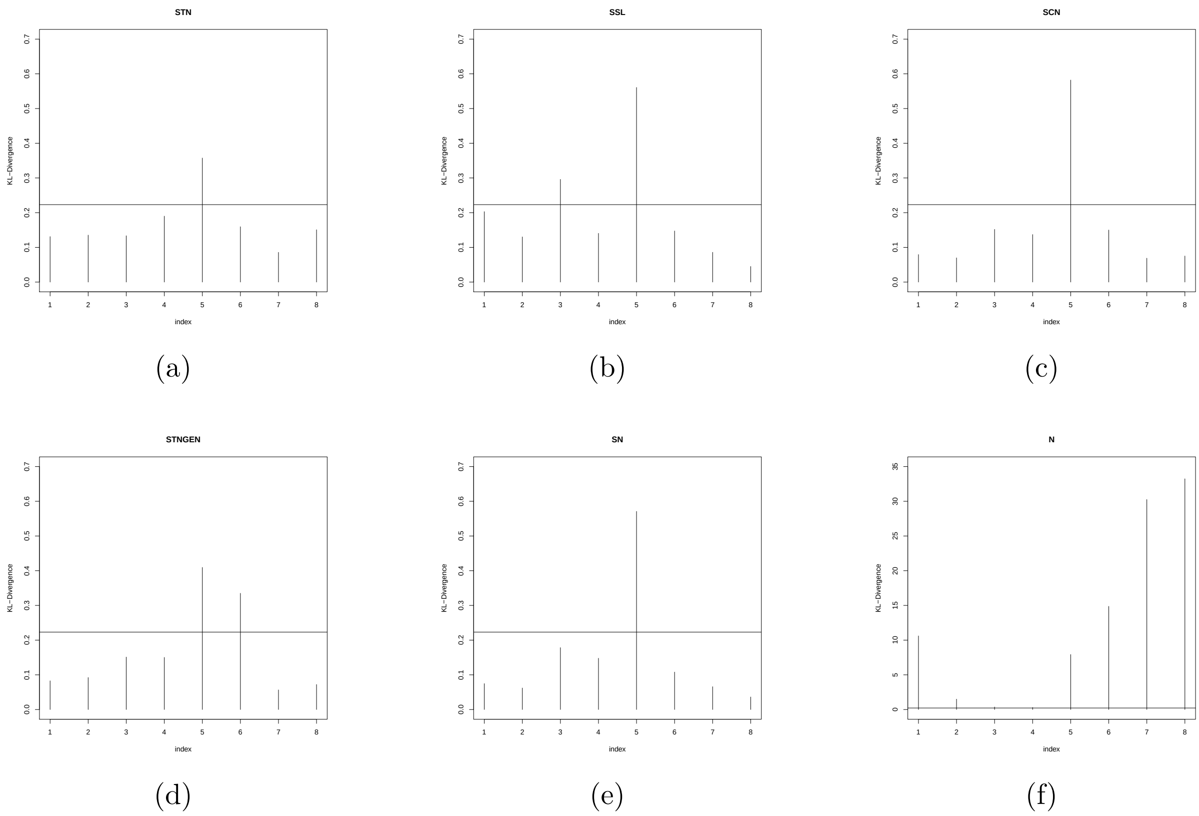


Figure 43 – Index plots of $K(P, P_{(i)})$ for the Beetle Mortality data using the link functions: skew-t (a), skew slash (b), skew contaminated normal (c), skew generalized t (d), skew normal (e) and normal (f).

2.6 Conclusions

In this chapter we proposed a new class of link functions based on the SMSN class under the centered parameterization. This class of link functions include symmetrical, asymmetrical and robust link functions. We performed Bayesian estimation using latent variables to described the binary model. Some methods of residual analysis for binary data was discussed, and simulation studies were performed, evaluating parameter recovery, residual and influence analysis. The first simulation study showed some problems in the accuracy of the estimates of ν for the binary model, specially for the skew contaminated normal and skew generalized t models. However for the skew-t and skew slash models, when true value of ν indicated heavy tails the parameter was appropriate recovered. Also, we noted as sample size increases, the estimates of all parameter tend to be closer to real values. In the second study, we simulated data using normal, skew-normal and skew-t link functions and we fitted the skew-t, skew-normal and normal models. For this study we observed that when data were simulated using an asymmetric and/or heavy tail link functions and the probit model was fitted, the estimates of $\boldsymbol{\beta}$ were biased. For the residual and influence analysis studies, when the data was simulated using a heavy tail distribution and the skew normal model is fitted, the residuals tend to lying outside the confidence bands and we observe some observations as influentials.

As in the linear model, an application was made in a study on the body fat percentage that indicated the skew and heavy tail link was preferred to the usual probit model. For the Beetle Mortality data we conclude that the specification of an asymmetric link provide better fit than the usual probit model.

3 Bivariate regression models for continuous-binary response based on the skew scale mixture of normal distribution class under the centered parameterization

3.1 Introduction

Situations where the response variable is either continuous or binary are quite common in several fields of knowledge, specially in social and health science. In general, in these studies, multiple responses are collected in order to characterize or evaluate their relationships with some covariates of interest. For example, to evaluate the efficacy of an experimental treatment on vision for macular degeneration, a study was performed in patients with age-related macular degeneration (see (GUYER et al., 1997)). For each patient it was evaluated their patient's visual acuity in the beginning and after one year of study. This acuity is measured by counting how many letters of a standardized vision chart are corrected read. These charts display line letters of decreasing size that the patient must read from the top (large letters) to bottom (small letters). In this study, two outcomes were obtained in order to evaluate the efficacy of the treatment: the binary outcome was defined as the loss of at least three lines of vision at one year compared with their baseline performance and the continuous outcome are defined as the difference between patient's visual acuity from one year and the beginning of the study.

The usual modeling strategy for this type of data is to perform separate analysis for each response variable. As noted by (TEIXEIRA-PINTO; NORMAND, 2009) this strategy is less efficient, since it ignores the extra information contained in the correlation among the outcomes. In the bivariate context, the most common approach is to model each variable separately, ignoring the potential correlation between them, or using the factorization methods as discussed in (COX; WERMUTH, 1992), which consists to write the likelihood as the product of the marginal distribution of one of the outcomes and the conditional distribution of second outcome given the first one. (FITZMAURICE; LAIRD, 1995) and (CATALANO; RYAN, 1992) extend this approach to situations with clustered data, in which the method proposed in (FITZMAURICE; LAIRD, 1995) is based on the general location model of (OLKIN; TATE, 1961).

Another approach to deal with bivariate outcomes is presented in (SAMMEL; RYAN; LEGLER, 1997), where both outcomes shared a random effect that induces the correlation between them, and conditionally to this random effect, the two outcomes are independent. (DUNSON, 2000) also used the latent variable modeling, but the covariates are not included in the model through the latent variable, as proposed in (SAMMEL; RYAN; LEGLER, 1997). (TEIXEIRA-PINTO; NORMAND, 2009), on the other hand, review the different approaches to deal with binary and continuous outcomes and proposed a new method, also based on the latent variable, which circumvents the identifiability problem present in Dunson's approach.

Recently, some authors developed models for bivariate discrete and continuous outcomes by using copulas. From this method the joint distribution is constructed by specifying marginal regression models for the outcomes and combining them via a copula. Recent references include (LEON; WU, 2011), (CHEN; HANSON, 2017) and (ZILKO; KUROWICKA, 2016).

All models cited earlier are based on the normality assumption for the continuous response and symmetrical link functions to the binary outcome. However, some data sets may not satisfy these assumptions. Motivated by these characteristics, (TEIMOURIAN et al., 2015) has developed a model considering an ordinal and a skewed continuous responses using the factorization method ((COX; WERMUTH, 1992)). For the ordinal response, the logit link function was considered. However, the model of (TEIMOURIAN et al., 2015) cannot handle with the possibility of asymmetrical link functions. All these models do not include the cases where the probability of success of the binary response increases at a different rate than decreases and continuous response with heavy tails. Then, we developed a regression model for bivariate continuous and binary responses using latent variable for incorporate the correlation between the responses, assuming the possibility of heavy tails and asymmetry for both continuous response and link function.

3.2 Model formulation

Consider a study with two responses of interest, in which one is continuous and the other is binary. To define the regression model consider $(Y_i, T_i)'$, $i = 1, \ldots, n$, where Y_i is the continuous response and $T_i \in \{0, 1\}$ the binary variable. Let $\boldsymbol{X}_c = (1, \boldsymbol{X}_{c1}, \boldsymbol{X}_{c2}, \ldots, \boldsymbol{X}_{c(p-1)})^t$ a $p \times n$ design matrix of fixed covariates associated with the binary response and $\boldsymbol{\beta}_c = (\beta_{c0}, \beta_{c1}, \ldots, \beta_{c(p-1)})^t$ a $p \times 1$ vector of regression coefficients associated to the continuous variable, $\boldsymbol{X}_b = (1, \boldsymbol{X}_{b1}, \boldsymbol{X}_{b2}, \ldots, \boldsymbol{X}_{b(p-1)})^t$ a $p \times n$ design matrix of fixed covariates associated with the binary response and $\boldsymbol{\beta}_b = (\beta_{b0}, \beta_{b1}, \ldots, \beta_{b(p-1)})^t$ a $p \times 1$ vector of regression coefficients associated to the binary variable. Let W_i a latent variable, so the bivariate regression model assumes that

$$
\begin{aligned}
P(T_i = 1) = p_i & \\
p_i = F(\eta_i + W_i) = F(X_{bi}^t \boldsymbol{\beta}_b + W_i), \quad i = 1, \ldots, n \\
Y_i = X_{ci}^t \boldsymbol{\beta}_c + W_i + \varepsilon_i \quad i = 1, \ldots, n
\end{aligned}
\tag{3.1}
$$

where $F(\cdot|\gamma_b, \boldsymbol{\nu}_b)$ is the cdf of a member of the SMSN class of distribution under the centered parameterization, as presented in Section 1.2.2, $\varepsilon_i \overset{iid}{\sim} SMSN_c(0, \sigma_c^2, \gamma_c, G, \boldsymbol{\nu}_c)$ and $W_i \overset{iid}{\sim} SMSN_c(0, \sigma_w^2, \gamma_w, G, \boldsymbol{\nu}_w)$.

The latent variable W_i induces the correlation and it is assumed that, conditioned to this variable, the two outcomes are independent. Using the formulation given in 3.1, and after some algebra, the correlation between the binary and continuous variable can be expressed as

$$
\begin{aligned}
Cor(T_i, Y_i) &= \frac{E(W_i F(\eta_i + W_i))}{\sqrt{Var(T_i)}\sqrt{\sigma_c^2 + \sigma_w^2}} \\
Var(T_i) &= E(F(\eta_i + W_i)(1 - F(\eta_i + W_i)) + Var(F(\eta_i + W_i))
\end{aligned}
\tag{3.2}
$$

This correlation has not a closed form, but it possible to obtain approximations through simulations. The two responses have lower correlations when σ_w^2 approaches to 0. On the other hand, as σ_w^2 increases the correlation between the continuous and binary response also increases.

Following the ideas presented in Sections 1.4 and 2.2, the model can be stochastically represented as

$$
\begin{aligned}
& Y_i \overset{d}{=} X_{ci}^t \boldsymbol{\beta}_c + W_i + U_{ci}{}^{-1/2}(\Delta_c(H_{ci} - b) + \sqrt{\tau_c} V_{ci}) \\
& T_i = I(Z_i > 0) = \begin{cases} 1 & \text{if } Z_i > 0 \\ 0 & \text{if } Z_i \leq 0 \end{cases} \\
& Z_i \overset{d}{=} X_{bi}^t \boldsymbol{\beta}_b + W_i + U_{bi}{}^{-1/2}(\Delta_b(b - H_{bi}) + \sqrt{\tau_b} V_{bi}) \\
& W_i \overset{d}{=} U_{wi}{}^{-1/2}(\Delta_w(H_{wi} - b) + \sqrt{\tau_w} V_{wi}) \\
& V_{ci} \sim N(0,1) \quad H_{ci} \sim HN(0,1) \quad U_{ci} \sim G(.|\boldsymbol{\nu}) \\
& V_{bi} \sim N(0,1) \quad H_{bi} \sim HN(0,1) \quad U_{bi} \sim G(.|\boldsymbol{\nu}) \\
& V_{wi} \sim N(0,1) \quad H_{wi} \sim HN(0,1) \quad U_{wi} \sim G(.|\boldsymbol{\nu}) \quad i = 1, \ldots, n
\end{aligned}
\tag{3.3}
$$

where $\lambda_j = \dfrac{s\gamma_j^{1/3}}{\sqrt{b^2 + s^2\gamma_j^{2/3}(b^2 - 1)}}$, $\delta_j = \dfrac{\lambda_j}{\sqrt{1 + \lambda_j^2}}$ for j=c,b or w, $\Delta_j = \dfrac{\sigma_j \delta_j}{\sqrt{1 - b^2\delta_j^2}}$ and $\tau_j = \dfrac{\sigma_j^2(1 - \delta_j^2)}{1 - b^2\delta_j^2}$ for j=c or w and $\Delta_b = \dfrac{\delta_b}{\sqrt{1 - b^2\delta_b^2}}$, $\tau_b = \dfrac{1 - \delta_b^2}{1 - b^2\delta_b^2}$, $b = \sqrt{\dfrac{2}{\pi}}$, $s = \left(\dfrac{2}{4 - \pi}\right)^{1/3}$.

3.3 Bayesian Inference

To use the Bayesian paradigm, it is essential to obtain the joint posterior distribution. However, since the necessary integrals are not easy to calculate, it is not possible to obtain such distribution, analytically. However, it is possible to obtain numerical approximation for the marginal posterior distributions of interest by using MCMC algorithms, see (GEMAN; GEMAN, 1984) and (HASTINGS, 1970).

To obtain the posterior distribution we need to consider the complete likelihood, that is,

$$
\begin{aligned}
L_c(\boldsymbol{\theta}|\boldsymbol{y},\boldsymbol{t},\boldsymbol{u}_c,\boldsymbol{u}_b,\boldsymbol{u}_w,\boldsymbol{h}_c,\boldsymbol{h}_b,\boldsymbol{h}_w) \propto & \prod_{i=1}^{n} \phi\left(y_i|\mu_{ci}+w_i+D_{ci},\tau_c u_{ci}^{-1}\right) f(h_{ci})h(u_{ci}|\boldsymbol{\nu}_c) \\
& \times \phi\left(z_i|\mu_{bi}+D_{bi},\tau_b u_{bi}^{-1}\right) I(z_i,t_i) f(h_{bi})h(u_{bi}|\boldsymbol{\nu}_b) \\
& \times \phi\left(w_i|D_{wi},\tau_b u_{wi}^{-1}\right) f(h_{wi})h(u_{wi}|\boldsymbol{\nu}_w),
\end{aligned}
$$

where $\mu_{ci} = \boldsymbol{X}_{ci}^t\boldsymbol{\beta}_c$ and $D_{ci} = \dfrac{\Delta_c}{\sqrt{u_{ci}}}(h_{ci}-b)$, $\mu_{bi} = \boldsymbol{X}_{bi}^t\boldsymbol{\beta}_b$ and $D_{bi} = \dfrac{\Delta_b}{\sqrt{u_{bi}}}(b-h_{bi})$ and $\boldsymbol{\theta} = (\boldsymbol{\beta}_c,\Delta_c,\tau_c,\boldsymbol{\nu}_c,\gamma_b,\delta_b,\boldsymbol{\nu}_b,\Delta_w,\tau_w,\boldsymbol{\nu}_w)$. Since we set $\sigma^2 = 1$ for the skew generalized t model, the MCMC algorithm will be slightly different from the other models. For this model, we have $\boldsymbol{\theta} = (\boldsymbol{\beta}_c,\delta_c,\boldsymbol{\nu}_c,\boldsymbol{\beta}_b,\gamma_b,\boldsymbol{\nu}_b,\delta_w,\boldsymbol{\nu}_w)$, therefore, Δ_j and τ_j, for $j = c, w$ are functions of only ν_j.

We need to consider a prior distribution for $\boldsymbol{\theta}$. We will assume an independence structure and conditional conjugate prior distributions for $\boldsymbol{\beta}_j$, τ_j^{-1}, Δ_j, $j = c, w$, and $\delta_c \in U(-1,1)$. For γ_b, we adopted the prior described in (AZEVEDO; BOLFARINE; ANDRADE, 2012), that is $\pi(\gamma_b) \propto (1+\gamma_b)^{\alpha_{\gamma_1}-1}(1-\gamma_b)^{\alpha_{\gamma_2}-1} I(\gamma_b \in A_{\gamma_b})$, where $A_{\gamma_b} = (-0.99527, 0.99527)$ On the other hand, for $\boldsymbol{\nu}_j$, $j = c, b$ or w, the choice of the prior distribution will depend on the model.

3.3.1 Full conditional distributions

In order to implement the MCMC algorithm, we have to simulate iteratively from the full conditionals described bellow.

Denoting by $\boldsymbol{\theta}_{-\theta_i}$ the parameter vector $\boldsymbol{\theta}$ without the component θ_i, the full conditional distributions are

For $\boldsymbol{\beta}_c$:

$$
\pi(\boldsymbol{\beta}_c|\boldsymbol{\theta}_{-\beta_c},\boldsymbol{y},\boldsymbol{w},\boldsymbol{u}_c,\boldsymbol{h}_c) \propto \exp\left\{-\frac{1}{2}\left(\boldsymbol{\beta}_c^t\Sigma_*^{-1}\boldsymbol{\beta}_c - 2\mu_*^t\Sigma_*^{-1}\boldsymbol{\beta}_c\right)\right\} I_{\mathbb{R}^p}(\boldsymbol{\beta}_c)
$$

which can be recognized as the kernel of p-variate normal distribution with variance $\boldsymbol{\Sigma_*} = \left(\frac{\sum_{i=1}^{n} u_{ci} x_i x_i^t}{\tau_c} + \boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{-1} \right)^{-1}$ and mean $\boldsymbol{\mu_*} = \left(\sum_{i=1}^{n} \frac{u_{ci}}{\tau_c} \left(y_i - w_i - D_{ci} \right) \boldsymbol{X}_{ci}^t + \boldsymbol{\mu}_{\boldsymbol{\beta}}^t \boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{-1} \right) \boldsymbol{\Sigma_*}$.

For $\boldsymbol{\beta}_b$:

$$\pi(\boldsymbol{\beta}_b | \boldsymbol{\theta}_{-\beta_b}, \boldsymbol{t}, \boldsymbol{z}, \boldsymbol{w}, \boldsymbol{u}_b, \boldsymbol{h}_b) \propto \exp\left\{ -\frac{1}{2} \left(\boldsymbol{\beta}_b^t \Sigma_*^{-1} \boldsymbol{\beta}_b - 2\mu_*^t \Sigma_*^{-1} \boldsymbol{\beta}_b \right) \right\} I_{\mathbb{R}^p}(\boldsymbol{\beta}_b)$$

which can be recognized as the kernel of p-variate normal distribution with variance $\boldsymbol{\Sigma_*} = \left(\frac{\sum_{i=1}^{n} u_{bi} x_{bi} x_{bi}^t}{\tau_b} + \boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{-1} \right)^{-1}$ and mean $\boldsymbol{\mu_*} = \left(\sum_{i=1}^{n} \frac{u_{bi}}{\tau_b} \left(z_i - w_i - \frac{\Delta_b}{\sqrt{u_{bi}}} (b - h_i) \right) \boldsymbol{x}_i^t + \boldsymbol{\mu}_{\boldsymbol{\beta}}^t \boldsymbol{\Sigma}_{\boldsymbol{\beta}}^{-1} \right) \boldsymbol{\Sigma_*}$.

For z_i:

$$f(z_i | \boldsymbol{\theta}, t_i, h_{bi}, u_{bi}, w_i) \propto \phi\left(z_i | \mu_{bi} + w_i + D_{bi}, \tau u_{bi}^{-1} \right) I(z_i, t_i)$$

Then,

$$z_i | \boldsymbol{\theta}, u_{bi}, h_{bi}, w_i, t_i = 1 \sim TN\left(\mu_{bi} + w_i + D_{bi}, \tau u_{bi}^{-1} \right) I(0, \infty)$$
$$z_i | \boldsymbol{\theta}, u_{bi}, h_{bi}, w_i, t_i = 0 \sim TN\left(\mu_{bi} + w_i + D_{bi}, \tau u_{bi}^{-1} \right) I(-\infty, 0)$$

For h_{ci}:

$$f(\boldsymbol{\theta}, u_{ci}, h_{ci} | y_i) \propto \exp\left\{ -\frac{1}{2} \left(\frac{\Delta_c^2 + \tau_c}{\tau_c} \right) \left[h_{ci}^2 - 2h_{ci} \left(\frac{\Delta_c^2 b + \Delta_c \sqrt{u_{ci}} (y_i - \mu_{ci} - w_i)}{\Delta_c^2 + \tau_c} \right) \right] \right\} I_{(0,\infty)}(h_{ci})$$

which can be recognized as the kernel of a truncated normal distribution, so

$$h_{ci} | \boldsymbol{\theta}, u_{ci}, y_i \sim TN\left(\frac{\Delta_c^2 b + \Delta_c \sqrt{u_{ci}} (y_i - \mu_{ci} - w_i)}{\Delta_c^2 + \tau_c}, \frac{\tau_c}{\Delta_c^2 + \tau_c} \right) I(0, \infty)$$

For h_{bi}:

$$f(\boldsymbol{\theta}, u_{bi}, h_{bi} | y_i) \propto \exp\left\{ -\frac{1}{2} \left(\frac{\Delta_b^2 + \tau_b}{\tau_c} \right) \left[h_{bi}^2 - 2h_{bi} \left(\frac{\Delta_b^2 b - \Delta_b \sqrt{u_{bi}} (z_i - \mu_{bi} - w_i)}{\Delta_b^2 + \tau_b} \right) \right] \right\} I_{(0,\infty)}(h_{bi})$$

which can be recognized as the kernel of a truncated normal distribution, so

$$h_{bi} | \boldsymbol{\theta}, u_{bi}, z_i \sim TN\left(\frac{\Delta_b^2 b - \Delta_b \sqrt{u_{bi}} (z_i - \mu_{bi} - w_i)}{\Delta_b^2 + \tau_b}, \frac{\tau_b}{\Delta_b^2 + \tau_b} \right) I(0, \infty)$$

For h_{wi}:

$$f(\boldsymbol{\theta}, u_{wi}, h_{wi}|w_i) \propto \exp\left\{-\frac{1}{2}\left(\frac{\Delta_w^2+\tau_w}{\tau_w}\right)\left[h_{wi}^2 - 2h_{wi}\left(\frac{\Delta_w}{\Delta_w^2+\tau_w}\right)(w_i+\Delta_w b)\right]\right\} I_{(0,\infty)}(h_{wi})$$

which can be recognized as the kernel of a truncated normal distribution, so

$$h_{bi}|\boldsymbol{\theta}, u_{bi}, w_i \sim TN\left(\left(\frac{\Delta_w}{\Delta_w^2+\tau_w}\right)(w_i+\Delta_w b), \frac{\tau_w}{\Delta_w^2+\tau_w}\right) I(0,\infty)$$

For u_{ci}:

- Skew slash:

$$f(\boldsymbol{\theta}, u_{ci}, h_{ci}|y_i) \propto u_{ci}^{\nu_c+1/2-1} \exp\left\{-\frac{u_{ci}}{2\tau_c}\left[(y_i-\mu_{ci}-w_i)^2 - 2\frac{\Delta_c}{\sqrt{u_{ci}}}(h_{ci}-b)(y_i-\mu_{ci}-w_i)\right]\right\} \times I_{(0,1)}(u_{ci})$$

- Skew-t:

$$f(\boldsymbol{\theta}, u_{ci}, h_{ci}|y_i) \propto u_{ci}^{\frac{\nu_c+1}{2}-1} \exp\left\{-\frac{u_{ci}}{2}\left[\frac{(y_i-\mu_{ci}-w_i)^2}{\tau_c}+\nu_c\right] + \frac{\Delta_c\sqrt{u_{ci}}}{\tau_c}(h_{ci}-b)(y_i-\mu_{ci}-w_i)\right\} \times I_{(0,\infty)}(u_{ci})$$

- Skew generalized t:

$$f(\boldsymbol{\theta}, u_{ci}, h_{ci}|y_i) \propto u_{ci}^{\frac{\nu_{c1}+1}{2}-1} \exp\left\{-\frac{u_{ci}}{2}\left[\frac{(y_i-\mu_{ci}-w_i)^2}{\tau_c}+\nu_{c2}\right] + \frac{\Delta_c\sqrt{u_{ci}}}{\tau_c}(h_{ci}-b)(y_i-\mu_{ci}-w_i)\right\} \times I_{(0,\infty)}(u_{ci})$$

- Skew-contaminated normal: the discrete conditional distribution of u_{ci} assumes ν_{c2} with probability $\dfrac{p_i}{p_i+q_i}$ and 1 with probability $\dfrac{q_i}{p_i+q_i}$ where

$$p_i = \nu_{c1}\sqrt{\nu_{c2}} \exp\left\{-\frac{\nu_{c2}}{2\tau_c}\left[(y_i-\mu_{ci}-w_i)^2 - 2\frac{\Delta_c}{\sqrt{\nu_{c2}}}(h_{ci}-b)(y_i-\mu_{ci}-w_i)\right]\right\}$$
$$q_i = (1-\nu_{c1}) \exp\left\{-\frac{1}{2\tau_c}\left[(y_i-\mu_{ci}-w_i)^2 - 2\Delta_c(h_{ci}-b)(y_i-\mu_{ci}-w_i)\right]\right\}$$

For u_{bi}:

- Skew slash:

$$f(\boldsymbol{\theta}, u_{bi}, h_{bi}|z_i) \propto u_{bi}^{\nu_b+1/2-1} \exp\left\{-\frac{u_{bi}}{2\tau_b}\left[(z_i-\mu_{bi}-w_i)^2 - 2\frac{\Delta_b}{\sqrt{u_{bi}}}(b-h_{bi})(z_i-\mu_{bi}-w_i)\right]\right\} \times I_{(0,1)}(u_{bi})$$

- Skew-t:

$$f(\boldsymbol{\theta}, u_{bi}, h_{bi}|z_i) \propto u_{bi}^{\frac{\nu_b+1}{2}-1} \exp\left\{-\frac{u_{bi}}{2}\left[\frac{(z_i-\mu_{bi}-w_i)^2}{\tau_b}+\nu_b\right]+\frac{\Delta_b\sqrt{u_{bi}}}{\tau_b}(b-h_{bi})(z_i-\mu_{bi}-w_i)\right\} \times I_{(0,\infty)}(u_{bi})$$

- Skew generalized t:

$$f(\boldsymbol{\theta}, u_{bi}, h_{bi}|z_i) \propto u_{bi}^{\frac{\nu_{b1}+1}{2}-1} \exp\left\{-\frac{u_{bi}}{2}\left[\frac{(z_i-\mu_{bi}-w_i)^2}{\tau_b}+\nu_{b2}\right]+\frac{\Delta_b\sqrt{u_{bi}}}{\tau_b}(b-h_{bi})(z_i-\mu_{bi}-w_i)\right\} \times I_{(0,\infty)}(u_{bi})$$

- Skew-contaminated normal: the discrete conditional distribution of u_{bi} assumes ν_{b2} with probability $\frac{p_i}{p_i+q_i}$ and 1 with probability $\frac{q_i}{p_i+q_i}$ where

$$p_i = \nu_{b1}\sqrt{\nu_{b2}}\exp\left\{-\frac{\nu_{b2}}{2\tau_b}\left[(z_i-\mu_{bi}-w_i)^2-2\frac{\Delta_b}{\sqrt{\nu_{b2}}}(b-h_{bi})(z_i-\mu_{bi}-w_i)\right]\right\}$$
$$q_i = (1-\nu_{b1})\exp\left\{-\frac{1}{2\tau_b}\left[(z_i-\mu_{bi}-w_i)^2-2\Delta_b(b-h_{bi})(z_i-\mu_{bi}-w_i)\right]\right\}$$

For u_{wi}:

- Skew slash:

$$f(\boldsymbol{\theta}, u_{wi}, h_{wi}|w_i) \propto u_{wi}^{\nu_w+1/2-1} \exp\left\{-\frac{u_{wi}}{2\tau_w}\left[w_i^2-2\frac{\Delta_w w_i}{\sqrt{u_{wi}}}(h_{wi}-b)\right]\right\} I_{(0,1)}(u_{wi})$$

- Skew-t:

$$f(\boldsymbol{\theta}, u_{wi}, h_{wi}|w_i) \propto u_{wi}^{\frac{\nu_w+1}{2}-1} \exp\left\{-\frac{u_{wi}}{2}\left[\frac{w_i^2}{\tau_w}+\nu_w\right]+\frac{\Delta_w w_i\sqrt{u_{wi}}}{\tau_w}(h_{wi}-b)\right\} I_{(0,\infty)}(u_{wi})$$

- Skew generalized t:

$$f(\boldsymbol{\theta}, u_{wi}, h_{wi}|w_i) \propto u_{wi}^{\frac{\nu_{b1}+1}{2}-1} \exp\left\{-\frac{u_{wi}}{2}\left[\frac{w_i^2}{\tau_w}+\nu_{b2}\right]+\frac{\Delta_w w_i\sqrt{u_{wi}}}{\tau_w}(h_{wi}-b)\right\} I_{(0,\infty)}(u_{wi})$$

- Skew-contaminated normal: the discrete conditional distribution of u_{wi} assumes ν_{b2} with probability $\frac{p_i}{p_i+q_i}$ and 1 with probability $\frac{q_i}{p_i+q_i}$ where

$$p_i = \nu_{b1}\sqrt{\nu_{b2}}\exp\left\{-\frac{\nu_{b2}}{2\tau_w}\left[w_i^2-2\frac{\Delta_w w_i}{\sqrt{\nu_{b2}}}(h_{wi}-b)\right]\right\}$$
$$q_i = (1-\nu_{b1})\exp\left\{-\frac{1}{2\tau_w}\left[w_i^2-2\Delta_w w_i(h_{wi}-b)\right]\right\}$$

For γ_b:

$$\pi(\gamma_b|\boldsymbol{\theta}_{-\gamma_b}, \boldsymbol{y}, \boldsymbol{u}_b, \boldsymbol{z}, \boldsymbol{h}_b) \propto \tau_b^{-n/2}\exp\left\{-\frac{1}{2\tau_b}\sum_{i=1}^{n}u_{bi}\left(z_i-\mu_{bi}-D_{bi}\right)^2\right\}(1+\gamma_b)^{\alpha_{\gamma_1}-1}(1-\gamma_b)^{\alpha_{\gamma_2}-1} \times I(\gamma_b \in A_{\gamma_b})$$

For Δ_c:

$$\pi(\Delta_c|\boldsymbol{\theta}_{-\Delta_c},\boldsymbol{y},\boldsymbol{u}_c,\boldsymbol{h}_c) \propto \exp\left\{-\frac{1}{2}\left(\frac{\sigma^2_{\Delta_c}\sum_{i=1}^n(h_{ci}-b)^2+\tau_c}{\tau_c\sigma^2_{\Delta_c}}\right)\left[\Delta_c^2-2\Delta_c m_{\Delta_c}\right]\right\} I_{\mathbb{R}}(\Delta_c)$$

where $m_{\Delta_c} = \dfrac{\sigma^2_{\Delta_c}\sum_{i=1}^n(h_{ci}-b)\sqrt{u_{ci}}(y_i-\mu_{ci}-w_i)+\mu_{\Delta_c}\tau_c}{\sigma^2_{\Delta_c}\sum_{i=1}^n(h_{ci}-b)^2+\tau_c}$. So,

$$\Delta_c|\boldsymbol{\theta}_{-\Delta_c},\boldsymbol{y},\boldsymbol{u}_c,\boldsymbol{h}_c \sim N\left(m_{\Delta_c},\frac{\tau_c\sigma^2_{\Delta_c}}{\sigma^2_{\Delta_c}\sum_{i=1}^n(h_{ci}-b)^2+\tau_c}\right)$$

For τ_c^{-1}:

$$\pi(\tau_c^{-1}|\boldsymbol{\theta}_{-\tau_c^{-1}},\boldsymbol{y},\boldsymbol{u}_c,\boldsymbol{h}_c) \propto (\tau_c^{-1})^{n/2+c-1}\exp\left\{-\tau_c^{-1}\left\{d+\sum_{i=1}^n\frac{u_{ci}}{2}\left(y_i-\left(w_i+\mu_{ci}+\frac{\Delta_c}{\sqrt{u_{ci}}}(h_{ci}-b)\right)\right)^2\right\}\right\}$$
$$\times I_{(0,\infty)}(\tau_c^{-1})$$

that can be recognized as the kernel of a gamma distribution. So, $\tau_c^{-1}|\boldsymbol{\theta}_{-\tau_c^{-1}},\boldsymbol{y},\boldsymbol{u}_c,\boldsymbol{h}_c \sim$ gamma $\left(n/2+c,d+\sum_{i=1}^n\frac{u_{ci}}{2}\left(y_i-\left(\mu_{ci}+w_i+\frac{\Delta_c}{\sqrt{u_{ci}}}(h_{ci}-b)\right)\right)^2\right)$

For Δ_w:

$$\pi(\Delta_w|\boldsymbol{\theta}_{-\Delta_w},\boldsymbol{w},\boldsymbol{u}_w,\boldsymbol{h}_w) \propto \exp\left\{-\frac{1}{2}\left(\frac{\sigma^2_{\Delta_w}\sum_{i=1}^n(h_{wi}-b)^2+\tau_w}{\tau_w\sigma^2_{\Delta_w}}\right)\left[\Delta_w^2-2\Delta_w m_{\Delta_w}\right]\right\} I_{\mathbb{R}}(\Delta_w)$$

where $m_{\Delta_w} = \dfrac{\sigma^2_{\Delta_w}\sum_{i=1}^n(h_{wi}-b)\sqrt{u_{wi}}w_i+\mu_{\Delta_w}\tau_w}{\sigma^2_{\Delta_w}\sum_{i=1}^n(h_{wi}-b)^2+\tau_w}$. So,

$$\Delta_w|\boldsymbol{\theta}_{-\Delta_w},\boldsymbol{y},\boldsymbol{u}_w,\boldsymbol{h}_w \sim N\left(m_{\Delta_w},\frac{\tau_w\sigma^2_{\Delta_w}}{\sigma^2_{\Delta_w}\sum_{i=1}^n(h_{wi}-b)^2+\tau_w}\right)$$

For τ_w^{-1}:

$$\pi(\tau_w^{-1}|\boldsymbol{\theta}_{-\tau_w^{-1}},\boldsymbol{w},\boldsymbol{u}_w,\boldsymbol{h}_w) \propto (\tau_w^{-1})^{n/2+c-1}\exp\left\{-\tau_w^{-1}\left\{d+\sum_{i=1}^n\frac{u_{wi}}{2}\left(w_i-\frac{\Delta_w}{\sqrt{u_{wi}}}(h_{wi}-b)\right)^2\right\}\right\}$$
$$\times I_{(0,\infty)}(\tau_w^{-1})$$

that can be recognized as the kernel of a gamma distribution. So,

$$\tau_w^{-1}|\boldsymbol{\theta}_{-\tau_w^{-1}},\boldsymbol{y},\boldsymbol{u}_w,\boldsymbol{h}_w \sim \text{gamma}\left(n/2+c,d+\sum_{i=1}^n\frac{u_{wi}}{2}\left(w_i-\frac{\Delta_w}{\sqrt{u_{wi}}}(h_{wi}-b)\right)^2\right)$$

For $\boldsymbol{\nu}_j, j = c$ or w:

- Skew slash: Considering a gamma distribution left truncated at 1 as prior with mean $\dfrac{\alpha_1}{\alpha_2}$ and variance $\dfrac{\alpha_1}{\alpha_2^2}$, it follows that

$$\pi(\nu_j|\boldsymbol{\theta}_{-\nu_j}, \boldsymbol{y}, \boldsymbol{u}_j, \boldsymbol{h}_j) \propto \nu_j^{n+\alpha_1-1} \exp\left\{-\nu_j(\alpha_2 - \sum_{i=1}^{n} ln(u_{ji}))\right\} I_{(1,\infty)}(\nu_j)$$

that is, $\nu_j|\boldsymbol{\theta}_{-\nu_j}, \boldsymbol{y}, \boldsymbol{u}_j, \boldsymbol{h}_j \sim TG(n+\alpha_1, \alpha_2 - \sum_{i=1}^{n} ln(u_{ji}))I(1,\infty)$, where TG denotes the Truncated Gamma distribution.

- Skew-t: For Skew-t, we have adopted a very useful hierarchical prior distribution as noted in (CABRAL; LACHOS; MADRUGA, 2012), which consists in $\nu_j|\lambda \sim exp(\lambda)I(\nu_j)_{(2,\infty)}$ and $\lambda \sim U(\rho_0, \rho_1)$ where $0 < \rho_0 < \rho_1$ are known. The exponential distribution is left truncated at 2 to insure finite variance. Then

$$\pi(\nu_j|\boldsymbol{\theta}_{-\nu_j}, \lambda, \boldsymbol{y}, \boldsymbol{u}_j, \boldsymbol{h}_j) \propto \frac{\frac{\nu_j}{2}^{\frac{n\nu_j}{2}}}{\Gamma(\nu_j/2)^n} \left(\prod_{i=1}^{n} u_{ji}\right)^{\nu_j/2-1} \exp\left\{-\nu_j\left(\frac{\sum_{i=1}^n u_{ji}}{2} + \lambda\right)\right\} I_{(2,\infty)}(\nu_j)$$

$$\pi(\lambda|\boldsymbol{\theta}, \boldsymbol{y}, \boldsymbol{u}_j, \boldsymbol{h}_j) \propto \lambda \exp -\lambda(\nu_j - 2) I_{(\rho_0,\rho_1)}(\lambda)$$

that is, $\lambda|\boldsymbol{\theta}, \boldsymbol{y}, \boldsymbol{u}_j, \boldsymbol{h}_j \sim TG(2, \nu_j - 2)I(\rho_0, \rho_1)$.

- Skew generalized t: Assuming $\nu_{j1}|\lambda_1 \sim exp(\lambda_1)I(\nu_{j1})_{(2,\infty)}$ and $\lambda_1 \sim U(\rho_0, \rho_1)$ where $0 < \rho_0 < \rho_1$ are known and $\nu_{j2}|\lambda_2 \sim exp(\lambda_2)$ and $\lambda_2 \sim U(\psi_0, \psi_1)$ where $0 < \psi_0 < \psi_1$ are known, we have

$$\pi(\nu_{j1}|\boldsymbol{\theta}_{-\nu_{j1}}, \lambda_1, \boldsymbol{y}, \boldsymbol{u}_j, \boldsymbol{h}_j) \propto \frac{\nu_{j2}/2^{n\nu_{j1}/2}}{\Gamma(\nu_{j1}/2)^n} \left(\prod_{i=1}^{n} u_{ji}\right)^{\nu_{j1}/2-1} \exp\left\{-\lambda_1(\nu_{j1} - 2)\right\} I_{(2,\infty)}(\nu_{j1})$$

$$\pi(\lambda_1|\boldsymbol{\theta}, \boldsymbol{y}, \boldsymbol{u}_j, \boldsymbol{h}_j) \propto \lambda_1 \exp\left\{-\lambda_1(\nu_{j1} - 2)\right\} I_{(\rho_0,\rho_1)}(\lambda_1)$$

that is, $\lambda_1|\boldsymbol{\theta}, \boldsymbol{y}, \boldsymbol{u}_j, \boldsymbol{h}_j \sim TG(2, \nu_{j1} - 2)I(\rho_0, \rho_1)$.

Also, we have that

$$\pi(\nu_{j2}|\boldsymbol{\theta}_{-\nu_{j2}}, \lambda_2, \boldsymbol{y}, \boldsymbol{u}_j, \boldsymbol{h}_j) \propto \nu_{j2}/2^{n\nu_{j1}/2} \exp\left\{-\nu_{j2}\left(\frac{\sum_{i=1}^n u_{ji}}{2} + \lambda_2\right)\right\} I_{(0,\infty)}(\nu_{j2})$$

and

$$\pi(\lambda_2|\boldsymbol{\theta}, \boldsymbol{y}, \boldsymbol{u}_j, \boldsymbol{h}_j) \propto \lambda_2 \exp\{-\lambda_2(\nu_{j2})\} I_{(\psi_0,\psi_1)}(\lambda_2)$$

that is, $\nu_{j2}|\boldsymbol{\theta}_{-\nu_{j2}}, \lambda_2, \boldsymbol{y}, \boldsymbol{u}_j, \boldsymbol{h}_j \sim \text{gamma}(\frac{n\nu_{j1}}{2} + 1, \frac{\sum_{i=1}^n u_{ji}}{2} + \lambda_2)$

and $\lambda_2|\boldsymbol{\theta}, \boldsymbol{y}, \boldsymbol{u}_j, \boldsymbol{h}_j \sim TG(2, \nu_{j2} - 2)I(\xi_0, \xi_1)$

- Skew-contaminated normal: Observe that distribution of U can be written as

$$h(u|\boldsymbol{\nu}_j) = \nu_{j1}^{\frac{1-u}{1-\nu_{j2}}}(1-\nu_{j1})^{\frac{u-\nu_{j2}}{1-\nu_{j2}}} I_{\{\nu_{j2},1\}}(u)$$

Setting as prior distributions $\nu_{j1} \sim beta(\alpha_1, \beta_1)$, $\nu_{j2} \sim beta(\alpha_2, \beta_2)$, it follows that the conditional distributions of ν_{j1} and ν_{j2} are

$$\pi(\nu_{j1}|\boldsymbol{\theta}_{-\nu_{j1}}, \boldsymbol{y}, \boldsymbol{u}_j, \boldsymbol{h}_j) \propto \nu_{j1}^{\frac{n-\sum_{i=1}^n u_{ji}}{1-\nu_{j2}}+\alpha_1-1}(1-\nu_{j1})^{\frac{\sum_{i=1}^n u_{ji}-n\nu_{j2}}{1-\nu_{j2}}+\beta_1-1} I_{(0,1)}(\nu_{j1})$$

which can be recognized as the kernel of a beta distribution. So, $\nu_{j1}|\boldsymbol{\theta}_{-\nu_{j1}}, \boldsymbol{y}, \boldsymbol{u}_j, \boldsymbol{h}_j \sim beta\left(\frac{n-\sum_{i=1}^n u_{ji}}{1-\nu_{j2}}+\alpha_1, \frac{\sum_{i=1}^n u_{ji}-n\nu_{j2}}{1-\nu_{j2}}+\beta_1\right)$ And

$$\pi(\nu_{j2}|\boldsymbol{\theta}_{-\nu_{j2}}, \boldsymbol{y}, \boldsymbol{u}_j, \boldsymbol{h}_j) \propto \nu_{j1}^{\frac{n-\sum_{i=1}^n u_{ji}}{1-\nu_{j2}}}(1-\nu_{j1})^{\frac{\sum_{i=1}^n u_{ji}-n\nu_{j2}}{1-\nu_{j2}}} \nu_{j2}^{\alpha_2-1}(1-\nu_{j2})^{(\beta_2-1)} I_{(0,1)}(\nu_{j2})$$

For $\boldsymbol{\nu}_b$:

- Skew slash: Considering a gamma distribution as prior with mean $\frac{\alpha_1}{\alpha_2}$ and variance $\frac{\alpha_1}{\alpha_2^2}$, it follows that

$$\pi(\nu_b|\boldsymbol{\theta}_{-\nu_b}, \boldsymbol{y}, \boldsymbol{u}_b, \boldsymbol{h}_b) \propto \nu_b^{n+\alpha_1-1}\exp\left\{-\nu_b(\alpha_2 - \sum_{i=1}^n ln(u_{bi}))\right\} I_{(0,\infty)}(\nu_b)$$

that is, $\nu_b|\boldsymbol{\theta}_{-\nu_b}, \boldsymbol{y}, \boldsymbol{u}_b, \boldsymbol{h}_b \sim \text{gamma}(n+\alpha_1, \alpha_2 - \sum_{i=1}^n ln(u_{bi}))$.

- Skew-t: For Skew-t, we have adopted a very useful hierarchical prior distribution as noted in (CABRAL; LACHOS; MADRUGA, 2012), which consists in $\nu_b|\lambda \sim exp(\lambda)$ and $\lambda \sim U(\rho_0, \rho_1)$ where $0 < \rho_0 < \rho_1$ are known. Then

$$\pi(\nu_b|\boldsymbol{\theta}_{-\nu_b}, \lambda, \boldsymbol{y}, \boldsymbol{u}_b, \boldsymbol{h}_b) \propto \frac{\frac{\nu_b}{2}^{\frac{n\nu_b}{2}}}{\Gamma(\nu_b/2)^n}\left(\prod_{i=1}^n u_{bi}\right)^{\nu_b/2-1}\exp\left\{-\nu_b\left(\frac{\sum_{i=1}^n u_{bi}}{2}+\lambda\right)\right\} I_{(0,\infty)}(\nu_b)$$

$$\pi(\lambda|\boldsymbol{\theta}, \boldsymbol{y}, \boldsymbol{u}_b, \boldsymbol{h}_b) \propto \lambda\exp\{-\lambda\nu_b\} I_{(\rho_0,\rho_1)}(\lambda)$$

that is, $\lambda|\boldsymbol{\theta}, \boldsymbol{y}, \boldsymbol{u}_b, \boldsymbol{h}_b \sim TG(2, \nu_b)I(\rho_0, \rho_1)$.

- Skew generalized t: Assuming $\nu_{b1}|\lambda_1 \sim exp(\lambda_1)$ and $\lambda_1 \sim U(\rho_0, \rho_1)$ where $0 < \rho_0 < \rho_1$ are known and $\nu_{b2}|\lambda_2 \sim exp(\lambda_2)$ and $\lambda_2 \sim U(\psi_0, \psi_1)$ where $0 < \psi_0 < \psi_1$ are known, we have

$$\pi(\nu_{b1}|\boldsymbol{\theta}_{-\nu_{b1}}, \lambda_1, \boldsymbol{y}, \boldsymbol{u}_b, \boldsymbol{h}_b) \propto \frac{\nu_{b2}/2^{n\nu_{b1}/2}}{\Gamma(\nu_{b1}/2)^n}\left(\prod_{i=1}^n u_{bi}\right)^{\nu_{b1}/2-1}\exp\{-\lambda_1(\nu_{b1}-2)\} I_{(0,\infty)}(\nu_{b1})$$

$$\pi(\lambda_1|\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{u}_b,\boldsymbol{h}_b)\propto\lambda_1\exp\left\{-\lambda_1(\nu_{b1})\right\}I_{(\rho_0,\rho_1)}(\lambda_1)$$

that is, $\lambda_1|\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{u}_b,\boldsymbol{h}_b\sim TG(2,\nu_{b1})I(\rho_0,\rho_1)$.

Also, we have that

$$\pi(\nu_{b2}|\boldsymbol{\theta}_{-\nu_{b2}},\lambda_2,\boldsymbol{y},\boldsymbol{u}_b,\boldsymbol{h}_b)\propto\nu_{b2}/2^{n\nu_{b1}/2}\exp\left\{-\nu_{b2}\left(\frac{\sum_{i=1}^n u_{bi}}{2}+\lambda_2\right)\right\}I_{(0,\infty)}(\nu_{b2})$$

and

$$\pi(\lambda_2|\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{u}_b,\boldsymbol{h}_b)\propto\lambda_2\exp\{-\lambda_2(\nu_{b2})\}I_{(\psi_0,\psi_1)}(\lambda_2)$$

that is, $\nu_{b2}|\boldsymbol{\theta}_{-\nu_{b2}},\lambda_2,\boldsymbol{y},\boldsymbol{u}_b,\boldsymbol{h}_b\sim\text{gamma}(\frac{n\nu_{b1}}{2}+1,\frac{\sum_{i=1}^n u_{bi}}{2}+\lambda_2)$
and $\lambda_2|\boldsymbol{\theta},\boldsymbol{y},\boldsymbol{u}_b,\boldsymbol{h}_b\sim TG(2,\nu_{b2}-2)I(\xi_0,\xi_1)$

- Skew-contaminated normal: Observe that distribution of U can be written as

$$h(u|\boldsymbol{\nu}_b)=\nu_{b1}^{\frac{1-u}{1-\nu_{b2}}}(1-\nu_{b1})^{\frac{u-\nu_{b2}}{1-\nu_{b2}}}I_{\{\nu_{b2},1\}}(u)$$

Setting as prior distributions $\nu_{b1}\sim beta(\alpha_1,\beta_1)$, $\nu_{b2}\sim beta(\alpha_2,\beta_2)$, it follows that the conditional distributions of ν_{b1} and ν_{b2} are

$$\pi(\nu_{b1}|\boldsymbol{\theta}_{-\nu_{b1}},\boldsymbol{y},\boldsymbol{u}_b,\boldsymbol{h}_b)\propto\nu_{b1}^{\frac{n-\sum_{i=1}^n u_{bi}}{1-\nu_{b2}}+\alpha_1-1}(1-\nu_{b1})^{\frac{\sum_{i=1}^n u_{bi}-n\nu_{b2}}{1-\nu_{b2}}+\beta_1-1}I_{(0,1)}(\nu_{b1})$$

which can be recognized as the kernel of a beta distribution. So,
$\nu_{b1}|\boldsymbol{\theta}_{-\nu_{b1}},\boldsymbol{y},\boldsymbol{u}_b,\boldsymbol{h}_b\sim beta\left(\frac{n-\sum_{i=1}^n u_{bi}}{1-\nu_{b2}}+\alpha_1,\frac{\sum_{i=1}^n u_{bi}-n\nu_{b2}}{1-\nu_{b2}}+\beta_1\right)$ And

$$\pi(\nu_{b2}|\boldsymbol{\theta}_{-\nu_{b2}},\boldsymbol{y},\boldsymbol{u}_b,\boldsymbol{h}_b)\propto\nu_{b1}^{\frac{n-\sum_{i=1}^n u_{bi}}{1-\nu_{b2}}}(1-\nu_{b1})^{\frac{\sum_{i=1}^n u_{bi}-n\nu_{b2}}{1-\nu_{b2}}}\nu_{b2}^{\alpha_2-1}(1-\nu_{b2})^{(\beta_2-1)}I_{(0,1)}(\nu_{b2})$$

Finally, for the skew generalized t distribution the conditional distributions of δ_c and δ_w are

For δ_c:

$$\pi(\delta_c|\boldsymbol{\theta}_{-\delta_c},\boldsymbol{y},\boldsymbol{u}_c,\boldsymbol{h}_c)\propto\left(\frac{\sqrt{1-b^2\delta_c^2}}{\sqrt{1-\delta_c^2}}\right)^n\exp\left\{-\frac{1-b^2\delta_c^2}{2(1-\delta_c^2)}\sum_{i=1}^n u_{ci}\left(y_i-\left(w_i+\mu_{ci}+\frac{\delta_c}{\sqrt{u_{ci}}\sqrt{1-b^2\delta_c^2}}(h_{ci}-b)\right)\right)\right\}$$
$$I_{(-1,1)}(\delta_c)$$

For δ_w:

$$\pi(\delta_w|\boldsymbol{\theta}_{-\delta_w},\boldsymbol{y},\boldsymbol{u}_w,\boldsymbol{h}_w)\propto\left(\frac{\sqrt{1-b^2\delta_w^2}}{\sqrt{1-\delta_w^2}}\right)^n\exp\left\{-\frac{1-b^2\delta_w^2}{2(1-\delta_w^2)}\sum_{i=1}^n u_{wi}\left(w_i-\frac{\delta_w}{\sqrt{u_{wi}}\sqrt{1-b^2\delta_w^2}}(h_{wi}-b)\right)\right\}$$
$$I_{(-1,1)}(\delta_w)$$

3.3.2 Residual analysis

Using the stochastic representation of the bivariate model given in Section 3.3, we can define the residual for the continuous variable as

$$\varepsilon_i = \frac{Y_i - \boldsymbol{X}_{ci}^t \hat{\boldsymbol{\beta}}_c - \hat{W}_i}{\sqrt{\hat{\sigma_c^2}}} \tag{3.4}$$

For the binary response, we can use the Deviance residual, defined as

$$d_i = sign(T_i - \hat{p}_i)\sqrt{-2\left(T_i log(\hat{p}_i) + (1 - T_i) log(1 - \hat{p}_i)\right)}. \tag{3.5}$$

where $p_i^{(m)} = F(\boldsymbol{X}_i^t \hat{\boldsymbol{\beta}}_b + \hat{W}_i | \hat{\gamma}_b, \hat{\boldsymbol{\nu}})$, $\hat{\boldsymbol{\beta}}_c, \hat{\boldsymbol{\beta}}_b, \hat{\sigma_c^2}, \hat{\gamma}_c, \hat{\gamma}_b, \hat{\boldsymbol{\nu}}_c$ and $\hat{\boldsymbol{\nu}}_b$ are the Bayesian estimates and $\hat{W}_i$ is the Bayesian estimate of the latent variable. We expected that the residuals in equation (3.4), a priori, approximately follows a Normal, SN, ST, SSL, SCN or SGT distribution, according to the respective adopted distribution, with $\boldsymbol{\nu}_c, \boldsymbol{\nu}_b$ and γ_c, γ_c equal to the Bayesian estimates.

For checking the goodness of fit, we can build envelope plots, using the above mentioned distributions to simulate the envelopes for the continuous response. For the binary response, the envelopes for the deviance residuals described in Section 2.3.2 are built as described in Appendix E.

3.4 Simulation study

We performed simulation studies in order to evaluate the performance of the model and the estimation method proposed in this work. All these models were implemented in JAGS ((PLUMMER, 2003)) through the interface provided by the rjags package ((PLUMMER, 2016)) available in R program ((R Development Core Team, 2008)). The codes are available from the authors upon request. We adopted weakly informative priors for all parameters, that is: $\beta_{b0} \sim N(0, 1000)$ and $\beta_{b1} \sim N(0, 1000)$, $\beta_{c0} \sim N(0, 1000)$ and $\beta_{c1} \sim N(0, 1000)$. For the γ_b, γ_c and γ_w we used the priors described in Section 3.3 with $\alpha_{\gamma_1} = \alpha_{\gamma_2} = 0.5$. To eliminate the effect of the initial values and to avoid correlations problems, we run a MCMC chain of size 600,000 with a burn-in of 100,000 and thin 1000, so we retain a valid MCMC chain of size 500. The values of the Gelman-Rubin statistics and the analyses of traceplots, Geweke and autocorrelation plots indicated that the MCMC algorithm converged and the autocorrelation were negligible.

3.4.1 Simulation study I

The objective of this simulation study is to analyze the quality of the estimates, in terms of bias and variance, when the two responses are correlated and the marginal and bivariate model are fitted. We considered two scenarios, varying the correlation magnitude between the two responses. Therefore, we simulate two data sets from the model

$$T_i = I(Z_i > 0) \tag{3.6}$$

$$Z_i = \beta_{b0} + \beta_{b1}x_i + w_i + \epsilon_i \quad i = 1, \ldots, n \tag{3.7}$$

$$Y_i = \beta_{c0} + \beta_{c1}x_i + w_i + \varepsilon_i \quad i = 1, \ldots, n \tag{3.8}$$

$$\tag{3.9}$$

where $\epsilon_i \sim SN_c(0, 1, -\gamma_b)$, $\varepsilon_i \sim SN_c(0, 1, \gamma_c)$ and $w_i \sim SN_c(0, 1, \gamma_w)$ (Scenario 1), $\epsilon_i \sim SN_c(0, 1, -\gamma_b)$, $\varepsilon_i \sim SN_c(0, 1, \gamma_c)$ and $w_i \sim SN_c(0, 10, \gamma_w)$ (Scenario 2) , $\boldsymbol{\beta} = (1, 2)^t$, $\gamma_b = \gamma_c = \gamma_w = 0.9$ and the covariate x_i was simulated from a $N(0, 1)$ distribution and centered in their respective mean. For Scenario 1, the correlation between the continuous and binary response is around .4 and for the second scenario, this correlation is approximately equal to .8.

We considered a sample size of n=500 and R=10 replicas were made. For each scenario, the bivariate model (considering the latent variable w_i) and the marginal model (without the w_i variable) were fitted and the mean of the estimate of parameter (Est), standard deviation of the estimates (SD), bias of the estimates (Bias), square root of the mean square error (RMSE), as described in section 1.5.1 were calculated.

Table 18 contains the simulation results for the bivariate and marginal models. We observe that for the bivariate model, all parameter are well recovered, with small standard deviations, bias, relative bias and RMSE. On the other hand, for the marginal model, the estimates of β_{b0} and β_{b1} had larger standard deviations compared with the bivariate model, and the γ_b estimate was not accurate. For σ^2_c the estimates presented higher bias than for the bivariate model. For the Scenario 2, presented in Table 19, the β_{b0}, β_{b1}, γ_b and σ^2_c estimates presented even worse results than what was observed in the first scenario.

Table 18 – Results of the simulation study: Scenario 1.

		Bivariate Model						Marginal Model					
Par	Real	Est.	SD	BIAS	RBIAS	RMSE	CR	Est.	SD	BIAS	RBIAS	RMSE	CR
β_{b0}	1.000	.946	.008	-0.054	.054	.054	1.000	.915	2.025	-0.085	.085	2.026	1.000
β_{b1}	2.000	1.849	.008	-0.151	.075	.151	1.000	2.745	11.412	.745	.373	11.436	1.000
β_{c0}	1.000	1.010	.002	.010	.010	.010	1.000	.956	.003	-0.044	.044	.044	1.000
β_{c1}	2.000	1.989	.002	-0.011	.006	.011	1.000	1.945	.002	-0.055	.028	.055	1.000
γ_b	.900	.923	.023	.023	.025	.032	1.000	.003	.123	-0.897	.997	.905	1.000
γ_c	.900	.858	.007	-0.042	.047	.043	1.000	.520	.003	-0.380	.423	.380	.000
γ_w	.900	.884	.006	-0.016	.017	.017	1.000	-	-	-	-	-	
σ^2_c	1.000	1.056	.008	.056	.056	.057	1.000	1.859	.004	.859	.859	.859	.000
σ^2_w	1.000	.873	.010	-0.127	.127	.128	1.000	-	-	-	-	-	

Table 19 – Results of the simulation study: Scenario 2.

		Bivariate Model						Marginal Model					
Par	Real	Est.	SD	BIAS	RBIAS	RMSE	CR	Est.	SD	BIAS	RBIAS	RMSE	CR
β_{b0}	1.000	0.845	.008	-0.155	.155	.155	1.000	.192	2.047	-0.808	.808	2.200	1.000
β_{b1}	2.000	2.143	.008	.143	.071	.143	1.000	-0.040	11.620	-2.040	1.020	11.798	1.000
β_{c0}	1.000	.968	.006	-0.032	.032	.033	1.000	1.066	.006	.066	.066	.067	1.000
β_{c1}	2.000	1.880	.003	-0.120	.060	.120	1.000	2.118	.002	.118	.059	.118	1.000
γ_b	.900	.799	.080	-0.101	.112	.129	1.000	.009	.099	-0.891	.990	.896	1.000
γ_c	.900	.607	.061	-0.293	.326	.299	1.000	.849	.002	-0.051	.057	.051	1.000
γ_w	.900	.862	.003	-0.038	.042	.038	1.000	-	-	-	-	-	
σ_c^2	1.000	.789	.011	-0.211	.211	.211	1.000	11.243	.027	10.243	10.243	10.243	.000
σ_w^2	10.000	9.820	.016	-0.180	.018	.180	1.000	-	-	-	-	-	

3.4.2 Simulation study II

The objective of this simulation study is to analyze the quality of the estimates for different sample sizes. We considered different scenarios based on some factors of interest. We generated data with samples of size n=50, 250, 500 and 1000, varying the values of γ_b, γ_c and γ_w, considering only one data set, from the following model

$$\begin{aligned} T_i &= I(Z_i > 0) \\ Z_i &= \beta_{b0} + \beta_{b1} x_i + w_i + \epsilon_i \quad i = 1, \dots, n \\ Y_i &= \beta_{c0} + \beta_{c1} x_i + w_i + \varepsilon_i \quad i = 1, \dots, n \end{aligned} \tag{3.10}$$

where $\epsilon_i \sim SN_c(0, 1, -\gamma_b)$, $\varepsilon_i \sim SN_c(0, 1, \gamma_c)$ and $w_i \sim SN_c(0, 1, \gamma_w)$, $\boldsymbol{\beta} = (1, 2)$, $\gamma_b \in \{-0.9, 0, 0.9\}$, $\gamma_c \in \{-0.9, 0, 0.9\}$, $\gamma_w \in \{-0.9, 0, 0.9\}$ and the covariate x_i was simulated from a $N(0, 1)$ distribution and centered in their respective mean.

Tables 20, 21 and 22 contain the simulation results for scenarios considering $\gamma_b = \gamma_c = \gamma_w = 0$, $\gamma_b = \gamma_c = \gamma_w = 0.9$ and $\gamma_b = \gamma_c = \gamma_w = -0.9$, respectively. The results from the other scenarios (other combinations of γ_b, γ_c and γ_w) are presented in Appendix G. It is possible to note that the estimates of $\boldsymbol{\beta}_b$, $\boldsymbol{\beta}_c$, σ_c^2 and σ_w^2 are accurate across all scenarios and sample sizes. As the sample size increases, the width of the credibility intervals become smaller. For the parameters $\gamma_b, \gamma_c, \gamma_w$, the estimates are accurate for large samples sizes. However, when $\gamma_j = 0$ j=b,c or w, the estimates present large credibility intervals.

Table 20 – Results of the simulation study under $\gamma_c = 0$, $\gamma_w = 0$ and $\gamma_b = 0$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	.970	.411	[.169; 1.786]	.959	.157	[.685; 1.305]	.977	.133	[.723; 1.222]	.960	.083	[.806; 1.133]
β_{b1}	2.000	1.915	.543	[1.050; 3.132]	1.767	.196	[1.355; 2.129]	1.997	.184	[1.647; 2.344]	1.879	.117	[1.639; 2.097]
β_{c0}	1.000	1.182	.223	[.723; 1.611]	1.050	.085	[.892; 1.220]	1.003	.074	[.866; 1.149]	1.015	.044	[.937; 1.107]
β_{c1}	2.000	1.788	.206	[.427; 2.226]	2.046	.084	[1.906; 2.221]	1.966	.068	[1.841; 2.105]	1.962	.042	[1.875; 2.035]
γ_b	.000	.087	.668	[-0.970; .995]	-0.328	.539	[-0.994; .717]	.532	.372	[-0.200; .994]	.318	.425	[-0.545; .993]
γ_c	.000	.180	.630	[-0.932; .991]	-0.171	.442	[-0.981; .644]	.322	.319	[-0.069; .984]	-0.357	.363	[-0.984; .113]
γ_w	.000	-0.221	.599	[-0.992; .921]	-0.280	.388	[-0.969; .235]	.561	.365	[-0.014; .990]	.265	.246	[-0.102; .720]
σ_c^2	1.000	.874	.514	[.018; 1.772]	.897	.225	[.449; 1.326]	1.359	.192	[1.000; 1.745]	.873	.114	[.656; 1.080]
σ_w^2	1.000	1.411	.608	[.303; 2.446]	1.054	.234	[.623; 1.483]	1.023	.207	[.638; 1.431]	1.174	.131	[.951; 1.462]

Table 21 – Results of the simulation study under $\gamma_c = 0.9$, $\gamma_w = 0.9$ and $\gamma_b = 0.9$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	1.600	.451	[.710; 2.507]	.992	.172	[.644; 1.325]	.898	.120	[.685; 1.143]	.969	.092	[.803; 1.163]
β_{b1}	2.000	1.460	.531	[.566; 2.530]	2.213	.275	[1.645; 2.684]	1.864	.158	[1.552; 2.163]	1.951	.120	[1.720; 2.178]
β_{c0}	1.000	1.130	.228	[.717; 1.568]	.786	.086	[.613; .957]	1.026	.062	[.901; 1.144]	.997	.043	[.919; 1.082]
β_{c1}	2.000	1.865	.188	[1.508; 2.221]	2.122	.076	[1.971; 2.254]	2.036	.044	[1.950; 2.129]	1.987	.036	[1.911; 2.053]
γ_b	.900	.038	.684	[-0.963; .994]	.742	.306	[.041; .995]	.201	.542	[-0.761; .994]	.822	.173	[.466; .995]
γ_c	.900	.034	.535	[-0.904; .979]	.353	.466	[-0.614; .982]	.916	.078	[.754; .992]	.842	.110	[.632; .991]
γ_w	.900	.546	.433	[-0.159; .992]	.814	.198	[.389; .993]	.894	.111	[.665; .991]	.782	.162	[.483; .990]
σ_c^2	1.000	.322	.341	[.005; .990]	.719	.216	[.367; 1.210]	1.214	.167	[.889; 1.517]	1.056	.110	[.812; 1.264]
σ_w^2	1.000	2.381	.680	[1.157; 3.942]	1.110	.248	[.624; 1.578]	.875	.162	[.568; 1.203]	.836	.112	[.638; 1.074]

Table 22 – Results of the simulation study under $\gamma_c = -0.9$, $\gamma_w = -0.9$ and $\gamma_b = -0.9$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	1.579	.444	[.789; 2.452]	1.271	.169	[.765; 1.392]	1.058	.115	[.718; 1.139]	.988	.080	[.821; 1.135]
β_{b1}	2.000	2.789	.826	[1.411; 4.527]	1.932	.244	[1.518; 2.451]	2.031	.181	[1.520; 2.170]	2.132	.128	[1.889; 2.359]
β_{c0}	1.000	.798	.223	[.372; 1.226]	1.238	.077	[1.089; 1.389]	.972	.061	[.865; 1.101]	1.013	.042	[.931; 1.086]
β_{c1}	2.000	2.002	.202	[1.546; 2.342]	1.907	.065	[1.795; 2.043]	1.940	.046	[1.850; 2.026]	1.980	.034	[1.913; 2.045]
γ_b	-0.900	.005	.661	[-0.956; .995]	-0.747	.515	[-0.594; .995]	-0.497	.445	[-0.994; .763]	-0.832	.173	[-0.995; -0.460]
γ_c	-0.900	-0.214	.622	[-0.990; .903]	-0.872	.203	[-0.991; -0.404]	-0.949	.045	[-0.993; -0.871]	-0.816	.122	[-0.984; -0.582]
γ_w	-0.900	.069	.660	[-0.948; .993]	-0.781	.165	[-0.991; -0.532]	-0.899	.099	[-0.989; -0.705]	-0.881	.100	[-0.989; -0.659]
σ_c^2	1.000	.678	.435	[.038; 1.506]	.734	.177	[.389; 1.044]	1.086	.173	[.817; 1.467]	.878	.098	[.693; 1.069]
σ_w^2	1.000	1.748	.648	[.639; 2.981]	.865	.193	[.500; 1.224]	.910	.173	[.547; 1.097]	.957	.108	[.726; 1.147]

3.5 Application

We analyze the dataset presented in (GAŁECKI; BURZYKOWSKI, 2013). The data consist on sample of 227 patients age-related macular degeneration (see (GUYER et al., 1997)). The objective was to evaluate the efficacy of an experimental treatment (interferon-α) with a corresponding placebo. For each patient it was evaluated his/her visual acuity in the beginning and after one year of study. This acuity is measured by counting how many letters of a standardized vision chart are corrected read. These charts display line letters of decreasing size that the patient must read from the top (large letters) to bottom (small letters). In this study, two outcomes were obtained in order to evaluate the efficacy of the treatment: the binary outcome T_i was defined as the loss at least three lines of vision at one year compared with their baseline performance and the continuous outcome Y_i are defined as the difference between patient's visual acuity from one year and the beginning of the study. The covariate considered in the analysis was the indicator variable $X_i = 1$ if the patient had received the treatment, or $X_i = 0$, if it was administrated the placebo.

We consider a regression model of the form $Y_i = \beta_{c0} + \beta_{c1}x_{1i} + w_i + \varepsilon_i$, for $i = 1, 2, \ldots, 227$, where w_i is the random effect. We also assume a latent variable $Z_i = \beta_{b0} + \beta_{b1}x_{1i} + w_i + \epsilon_i$, for $i = 1, 2, \ldots, 227$, such that $T_i = I(Z_i > 0)$, where $T_i = 1$ if the patient have lost loss at least three lines of vision at one year compared with their baseline performance.

We fitted eight models, assuming that: $\varepsilon \overset{iid}{\sim} SN_c(0, \sigma_c^2, \gamma_c)$, or $\varepsilon \overset{iid}{\sim} N(0, \sigma_c^2)$, $\epsilon \overset{iid}{\sim} SN_c(0, 1, -\gamma_b)$, or $\epsilon \overset{iid}{\sim} N(0, 1)$, $w_i \overset{iid}{\sim} SN_c(0, \sigma_w^2, \gamma_w)$, or $w_i \overset{iid}{\sim} N(0, \sigma_w^2)$. The summary of the results can be found in Table 23.

Table 23 – Distributions for ε_i, ϵ_i and w_i, respectively, for the vision dataset.

Model	Distribution		
M1	$SN_c(0, \sigma_c^2, \gamma_c)$	$SN_c(0, 1, -\gamma_b)$	$SN_c(0, \sigma_w^2, \gamma_w)$
M2	$SN_c(0, \sigma_c^2, \gamma_c)$	$SN_c(0, 1, -\gamma_b)$	$N(0, \sigma_w^2)$
M3	$SN_c(0, \sigma_c^2, \gamma_c)$	$N(0, 1)$	$SN_c(0, \sigma_w^2, \gamma_w)$
M4	$N(0, \sigma_c^2)$	$SN_c(0, 1, -\gamma_b)$	$SN_c(0, \sigma_w^2, \gamma_w)$
M5	$SN_c(0, \sigma_c^2, \gamma_c)$	$N(0, 1)$	$N(0, \sigma_w^2)$
M6	$N(0, \sigma_c^2)$	$SN_c(0, 1, -\gamma_b)$	$N(0, \sigma_w^2)$
M7	$N(0, \sigma_c^2)$	$N(0, 1)$	$SN_c(0, \sigma_w^2, \gamma_w)$
M8	$N(0, \sigma_c^2)$	$N(0, 1)$	$N(0, \sigma_w^2)$

The values for the MCMC algorithm were the same used in the simulation study available in Section 3.4.

Similarly to Section 1.4.4, the model comparison criteria was calculated used the likelihood

$$L(\boldsymbol{\beta}_b, \gamma_b, \boldsymbol{\nu}_b, \boldsymbol{\beta}_c, \gamma_c, \boldsymbol{\nu}_c|\boldsymbol{y}) = \prod_{i=1}^{n} \int_{-\infty}^{\infty} (F(\boldsymbol{X}_{bi}^t\boldsymbol{\beta}_b + w_i|\gamma_b, \boldsymbol{\nu}_b))^{t_i}(1 - F(\boldsymbol{X}_{bi}^t\boldsymbol{\beta}_b + w_i|\gamma_b, \boldsymbol{\nu}_b))^{1-t_i}$$
$$\times f(y_i|\boldsymbol{X}_{ci}^t\boldsymbol{\beta}_c + w_i, \sigma_c^2, \gamma_c, G, \boldsymbol{\nu}_c) f(w_i|0, \sigma_w^2, \gamma_w, G, \boldsymbol{\nu}_w) dw_i$$

Table 24 presents the statistics for model comparison. The model M3, that is, considering $\gamma_c \sim SN_c(0, \sigma_c^2, \gamma_c)$, $\gamma_b \sim N(0, 1)$ and $\gamma_w \sim SN_c(0, \sigma_w^2, \gamma_w)$ was selected by all criteria. Another model, with the same distributions, but not considering the correlation between the response was fitted and they were compared. For the marginal model the EAIC, EBIC, DIC and LPML criteria were, respectively, 1289.38, 1313.35, 1282.61 and -643.71. Therefore, we can conclude that the bivariate model outperforms the marginal model.

From Table 25, we can see the estimates for the bivariate and marginal model. Comparing the estimates of $\boldsymbol{\beta}_b$ for the bivariate and marginal model we noted that the estimates do not present many differences between them 25. From the bivariate model we can conclude that the treatment has almost no effect in the probability of loss at least three lines of vision but has a positive effect on the patient's acuity.

3.6 Conclusions

In this chapter, we developed a bivariate model for continuous and binary responses, using the SMSN class under the centered parameterization. We performed

Table 24 – **Vision dataset**: Statistics for model comparison

	Model							
criterion	M1	M2	M3	M4	M5	M6	M7	M8
EAIC	706.81	1883.50	695.27	728.89	1706.74	1776.18	797.67	2184.85
EBIC	737.63	1907.47	726.10	759.72	1730.71	1800.15	828.49	2208.82
DIC	694.41	2153.29	660.48	825.85	1766.64	1983.98	726.77	2257.07
LPML	-386.38	-2759.97	-376.17	-402.22	-1234.15	-2728.56	-438.81	-1676.89

Table 25 – **Vision dataset**: Posterior parameter estimates for the selected bivariate model and the marginal model.

	Model					
	Bivariate			Marginal		
	Est	SD	%95 HPD	Est	SD	%95 HPD
β_{b0}	-2.063	.779	[-3.519; -0.728]	-0.71	.129	[-0.959; -0.461]
β_{b1}	.271	.5	[-0.732; 1.229]	.106	.183	[-0.25; .449]
β_{c0}	2.446	1.048	[0.333; 4.326]	2.463	1.028	[0.537; 4.462]
β_{c1}	3.787	1.422	[1.225; 6.724]	3.801	1.438	[1.039; 6.605]
γ_c	.519	.113	[0.302; .742]	.472	.096	[0.276; .638]
γ_w	-0.167	.792	[-0.995; .989]	-	-	-
σ_c^2	128.894	13.89	[101.176; 154.289]	133.459	14.027	[105.457; 158.981]
σ_w^2	7.906	5.072	[0.545; 17.726]	-	-	-

Bayesian estimation and the method was applied in a real dataset. The simulation study performed in this chapter indicated that considering the bivariate model when the two outcomes are correlated can improve substantially the accuracy of the estimates. In addition, simulation studies indicated that the models considering asymmetry and/or heavy tails when the data have these characteristics present more accurate results than those models with the assumption of normality of the data.

4 Conclusion

In this work we developed a scale mixture of skew-normal distribution under the centered parameterization class of probability distributions as an alternative to the parameterization used in (FERREIRA; BOLFARINE; LACHOS, 2011). It was decided to use a new parameterization for this class for several reasons, among them, the simplicity of parameter interpretation compared to the parameterization used in (FERREIRA; BOLFARINE; LACHOS, 2011). Another motivation was the issues related to the estimation process of parameter λ in the direct parameterization. We have showed, through profiled log-likelihood, that the SMSN class under direct parameterization can heritage the problem caused by the non quadratic likelihood shape of the direct parameterization.

A class of linear regression models based on the SMSN family under the centered parameterization was introduced, and we developed the Bayesian estimation approach. Also, we described model comparison criteria, and we developed analysis of influential observations and residual analysis. Simulation studies were performed in order to evaluate the parameter recovery under different scenarios. We concluded that for values of $\boldsymbol{\nu}$ that generate distributions with heavy tails the estimates are very accurate. On the other hand, for values of $\boldsymbol{\nu}$ close to the skew normal (or the symmetric) model, the estimates tend to be biased and the credibility intervals to be large. However, as the sample size increases, the estimates are improved. An application of the proposed model in a real dataset was performed in order to show that heavy tails models (special cases of the developed class of linear regression model) provide better fits than the skew normal linear regression.

In the context of binary regression, we proposed a new class of link functions based on the SMSN class under the centered parameterization. This class of link functions include symmetrical, asymmetrical and robust link functions. We performed Bayesian estimation using latent variables to described the binary model. Some methods of residual analysis for binary data was discussed, and simulation studies were performed, evaluating parameter recovery, residual and influence analysis. The first simulation study showed some problems in the accuracy of the estimates of ν for the binary model, specially for the skew contaminated normal and skew generalized t models. However for the skew-t and skew slash models, when true value of ν indicated heavy tails the parameter was appropriate recovered. Also, we noted as sample size increases, the estimates of all parameter tend to be closer to real values. In the second study, we simulated data using normal, skew-normal and skew-t link functions and we fitted the skew-t, skew-normal and normal models. For this study we observed that when data were simulated using an asymmetric and/or heavy tail link functions and the probit model was fitted, the estimates of $\boldsymbol{\beta}$ were biased. For the residual and influence analysis studies, when the data was simulated using a heavy tail distribution and the skew normal model is fitted, the residuals tend to lying outside the confidence bands and we observe some observations as influentials. As in the linear model, an application was made in a study on the body fat percentage that indicated the skew and heavy tail link was preferred to the usual probit model. For the Beetle Mortality data we conclude that the specification of an asymmetric link provide better fit than the usual probit model.

Finally, we developed a bivariate model for continuous and binary responses, using the SMSN class under the centered parameterization. We performed Bayesian estimation and the method was applied in a real dataset. The simulation study performed in this chapter indicated that considering the bivariate model when the two outcomes are correlated can improve substantially the accuracy of the estimates. In addition, simulation studies indicated that the models considering asymmetry and/or heavy tails when the data have these characteristics present more accurate results than those models with the assumption of normality of the data.

As discussed in this work, the estimation of the parameter $\boldsymbol{\nu}$ for the SMSN class is a complicated task and more study should be performed. For future works, we intend to improve these estimates considering appropriate prior distributions. According to some tests performed by the authors, the use of Jeffreys' prior and the direct implementation of the Bayesian algorithm can produce better estimates, since the usual MCMC programs as WinBUGS, OpenBUGS, JAGS, do not allow to consider some kind of priors. Also, we intend to perform residual and influence analysis for the bivariate model. It may also be interesting to extend this model for the longitudinal case, as well for the multivariate regression model, considering multiple continuous and categorical responses.

Bibliography

ALBERT, J.; CHIB, S. Bayesian residual analysis for binary response regression models. v. 82, 02 1995. Citado na página 78.

ALBERT, J. H.; CHIB, S. Bayesian analysis of binary and polychotomous response data. *Journal of the American Statistical Association*, [American Statistical Association, Taylor & Francis, Ltd.], v. 88, n. 422, p. 669–679, 1993. ISSN 01621459. Disponível em: <http://www.jstor.org/stable/2290350>. Citado na página 74.

ANDO, T. Bayesian predictive information criterion for the evaluation of hierarchical bayesian and empirical bayes models. *Biometrika*, [Oxford University Press, Biometrika Trust], v. 94, n. 2, p. 443–458, 2007. ISSN 00063444. Disponível em: <http://www.jstor.org/stable/20441383>. Citado na página 44.

ANDREWS, D. F.; MALLOWS, C. L. Scale mixtures of normal distributions. *Journal of the Royal Statistical Society. Series B (Methodological)*, [Royal Statistical Society, Wiley], v. 36, n. 1, p. 99–102, 1974. ISSN 00359246. Disponível em: <http://www.jstor.org/stable/2984774>. Citado 4 vezes nas páginas 22, 24, 27, and 28.

ARELLANO-VALLE, R. B.; AZZALINI, A. The centred parametrization for the multivariate skew-normal distribution. *Journal of Multivariate Analysis*, v. 99, n. 7, p. 1362 – 1382, 2008. ISSN 0047-259X. Special Issue: Multivariate Distributions, Inference and Applications in Memory of Norman L. Johnson. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0047259X08000341>. Citado 3 vezes nas páginas 21, 25, and 35.

ARELLANO-VALLE, R. B.; GENTON, M. G.; LOSCHI, R. H. Shape mixtures of multivariate skew-normal distributions. *Journal of Multivariate Analysis*, v. 100, n. 1, p. 91 – 101, 2009. ISSN 0047-259X. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0047259X08000985>. Citado na página 24.

AZEVEDO, C. L.; BOLFARINE, H.; ANDRADE, D. F. Parameter recovery for a skew-normal irt model under a bayesian approach: hierarchical framework, prior and kernel sensitivity and sample size. *Journal of Statistical Computation and Simulation*, Taylor & Francis, v. 82, n. 11, p. 1679–1699, 2012. Disponível em: <https://doi.org/10.1080/00949655.2011.591798>. Citado 3 vezes nas páginas 75, 80, and 107.

AZZALINI, A. A class of distributions which includes the normal ones. *Scandinavian Journal of Statistics*, v. 12, p. 171–178, 1985. Citado 5 vezes nas páginas 21, 22, 23, 24, and 25.

_____. The skew-normal distribution and related multivariate families*. *Scandinavian Journal of Statistics*, Blackwell Publishing Ltd, v. 32, n. 2, p. 159–188, 2005. ISSN 1467-9469. Disponível em: <http://dx.doi.org/10.1111/j.1467-9469.2005.00426.x>. Citado na página 25.

AZZALINI, A.; CAPITANIO, A. *The Skew-Normal and Related Families:*. [S.l.]: Cambridge University Press, 2013. (Institute of Mathematical Statistics Monographs). Citado 2 vezes nas páginas 24 and 25.

BAZáN, J. L.; BOLFARINE, H.; BRANCO, M. D. A framework for skew-probit links in binary regression. *Communications in Statistics - Theory and Methods*, Taylor & Francis, v. 39, n. 4, p. 678–697, 2010. Disponível em: <https://doi.org/10.1080/03610920902783849>. Citado 4 vezes nas páginas 22, 69, 74, and 99.

BAZáN, J. L.; ROMEO, J. S.; RODRIGUES, J. Bayesian skew-probit regression for binary response data. *Braz. J. Probab. Stat.*, Brazilian Statistical Association, v. 28, n. 4, p. 467–482, 11 2014. Disponível em: <https://doi.org/10.1214/13-BJPS218>. Citado 2 vezes nas páginas 22 and 70.

BRANCO, M. D.; DEY, D. K. A general class of multivariate skew-elliptical distributions. *Journal of Multivariate Analysis*, v. 79, n. 1, p. 99 – 113, 2001. ISSN 0047-259X. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0047259X00919602>. Citado 2 vezes nas páginas 22 and 27.

BURR, I. W. Cumulative frequency functions. *Ann. Math. Statist.*, The Institute of Mathematical Statistics, v. 13, n. 2, p. 215–232, 06 1942. Disponível em: <https://doi.org/10.1214/aoms/1177731607>. Citado 2 vezes nas páginas 22 and 70.

CABRAL, C. R. B.; LACHOS, V. H.; MADRUGA, M. R. Bayesian analysis of skew-normal independent linear mixed models with heterogeneity in the random-effects population. *Journal of Statistical Planning and Inference*, v. 142, n. 1, p. 181 – 200, 2012. ISSN 0378-3758. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0378375811002771>. Citado 4 vezes nas páginas 42, 77, 112, and 113.

CATALANO, P. J.; RYAN, L. M. Bivariate latent variable models for clustered discrete and continuous outcomes. *Journal of the American Statistical Association*, [American Statistical Association, Taylor & Francis, Ltd.], v. 87, n. 419, p. 651–658, 1992. ISSN 01621459. Disponível em: <http://www.jstor.org/stable/2290200>. Citado 2 vezes nas páginas 21 and 104.

CHAVES, N. L. *Birnbaum-Saunders regression models based on skew-normal centered distribution*. Dissertação (Mestrado) — IMECC/UNICAMP, 2015. Citado 3 vezes nas páginas 35, 45, and 47.

CHEN, M.-H.; DEY, D. K.; SHAO, Q.-M. A new skewed link model for dichotomous quantal response data. *Journal of the American Statistical Association*, [American Statistical Association, Taylor & Francis, Ltd.], v. 94, n. 448, p. 1172–1186, 1999. ISSN 01621459. Disponível em: <http://www.jstor.org/stable/2669933>. Citado 2 vezes nas páginas 22 and 69.

CHEN, Y.; HANSON, T. Copula regression models for discrete and mixed bivariate responses. *Journal of Statistical Theory and Practice*, Taylor & Francis, v. 11, n. 4, p. 515–530, 2017. Disponível em: <https://doi.org/10.1080/15598608.2016.1278059>. Citado na página 105.

CHO, H.; IBRAHIM, J. G.; SINHA, D.; ZHU, H. Bayesian case influence diagnostics for survival models. *Biometrics*, [Wiley, International Biometric Society], v. 65, n. 1, p. 116–124, 2009. ISSN 0006341X, 15410420. Disponível em: <http://www.jstor.org/stable/25502250>. Citado 2 vezes nas páginas 44 and 45.

COLLETT, D. *Modelling Binary Data, Second Edition*. Taylor & Francis, 2002. (Chapman & Hall/CRC Texts in Statistical Science). ISBN 9781584883241. Disponível em: <https://books.google.com.br/books?id=LMRAIBEbdqsC>. Citado 2 vezes nas páginas 22 and 69.

COX, D. R.; WERMUTH, N. Response models for mixed binary and quantitative variables. *Biometrika*, [Oxford University Press, Biometrika Trust], v. 79, n. 3, p. 441–461, 1992. ISSN 00063444. Disponível em: <http://www.jstor.org/stable/2336778>. Citado 3 vezes nas páginas 21, 104, and 105.

CZADO, C. Bayesian inference of binary regression models with parametric link. *Journal of Statistical Planning and Inference*, v. 41, n. 2, p. 121 – 140, 1994. ISSN 0378-3758. Disponível em: <http://www.sciencedirect.com/science/article/pii/0378375894901589>. Citado na página 99.

CZADO, C.; SANTNER, T. J. The effect of link misspecification on binary regression inference. *Journal of Statistical Planning and Inference*, v. 33, n. 2, p. 213 – 231, 1992. ISSN 0378-3758. Disponível em: <http://www.sciencedirect.com/science/article/pii/037837589290069S>. Citado 2 vezes nas páginas 22 and 69.

DUNSON, D. B. Bayesian latent variable models for clustered mixed outcomes. *Journal of the Royal Statistical Society. Series B (Statistical Methodology)*, [Royal Statistical Society, Wiley], v. 62, n. 2, p. 355–366, 2000. ISSN 13697412, 14679868. Disponível em: <http://www.jstor.org/stable/3088864>. Citado na página 105.

FARIAS, R. B. A.; BRANCO, M. D. Latent residual analysis in binary regression with skewed link. *Braz. J. Probab. Stat.*, Brazilian Statistical Association, v. 26, n. 4, p. 344–357, 11 2012. Disponível em: <https://doi.org/10.1214/11-BJPS143>. Citado na página 78.

FERREIRA, C. da S.; BOLFARINE, H.; LACHOS, V. H. Skew scale mixtures of normal distributions: Properties and estimation. *Statistical Methodology*, v. 8, n. 2, p. 154 – 171, 2011. ISSN 1572-3127. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1572312710000948>. Citado 8 vezes nas páginas 22, 24, 27, 28, 35, 68, 121, and 130.

FITZMAURICE, G. M.; LAIRD, N. M. Regression models for a bivariate discrete and continuous outcome with clustering. *Journal of the American Statistical Association*, Taylor & Francis, v. 90, n. 431, p. 845–852, 1995. Disponível em: <http://www.tandfonline.com/doi/abs/10.1080/01621459.1995.10476583>. Citado 2 vezes nas páginas 21 and 104.

GAŁECKI, A.; BURZYKOWSKI, T. *Linear Mixed-Effects Models Using R: A Step-by-Step Approach*. Springer New York, 2013. (Springer Texts in Statistics). ISBN 9781461439004. Disponível em: <https://books.google.com.br/books?id=rbk_AAAAQBAJ>. Citado na página 118.

GARAY, A. M.; BOLFARINE, H.; LACHOS, V. H.; CABRAL, C. R. Bayesian analysis of censored linear regression models with scale mixtures of normal distributions. *Journal of Applied Statistics*, Taylor & Francis, v. 42, n. 12, p. 2694–2714, 2015. Disponível em: <http://dx.doi.org/10.1080/02664763.2015.1048671>. Citado 2 vezes nas páginas 44 and 45.

GARAY, A. M.; LACHOS, V. H.; ABANTO-VALLE, C. A. Nonlinear regression models based on scale mixtures of skew-normal distributions. *Journal of the Korean Statistical Society*, v. 40, n. 1, p. 115 – 124, 2011. ISSN 1226-3192. Disponível em: <http://www.sciencedirect.com/science/article/pii/S1226319210000736>. Citado 2 vezes nas páginas 31 and 44.

GELMAN, A. Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis*, v. 1, p. 1–19, 2006. Citado 2 vezes nas páginas 40 and 75.

GEMAN, S.; GEMAN, D. Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Trans. Pattern Anal. Mach. Intell.*, IEEE Computer Society, Washington, DC, USA, v. 6, n. 6, p. 721–741, nov. 1984. ISSN 0162-8828. Disponível em: <http://dx.doi.org/10.1109/TPAMI.1984.4767596>. Citado 3 vezes nas páginas 40, 75, and 107.

GENTON, M. *Skew-Elliptical Distributions and Their Applications: A Journey Beyond Normality*. CRC Press, 2004. ISBN 9780203492000. Disponível em: <https://books.google.com.br/books?id=M8yIk467vjkC>. Citado 2 vezes nas páginas 24 and 25.

GUERRERO, V. M.; JOHNSON, R. A. Use of the box-cox transformation with binary response models. *Biometrika*, [Oxford University Press, Biometrika Trust], v. 69, n. 2, p. 309–314, 1982. ISSN 00063444. Disponível em: <http://www.jstor.org/stable/2335404>. Citado 2 vezes nas páginas 22 and 69.

GUPTA, A. K.; NGUYEN, T. T.; SANQUI, J. A. T. Characterization of the skew-normal distribution. *Annals of the Institute of Statistical Mathematics*, v. 56, n. 2, p. 351–360, 2004. Citado 2 vezes nas páginas 21 and 24.

GUYER, D.; ADAMIS, A.; YANNUZZI, L.; GRAGOUDAS, E.; COSCAS, G.; BRANCATO, R.; BIRD, A.; SOUBRANE, G.; MILLER, J.; BRESSLER, N.; SLAKTER, J.; BUYSE, M.; O'SHAUGHNESSY, D.; DEROSA, J. Interferon alfa-2a is ineffective for patients with choroidal neovascularization secondary to age-related macular degeneration: Results of a prospective randomized placebo-controlled clinical trial. *JAMA Ophthalmology*, American Medical Association, v. 115, n. 7, p. 865–872, 1997. ISSN 2168-6165. Citado 3 vezes nas páginas 21, 104, and 118.

HASTINGS, W. K. Monte carlo sampling methods using markov chains and their applications. *Biometrika*, [Oxford University Press, Biometrika Trust], v. 57, n. 1, p. 97–109, 1970. ISSN 00063444. Disponível em: <http://www.jstor.org/stable/2334940>. Citado 3 vezes nas páginas 40, 75, and 107.

HENZE, N. A probabilistic representation of the 'skew-normal' distribution. *Scandinavian Journal of Statistics*, [Board of the Foundation of the Scandinavian Journal of Statistics, Wiley], v. 13, n. 4, p. 271–275, 1986. ISSN 03036898, 14679469. Disponível em: <http://www.jstor.org/stable/4616036>. Citado na página 26.

JOHNSON, R. W. Fitting percentage of body fat to simple body measurements. *Journal of Statistics Education*, Taylor & Francis, v. 4, n. 1, p. null, 1996. Disponível em: <https://doi.org/10.1080/10691898.1996.11910505>. Citado na página 93.

KIM, H.-J. Binary regression with a class of skewed t link models. *Communications in Statistics - Theory and Methods*, Taylor & Francis, v. 31, n. 10, p. 1863–1886, 2002. Disponível em: <https://doi.org/10.1081/STA-120014917>. Citado na página 93.

KIM, S.; CHEN, M.-H.; DEY, D. K. Flexible generalized t-link models for binary response data. *Biometrika*, v. 95, n. 1, p. 93–106, 2008. Disponível em: <+http://dx.doi.org/10.1093/biomet/asm079>. Citado 2 vezes nas páginas 23 and 70.

LEON, A. R. de; WU, B. Copula-based regression models for a bivariate mixed discrete and continuous outcome. *Statistics in Medicine*, Wiley Online Library, v. 30, n. 2, p. 175–185, 2011. Citado na página 105.

LIU, C. Robit regression: A simple robust alternative to logistic and probit regression. In: ______. *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives*. John Wiley & Sons, Ltd, 2005. p. 227–238. ISBN 9780470090459. Disponível em: <http://dx.doi.org/10.1002/0470090456.ch21>. Citado 2 vezes nas páginas 23 and 70.

MCCULLAGH, P.; NELDER, J. *Generalized Linear Models, Second Edition*. Taylor & Francis, 1989. (Chapman & Hall/CRC Monographs on Statistics & Applied Probability). ISBN 9780412317606. Disponível em: <https://books.google.com.br/books?id=h9kFH2_FfBkC>. Citado na página 78.

NAGLER, J. Scobit: An alternative estimator to logit and probit. *American Journal of Political Science*, [Midwest Political Science Association, Wiley], v. 38, n. 1, p. 230–255, 1994. ISSN 00925853, 15405907. Disponível em: <http://www.jstor.org/stable/2111343>. Citado 2 vezes nas páginas 22 and 70.

OLKIN, I.; TATE, R. F. Multivariate correlation models with mixed discrete and continuous variables. *Ann. Math. Statist.*, The Institute of Mathematical Statistics, v. 32, n. 2, p. 448–465, 06 1961. Disponível em: <https://doi.org/10.1214/aoms/1177705052>. Citado 2 vezes nas páginas 21 and 104.

PEWSEY, A. Problems of inference for azzalini's skewnormal distribution. *Journal of Applied Statistics*, v. 27, n. 7, p. 859–870, 2000. Disponível em: <http://dx.doi.org/10.1080/02664760050120542>. Citado 2 vezes nas páginas 22 and 25.

PLUMMER, M. Jags: A program for analysis of bayesian graphical models using gibbs sampling. v. 124, 04 2003. Citado 3 vezes nas páginas 46, 79, and 115.

______. *rjags: Bayesian Graphical Models using MCMC*. [S.l.], 2016. R package version 4-6. Disponível em: <https://CRAN.R-project.org/package=rjags>. Citado 3 vezes nas páginas 46, 79, and 115.

R Development Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2008. ISBN 3-900051-07-0. Disponível em: <http://www.R-project.org>. Citado 3 vezes nas páginas 46, 79, and 115.

SAMMEL, M. D.; RYAN, L. M.; LEGLER, J. M. Latent variable models for mixed discrete and continuous outcomes. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Blackwell Publishers Ltd., v. 59, n. 3, p. 667–678, 1997. ISSN 1467-9868. Disponível em: <http://dx.doi.org/10.1111/1467-9868.00090>. Citado 2 vezes nas páginas 104 and 105.

SANTOS, J. R. S. dos. *A multiple group IRT model with skew-normal latent trait distribution under the centred parametrization*. Dissertação (Mestrado) — IMECC/UNICAMP, 2012. Citado 2 vezes nas páginas 25 and 35.

SPIEGELHALTER, D. J.; BEST, N. G.; CARLIN, B. P.; LINDE, A. V. D. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, Blackwell Publishers, v. 64, n. 4, p. 583–639, 2002. ISSN 1467-9868. Disponível em: <http://dx.doi.org/10.1111/1467-9868.00353>. Citado na página 44.

STUKEL, T. A. Generalized logistic models. *Journal of the American Statistical Association*, [American Statistical Association, Taylor & Francis, Ltd.], v. 83, n. 402, p. 426–431, 1988. ISSN 01621459. Disponível em: <http://www.jstor.org/stable/2288858>. Citado 2 vezes nas páginas 22 and 69.

TEIMOURIAN, M.; BAGHFALAKI, T.; GANJALI, M.; BERRIDGE, D. Joint modeling of mixed skewed continuous and ordinal longitudinal responses: a bayesian approach. *Journal of Applied Statistics*, Taylor & Francis, v. 42, n. 10, p. 2233–2256, 2015. Disponível em: <https://doi.org/10.1080/02664763.2015.1023557>. Citado na página 105.

TEIXEIRA-PINTO, A.; NORMAND, S.-L. T. Correlated bivariate continuous and binary outcomes: Issues and applications. *Statistics in Medicine*, John Wiley & Sons, Ltd., v. 28, n. 13, p. 1753–1773, 2009. ISSN 1097-0258. Disponível em: <http://dx.doi.org/10.1002/sim.3588>. Citado 3 vezes nas páginas 21, 104, and 105.

VILCA, F.; AZEVEDO, C. L. N.; BALAKRISHNAN, N. Bayesian inference for sinh-normal/independent nonlinear regression models. *Journal of Applied Statistics*, Taylor & Francis, v. 44, n. 11, p. 2052–2074, 2017. Disponível em: <https://doi.org/10.1080/02664763.2016.1238058>. Citado na página 35.

WEISBERG, S. *Applied Linear Regression*. Third. Hoboken NJ: Wiley, 2005. Disponível em: <http://www.stat.umn.edu/alr>. Citado na página 63.

WEST, M. On scale mixtures of normal distributions. *Biometrika*, [Oxford University Press, Biometrika Trust], v. 74, n. 3, p. 646–648, 1987. ISSN 00063444. Disponível em: <http://www.jstor.org/stable/2336707>. Citado na página 27.

ZILKO, A. A.; KUROWICKA, D. Copula in a multivariate mixed discrete–continuous model. *Computational Statistics & Data Analysis*, v. 103, p. 28 – 55, 2016. ISSN 0167-9473. Disponível em: <http://www.sciencedirect.com/science/article/pii/S0167947316300895>. Citado na página 105.

Appendix

APPENDIX A – Log-likelihoods for SMSN family under the direct and centered parametrization for section 1.3

Consider a random sample $\boldsymbol{Y} = (Y_1, Y_1, \ldots, Y_n)^t$ from the SMSN family under direct and centered parameterization and the respective observed values $\boldsymbol{y} = (y_1, y_1, \ldots, y_n)^t$. We shall denote the log-likelihood under the centered parametrization by $l(\theta|y)_{CP}$. Under direct parameterization we shall use the density functions as described in (FERREIRA; BOLFARINE; LACHOS, 2011) and denote the respective log-likelihoods by $l(\theta|y)_{DP}$. Let $\theta = (\mu, \sigma, \lambda, \boldsymbol{\nu})$, then, the respective functions are:

A.1 Skew-t distribution:

$$\begin{aligned} l(\theta|y)_{DP} = nln(2) - \frac{n}{2}ln(\sigma^2) - \frac{n}{2}ln(\pi\nu) + nln\left(\Gamma\left(\frac{\nu+1}{2}\right)\right) - nln\left(\Gamma\left(\frac{\nu}{2}\right)\right) \\ - \left(\frac{\nu+1}{2}\right)\sum_{i=1}^{n} ln\left(1 + \frac{\delta_i}{\nu}\right) + \sum_{i=1}^{n} ln\left(\Phi\left(\lambda\frac{y_i-\mu}{\sigma}\right)\right) \end{aligned} \tag{A.1}$$

where $\delta_i = \dfrac{(y_i-\mu)^2}{\sigma^2}$

$$\begin{aligned} l(\theta|y)_{CP} = nln(2) + \frac{n\nu}{2}ln\left(\frac{\nu}{2}\right) - \frac{n}{2}ln(\sigma^2) - nln(\omega_1) - \frac{n}{2}ln(2\pi) - nln\left(\Gamma\left(\frac{\nu}{2}\right)\right) - \frac{n\xi_1^2}{2\omega_1^2} \\ + \sum_{i=1}^{n} ln\left(\int_0^\infty u_i^{\frac{\nu_1+1}{2}-1} \exp\left\{-\frac{1}{2}\left[u_i(d_i^2+\nu) - 2\sqrt{u_i}d_i\frac{\xi_1}{\omega_1}\right]\right\}\Phi\left(\lambda\left(\sqrt{u_i}d_i - \frac{\xi_1}{\omega_1}\right)\right)du_i\right) \end{aligned} \tag{A.2}$$

where $d_i = \dfrac{(y_i-\mu)^2}{\omega_1^2\sigma^2}$

A.2 Skew-slash distribution:

$$l(\theta|y)_{DP} = nln(2) - nln(\nu)\sum_{i=1}^{n} ln\left(\Phi\left(\lambda\frac{y_i-\mu}{\sigma}\right)\right)\sum_{i=1}^{n} ln\left(\int_0^1 u_i^{\nu-1}\phi(y_i|\mu, \sigma^2/u_i)du_i\right) \tag{A.3}$$

$$l(\theta|y)_{CP} = nln(2) + nln(\nu) - \frac{n}{2}ln(\sigma^2) - nln(\omega_1) - \frac{n}{2}ln(2\pi) - \frac{n\xi_1^2}{2\omega_1^2}$$
$$+\sum_{i=1}^{n} ln\left(\int_0^1 u_i^{\nu+\frac{1}{2}-1}\exp\left\{-\frac{1}{2}\left[u_i(d_i^2+\nu) - 2\sqrt{u_i}d\frac{\xi_1}{\omega_1}\right]\right\}\Phi\left(\lambda\left(\sqrt{u_i}d_i - \frac{\xi_1}{\omega_1}\right)\right)du_i\right) \quad (A.4)$$

where $d_i = \frac{(y_i-\mu)^2}{\omega_1^2\sigma^2}$

A.3 Skew-contaminated normal distribution:

$$l(\theta|y)_{DP} = nln(2) + \sum_{i=1}^{n} ln\left(\Phi\left(\lambda\frac{y_i-\mu}{\sigma}\right)\right) + \sum_{i=1}^{n} ln\left(\nu_1\phi(y_i|\mu,\sigma^2/\nu_2) + (1-\nu_1)\phi(y_i|\mu,\sigma^2)\right) \quad (A.5)$$

$$\begin{aligned} l(\theta|y)_{CP} &= nln(2) \\ &+ \sum_{i=1}^{n} log\left[\nu_1\frac{\sqrt{\nu_2}}{\sqrt{2\pi\sigma^2}\omega_1}\exp\left\{-\frac{1}{2}\left(\sqrt{\nu_2}d_i - \frac{\xi_1}{\omega_1}\right)^2\right\}\Phi\left(\lambda\left(\sqrt{\nu_2}d_i - \frac{\xi_1}{\omega_1}\right)\right)\right. \\ &\left. + (1-\nu_1)\frac{1}{\sqrt{2\pi\sigma^2}\omega_1}\exp\left\{-\frac{1}{2}\left(d_i - \frac{\xi_1}{\omega_1}\right)^2\right\}\Phi\left(\lambda\left(d_i - \frac{\xi_1}{\omega_1}\right)\right)\right] \end{aligned} \quad (A.6)$$

where $d_i = \frac{(y_i-\mu)^2}{\omega_1^2\sigma^2}$

A.4 Skew generalized t distribution:

$$\begin{aligned} l(\theta|y)_{DP} = nln(2) - \frac{n}{2}ln(\pi\nu_2) + nln\left(\Gamma\left(\frac{\nu_1+1}{2}\right)\right) - nln\left(\Gamma\left(\frac{\nu_1}{2}\right)\right) \\ - \left(\frac{\nu_1+1}{2}\right)\sum_{i=1}^{n} ln\left(1+\frac{\delta_i}{\nu_2}\right) + \sum_{i=1}^{n} ln\left(\Phi\left(\lambda(y_i-\mu)\right)\right) \end{aligned} \quad (A.7)$$

where $\delta_i = (y_i-\mu)^2$

$$l(\theta|y)_{CP} = nln(2) + \frac{n\nu_1}{2}ln\left(\frac{\nu_2}{2}\right) - nln(\omega_1) - \frac{n}{2}ln(2\pi) - nln\left(\Gamma\left(\frac{\nu_1}{2}\right)\right) - \frac{n\xi_1^2}{2\omega_1^2}$$
$$+\sum_{i=1}^{n} ln\left(\int_0^\infty u_i^{\frac{\nu_1+1}{2}-1}\exp\left\{-\frac{1}{2}\left[u_i(d_i^2+\nu_2) - 2\sqrt{u_i}d_i\frac{\xi_1}{\omega_1}\right]\right\}\Phi\left(\lambda\left(\sqrt{u_i}d_i - \frac{\xi_1}{\omega_1}\right)\right)du_i\right) \quad (A.8)$$

where $d_i = \dfrac{(y_i - \mu)^2}{\omega_1^2}$

APPENDIX B – Results of the simulations study: parameter recovery for section 1.5.1

Here we presented the tables with the results for the simulation study in Section 1.5.1 containing the scenarios not presented in this section.

B.1 Skew-t

Table 26 – Results of the simulation study for the skew-t model with $\nu = 3$ and $\gamma = 0$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	1.1989	.0001	.1989	.1989	.1989	1.0000	.9045
	β_1	2.0000	2.4244	.0002	.4244	.2122	.4244	1.0000	1.2438
	γ	.0000	.6855	.0005	.6855	-	.6855	1.0000	1.3258
	σ^2	1.0000	1.9017	.0006	.9017	.9017	.9017	.9200	2.2816
	ν	3.0000	11.1746	.6529	8.1746	2.7249	8.2007	1.0000	43.2765
250	β_0	1.0000	.9927	.0001	-0.0073	.0073	.0073	1.0000	.3980
	β_1	2.0000	2.1147	.0000	.1147	.0574	.1147	1.0000	.2830
	γ	.0000	-0.1937	.0013	-0.1937	-	.1938	1.0000	1.2958
	σ^2	1.0000	.9713	.0000	-0.0287	.0287	.0287	1.0000	.6195
	ν	3.0000	2.7525	.0005	-0.2475	.0825	.2475	1.0000	1.9546
500	β_0	1.0000	.9783	.0001	-0.0217	.0217	.0217	1.0000	.2638
	β_1	2.0000	2.1271	.0000	.1271	.0636	.1271	.0000	.1960
	γ	.0000	.2167	.0021	.2167	-	.2167	1.0000	.9638
	σ^2	1.0000	.9395	.0000	-0.0605	.0605	.0605	1.0000	.4224
	ν	3.0000	2.8869	.0005	-0.1131	.0377	.1131	1.0000	1.5700
1000	β_0	1.0000	1.1134	.0000	.1134	.1134	.1134	.0000	.1823
	β_1	2.0000	2.0789	.0000	.0789	.0395	.0789	.5800	.1605
	γ	.0000	.0829	.0001	.0829	-	.0829	1.0000	.5268
	σ^2	1.0000	1.1551	.0000	.1551	.1551	.1551	1.0000	.3662
	ν	3.0000	3.5113	.0002	.5113	.1704	.5113	1.0000	1.6394

Table 27 – Results of the simulation study for the skew-t model with $\nu = 3$ and $\gamma = 0.9$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
	β_0	1.0000	.8729	.0000	-0.1271	.1271	.1271	1.0000	.6854
	β_1	2.0000	2.0386	.0000	.0386	.0193	.0386	1.0000	.5787
50	γ	.9000	.5893	.0007	-0.3107	.3452	.3107	1.0000	1.5500
	σ^2	1.0000	.8835	.0002	-0.1165	.1165	.1165	1.0000	1.2310
	ν	3.0000	5.0219	.3255	2.0219	.6740	2.0480	1.0000	12.3533
	β_0	1.0000	.9390	.0000	-0.0610	.0610	.0610	1.0000	.3708
	β_1	2.0000	2.0527	.0000	.0527	.0263	.0527	1.0000	.2828
250	γ	.9000	.7117	.0005	-0.1883	.2092	.1883	1.0000	1.0066
	σ^2	1.0000	.9624	.0000	-0.0376	.0376	.0376	1.0000	.5687
	ν	3.0000	2.6250	.0002	-0.3750	.1250	.3750	1.0000	1.4837
	β_0	1.0000	.8876	.0000	-0.1124	.1124	.1124	.5800	.2302
	β_1	2.0000	1.8935	.0000	-0.1065	.0532	.1065	.0400	.1956
500	γ	.9000	.8008	.0003	-0.0992	.1103	.0992	1.0000	.5600
	σ^2	1.0000	.8368	.0000	-0.1632	.1632	.1632	1.0000	.3815
	ν	3.0000	3.0511	.0007	.0511	.0170	.0511	1.0000	1.8051
	β_0	1.0000	1.0332	.0000	.0332	.0332	.0332	1.0000	.1672
	β_1	2.0000	1.9943	.0000	-0.0057	.0028	.0057	1.0000	.1291
1000	γ	.9000	.8931	.0001	-0.0069	.0076	.0069	1.0000	.3058
	σ^2	1.0000	.9188	.0000	-0.0812	.0812	.0812	1.0000	.3014
	ν	3.0000	2.7554	.0001	-0.2446	.0815	.2446	1.0000	1.0133

Table 28 – Results from the simulation using the skew-t distribution with $\nu = 10$ and $\gamma = -0.9$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
	β_0	1.0000	1.0790	.0001	.0790	.0790	.0790	1.0000	.6424
	β_1	2.0000	2.0140	$<$ .0001	.0140	.0070	.0140	1.0000	.4660
50	γ	-0.9000	-0.9078	$<$ .0001	-0.0078	.0087	.0078	1.0000	.3910
	σ^2	1.0000	1.0124	.0001	.0124	.0124	.0124	1.0000	1.1266
	ν	10.0000	15.2063	.6516	5.2063	.5206	5.2469	1.0000	56.3893
	β_0	1.0000	.9673	$<$ .0001	-0.0327	.0327	.0327	1.0000	.2755
	β_1	2.0000	1.9188	$<$ .0001	-0.0812	.0406	.0812	1.0000	.1932
250	γ	-0.9000	-0.9320	$<$ .0001	-0.0320	.0355	.0320	1.0000	.1860
	σ^2	1.0000	1.0908	$<$ .0001	.0908	.0908	.0908	1.0000	.5211
	ν	10.0000	18.2538	.9761	8.2538	.8254	8.3113	1.0000	51.5009
	β_0	1.0000	1.0257	$<$ .0001	.0257	.0257	.0257	1.0000	.1852
	β_1	2.0000	2.0335	$<$ .0001	.0335	.0168	.0335	1.0000	.1449
500	γ	-0.9000	-0.9299	$<$ .0001	-0.0299	.0332	.0299	1.0000	.1950
	σ^2	1.0000	.9048	$<$ .0001	-0.0952	.0952	.0952	1.0000	.3326
	ν	10.0000	10.2599	.0756	.2599	.0260	.2706	1.0000	13.1968
	β_0	1.0000	.9874	$<$ .0001	-0.0126	.0126	.0126	1.0000	.1414
	β_1	2.0000	1.9862	$<$ .0001	-0.0138	.0069	.0138	1.0000	.1161
1000	γ	-0.9000	-0.9068	.0001	-0.0068	.0075	.0068	1.0000	.1936
	σ^2	1.0000	1.1018	$<$ .0001	.1018	.1018	.1018	1.0000	.2980
	ν	10.0000	13.3918	.4368	3.3918	.3392	3.4198	1.0000	22.5519

Table 29 – Results of the simulation study for the skew-t model with $\nu = 10$ and $\gamma = 0$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
	β_0	1.0000	.8560	.0001	-0.1440	.1440	.1440	1.0000	.7409
	β_1	2.0000	1.9019	.0000	-0.0981	.0490	.0981	1.0000	.5369
50	γ	.0000	-0.3710	.0034	-0.3710	-	.3710	1.0000	1.6774
	σ^2	1.0000	1.2724	.0002	.2724	.2724	.2724	1.0000	1.4413
	ν	10.0000	10.2194	.4003	.2194	.0219	.4565	1.0000	41.7672
	β_0	1.0000	1.0065	.0000	.0065	.0065	.0065	1.0000	.2657
	β_1	2.0000	2.0333	.0000	.0333	.0166	.0333	1.0000	.2712
250	γ	.0000	.1646	.0002	.1646	-	.1646	1.0000	.7589
	σ^2	1.0000	.9967	.0000	-0.0033	.0033	.0033	1.0000	.4578
	ν	10.0000	22.6835	2.5129	12.6835	1.2683	12.9300	1.0000	74.1976
	β_0	1.0000	.9783	.0000	-0.0217	.0217	.0217	1.0000	.2025
	β_1	2.0000	2.0090	.0000	.0090	.0045	.0090	1.0000	.2014
500	γ	.0000	-0.0300	.0000	-0.0300	-	.0300	1.0000	.5118
	σ^2	1.0000	1.0818	.0000	.0818	.0818	.0818	1.0000	.3887
	ν	10.0000	9.9178	.0376	-0.0822	.0082	.0904	1.0000	13.8089
	β_0	1.0000	.9921	.0000	-0.0079	.0079	.0079	1.0000	.1313
	β_1	2.0000	1.9664	.0000	-0.0336	.0168	.0336	1.0000	.1314
1000	γ	.0000	.0047	.0000	.0047	-	.0047	1.0000	.3119
	σ^2	1.0000	.9467	.0000	-0.0533	.0533	.0533	1.0000	.2642
	ν	10.0000	10.7789	.0385	.7789	.0779	.7799	1.0000	13.6974

Table 30 – Results of the simulation study for the skew-t model with $\nu = 10$ and $\gamma = 0.9$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
	β_0	1.0000	.9259	.0000	-0.0741	.0741	.0741	1.0000	.6528
	β_1	2.0000	1.8678	.0000	-0.1322	.0661	.1322	1.0000	.4510
50	γ	.9000	.7637	.0002	-0.1363	.1515	.1363	1.0000	1.0649
	σ^2	1.0000	.9772	.0001	-0.0228	.0228	.0228	1.0000	1.1601
	ν	10.0000	8.0417	.1557	-1.9583	.1958	1.9644	1.0000	22.3129
	β_0	1.0000	.9586	.0000	-0.0414	.0414	.0414	1.0000	.2702
	β_1	2.0000	1.9494	.0000	-0.0506	.0253	.0506	1.0000	.1915
250	γ	.9000	.9225	.0000	.0225	.0250	.0225	1.0000	.2258
	σ^2	1.0000	1.0524	.0000	.0524	.0524	.0524	1.0000	.4979
	ν	10.0000	16.5037	.6028	6.5037	.6504	6.5316	1.0000	41.6902
	β_0	1.0000	.9834	.0000	-0.0166	.0166	.0166	1.0000	.2032
	β_1	2.0000	2.0249	.0000	.0249	.0124	.0249	1.0000	.1576
500	γ	.9000	.9465	.0000	.0465	.0517	.0465	1.0000	.1467
	σ^2	1.0000	1.0027	.0000	.0027	.0027	.0027	1.0000	.4016
	ν	10.0000	8.0917	.0117	-1.9083	.1908	1.9084	1.0000	8.5673
	β_0	1.0000	1.0405	.0000	.0405	.0405	.0405	1.0000	.1436
	β_1	2.0000	1.9981	.0000	-0.0019	.0010	.0019	1.0000	.1070
1000	γ	.9000	.9531	.0000	.0531	.0590	.0531	1.0000	.1188
	σ^2	1.0000	1.0597	.0000	.0597	.0597	.0597	1.0000	.2781
	ν	10.0000	8.8144	.0099	-1.1856	.1186	1.1856	1.0000	6.5708

Table 31 – Results from the simulation using the skew-t distribution with $\nu = 30$ and $\gamma = -0.9$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	1.1529	< .0001	.1529	.1529	.1529	1.0000	.5598
	β_1	2.0000	1.9910	< .0001	-0.0090	.0045	.0090	1.0000	.4742
	γ	-0.9000	-0.9040	< .0001	-0.0040	.0045	.0040	1.0000	.4027
	σ^2	1.0000	.8132	.0001	-0.1868	.1868	.1868	1.0000	.8820
	ν	30.0000	18.5535	1.2958	-11.4465	.3815	11.5196	1.0000	72.9905
250	β_0	1.0000	1.1014	< .0001	.1014	.1014	.1014	1.0000	.2515
	β_1	2.0000	1.9180	< .0001	-0.0820	.0410	.0820	1.0000	.1813
	γ	-0.9000	-0.9296	< .0001	-0.0296	.0329	.0296	1.0000	.1982
	σ^2	1.0000	.8723	< .0001	-0.1277	.1277	.1277	1.0000	.4380
	ν	30.0000	16.7695	.8023	-13.2305	.4410	13.2548	1.0000	45.2547
500	β_0	1.0000	1.0478	< .0001	.0478	.0478	.0478	1.0000	.1749
	β_1	2.0000	2.0074	< .0001	.0074	.0037	.0074	1.0000	.1320
	γ	-0.9000	-0.9457	< .0001	-0.0457	.0508	.0457	1.0000	.1336
	σ^2	1.0000	.9220	< .0001	-0.0780	.0780	.0780	1.0000	.3025
	ν	30.0000	25.6242	2.0447	-4.3758	.1459	4.8300	1.0000	61.5862
1000	β_0	1.0000	1.0298	< .0001	.0298	.0298	.0298	1.0000	.1279
	β_1	2.0000	2.0036	< .0001	.0036	.0018	.0036	1.0000	.0954
	γ	-0.9000	-0.9149	< .0001	-0.0149	.0166	.0149	1.0000	.1498
	σ^2	1.0000	1.0089	< .0001	.0089	.0089	.0089	1.0000	.2355
	ν	30.0000	33.7322	6.5150	3.7322	.1244	7.5082	1.0000	88.1083

Table 32 – Results of the simulation study for the skew-t model with $\nu = 30$ and $\gamma = 0$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	1.0260	.0000	.0260	.0260	.0260	1.0000	.6585
	β_1	2.0000	1.8874	.0000	-0.1126	.0563	.1126	1.0000	.5486
	γ	.0000	-0.1762	.0016	-0.1762	-	.1762	1.0000	1.7687
	σ^2	1.0000	.9783	.0001	-0.0217	.0217	.0217	1.0000	1.1245
	ν	30.0000	10.5808	.4118	-19.4192	.6473	19.4235	1.0000	43.0512
250	β_0	1.0000	1.0229	.0000	.0229	.0229	.0229	1.0000	.2459
	β_1	2.0000	1.9831	.0000	-0.0169	.0084	.0169	1.0000	.2425
	γ	.0000	.0101	.0001	.0101	-	.0101	1.0000	.6273
	σ^2	1.0000	.8583	.0000	-0.1417	.1417	.1417	1.0000	.3947
	ν	30.0000	24.6001	3.0313	-5.3999	.1800	6.1926	1.0000	82.3760
500	β_0	1.0000	1.0058	.0000	.0058	.0058	.0058	1.0000	.1709
	β_1	2.0000	2.0134	.0000	.0134	.0067	.0134	1.0000	.1748
	γ	.0000	.0170	.0000	.0170	-	.0170	1.0000	.3795
	σ^2	1.0000	.8905	.0000	-0.1095	.1095	.1095	1.0000	.2671
	ν	30.0000	39.1675	6.9819	9.1675	.3056	11.5235	1.0000	112.8711
1000	β_0	1.0000	1.0028	.0000	.0028	.0028	.0028	1.0000	.1252
	β_1	2.0000	2.0476	.0000	.0476	.0238	.0476	1.0000	.1263
	γ	.0000	.0423	.0000	.0423	-	.0423	1.0000	.2903
	σ^2	1.0000	.9371	.0000	-0.0629	.0629	.0629	1.0000	.2292
	ν	30.0000	32.0601	5.8104	2.0601	.0687	6.1648	1.0000	86.8478

Table 33 – Results of the simulation study for the skew-t model with $\nu = 30$ and $\gamma = 0.9$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	.8035	.0000	-0.1965	.1965	.1965	1.0000	.5305
	β_1	2.0000	1.9553	.0000	-0.0447	.0224	.0447	1.0000	.6065
	γ	.9000	.7692	.0001	-0.1308	.1454	.1308	1.0000	.9964
	σ^2	1.0000	.7515	.0000	-0.2485	.2485	.2485	1.0000	.7929
	ν	30.0000	17.2585	.4922	-12.7415	.4247	12.7510	1.0000	68.2674
250	β_0	1.0000	.9967	.0000	-0.0033	.0033	.0033	1.0000	.2733
	β_1	2.0000	2.0525	.0000	.0525	.0262	.0525	1.0000	.2086
	γ	.9000	.9558	.0000	.0558	.0620	.0558	1.0000	.1237
	σ^2	1.0000	1.0819	.0001	.0819	.0819	.0819	1.0000	.5175
	ν	30.0000	22.6925	.6410	-7.3075	.2436	7.3356	1.0000	58.6967
500	β_0	1.0000	.9754	.0000	-0.0246	.0246	.0246	1.0000	.1719
	β_1	2.0000	2.0793	.0000	.0793	.0397	.0793	.0000	.1430
	γ	.9000	.8428	.0000	-0.0572	.0636	.0572	1.0000	.2812
	σ^2	1.0000	.8935	.0000	-0.1065	.1065	.1065	1.0000	.2895
	ν	30.0000	37.7354	6.9979	7.7354	.2578	10.4310	1.0000	109.3543
1000	β_0	1.0000	1.0465	.0000	.0465	.0465	.0465	1.0000	.1269
	β_1	2.0000	2.0487	.0000	.0487	.0243	.0487	.9000	.1031
	γ	.9000	.8514	.0000	-0.0486	.0540	.0486	1.0000	.2054
	σ^2	1.0000	.9754	.0000	-0.0246	.0246	.0246	1.0000	.2344
	ν	30.0000	33.6360	9.4285	3.6360	.1212	10.1053	1.0000	88.7431

B.2 Skew slash

Table 34 – Results of the simulation study for the skew slash model with $\nu = 3$ and $\gamma = -0.9$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	1.0439	.0001	.0439	.0439	.0439	1.0000	.8320
	β_1	2.0000	1.8084	.0001	-0.1916	.0958	.1916	1.0000	.7521
	γ	-0.9000	-0.8254	.0002	.0746	.0829	.0746	1.0000	.8898
	σ^2	1.0000	1.3463	.0033	.3463	.3463	.3463	1.0000	2.0684
	ν	3.0000	9.2591	5.9205	6.2591	2.0864	8.6156	1.0000	34.7947
250	β_0	1.0000	1.1148	$< .0001$	.1148	.1148	.1148	1.0000	.3009
	β_1	2.0000	1.9995	$< .0001$	-0.0005	.0002	.0005	1.0000	.2437
	γ	-0.9000	-0.7799	.0002	.1201	.1334	.1201	1.0000	.5042
	σ^2	1.0000	.8453	.0005	-0.1547	.1547	.1547	1.0000	.6864
	ν	3.0000	2.7430	.6948	-0.2570	.0857	.7408	1.0000	8.9155
500	β_0	1.0000	1.0723	$< .0001$	.0723	.0723	.0723	1.0000	.2168
	β_1	2.0000	2.0028	$< .0001$	.0028	.0014	.0028	1.0000	.1403
	γ	-0.9000	-0.9523	$< .0001$	-0.0523	.0581	.0523	1.0000	.1104
	σ^2	1.0000	.9421	$< .0001$	-0.0579	.0579	.0579	1.0000	.4211
	ν	3.0000	2.5603	.0011	-0.4397	.1466	.4397	1.0000	1.8669
1000	β_0	1.0000	1.0405	$< .0001$	.0405	.0405	.0405	1.0000	.1493
	β_1	2.0000	2.0367	$< .0001$	.0367	.0184	.0367	1.0000	.1156
	γ	-0.9000	-0.9120	$< .0001$	-0.0120	.0133	.0120	1.0000	.1437
	σ^2	1.0000	1.0203	$< .0001$	.0203	.0203	.0203	1.0000	.3368
	ν	3.0000	3.1046	.0025	.1046	.0349	.1046	1.0000	2.3084

Table 35 – Results of the simulation study for the skew slash model with $\nu = 3$ and $\gamma = 0$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
	β_0	1.0000	1.1448	.0000	.1448	.1448	.1448	1.0000	.7575
	β_1	2.0000	1.8906	.0000	-0.1094	.0547	.1094	1.0000	.6604
50	γ	.0000	.1645	.0007	.1645	-	.1645	1.0000	1.5081
	σ^2	1.0000	1.0611	.0010	.0611	.0611	.0611	1.0000	1.8622
	ν	3.0000	4.9887	.3055	1.9887	.6629	2.0120	1.0000	24.8208
	β_0	1.0000	1.0014	.0000	.0014	.0014	.0014	1.0000	.3287
	β_1	2.0000	1.9107	.0000	-0.0893	.0446	.0893	1.0000	.2842
250	γ	.0000	.2826	.0001	.2826	-	.2826	1.0000	.8198
	σ^2	1.0000	.8388	.0002	-0.1612	.1612	.1612	1.0000	.8051
	ν	3.0000	2.0149	.0213	-0.9851	.3284	.9853	1.0000	3.5538
	β_0	1.0000	1.0590	.0000	.0590	.0590	.0590	1.0000	.2264
	β_1	2.0000	1.9911	.0000	-0.0089	.0044	.0089	1.0000	.2277
500	γ	.0000	.1618	.0001	.1618	-	.1618	1.0000	.5037
	σ^2	1.0000	1.0624	.0003	.0624	.0624	.0624	1.0000	.6736
	ν	3.0000	3.1657	.2894	.1657	.0552	.3335	1.0000	7.4000
	β_0	1.0000	1.0095	.0000	.0095	.0095	.0095	1.0000	.1521
	β_1	2.0000	2.0961	.0000	.0961	.0481	.0961	.0000	.1515
1000	γ	.0000	-0.0523	.0000	-0.0523	-	.0523	1.0000	.3238
	σ^2	1.0000	.9741	.0000	-0.0259	.0259	.0259	1.0000	.3803
	ν	3.0000	2.7235	.0012	-0.2765	.0922	.2765	1.0000	2.1902

Table 36 – Results of the simulation study for the skew slash model with $\nu = 3$ and $\gamma = 0.9$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
	β_0	1.0000	.9277	.0001	-0.0723	.0723	.0723	1.0000	.8280
	β_1	2.0000	2.0441	.0000	.0441	.0220	.0441	1.0000	.6772
50	γ	.9000	.9088	.0000	.0088	.0098	.0088	1.0000	.3381
	σ^2	1.0000	1.5583	.0010	.5583	.5583	.5583	1.0000	2.0341
	ν	3.0000	8.9112	.2351	5.9112	1.9704	5.9159	1.0000	32.2555
	β_0	1.0000	.8935	.0000	-0.1065	.1065	.1065	1.0000	.2949
	β_1	2.0000	2.0203	.0000	.0203	.0102	.0203	1.0000	.2573
250	γ	.9000	.8717	.0000	-0.0283	.0315	.0283	1.0000	.3570
	σ^2	1.0000	.8480	.0001	-0.1520	.1520	.1520	1.0000	.5837
	ν	3.0000	2.5599	.0075	-0.4401	.1467	.4401	1.0000	3.1564
	β_0	1.0000	.9419	.0000	-0.0581	.0581	.0581	1.0000	.2252
	β_1	2.0000	1.9697	.0000	-0.0303	.0152	.0303	1.0000	.1729
500	γ	.9000	.9214	.0000	.0214	.0238	.0214	1.0000	.1727
	σ^2	1.0000	1.0295	.0001	.0295	.0295	.0295	1.0000	.4801
	ν	3.0000	2.7052	.0019	-0.2948	.0983	.2948	1.0000	2.2723
	β_0	1.0000	1.0035	.0000	.0035	.0035	.0035	1.0000	.1493
	β_1	2.0000	1.9805	.0000	-0.0195	.0098	.0195	1.0000	.1118
1000	γ	.9000	.9459	.0000	.0459	.0510	.0459	1.0000	.1158
	σ^2	1.0000	.9762	.0000	-0.0238	.0238	.0238	1.0000	.2942
	ν	3.0000	2.7582	.0015	-0.2418	.0806	.2418	1.0000	1.5129

Table 37 – Results of the simulation study for the skew slash model with $\nu = 10$ and $\gamma = -0.9$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
	β_0	1.0000	.9123	.0001	-0.0877	.0877	.0877	1.0000	.6881
	β_1	2.0000	2.0093	.0001	.0093	.0047	.0093	1.0000	.5946
50	γ	-0.9000	-0.8936	$<$ .0001	.0064	.0071	.0064	1.0000	.4013
	σ^2	1.0000	1.3414	.0062	.3414	.3414	.3415	1.0000	1.4490
	ν	10.0000	21.0701	92.3600	11.0701	1.1070	93.0210	1.0000	64.0377
	β_0	1.0000	1.0590	$<$ .0001	.0590	.0590	.0590	1.0000	.2560
	β_1	2.0000	2.0443	$<$ .0001	.0443	.0221	.0443	1.0000	.1912
250	γ	-0.9000	-0.8975	$<$ .0001	.0025	.0027	.0025	1.0000	.2169
	σ^2	1.0000	.9249	.0003	-0.0751	.0751	.0751	1.0000	.5545
	ν	10.0000	17.2884	8.8470	7.2884	.7288	11.4625	1.0000	71.6213
	β_0	1.0000	.9387	$<$ .0001	-0.0613	.0613	.0613	1.0000	.1932
	β_1	2.0000	2.1000	$<$ .0001	.1000	.0500	.1000	.0000	.1466
500	γ	-0.9000	-0.9117	$<$ .0001	-0.0117	.0130	.0117	1.0000	.1529
	σ^2	1.0000	1.1416	.0001	.1416	.1416	.1416	1.0000	.4077
	ν	10.0000	14.1689	1.4274	4.1689	.4169	4.4065	1.0000	36.7813
	β_0	1.0000	1.0361	$<$ .0001	.0361	.0361	.0361	1.0000	.1272
	β_1	2.0000	1.9814	$<$ .0001	-0.0186	.0093	.0186	1.0000	.0925
1000	γ	-0.9000	-0.9286	$<$ .0001	-0.0286	.0318	.0286	1.0000	.0839
	σ^2	1.0000	1.0012	.0001	.0012	.0012	.0012	1.0000	.2694
	ν	10.0000	15.5633	1.8383	5.5633	.5563	5.8592	1.0000	37.9126

Table 38 – Results of the simulation study for the skew slash model with $\nu = 10$ and $\gamma = 0$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
	β_0	1.0000	1.1015	.0000	.1015	.1015	.1015	1.0000	.6024
	β_1	2.0000	1.8659	.0000	-0.1341	.0671	.1341	1.0000	.5597
50	γ	.0000	-0.5325	.0003	-0.5325	-	.5325	1.0000	1.0413
	σ^2	1.0000	1.0672	.0015	.0672	.0672	.0673	1.0000	1.0497
	ν	10.0000	26.3217	60.0651	16.3217	1.6322	62.2432	1.0000	78.0328
	β_0	1.0000	.9604	.0000	-0.0396	.0396	.0396	1.0000	.2679
	β_1	2.0000	2.0538	.0000	.0538	.0269	.0538	1.0000	.2547
250	γ	.0000	.0222	.0000	.0222	-	.0222	1.0000	.5565
	σ^2	1.0000	.9726	.0007	-0.0274	.0274	.0274	1.0000	.6773
	ν	10.0000	10.4813	10.8731	.4813	.0481	10.8837	1.0000	35.4027
	β_0	1.0000	.9689	.0000	-0.0311	.0311	.0311	1.0000	.1781
	β_1	2.0000	1.9971	.0000	-0.0029	.0014	.0029	1.0000	.1874
500	γ	.0000	.0366	.0000	.0366	-	.0366	1.0000	.3700
	σ^2	1.0000	.9037	.0008	-0.0963	.0963	.0963	1.0000	.4403
	ν	10.0000	13.0671	26.5254	3.0671	.3067	26.7021	1.0000	38.7059
	β_0	1.0000	.9356	.0000	-0.0644	.0644	.0644	.4400	.1279
	β_1	2.0000	1.9714	.0000	-0.0286	.0143	.0286	1.0000	.1277
1000	γ	.0000	-0.0887	.0000	-0.0887	-	.0887	1.0000	.3187
	σ^2	1.0000	.8709	.0003	-0.1291	.1291	.1291	1.0000	.3830
	ν	10.0000	8.6755	4.7479	-1.3245	.1325	4.9291	1.0000	27.7433

Table 39 – Results of the simulation study for the skew slash model with $\nu = 30$ and $\gamma = -0.9$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
	β_0	1.0000	.9245	.0001	-0.0755	.0755	.0755	1.0000	.6557
	β_1	2.0000	2.0049	< .0001	.0049	.0024	.0049	1.0000	.5195
50	γ	-0.9000	-0.8997	< .0001	.0003	.0003	.0003	1.0000	.3616
	σ^2	1.0000	1.2703	.0025	.2703	.2703	.2703	1.0000	1.2697
	ν	30.0000	28.2564	45.2256	-1.7436	.0581	45.2592	1.0000	81.5774
	β_0	1.0000	1.0634	< .0001	.0634	.0634	.0634	1.0000	.2475
	β_1	2.0000	2.0114	< .0001	.0114	.0057	.0114	1.0000	.1549
250	γ	-0.9000	-0.9016	< .0001	-0.0016	.0017	.0016	1.0000	.2064
	σ^2	1.0000	.8984	.0002	-0.1016	.1016	.1016	1.0000	.4995
	ν	30.0000	20.6677	10.3197	-9.3323	.3111	13.9136	1.0000	79.8411
	β_0	1.0000	.9353	< .0001	-0.0647	.0647	.0647	1.0000	.1808
	β_1	2.0000	1.9867	< .0001	-0.0133	.0067	.0133	1.0000	.1461
500	γ	-0.9000	-0.8235	< .0001	.0765	.0850	.0765	1.0000	.2070
	σ^2	1.0000	1.0509	.0001	.0509	.0509	.0509	1.0000	.3309
	ν	30.0000	34.5914	17.9223	4.5914	.1530	18.5011	1.0000	88.6817
	β_0	1.0000	1.0415	< .0001	.0415	.0415	.0415	1.0000	.1226
	β_1	2.0000	2.0036	< .0001	.0036	.0018	.0036	1.0000	.0815
1000	γ	-0.9000	-0.9152	< .0001	-0.0152	.0169	.0152	1.0000	.0876
	σ^2	1.0000	.9616	.0002	-0.0384	.0384	.0384	1.0000	.2564
	ν	30.0000	29.2336	44.2209	-0.7664	.0255	44.2275	1.0000	80.7713

Table 40 – Results of the simulation study for the skew slash model with $\nu = 30$ and $\gamma = 0$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
	β_0	1.0000	1.0321	.0000	.0321	.0321	.0321	1.0000	.5771
	β_1	2.0000	1.8565	.0000	-0.1435	.0718	.1435	1.0000	.5817
50	γ	.0000	-0.4025	.0014	-0.4025	-	.4025	1.0000	1.1384
	σ^2	1.0000	.9263	.0027	-0.0737	.0737	.0738	1.0000	.9845
	ν	30.0000	18.8326	89.9383	-11.1674	.3722	90.6290	1.0000	58.8714
	β_0	1.0000	.9318	.0000	-0.0682	.0682	.0682	1.0000	.2619
	β_1	2.0000	1.9970	.0000	-0.0030	.0015	.0030	1.0000	.2489
250	γ	.0000	.0376	.0000	.0376	-	.0376	1.0000	.5281
	σ^2	1.0000	.9582	.0017	-0.0418	.0418	.0418	1.0000	.6035
	ν	30.0000	14.4251	51.0512	-15.5749	.5192	53.3742	.9800	45.2735
	β_0	1.0000	.9679	.0000	-0.0321	.0321	.0321	1.0000	.1679
	β_1	2.0000	1.9610	.0000	-0.0390	.0195	.0390	1.0000	.1662
500	γ	.0000	-0.0091	.0000	-0.0091	-	.0091	1.0000	.3444
	σ^2	1.0000	.8663	.0002	-0.1337	.1337	.1337	.6200	.2795
	ν	30.0000	30.4408	34.8418	.4408	.0147	34.8446	1.0000	83.2606
	β_0	1.0000	1.0042	.0000	.0042	.0042	.0042	1.0000	.1204
	β_1	2.0000	2.0220	.0000	.0220	.0110	.0220	1.0000	.1206
1000	γ	.0000	-0.0419	.0000	-0.0419	-	.0419	1.0000	.2583
	σ^2	1.0000	.9121	.0003	-0.0879	.0879	.0879	.9600	.2276
	ν	30.0000	29.6673	55.8715	-0.3327	.0111	55.8725	1.0000	77.0663

B.3 Skew contaminated normal

Table 41 – Results of the simulation study for the skew contaminated normal model with $\nu_1 = 0.1$ and $\nu_2 = 0.1$ and $\gamma = -0.9$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	.6754	.0001	-0.3246	.3246	.3246	1.0000	.8966
	β_1	2.0000	1.7904	$<$ .0001	-0.2096	.1048	.2096	1.0000	.8457
	γ	-0.9000	-0.9273	.0001	-0.0273	.0303	.0273	1.0000	1.0978
	σ^2	1.0000	1.5814	.0002	.5814	.5814	.5814	1.0000	2.2181
	ν_1	0.1000	.0835	.0001	-0.0165	.1655	.0165	1.0000	.8557
	ν_2	0.1000	.1558	$<$ .0001	.0558	.5584	.0558	1.0000	.8468
250	β_0	1.0000	1.2977	.0008	.2977	.2977	.2977	1.0000	.5265
	β_1	2.0000	2.0616	$<$ .0001	.0616	.0308	.0616	1.0000	.2253
	γ	-0.9000	-0.3045	.0089	.5955	.6617	.5956	.5400	1.6285
	σ^2	1.0000	.2740	.0027	-0.7260	.7260	.7260	.9200	.8807
	ν_1	0.1000	.4150	.0012	.3150	3.1497	.3150	.1200	.5786
	ν_2	0.1000	.0925	$<$ .0001	-0.0075	.0750	.0075	.1200	.1480
500	β_0	1.0000	.9943	$<$ .0001	-0.0057	.0057	.0057	1.0000	.2243
	β_1	2.0000	1.9407	$<$ .0001	-0.0593	.0297	.0593	1.0000	.1543
	γ	-0.9000	-0.8794	$<$ .0001	.0206	.0229	.0206	1.0000	.2193
	σ^2	1.0000	1.0612	.0001	.0612	.0612	.0612	1.0000	.4275
	ν_1	0.1000	.1470	$<$ .0001	.0470	.4702	.0470	.4000	.1136
	ν_2	0.1000	.0999	$<$ .0001	-0.0001	.0006	.0001	.4000	.0613
1000	β_0	1.0000	1.1147	$<$ .0001	.1147	.1147	.1147	.0000	.1294
	β_1	2.0000	2.0145	$<$ .0001	.0145	.0072	.0145	1.0000	.0917
	γ	-0.9000	-0.8809	$<$ .0001	.0191	.0212	.0191	1.0000	.1371
	σ^2	1.0000	.8180	.0001	-0.1820	.1820	.1820	.0000	.2017
	ν_1	0.1000	.1403	$<$ .0001	.0403	.4030	.0403	.0000	.0714
	ν_2	0.1000	.0908	$<$ .0001	-0.0092	.0919	.0092	1.0000	.0402

Table 42 – Results of the simulation study for the skew contaminated normal model with $\nu_1 = 0.1$ and $\nu_2 = 0.1$ and $\gamma = 0$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	1.0037	.0000	.0037	.0037	.0037	1.0000	.6766
	β_1	2.0000	2.0048	.0001	.0048	.0024	.0048	1.0000	.7115
	γ	.0000	.0095	.0006	.0095	-	.0096	1.0000	1.4488
	σ^2	1.0000	.8866	.0001	-0.1134	.1134	.1134	1.0000	1.2433
	ν_1	0.1000	.0662	.0000	-0.0338	.3383	.0338	1.0000	.3516
	ν_2	0.1000	.0269	.0000	-0.0731	.7305	.0731	1.0000	.2002
250	β_0	1.0000	1.1387	.0000	.1387	.1387	.1387	.9800	.3193
	β_1	2.0000	1.9476	.0000	-0.0524	.0262	.0524	1.0000	.2771
	γ	.0000	.3497	.0001	.3497	-	.3497	1.0000	.7335
	σ^2	1.0000	.9644	.0000	-0.0356	.0356	.0356	1.0000	.8105
	ν_1	0.1000	.1365	.0000	.0365	.3649	.0365	1.0000	.3643
	ν_2	0.1000	.1308	.0000	.0308	.3077	.0308	1.0000	.1814
500	β_0	1.0000	1.0257	.0000	.0257	.0257	.0257	1.0000	.2054
	β_1	2.0000	1.9513	.0000	-0.0487	.0243	.0487	1.0000	.1983
	γ	.0000	-0.0018	.0000	-0.0018	-	.0018	1.0000	.4847
	σ^2	1.0000	1.0829	.0000	.0829	.0829	.0829	1.0000	.3743
	ν_1	0.1000	.0796	.0000	-0.0204	.2044	.0204	1.0000	.1069
	ν_2	0.1000	.0829	.0000	-0.0171	.1707	.0171	1.0000	.0972
1000	β_0	1.0000	1.0231	.0000	.0231	.0231	.0231	1.0000	.1391
	β_1	2.0000	2.0520	.0000	.0520	.0260	.0520	1.0000	.1377
	γ	.0000	-0.0009	.0000	-0.0009	-	.0009	1.0000	.3483
	σ^2	1.0000	.9734	.0000	-0.0266	.0266	.0266	1.0000	.2599
	ν_1	0.1000	.0911	.0000	-0.0089	.0890	.0089	1.0000	.0942
	ν_2	0.1000	.0969	.0000	-0.0031	.0310	.0031	1.0000	.0777

Table 43 – Results of the simulation study for the skew contaminated normal model with $\nu_1 = 0.1$ and $\nu_2 = 0.1$ and $\gamma = 0.9$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	1.3744	.0001	.3744	.3744	.3744	1.0000	.9890
	β_1	2.0000	2.1473	.0001	.1473	.0736	.1473	1.0000	.8513
	γ	.9000	.9356	.0002	.0356	.0395	.0356	1.0000	1.0150
	σ^2	1.0000	1.2991	.0006	.2991	.2991	.2991	1.0000	2.5353
	ν_1	0.1000	.1790	.0000	.0790	.7896	.0790	1.0000	.6967
	ν_2	0.1000	.1419	.0000	.0419	.4191	.0419	1.0000	.6689
250	β_0	1.0000	.9637	.0000	-0.0363	.0363	.0363	1.0000	.3101
	β_1	2.0000	1.9397	.0000	-0.0603	.0301	.0603	1.0000	.2443
	γ	.9000	.7686	.0001	-0.1314	.1460	.1314	.6600	.7772
	σ^2	1.0000	.8350	.0000	-0.1650	.1650	.1650	1.0000	.5994
	ν_1	0.1000	.1547	.0000	.0547	.5471	.0547	1.0000	.3168
	ν_2	0.1000	.1276	.0000	.0276	.2761	.0276	1.0000	.1519
500	β_0	1.0000	1.0199	.0001	.0199	.0199	.0199	1.0000	.2150
	β_1	2.0000	1.9875	.0000	-0.0125	.0062	.0125	1.0000	.1384
	γ	.9000	.8978	.0000	-0.0022	.0024	.0022	1.0000	.1614
	σ^2	1.0000	1.0439	.0006	.0439	.0439	.0439	1.0000	.3913
	ν_1	0.1000	.1151	.0000	.0151	.1507	.0151	1.0000	.1004
	ν_2	0.1000	.1030	.0000	.0030	.0299	.0030	1.0000	.0646
1000	β_0	1.0000	.9145	.0000	-0.0855	.0855	.0855	.0000	.1285
	β_1	2.0000	2.0247	.0000	.0247	.0124	.0247	1.0000	.0854
	γ	.9000	.8800	.0000	-0.0200	.0222	.0200	1.0000	.1134
	σ^2	1.0000	.8441	.0002	-0.1559	.1559	.1559	.0000	.2002
	ν_1	0.1000	.1105	.0000	.0105	.1047	.0105	1.0000	.0640
	ν_2	0.1000	.0888	.0000	-0.0112	.1118	.0112	1.0000	.0437

Table 44 – Results of the simulation study for the skew contaminated normal model with $\nu_1 = 0.9$ and $\nu_2 = 0.9$ and $\gamma = -0.9$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	.6669	$<$ .0001	-0.3331	.3331	.3331	.0000	.7176
	β_1	2.0000	1.7910	$<$ .0001	-0.2090	.1045	.2090	1.0000	.6975
	γ	-0.9000	-0.9742	$<$ .0001	-0.0742	.0824	.0742	1.0000	.2703
	σ^2	1.0000	1.3736	.0003	.3736	.3736	.3736	1.0000	1.7478
	ν_1	0.9000	.2428	.0002	-0.6572	.7303	.6572	1.0000	.9577
	ν_2	0.9000	.8727	$<$ .0001	-0.0273	.0304	.0273	1.0000	.7533
250	β_0	1.0000	1.0762	$<$ .0001	.0762	.0762	.0762	1.0000	.2318
	β_1	2.0000	2.0707	$<$ .0001	.0707	.0353	.0707	1.0000	.1827
	γ	-0.9000	-0.8920	$<$ .0001	.0080	.0089	.0080	1.0000	.2810
	σ^2	1.0000	.7957	$<$ .0001	-0.2043	.2043	.2043	.9000	.5465
	ν_1	0.9000	.1063	.0003	-0.7937	.8819	.7937	.9500	.9231
	ν_2	0.9000	.4681	$<$ .0001	-0.4319	.4799	.4319	.9500	.7231
500	β_0	1.0000	1.0005	$<$ .0001	.0005	.0005	.0005	1.0000	.1926
	β_1	2.0000	1.9703	$<$ .0001	-0.0297	.0148	.0297	1.0000	.1408
	γ	-0.9000	-0.9072	$<$ .0001	-0.0072	.0080	.0072	1.0000	.1578
	σ^2	1.0000	1.1531	.0003	.1531	.1531	.1531	1.0000	.8555
	ν_1	0.9000	.5389	.0006	-0.3611	.4013	.3611	1.0000	.9536
	ν_2	0.9000	.7519	.0001	-0.1481	.1646	.1481	1.0000	.6386
1000	β_0	1.0000	1.0850	$<$ .0001	.0850	.0850	.0850	.0000	.1217
	β_1	2.0000	2.0100	$<$ .0001	.0100	.0050	.0100	1.0000	.0822
	γ	-0.9000	-0.9088	$<$ .0001	-0.0088	.0097	.0088	1.0000	.1050
	σ^2	1.0000	.9474	.0001	-0.0526	.0526	.0526	1.0000	.5891
	ν_1	0.9000	.5877	.0006	-0.3123	.3470	.3123	1.0000	.9570
	ν_2	0.9000	.8478	.0001	-0.0522	.0580	.0522	1.0000	.5746

Table 45 – Results of the simulation study for the skew contaminated normal model with $\nu_1 = 0.9$ and $\nu_2 = 0.9$ and $\gamma = 0$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	1.0664	.0000	.0664	.0664	.0664	1.0000	.6519
	β_1	2.0000	2.0216	.0000	.0216	.0108	.0216	1.0000	.7067
	γ	.0000	.7486	.0003	.7486	-	.7486	1.0000	1.3275
	σ^2	1.0000	.8440	.0003	-0.1560	.1560	.1560	1.0000	1.5488
	ν_1	0.9000	.4090	.0001	-0.4910	.5455	.4910	1.0000	.9231
	ν_2	0.9000	.2164	.0001	-0.6836	.7595	.6836	1.0000	.8970
250	β_0	1.0000	1.1036	.0000	.1036	.1036	.1036	1.0000	.2535
	β_1	2.0000	1.9461	.0000	-0.0539	.0270	.0539	1.0000	.2405
	γ	.0000	.0042	.0000	.0042	-	.0042	1.0000	.5751
	σ^2	1.0000	.8965	.0002	-0.1035	.1035	.1035	1.0000	.9426
	ν_1	0.9000	.6219	.0002	-0.2781	.3089	.2781	1.0000	.9286
	ν_2	0.9000	.4104	.0001	-0.4896	.5440	.4896	1.0000	.8107
500	β_0	1.0000	1.0259	.0000	.0259	.0259	.0259	1.0000	.1862
	β_1	2.0000	1.9692	.0000	-0.0308	.0154	.0308	1.0000	.1855
	γ	.0000	-0.0001	.0000	-0.0001	-	.0001	1.0000	.3513
	σ^2	1.0000	1.0628	.0001	.0628	.0628	.0628	1.0000	.7767
	ν_1	0.9000	.4784	.0008	-0.4216	.4685	.4216	1.0000	.9705
	ν_2	0.9000	.8906	.0001	-0.0094	.0104	.0094	1.0000	.6792
1000	β_0	1.0000	1.0206	.0000	.0206	.0206	.0206	1.0000	.1290
	β_1	2.0000	2.0590	.0000	.0590	.0295	.0590	1.0000	.1282
	γ	.0000	-0.0036	.0000	-0.0036	-	.0036	1.0000	.2912
	σ^2	1.0000	1.0101	.0008	.0101	.0101	.0101	1.0000	.8675
	ν_1	0.9000	.8402	.0021	-0.0598	.0664	.0598	1.0000	.9536
	ν_2	0.9000	.8313	.0003	-0.0687	.0763	.0687	1.0000	.7721

Table 46 – Results of the simulation study for the skew contaminated normal model with $\nu_1 = 0.9$ and $\nu_2 = 0.9$ and $\gamma = 0.9$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	1.3533	.0001	.3533	.3533	.3533	.0000	.7273
	β_1	2.0000	2.1314	.0000	.1314	.0657	.1314	1.0000	.6661
	γ	.9000	.9587	.0000	.0587	.0652	.0587	1.0000	.4924
	σ^2	1.0000	1.3316	.0004	.3316	.3316	.3316	1.0000	1.8164
	ν_1	0.9000	.1430	.0001	-0.7570	.8411	.7570	1.0000	.9543
	ν_2	0.9000	.7594	.0000	-0.1406	.1562	.1406	1.0000	.7777
250	β_0	1.0000	.9681	.0000	-0.0319	.0319	.0319	1.0000	.2443
	β_1	2.0000	1.9736	.0000	-0.0264	.0132	.0264	1.0000	.2094
	γ	.9000	.8438	.0000	-0.0562	.0624	.0562	1.0000	.3459
	σ^2	1.0000	.8982	.0002	-0.1018	.1018	.1018	1.0000	.8382
	ν_1	0.9000	.7960	.0004	-0.1040	.1156	.1040	1.0000	.9603
	ν_2	0.9000	.8712	.0002	-0.0288	.0320	.0288	1.0000	.7405
500	β_0	1.0000	1.0192	.0000	.0192	.0192	.0192	1.0000	.1855
	β_1	2.0000	2.0123	.0000	.0123	.0061	.0123	1.0000	.1301
	γ	.9000	.9277	.0000	.0277	.0308	.0277	1.0000	.1320
	σ^2	1.0000	1.0692	.0004	.0692	.0692	.0692	1.0000	.7915
	ν_1	0.9000	.7296	.0008	-0.1704	.1893	.1704	1.0000	.9547
	ν_2	0.9000	.7533	.0002	-0.1467	.1630	.1467	1.0000	.6264
1000	β_0	1.0000	.9745	.0000	-0.0255	.0255	.0255	1.0000	.1270
	β_1	2.0000	2.0494	.0000	.0494	.0247	.0494	.0200	.0879
	γ	.9000	.9010	.0000	.0010	.0011	.0010	1.0000	.1190
	σ^2	1.0000	1.0295	.0009	.0295	.0295	.0295	1.0000	.7288
	ν_1	0.9000	.8098	.0020	-0.0902	.1002	.0902	1.0000	.9585
	ν_2	0.9000	.9086	.0005	.0086	.0096	.0086	1.0000	.6458

Table 47 – Results of the simulation study for the skew contaminated normal model with $\nu_1 = 0.9$ and $\nu_2 = 0.1$ and $\gamma = 0$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	1.1867	.0007	.1867	.1867	.1867	1.0000	1.8168
	β_1	2.0000	2.0642	.0007	.0642	.0321	.0642	1.0000	1.9578
	γ	.0000	.7977	.0015	.7977	-	.7977	1.0000	1.6262
	σ^2	1.0000	3.0416	.0617	2.0416	2.0416	2.0425	1.0000	11.6915
	ν_1	0.9000	.4659	.0002	-0.4341	.4823	.4341	.9400	.8696
	ν_2	0.1000	.1194	.0003	.0194	.1938	.0194	.9400	.8894
250	β_0	1.0000	1.3618	.0001	.3618	.3618	.3618	.4800	.7448
	β_1	2.0000	1.8299	.0000	-0.1701	.0850	.1701	1.0000	.6965
	γ	.0000	.0069	.0000	.0069	-	.0069	1.0000	.6287
	σ^2	1.0000	5.3739	.0249	4.3739	4.3739	4.3740	.3000	8.3560
	ν_1	0.9000	.6502	.0003	-0.2498	.2776	.2498	1.0000	.8800
	ν_2	0.1000	.3005	.0001	.2005	2.0051	.2005	1.0000	.8362
500	β_0	1.0000	1.0863	.0000	.0863	.0863	.0863	1.0000	.5312
	β_1	2.0000	1.9209	.0000	-0.0791	.0395	.0791	1.0000	.5303
	γ	.0000	.0002	.0000	.0002	-	.0002	1.0000	.3552
	σ^2	1.0000	8.6196	.0627	7.6196	7.6196	7.6198	.1800	8.1966
	ν_1	0.9000	.7376	.0010	-0.1624	.1804	.1624	1.0000	.9600
	ν_2	0.1000	.8244	.0004	.7244	7.2440	.7244	1.0000	.8199
1000	β_0	1.0000	1.0394	.0000	.0394	.0394	.0394	1.0000	.3592
	β_1	2.0000	2.1576	.0000	.1576	.0788	.1576	1.0000	.3588
	γ	.0000	-0.0094	.0000	-0.0094	-	.0094	1.0000	.3481
	σ^2	1.0000	1.7595	.0251	.7595	.7595	.7599	1.0000	6.1449
	ν_1	0.9000	.8188	.0002	-0.0812	.0902	.0812	1.0000	.5172
	ν_2	0.1000	.1729	.0002	.0729	.7290	.0729	1.0000	.5540
2000	β_0	1.0000	.9685	.0000	-0.0315	.0315	.0315	1.0000	.2429
	β_1	2.0000	1.9441	.0000	-0.0559	.0279	.0559	1.0000	.2530
	γ	.0000	-0.0016	.0000	-0.0016	-	.0016	1.0000	.1998
	σ^2	1.0000	.7270	.0039	-0.2730	.2730	.2731	1.0000	2.1762
	ν_1	0.9000	.8909	.0000	-0.0091	.0101	.0091	1.0000	.1474
	ν_2	0.1000	.0753	.0000	-0.0247	.2465	.0247	1.0000	.2143

Table 48 – Results of the simulation study for the skew contaminated normal model with $\nu_1 = 0.9$ and $\nu_2 = 0.1$ and $\gamma = 0.9$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	2.1699	.0008	1.1699	1.1699	1.1699	.0000	2.1216
	β_1	2.0000	2.2053	.0006	.2053	.1027	.2053	1.0000	2.0829
	γ	.9000	.9685	.0001	.0685	.0761	.0685	1.0000	.5598
	σ^2	1.0000	10.9945	.0618	9.9945	9.9945	9.9947	.0200	16.2746
	ν_1	0.9000	.2879	.0002	-0.6121	.6801	.6121	1.0000	.9439
	ν_2	0.1000	.5436	.0002	.4436	4.4362	.4436	1.0000	.8098
250	β_0	1.0000	.9491	.0004	-0.0509	.0509	.0509	1.0000	.7275
	β_1	2.0000	1.9157	.0003	-0.0843	.0422	.0843	1.0000	.6204
	γ	.9000	.8877	.0004	-0.0123	.0137	.0123	1.0000	.4164
	σ^2	1.0000	3.4500	.4036	2.4500	2.4500	2.4831	1.0000	8.6779
	ν_1	0.9000	.8098	.0015	-0.0902	.1002	.0902	1.0000	.9266
	ν_2	0.1000	.3182	.0042	.2182	2.1816	.2182	1.0000	.9273
500	β_0	1.0000	1.0487	.0000	.0487	.0487	.0487	1.0000	.5453
	β_1	2.0000	2.0389	.0000	.0389	.0195	.0389	1.0000	.4162
	γ	.9000	.9308	.0000	.0308	.0343	.0308	1.0000	.1819
	σ^2	1.0000	7.8386	.0946	6.8386	6.8386	6.8393	.0400	8.0300
	ν_1	0.9000	.7193	.0008	-0.1807	.2008	.1807	1.0000	.9244
	ν_2	0.1000	.4507	.0008	.3507	3.5072	.3507	1.0000	.7665
1000	β_0	1.0000	.9724	.0005	-0.0276	.0276	.0276	1.0000	.3712
	β_1	2.0000	2.0397	.0001	.0397	.0198	.0397	1.0000	.2800
	γ	.9000	.9217	.0001	.0217	.0241	.0217	1.0000	.1023
	σ^2	1.0000	2.3584	4.6594	1.3584	1.3584	4.8533	.9000	6.5336
	ν_1	0.9000	.8796	.0144	-0.0204	.0227	.0250	1.0000	.6377
	ν_2	0.1000	.1919	.0448	.0919	.9186	.1022	1.0000	.6783
2000	β_0	1.0000	.8996	.0001	-0.1004	.1004	.1004	.9800	.2484
	β_1	2.0000	1.9523	.0000	-0.0477	.0239	.0477	1.0000	.1804
	γ	.9000	.9077	.0000	.0077	.0085	.0077	1.0000	.0739
	σ^2	1.0000	1.0710	.0551	.0710	.0710	.0899	.9600	1.7487
	ν_1	0.9000	.8767	.0001	-0.0233	.0259	.0233	1.0000	.1317
	ν_2	0.1000	.1105	.0006	.0105	.1049	.0105	1.0000	.1788

Table 49 – Results of the simulation study for the skew contaminated normal model with $\nu_1 = 0.5$ and $\nu_2 = 0.5$ and $\gamma = -0.9$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	.6574	< .0001	-0.3426	.3426	.3426	.3500	.7810
	β_1	2.0000	1.8361	< .0001	-0.1639	.0820	.1639	1.0000	.8651
	γ	-0.9000	-0.9689	< .0001	-0.0689	.0766	.0689	1.0000	.3630
	σ^2	1.0000	1.6157	.0003	.6157	.6157	.6157	1.0000	2.1047
	ν_1	0.5000	.2320	.0001	-0.2680	.5360	.2680	1.0000	.9556
	ν_2	0.5000	.8333	< .0001	.3333	.6667	.3333	1.0000	.7531
250	β_0	1.0000	1.0884	< .0001	.0884	.0884	.0884	1.0000	.2745
	β_1	2.0000	2.0791	< .0001	.0791	.0396	.0791	1.0000	.2423
	γ	-0.9000	-0.7803	< .0001	.1197	.1330	.1197	1.0000	.3886
	σ^2	1.0000	1.0130	.0003	.0130	.0130	.0130	1.0000	.9753
	ν_1	0.5000	.4365	.0003	-0.0635	.1270	.0635	1.0000	.9163
	ν_2	0.5000	.4158	.0001	-0.0842	.1685	.0842	1.0000	.7268
500	β_0	1.0000	.9854	< .0001	-0.0146	.0146	.0146	1.0000	.2259
	β_1	2.0000	1.9484	< .0001	-0.0516	.0258	.0516	1.0000	.1781
	γ	-0.9000	-0.8490	< .0001	.0510	.0566	.0510	1.0000	.2200
	σ^2	1.0000	1.4242	.0004	.4242	.4242	.4242	1.0000	1.1116
	ν_1	0.5000	.2531	.0003	-0.2469	.4939	.2469	1.0000	.9335
	ν_2	0.5000	.5313	.0001	.0313	.0627	.0313	1.0000	.6721
1000	β_0	1.0000	1.0935	< .0001	.0935	.0935	.0935	0.0000	.1431
	β_1	2.0000	2.0311	< .0001	.0311	.0155	.0311	1.0000	.1068
	γ	-0.9000	-0.8415	< .0001	.0585	.0650	.0585	1.0000	.1754
	σ^2	1.0000	1.2042	.0007	.2042	.2042	.2042	1.0000	.7889
	ν_1	0.5000	.4272	.0008	-0.0728	.1456	.0728	1.0000	.9339
	ν_2	0.5000	.5932	.0003	.0932	.1865	.0932	1.0000	.6001

Table 50 – Results of the simulation study for the skew contaminated normal model with $\nu_1 = 0.5$ and $\nu_2 = 0.5$ and $\gamma = 0$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	1.0666	.0000	.0666	.0666	.0666	1.0000	.7541
	β_1	2.0000	2.0383	.0000	.0383	.0192	.0383	1.0000	.8111
	γ	.0000	.0299	.0003	.0299	-	.0299	1.0000	1.4477
	σ^2	1.0000	.7446	.0003	-0.2554	.2554	.2554	1.0000	1.9662
	ν_1	0.5000	.2714	.0001	-0.2286	.4572	.2286	1.0000	.8584
	ν_2	0.5000	.1614	.0001	-0.3386	.6772	.3386	1.0000	.8756
250	β_0	1.0000	1.1428	.0000	.1428	.1428	.1428	.7600	.3000
	β_1	2.0000	1.9426	.0000	-0.0574	.0287	.0574	1.0000	.2712
	γ	.0000	.3346	.0001	.3346	-	.3346	1.0000	.6863
	σ^2	1.0000	.6068	.0007	-0.3932	.3932	.3932	1.0000	1.3681
	ν_1	0.5000	.7253	.0002	.2253	.4505	.2253	1.0000	.9025
	ν_2	0.5000	.2435	.0002	-0.2565	.5130	.2565	1.0000	.8607
500	β_0	1.0000	1.0166	.0000	.0166	.0166	.0166	1.0000	.2160
	β_1	2.0000	1.9607	.0000	-0.0393	.0197	.0393	1.0000	.2141
	γ	.0000	-0.0003	.0000	-0.0003	-	.0003	1.0000	.3600
	σ^2	1.0000	1.4013	.0004	.4013	.4013	.4013	1.0000	1.1050
	ν_1	0.5000	.4316	.0005	-0.0684	.1368	.0684	1.0000	.9619
	ν_2	0.5000	.8144	.0001	.3144	.6287	.3144	1.0000	.7144
1000	β_0	1.0000	1.0265	.0000	.0265	.0265	.0265	1.0000	.1489
	β_1	2.0000	2.0719	.0000	.0719	.0359	.0719	.7000	.1485
	γ	.0000	-0.0007	.0000	-0.0007	-	.0007	1.0000	.2763
	σ^2	1.0000	.5312	.0018	-0.4688	.4688	.4688	1.0000	1.2221
	ν_1	0.5000	.7970	.0009	.2970	.5940	.2970	1.0000	.8393
	ν_2	0.5000	.3133	.0004	-0.1867	.3735	.1867	1.0000	.7940

Table 51 – Results of the simulation study for the skew contaminated normal model with $\nu_1 = 0.5$ and $\nu_2 = 0.5$ and $\gamma = 0.9$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	1.3644	.0000	.3644	.3644	.3644	.2800	.8170
	β_1	2.0000	2.1895	.0000	.1895	.0948	.1895	1.0000	.8717
	γ	.9000	.9207	.0001	.0207	.0230	.0207	1.0000	.9849
	σ^2	1.0000	1.6647	.0006	.6647	.6647	.6647	1.0000	2.3233
	ν_1	0.5000	.1422	.0002	-0.3578	.7156	.3578	1.0000	.9528
	ν_2	0.5000	.6411	.0001	.1411	.2822	.1411	1.0000	.7972
250	β_0	1.0000	.9797	.0000	-0.0203	.0203	.0203	1.0000	.2860
	β_1	2.0000	1.9492	.0000	-0.0508	.0254	.0508	1.0000	.2409
	γ	.9000	.8021	.0001	-0.0979	.1087	.0979	1.0000	.3961
	σ^2	1.0000	1.1503	.0008	.1503	.1503	.1503	1.0000	1.2062
	ν_1	0.5000	.6530	.0004	.1530	.3060	.1530	1.0000	.9372
	ν_2	0.5000	.3590	.0003	-0.1410	.2820	.1410	1.0000	.8093
500	β_0	1.0000	1.0265	.0000	.0265	.0265	.0265	1.0000	.2144
	β_1	2.0000	2.0265	.0000	.0265	.0132	.0265	1.0000	.1643
	γ	.9000	.8762	.0000	-0.0238	.0265	.0238	1.0000	.1862
	σ^2	1.0000	1.4059	.0013	.4059	.4059	.4059	1.0000	1.1250
	ν_1	0.5000	.7262	.0012	.2262	.4524	.2262	1.0000	.9465
	ν_2	0.5000	.6675	.0003	.1675	.3349	.1675	1.0000	.6873
1000	β_0	1.0000	1.0650	.0000	.0650	.0650	.0650	1.0000	.1557
	β_1	2.0000	1.9866	.0000	-0.0134	.0067	.0134	1.0000	.1172
	γ	.9000	.8278	.0000	-0.0722	.0803	.0722	1.0000	.1995
	σ^2	1.0000	1.3148	.0009	.3148	.3148	.3148	1.0000	1.0044
	ν_1	0.5000	.5478	.0011	.0478	.0957	.0478	1.0000	.9176
	ν_2	0.5000	.4966	.0002	-0.0034	.0067	.0034	1.0000	.6305

B.4 Skew generalized t

Table 52 – Results of the simulation study for the skew generalized t model with $\nu_1 = 15$ and $\nu_2 = 5$ and $\gamma = -0.9$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	1.0702	< .0001	.0702	.0702	.0702	1.0000	.3291
	β_1	2.0000	2.0199	< .0001	.0199	.0100	.0199	1.0000	.3001
	γ	-0.9000	-0.7604	.0039	.1396	.1551	.1397	1.0000	.8963
	ν_1	15.0000	12.3660	3.7611	-2.6340	.1756	4.5917	1.0000	32.8329
	ν_2	5.0000	5.6363	1.2700	.6363	.1273	1.4205	1.0000	17.7954
250	β_0	1.0000	1.0628	< .0001	.0628	.0628	.0628	1.0000	.1496
	β_1	2.0000	2.0161	< .0001	.0161	.0081	.0161	1.0000	.1195
	γ	-0.9000	-0.8865	< .0001	.0135	.0150	.0135	1.0000	.2981
	ν_1	15.0000	11.6896	.2553	-3.3104	.2207	3.3203	1.0000	21.2921
	ν_2	5.0000	5.3889	.0697	.3889	.0778	.3951	1.0000	11.1118
500	β_0	1.0000	1.0234	< .0001	.0234	.0234	.0234	1.0000	.1011
	β_1	2.0000	1.9857	< .0001	-0.0143	.0072	.0143	1.0000	.0836
	γ	-0.9000	-0.9116	< .0001	-0.0116	.0129	.0116	1.0000	.1855
	ν_1	15.0000	15.7367	2.4037	.7367	.0491	2.5140	1.0000	24.4569
	ν_2	5.0000	7.2585	.6079	2.2585	.4517	2.3389	1.0000	12.4724
1000	β_0	1.0000	.9933	< .0001	-0.0067	.0067	.0067	1.0000	.0783
	β_1	2.0000	1.9977	< .0001	-0.0023	.0011	.0023	1.0000	.0642
	γ	-0.9000	-0.9375	< .0001	-0.0375	.0417	.0375	1.0000	.1302
	ν_1	15.0000	11.6692	.1879	-3.3308	.2221	3.3361	1.0000	12.3566
	ν_2	5.0000	6.1734	.0640	1.1734	.2347	1.1752	1.0000	7.3765

Table 53 – Results of the simulation study for the skew generalized t model with $\nu_1 = 15$ and $\nu_2 = 5$ and $\gamma = 0.9$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	.9426	.0000	-0.0574	.0574	.0574	1.0000	.3463
	β_1	2.0000	2.0217	.0000	.0217	.0109	.0217	1.0000	.2678
	γ	.9000	.3832	.0003	-0.5168	.5742	.5168	1.0000	1.1461
	ν_1	15.0000	10.8288	.2093	-4.1712	.2781	4.1764	1.0000	23.0118
	ν_2	5.0000	4.4050	.0450	-0.5950	.1190	.5967	1.0000	10.3769
250	β_0	1.0000	.9690	.0000	-0.0310	.0310	.0310	1.0000	.1467
	β_1	2.0000	1.9733	.0000	-0.0267	.0134	.0267	1.0000	.1018
	γ	.9000	.9405	.0000	.0405	.0450	.0405	1.0000	.1603
	ν_1	15.0000	20.1713	12.3478	5.1713	.3448	13.3870	1.0000	43.6633
	ν_2	5.0000	10.2310	3.5553	5.2310	1.0462	6.3248	1.0000	24.1319
500	β_0	1.0000	1.0123	.0000	.0123	.0123	.0123	1.0000	.1058
	β_1	2.0000	1.9945	.0000	-0.0055	.0027	.0055	1.0000	.0945
	γ	.9000	.8249	.0000	-0.0751	.0834	.0751	1.0000	.3141
	ν_1	15.0000	14.0663	1.7582	-0.9337	.0622	1.9907	1.0000	25.2807
	ν_2	5.0000	6.7107	.4678	1.7107	.3421	1.7735	1.0000	13.0913
1000	β_0	1.0000	1.0255	.0000	.0255	.0255	.0255	1.0000	.0811
	β_1	2.0000	1.9834	.0000	-0.0166	.0083	.0166	1.0000	.0624
	γ	.9000	.9676	.0000	.0676	.0751	.0676	.0400	.0785
	ν_1	15.0000	12.9975	.3901	-2.0025	.1335	2.0401	1.0000	14.8208
	ν_2	5.0000	7.3505	.1595	2.3505	.4701	2.3559	1.0000	9.5126

Table 54 – Results of the simulation study for the skew generalized t model with $\nu_1 = 5$ and $\nu_2 = 15$ and $\gamma = -0.9$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	1.4877	.0001	.4877	.4877	.4877	1.0000	1.1344
	β_1	2.0000	2.2021	$< .0001$	.2021	.1010	.2021	1.0000	1.1828
	γ	-0.9000	-0.2562	.0001	.6438	.7153	.6438	1.0000	1.1821
	ν_1	5.0000	4.3446	.0211	-0.6554	.1311	.6557	1.0000	8.7688
	ν_2	15.0000	14.3651	.5067	-0.6349	.0423	.8123	1.0000	47.3544
250	β_0	1.0000	1.1469	$< .0001$	.1469	.1469	.1469	1.0000	.5454
	β_1	2.0000	2.0418	$< .0001$	.0418	.0209	.0418	1.0000	.4103
	γ	-0.9000	-0.8788	$< .0001$	.0212	.0235	.0212	1.0000	.3753
	ν_1	5.0000	5.0572	.0035	.0572	.0114	.0573	1.0000	5.7130
	ν_2	15.0000	24.7713	.1813	9.7713	.6514	9.7730	1.0000	39.0705
500	β_0	1.0000	.9996	$< .0001$	-0.0004	.0004	.0004	1.0000	.3976
	β_1	2.0000	1.9729	$< .0001$	-0.0271	.0136	.0271	1.0000	.3269
	γ	-0.9000	-0.7920	.0001	.1080	.1200	.1080	1.0000	.4388
	ν_1	5.0000	4.3893	.0017	-0.6107	.1221	.6107	1.0000	3.4331
	ν_2	15.0000	19.6827	.0871	4.6827	.3122	4.6835	1.0000	22.0990
1000	β_0	1.0000	1.0978	$< .0001$	.0978	.0978	.0978	1.0000	.2756
	β_1	2.0000	1.8977	$< .0001$	-0.1023	.0511	.1023	1.0000	.2247
	γ	-0.9000	-0.8696	$< .0001$	.0304	.0338	.0304	1.0000	.2417
	ν_1	5.0000	4.5483	.0012	-0.4517	.0903	.4517	1.0000	2.5821
	ν_2	15.0000	21.1129	.0404	6.1129	.4075	6.1130	1.0000	16.3782

Table 55 – Results of the simulation study for the skew generalized t model with $\nu_1 = 5$ and $\nu_2 = 15$ and $\gamma = 0$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	.7419	.0001	-0.2581	.2581	.2581	1.0000	1.1925
	β_1	2.0000	1.9481	.0001	-0.0519	.0260	.0519	1.0000	1.3693
	γ	.0000	.0007	.0000	.0007	-	.0007	1.0000	.8963
	ν_1	5.0000	4.5297	.0256	-0.4703	.0941	.4710	1.0000	9.6611
	ν_2	15.0000	16.8329	.7727	1.8329	.1222	1.9891	1.0000	51.8011
250	β_0	1.0000	.9022	.0000	-0.0978	.0978	.0978	1.0000	.4789
	β_1	2.0000	2.0080	.0000	.0080	.0040	.0080	1.0000	.4685
	γ	.0000	.0190	.0000	.0190	-	.0190	1.0000	.6603
	ν_1	5.0000	4.0998	.0024	-0.9002	.1800	.9002	1.0000	4.5467
	ν_2	15.0000	11.4073	.0440	-3.5927	.2395	3.5930	1.0000	18.9437
500	β_0	1.0000	.8728	.0000	-0.1272	.1272	.1272	1.0000	.3387
	β_1	2.0000	1.9146	.0000	-0.0854	.0427	.0854	1.0000	.3353
	γ	.0000	.0128	.0000	.0128	-	.0128	1.0000	.4844
	ν_1	5.0000	5.4850	.0113	.4850	.0970	.4851	1.0000	5.5162
	ν_2	15.0000	16.3160	.1897	1.3160	.0877	1.3296	1.0000	22.4285
1000	β_0	1.0000	1.0983	.0000	.0983	.0983	.0983	1.0000	.2735
	β_1	2.0000	1.8829	.0000	-0.1171	.0586	.1171	.5400	.2345
	γ	.0000	.2922	.0007	.2922	-	.2922	1.0000	.6480
	ν_1	5.0000	5.4216	.0072	.4216	.0843	.4217	1.0000	3.7159
	ν_2	15.0000	19.1846	.3341	4.1846	.2790	4.1979	1.0000	20.3300

Table 56 – Results of the simulation study for the skew generalized t model with $\nu_1 = 5$ and $\nu_2 = 15$ and $\gamma = 0.9$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	.4429	.0001	-0.5571	.5571	.5571	.9800	1.1772
	β_1	2.0000	1.7195	.0001	-0.2805	.1402	.2805	1.0000	1.2659
	γ	.9000	.1873	.0002	-0.7127	.7919	.7127	1.0000	1.1694
	ν_1	5.0000	5.2312	.0428	.2312	.0462	.2352	1.0000	11.3928
	ν_2	15.0000	19.0628	1.1988	4.0628	.2709	4.2360	1.0000	58.9464
250	β_0	1.0000	.8250	.0001	-0.1750	.1750	.1750	1.0000	.5611
	β_1	2.0000	1.9118	.0000	-0.0882	.0441	.0882	1.0000	.4623
	γ	.9000	.7831	.0006	-0.1169	.1298	.1169	1.0000	.6010
	ν_1	5.0000	6.0773	.0342	1.0773	.2155	1.0778	1.0000	8.9527
	ν_2	15.0000	30.8466	1.4614	15.8466	1.0564	15.9139	1.0000	59.1671
500	β_0	1.0000	.9549	.0000	-0.0451	.0451	.0451	1.0000	.3836
	β_1	2.0000	2.0017	.0000	.0017	.0008	.0017	1.0000	.3253
	γ	.9000	.8870	.0000	-0.0130	.0144	.0130	1.0000	.2885
	ν_1	5.0000	5.0246	.0046	.0246	.0049	.0250	1.0000	3.9972
	ν_2	15.0000	24.3771	.1915	9.3771	.6251	9.3791	1.0000	26.5736
1000	β_0	1.0000	.9545	.0000	-0.0455	.0455	.0455	1.0000	.2697
	β_1	2.0000	2.0440	.0000	.0440	.0220	.0440	1.0000	.2205
	γ	.9000	.8916	.0000	-0.0084	.0094	.0084	1.0000	.2193
	ν_1	5.0000	5.4574	.0054	.4574	.0915	.4574	1.0000	3.4592
	ν_2	15.0000	26.5296	.1629	11.5296	.7686	11.5307	1.0000	21.4267

Table 57 – Results of the simulation study for the skew generalized t model with $\nu_1 = 30$ and $\nu_2 = 30$ and $\gamma = -0.9$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	1.2513	$< .0001$	.2513	.2513	.2513	.9500	.5399
	β_1	2.0000	1.9577	$< .0001$	-0.0423	.0212	.0423	1.0000	.5630
	γ	-0.9000	-0.5381	.0026	.3619	.4021	.3619	1.0000	1.0892
	ν_1	30.0000	7.6424	.1972	-22.3576	.7453	22.3584	.1000	19.0875
	ν_2	30.0000	7.3182	.3618	-22.6818	.7561	22.6847	.1000	23.1226
250	β_0	1.0000	1.1208	$< .0001$	.1208	.1208	.1208	.7000	.2454
	β_1	2.0000	1.9148	$< .0001$	-0.0852	.0426	.0852	.8500	.1875
	γ	-0.9000	-0.9326	$< .0001$	-0.0326	.0362	.0326	1.0000	.1944
	ν_1	30.0000	12.3288	.5415	-17.6712	.5890	17.6795	.4000	26.3319
	ν_2	30.0000	15.8351	1.1933	-14.1649	.4722	14.2150	.4000	38.2236
500	β_0	1.0000	1.0588	$< .0001$	.0588	.0588	.0588	1.0000	.1701
	β_1	2.0000	2.0089	$< .0001$	.0089	.0044	.0089	1.0000	.1333
	γ	-0.9000	-0.9620	$< .0001$	-0.0620	.0689	.0620	.8000	.0987
	ν_1	30.0000	17.9127	1.1854	-12.0873	.4029	12.1453	.7500	26.0248
	ν_2	30.0000	24.7469	2.6667	-5.2531	.1751	5.8912	.7500	40.3640
1000	β_0	1.0000	1.0344	$< .0001$	.0344	.0344	.0344	1.0000	.1276
	β_1	2.0000	2.0050	$< .0001$	.0050	.0025	.0050	1.0000	.0970
	γ	-0.9000	-0.9456	$< .0001$	-0.0456	.0507	.0456	1.0000	.1008
	ν_1	30.0000	23.0520	7.8809	-6.9480	.2316	10.5063	1.0000	34.4769
	ν_2	30.0000	35.1296	20.1995	5.1296	.1710	20.8406	1.0000	56.0644

Table 58 – Results of the simulation study for the skew generalized t model with $\nu_1 = 30$ and $\nu_2 = 30$ and $\gamma = 0$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	.8380	.0000	-0.1620	.1620	.1620	1.0000	.6343
	β_1	2.0000	2.0105	.0000	.0105	.0052	.0105	1.0000	.7601
	γ	.0000	.0153	.0000	.0153	-	.0153	1.0000	1.0411
	ν_1	30.0000	5.9724	.0638	-24.0276	.8009	24.0277	.0000	13.5852
	ν_2	30.0000	6.5047	.1222	-23.4953	.7832	23.4956	.0000	18.7696
250	β_0	1.0000	.9514	.0000	-0.0486	.0486	.0486	1.0000	.2611
	β_1	2.0000	1.9845	.0000	-0.0155	.0077	.0155	1.0000	.2494
	γ	.0000	.0182	.0000	.0182	-	.0182	1.0000	.5839
	ν_1	30.0000	7.9691	.2029	-22.0309	.7344	22.0319	.0200	14.3909
	ν_2	30.0000	7.7007	.2911	-22.2993	.7433	22.3012	.0200	17.3357
500	β_0	1.0000	1.0426	.0000	.0426	.0426	.0426	1.0000	.1744
	β_1	2.0000	2.0820	.0000	.0820	.0410	.0820	.9800	.1884
	γ	.0000	-0.0117	.0000	-0.0117	-	.0117	1.0000	.3726
	ν_1	30.0000	21.3851	16.7628	-8.6149	.2872	18.8470	1.0000	42.8951
	ν_2	30.0000	20.1523	17.3432	-9.8477	.3283	19.9441	1.0000	43.6459
1000	β_0	1.0000	.9803	.0000	-0.0197	.0197	.0197	1.0000	.1273
	β_1	2.0000	2.0269	.0000	.0269	.0134	.0269	1.0000	.1244
	γ	.0000	.1189	.0001	.1189	-	.1189	1.0000	.4548
	ν_1	30.0000	30.0794	41.4154	.0794	.0026	41.4155	1.0000	50.3727
	ν_2	30.0000	32.5994	57.0358	2.5994	.0866	57.0950	1.0000	58.8200

Table 59 – Results of the simulation study for the skew generalized t model with $\nu_1 = 30$ and $\nu_2 = 30$ and $\gamma = 0.9$

n	Parameter	Real	Est	Var	Bias	Rel Bias	RMSE	CR	length CI
50	β_0	1.0000	.6969	.0000	-0.3031	.3031	.3031	.1200	.5713
	β_1	2.0000	2.0877	.0000	.0877	.0438	.0877	1.0000	.6632
	γ	.9000	.2253	.0003	-0.6747	.7497	.6747	1.0000	1.1824
	ν_1	30.0000	7.6380	.6413	-22.3620	.7454	22.3712	.0200	19.3999
	ν_2	30.0000	7.3938	.6959	-22.6062	.7535	22.6169	.0200	22.5823
250	β_0	1.0000	.8599	.0000	-0.1401	.1401	.1401	.0400	.2493
	β_1	2.0000	2.0666	.0000	.0666	.0333	.0666	1.0000	.1958
	γ	.9000	.8833	.0000	-0.0167	.0186	.0167	1.0000	.3176
	ν_1	30.0000	10.2269	.6962	-19.7731	.6591	19.7853	.1200	19.8956
	ν_2	30.0000	12.5730	1.4558	-17.4270	.5809	17.4877	.1200	28.7830
500	β_0	1.0000	.9144	.0000	-0.0856	.0856	.0856	.9000	.1778
	β_1	2.0000	1.9606	.0000	-0.0394	.0197	.0394	1.0000	.1482
	γ	.9000	.9295	.0000	.0295	.0328	.0295	1.0000	.1664
	ν_1	30.0000	15.6277	12.6373	-14.3723	.4791	19.1381	.6600	30.4127
	ν_2	30.0000	21.7712	28.1327	-8.2288	.2743	29.3114	.6600	46.1574
1000	β_0	1.0000	.9899	.0000	-0.0101	.0101	.0101	1.0000	.1259
	β_1	2.0000	2.0065	.0000	.0065	.0032	.0065	1.0000	.0954
	γ	.9000	.9450	.0000	.0450	.0500	.0450	.7800	.0936
	ν_1	30.0000	29.1444	22.4531	-0.8556	.0285	22.4694	1.0000	45.3264
	ν_2	30.0000	45.1337	58.2821	15.1337	.5045	60.2149	1.0000	73.5480

APPENDIX C – Average of the values of the model selection criteria of Section 1.5.2

Table 60 – Average of the values of the model selection criteria, for the skew-t model

		Model				
n	criterion	SCN	SSL1	SSL2	ST	SGT
50	EAIC	191.53	191.99	192.28	191.48	193.61
	EBIC	203.00	201.55	201.84	201.04	203.17
	DIC	184.06	185.28	185.33	185.06	186.48
	LPML	-93.01	-93.47	-93.71	-93.24	-93.96
250	EAIC	1158.73	936.71	936.81	937.56	946.51
	EBIC	1179.86	954.32	954.42	955.16	964.12
	DIC	1151.17	930.67	930.80	931.79	942.22
	LPML	-466.31	-465.67	-465.75	-466.00	-471.43
500	EAIC	1747.36	1738.86	1738.89	1737.36	1757.06
	EBIC	1772.64	1759.93	1759.96	1758.43	1778.14
	DIC	1741.32	1733.61	1733.64	1731.74	1750.98
	LPML	-871.30	-866.76	-866.78	-865.89	-875.39
1000	EAIC	3559.98	3523.82	3523.90	3515.58	3515.58
	EBIC	3589.43	3548.36	3548.44	3540.12	3540.11
	DIC	3553.91	3518.43	3518.56	3509.86	3509.89
	LPML	-1781.08	-1759.19	-1759.26	-1754.88	-1754.90

Table 61 – Average of the values of the model selection criteria, for the skew slash model

		Model				
n	criterion	SCN	SSL1	SSL2	ST	SGT
50	EAIC	191.34	174.83	175.33	173.61	176.50
	EBIC	202.81	184.39	184.89	183.17	186.06
	DIC	185.91	167.08	168.10	166.81	169.48
	LPML	-84.81	-84.77	-84.90	-83.99	-85.40
250	EAIC	799.51	797.60	798.17	797.01	815.84
	EBIC	820.64	815.21	815.78	814.61	833.44
	DIC	793.59	792.53	789.11	791.32	809.06
	LPML	-396.50	-396.25	-396.85	-395.56	-404.47
500	EAIC	1572.75	1570.14	1570.74	1571.33	1610.50
	EBIC	1598.04	1591.21	1591.82	1592.40	1631.57
	DIC	1565.26	1562.87	1563.35	1566.03	1604.15
	LPML	-783.31	-782.84	-783.15	-783.08	-801.38
1000	EAIC	3003.31	3002.54	3002.83	3005.38	3109.88
	EBIC	3032.76	3027.08	3027.37	3029.92	3134.42
	DIC	2996.96	2997.33	2995.06	2999.58	3101.30
	LPML	-1499.10	-1498.81	-1499.08	-1499.98	-1550.92

Table 62 – Average of the values of the model selection criteria, for the skew contaminated model

		Model				
n	criterion	SCN	SSL1	SSL2	ST	SGT
50	EAIC	183.66	182.13	183.74	182.33	188.01
	EBIC	195.13	191.69	193.30	191.89	197.57
	DIC	176.35	174.71	177.55	175.70	181.41
	LPML	-87.76	-88.53	-89.73	-87.96	-90.73
250	EAIC	782.63	781.07	781.86	781.39	791.48
	EBIC	803.76	798.67	799.46	799.00	809.08
	DIC	776.26	774.85	769.53	775.99	788.54
	LPML	-389.11	-387.95	-388.81	-387.87	-393.27
500	EAIC	1659.63	1666.98	1667.05	1667.37	1693.08
	EBIC	1684.92	1688.05	1688.12	1688.45	1714.16
	DIC	1653.64	1661.72	1661.81	1661.71	1686.29
	LPML	-827.00	-830.77	-830.83	-830.82	-842.99
1000	EAIC	3069.47	3080.84	3080.80	3084.65	3163.52
	EBIC	3098.91	3105.38	3105.34	3109.19	3188.06
	DIC	3063.24	3075.30	3075.22	3078.65	3155.23
	LPML	-1531.44	-1537.50	-1537.47	-1539.25	-1577.89

Table 63 – Average of the values of the model selection criteria, for the skew generalized t model

		Model				
n	criterion	SCN	SSL1	SSL2	ST	SGT
50	EAIC	147.47	147.04	135.84	146.69	147.78
	EBIC	158.94	156.60	145.40	156.25	157.34
	DIC	141.21	140.99	126.19	140.73	141.45
	LPML	-70.93	-71.08	-63.13	-70.87	-71.44
250	EAIC	668.30	667.53	667.72	668.15	678.68
	EBIC	689.43	685.14	685.32	685.75	696.29
	DIC	661.96	661.78	662.04	662.54	674.84
	LPML	-331.38	-331.03	-331.17	-331.32	-337.61
500	EAIC	1290.39	1293.93	1293.94	1299.32	1319.59
	EBIC	1315.68	1315.00	1315.01	1320.40	1340.66
	DIC	1284.09	1288.94	1288.95	1294.14	1318.09
	LPML	-642.28	-644.49	-644.50	-647.28	-658.72
1000	EAIC	2560.07	2549.88	2589.64	2549.73	2546.39
	EBIC	2589.52	2574.41	2614.18	2574.27	2570.93
	DIC	2554.39	2544.79	2589.28	2544.49	2540.81
	LPML	-1278.62	-1272.59	-1295.38	-1272.62	-1270.75

APPENDIX D – Proof of the preposition 2.2.1

Proof of the preposition 2.2.1 presented in section 2.2

Proof. Consider $W_i \sim SMSN_c(0, 1, \gamma, G, \boldsymbol{\nu})$. From the stochastic representation of the skew-normal we have that

$$W_i = U_i^{-1/2}(\Delta(H_i - b) + \sqrt{\tau}T_i) \quad i = 1, \dots, n$$

then,

$$\begin{aligned} W_i|H_i, U_i &\sim N(u_i^{-1/2}\Delta(h_i - b), u_i^{-1}\tau) \\ H_i &\sim HN(0, 1) \\ U_i &\sim G(.|\boldsymbol{\nu}) \end{aligned} \tag{D.1}$$

Therefore, considering $T_i \sim N(0, 1)$, $\eta_i = {X_i}^t\boldsymbol{\beta}$ and denoting the cdf of H_i as $F(h_i)$, then

$$\begin{aligned} p_i &= \int\int P\left(T_i \leq \frac{\eta_i - u_i^{-1/2}\Delta(h_i - b)}{u_i^{-1/2}\sqrt{\tau}}|H_i, U_i\right) dG(u_i|\boldsymbol{\nu})dF(h_i) \\ &= \int\int P\left(T_i > \frac{-\eta_i + u_i^{-1/2}\Delta(h_i - b)}{u_i^{-1/2}\sqrt{\tau}}|H_i, U_i\right) dG(u_i|\boldsymbol{\nu})dF(h_i) \\ &= \int\int P\left(u_i^{-1/2}\sqrt{\tau}T_i + \eta_i - u_i^{-1/2}\Delta(h_i - b) > 0|H_i, U_i\right) dG(u_i|\boldsymbol{\nu})dF(h_i) \\ &= \int\int P\left(Z_i > 0|H_i, U_i\right) dG(u_i|\boldsymbol{\nu})dF(h_i) \\ &= P(Z_i > 0), \end{aligned} \tag{D.2}$$

where $Z_i \sim SMSN_c(\eta_i, 1, -\gamma, G, \boldsymbol{\nu})$. Therefore, we found that considering $y_i = I(Z_i > 0)$ with $Z_i = {X_i}^t\boldsymbol{\beta} + {U_i}^{-1/2}(\Delta(b - H_i) + \sqrt{\tau}T_i$ is equivalent to the binary model $Y_i \sim$ bernoulli(p_i) with $p_i = F({X_i}^t\boldsymbol{\beta}|\gamma, \boldsymbol{\nu})$. $\square$

APPENDIX E – Construction of the simulated envelopes

The envelopes are constructed (for the deviance residuals) in the following way:

1. From the results of fitting the binary response with link function based on the SMSN class, a sample of size n is generated considering appropriate estimates (e.g., posterior expectation) for the parameters.
2. We fit the model used to generate the sample in (1) in this dataset, generating a valid MCMC sample of size m to the model parameters.
3. For $i = 1, \ldots, n$ and $j = 1, \ldots, m$ we calculate the deviance residual (d^*_{ij}) using the j-th values of the valid MCMC sample for $\boldsymbol{\beta}$,γ and $\boldsymbol{\nu}$.
4. From the previous step, we will have a matrix with the residuals
$$\begin{bmatrix} d^*_{11} & d^*_{12} & \cdots & d^*_{1m} \\ d^*_{21} & d^*_{22} & \cdots & d^*_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ d^*_{n1} & d^*_{n2} & & d^*_{nm} \end{bmatrix}$$
5. From each sample, we sort the residuals into ascending order, that we shall denote by ($d^*_{(i)j}$) the i-th order residual.
$$\begin{bmatrix} d^*_{(1)1} & d^*_{(1)2} & \cdots & d^*_{(1)m} \\ d^*_{(2)1} & d^*_{(2)2} & \cdots & d^*_{(2)m} \\ \vdots & \vdots & \ddots & \vdots \\ d^*_{(n)1} & d^*_{(n)2} & & d^*_{(n)m} \end{bmatrix}$$
6. Then, we calculate the limits $d^*_{(i)L} = \dfrac{d^*_{(i)(2)} + d^*_{(i)(3)}}{2}$ and $d^*_{(i)U} = \dfrac{d^*_{(i)(m-2)} + d^*_{(i)(m-1)}}{2}$, where $d^*_{(i)(r)}$ is the r-th order statistic for the i-th line. We also consider the reference line $d^*_{(i)} = \dfrac{1}{m}\sum_{j=1}^{m} d^*_{(i)j}$, for $i = 1, 2, \ldots, n$.

APPENDIX F – Results of the simulations study: parameter recovery for section 2.4.1

We present the results related to the Section 2.4.1 not presented in the main document.

Table 64 – Results of the simulation study for the skew-t model with $\nu = 10$.

		50			250			500			1000		
		Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
	$\gamma = -0.9$	1.459	.76	[0.272;2.911]	1.04	.17	[0.746;1.38]	.981	.152	[0.716;1.285]	1.027	.08	[0.877;1.184]
β_0	$\gamma = 0$	.782	.357	[0.128;1.427]	.954	.211	[0.592;1.391]	.917	.148	[0.663;1.211]	1.042	.138	[0.797;1.302]
	$\gamma = .9$	.578	.343	[-0.088;1.23]	1.32	.317	[0.803;2.012]	1.08	.217	[0.717;1.538]	.93	.132	[0.718;1.226]
	$\gamma = -0.9$	3.271	1.452	[1.015;5.949]	1.96	.371	[1.333;2.683]	1.949	.303	[1.473;2.586]	1.806	.162	[1.509;2.14]
β_1	$\gamma = 0$	1.929	.725	[0.836;3.421]	2.184	.456	[1.417;3.193]	2.059	.296	[1.592;2.672]	2.207	.291	[1.76;2.846]
	$\gamma = .9$	2.068	.593	[1.029;3.239]	2.432	.57	[1.524;3.61]	2.055	.329	[1.553;2.786]	1.95	.219	[1.611;2.461]
	$\gamma = -0.9$	.868	.582	[-0.862;0.995]	-0.946	.244	[-0.995;-0.259]	-0.936	.2	[-0.995;-0.37]	-0.976	.082	[-0.995;-0.739]
γ	$\gamma = 0$	.803	.627	[-0.995;0.937]	-0.849	.472	[-0.995;0.484]	.107	.442	[-0.678;0.989]	-0.054	.368	[-0.946;0.493]
	$\gamma = .9$	.87	.653	[-0.94;0.995]	.003	.519	[-0.995;0.753]	.969	.157	[0.547;0.995]	.975	.115	[0.644;0.995]
	$\gamma = -0.9$	18.503	19.679	[2.003;57.66]	19.883	18.559	[2.036;57.776]	16.213	18.016	[2.075;51.191]	21.364	16.022	[2.871;49.007]
ν	$\gamma = 0$	15.456	16.474	[2.009;51.824]	16.973	18.725	[2.013;49.602]	16.803	16.749	[2.107;53.724]	13.585	15.06	[2.013;42.553]
	$\gamma = .9$	21.427	19.925	[2.043;62.099]	14.507	17.319	[2.006;46.514]	17.057	18.171	[2.003;51.503]	18.623	16.708	[2.145;51.55]

Table 65 – Results of the simulation study for the skew-t model with $\nu = 50$.

		50			250			500			1000		
		Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
	$\gamma = -0.9$	1.461	.472	[0.701;2.525]	1.224	.199	[0.826;1.608]	1.149	.133	[0.898;1.402]	1.046	.081	[0.891;1.204]
β_0	$\gamma = 0$	1.419	.637	[0.411;2.773]	1.49	.312	[0.966;2.161]	1.036	.124	[0.828;1.307]	1.07	.096	[0.885;1.251]
	$\gamma = .9$	1.82	.832	[0.599;3.423]	1.314	.245	[0.863;1.797]	1.121	.154	[0.848;1.439]	1.058	.123	[0.83;1.31]
	$\gamma = -0.9$	2.157	.693	[1.014;3.596]	2.607	.402	[1.872;3.374]	2.273	.261	[1.823;2.831]	2.083	.18	[1.709;2.407]
β_1	$\gamma = 0$	3.019	1.093	[1.273;5.346]	3.092	.545	[2.125;4.122]	2.057	.228	[1.601;2.476]	2.257	.181	[1.938;2.614]
	$\gamma = .9$	4.61	2.049	[1.618;8.616]	2.406	.382	[1.751;3.197]	2.109	.213	[1.721;2.548]	2.228	.196	[1.871;2.599]
	$\gamma = -0.9$	-0.851	.649	[-0.995;0.948]	-0.891	.585	[-0.995;0.819]	-0.9	.354	[-0.995;0.137]	-0.97	.117	[-0.995;-0.629]
γ	$\gamma = 0$	.857	.636	[-0.911;0.995]	.905	.404	[-0.275;0.995]	-0.074	.476	[-0.993;0.744]	.044	.344	[-0.797;0.516]
	$\gamma = .9$	-0.871	.643	[-0.995;0.91]	.884	.477	[-0.501;0.995]	.968	.184	[0.429;0.995]	.975	.128	[0.607;0.995]
	$\gamma = -0.9$	21.334	19.534	[2.006;62.143]	27.961	21.746	[2.65;69.919]	26.16	19.555	[2.952;66.408]	27.988	19.734	[3.8;67.138]
ν	$\gamma = 0$	19.312	19.843	[2.013;60.652]	22.107	19.559	[2.024;62.295]	26.125	20.647	[2.635;67.37]	30.316	21.405	[4.238;71.885]
	$\gamma = .9$	20.136	20.446	[2.013;59.708]	21.829	19.391	[2.016;61.454]	25.125	20.635	[2.268;63.723]	23.637	18.864	[3.15;61.944]

Table 66 – Results of the simulation study for the skew slash model with $\nu = 10$.

		50			250			500			1000		
		Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
	$\gamma = -0.9$	1.087	.509	[0.346;2.099]	.84	.131	[0.586;1.099]	.941	.134	[0.721;1.214]	.975	.079	[0.843;1.135]
β_0	$\gamma = 0$	1.615	.649	[0.523;2.994]	.81	.143	[0.565;1.077]	.895	.158	[0.625;1.201]	1.038	.165	[0.807;1.419]
	$\gamma = .9$	1.653	.618	[0.574;2.937]	.828	.228	[0.435;1.301]	.912	.137	[0.651;1.18]	.946	.114	[0.753;1.193]
	$\gamma = -0.9$	2.219	.773	[0.857;3.679]	1.567	.265	[1.067;2.037]	2.192	.31	[1.59;2.721]	1.965	.16	[1.663;2.28]
β_1	$\gamma = 0$	2.925	.999	[1.312;5.055]	1.614	.298	[1.061;2.18]	2.092	.304	[1.62;2.707]	2.075	.324	[1.693;2.864]
	$\gamma = .9$	2.712	.875	[1.181;4.512]	1.975	.488	[1.258;3.07]	2.014	.218	[1.618;2.413]	1.977	.197	[1.653;2.384]
	$\gamma = -0.9$	.774	.595	[-0.895;0.995]	-0.942	.273	[-0.995;-0.15]	-0.901	.257	[-0.995;-0.154]	-0.963	.119	[-0.995;-0.629]
γ	$\gamma = 0$	-0.868	.692	[-0.995;0.968]	-0.858	.355	[-0.995;0.132]	.705	.284	[0.045;0.994]	-0.06	.274	[-0.563;0.483]
	$\gamma = .9$	.862	.704	[-0.976;0.992]	.515	.426	[-0.473;0.993]	.975	.128	[0.622;0.995]	.966	.11	[0.664;0.995]
	$\gamma = -0.9$	18.649	18.897	[1.01;57.742]	21.338	19.08	[1.007;61.975]	20.356	18.757	[1.011;58.968]	24.594	20.499	[1.5;64.689]
ν	$\gamma = 0$	20.026	19.687	[1.004;61.925]	20.146	19.939	[1.015;61.305]	18.054	16.955	[1.043;54.554]	15.365	15.435	[1.038;48.75]
	$\gamma = .9$	19.052	18.315	[1.006;60.774]	12.571	15.349	[1.002;44.78]	21.171	18.923	[1.132;57.577]	17.942	16.822	[1.352;52.029]

Table 67 – Results of the simulation study for the skew slash model with $\nu = 50$.

		50			250			500			1000		
		Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_0	$\gamma = -0.9$	1.974	.968	[0.546;3.828]	1.214	.199	[0.838;1.573]	1.188	.136	[0.954;1.473]	1.053	.097	[0.868;1.241]
	$\gamma = 0$	1.043	.406	[0.331;1.814]	1.076	.235	[0.7;1.598]	1.119	.203	[0.754;1.506]	1.081	.109	[0.88;1.297]
	$\gamma = .9$	2.645	1.279	[0.986;5.197]	1.321	.321	[0.782;1.971]	.993	.147	[0.708;1.271]	1.088	.106	[0.886;1.287]
β_1	$\gamma = -0.9$	4.456	1.995	[1.782;8.181]	2.387	.407	[1.681;3.17]	2.663	.311	[2.099;3.253]	2.12	.207	[1.764;2.502]
	$\gamma = 0$	2.452	.779	[1.181;3.834]	2.481	.487	[1.704;3.519]	2.274	.365	[1.77;3.147]	2.274	.22	[1.908;2.77]
	$\gamma = .9$	4.044	1.846	[1.71;7.65]	2.317	.473	[1.626;3.299]	2.106	.228	[1.726;2.569]	2.258	.172	[1.933;2.575]
γ	$\gamma = -0.9$	-0.837	.628	[-0.995;0.922]	-0.921	.249	[-0.995;-0.21]	-0.939	.249	[-0.995;-0.229]	-0.94	.139	[-0.995;-0.571]
	$\gamma = 0$	-0.848	.601	[-0.995;0.868]	-0.315	.43	[-0.969;0.583]	.737	.288	[0.063;0.993]	-0.192	.297	[-0.716;0.406]
	$\gamma = .9$	.829	.621	[-0.932;0.995]	.907	.327	[-0.016;0.995]	.964	.161	[0.534;0.995]	.959	.159	[0.5;0.995]
ν	$\gamma = -0.9$	17.553	19.273	[1.009;54.037]	20.305	21.095	[1.109;62.478]	22.875	21.128	[1.816;65.203]	19.926	18.295	[1.667;55.529]
	$\gamma = 0$	18.92	19.105	[1;57.34]	18.672	19.015	[1.015;57.292]	17.591	19.289	[1.032;51.412]	23.711	20.239	[1.503;67.321]
	$\gamma = .9$	17.761	20.433	[1.004;57.001]	18.277	19.319	[1.01;57.119]	19.064	17.368	[1.302;54.553]	22.458	18.662	[1.828;56.333]

Table 68 – Results of the simulation study for the skew generalized t model with $\nu_1 = 5$ and $\nu_2 = 15$.

		50			250			500			1000		
		Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_0	$\gamma = -0.9$	.738	.83	[-0.107;2.354]	1.342	.643	[0.391;2.765]	1.172	.803	[0.311;2.758]	.167	.046	[0.106;0.264]
	$\gamma = 0$	1.395	1.029	[0.191;3.445]	.664	.452	[0.11;1.606]	.832	.424	[0.236;1.667]	.967	.448	[0.359;1.745]
	$\gamma = .9$	1.722	1.268	[0.158;3.87]	1.212	.557	[0.287;2.273]	1.232	.49	[0.402;2.125]	.995	.711	[0.251;2.533]
β_1	$\gamma = -0.9$	2.04	1.851	[0.177;5.007]	2.092	1.002	[0.607;3.863]	2.221	1.52	[0.534;5.408]	.332	.086	[0.22;0.517]
	$\gamma = 0$	1.955	1.537	[0.361;5.153]	1.832	.99	[0.512;3.69]	1.508	.763	[0.535;2.911]	1.655	.762	[0.638;3.168]
	$\gamma = .9$	1.992	1.429	[0.362;4.505]	2.032	.904	[0.474;3.701]	2.967	1.168	[0.856;4.779]	1.69	1.143	[0.438;4.163]
γ	$\gamma = -0.9$	-0.597	.638	[-0.947;0.989]	-0.853	.578	[-0.995;0.751]	-0.402	.438	[-0.992;0.433]	-0.638	.251	[-0.992;-0.158]
	$\gamma = 0$	.789	.668	[-0.957;0.99]	.811	.595	[-0.955;0.995]	-0.758	.542	[-0.994;0.792]	-0.609	.272	[-0.995;-0.074]
	$\gamma = .9$	.866	.695	[-0.974;0.995]	.849	.562	[-0.72;0.994]	.882	.375	[-0.146;0.995]	.887	.418	[-0.305;0.993]
ν_1	$\gamma = -0.9$	10.55	12.746	[2.002;39.722]	4.124	2.553	[2.003;9.845]	6.245	3.979	[2.019;13.897]	10.43	6.667	[2.383;25.299]
	$\gamma = 0$	8.326	9.851	[2.01;24.707]	4.739	3.15	[2.002;11.469]	7.023	5.156	[2.024;16.659]	4.532	2.668	[2.009;12.132]
	$\gamma = .9$	9.014	6.968	[2.081;24.288]	5.312	3.317	[2.023;12.27]	3.218	1.651	[2.029;6.238]	9.137	6.517	[2.131;21.944]
ν_2	$\gamma = -0.9$	14.986	17.07	[0.369;50.12]	17.762	16.187	[1.714;50.932]	12.883	15.704	[0.495;53.266]	1.032	.734	[0.203;2.417]
	$\gamma = 0$	16.653	14.607	[0.974;48.638]	11.27	10.744	[0.708;31.321]	13.037	13.561	[0.714;36.443]	10.914	9.809	[1.309;31.425]
	$\gamma = .9$	14.144	14.144	[0.685;45.736]	17.864	16.819	[1.999;48.058]	15.2	12.497	[0.686;42.56]	15.361	14.412	[0.665;52.763]

Table 69 – Results of the simulation study for the skew generalized t model with $\nu_1 = 15$ and $\nu_2 = 5$.

		50			250			500			1000		
		Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_0	$\gamma = -0.9$	2.218	1.603	[0.327;5.8]	1.806	.881	[0.632;3.642]	3.4	1.347	[1.433;6.455]	1.139	.648	[0.473;2.692]
	$\gamma = 0$	2.844	1.515	[0.609;5.628]	1.529	.751	[0.287;2.89]	2.465	.974	[0.944;4.496]	3.573	1.052	[1.939;5.664]
	$\gamma = .9$	3.07	1.652	[0.864;6.584]	2.257	1.148	[0.538;4.586]	2.292	.705	[0.963;3.582]	2.192	.781	[1.056;3.646]
β_1	$\gamma = -0.9$	6.15	4.025	[0.626;14.212]	3.45	1.601	[1.332;6.574]	6.66	2.651	[2.667;12.364]	2.361	1.354	[0.932;5.596]
	$\gamma = 0$	7.756	3.788	[2.246;15.09]	2.908	1.422	[0.655;5.444]	4.694	1.821	[1.749;8.17]	7.162	2.015	[3.817;10.969]
	$\gamma = .9$	7.257	3.803	[1.892;14.288]	4.186	2.093	[1.113;8.734]	4.5	1.356	[2.046;6.868]	4.529	1.598	[2.204;7.446]
γ	$\gamma = -0.9$	.812	.691	[-0.953;0.994]	-0.684	.541	[-0.994;0.846]	-0.836	.565	[-0.991;0.841]	-0.455	.303	[-0.759;0.299]
	$\gamma = 0$	.827	.679	[-0.944;0.995]	.875	.415	[-0.385;0.993]	.296	.446	[-0.606;0.989]	-0.025	.412	[-0.753;0.809]
	$\gamma = .9$	.778	.671	[-0.973;0.995]	.926	.291	[0.076;0.994]	.945	.21	[0.279;0.995]	.925	.194	[0.417;0.994]
ν_1	$\gamma = -0.9$	13.455	13.771	[2.137;41.195]	13.326	9.45	[2.385;32.589]	8.047	6.478	[2.653;16.97]	23.721	14.526	[2.93;52.123]
	$\gamma = 0$	10.378	10.993	[2.025;33.371]	15.17	9.846	[2.399;36.349]	8.428	5.127	[2.3;17.531]	6.1	3.392	[2.192;14.154]
	$\gamma = .9$	8.807	7.222	[2.025;24.184]	11.768	9.519	[2.163;31.956]	12.5	7.591	[3.147;28.531]	6.226	3.885	[2.043;15.017]
ν_2	$\gamma = -0.9$	9.738	10.011	[0.42;27.038]	10.724	11.802	[0.643;30.807]	24.23	24.093	[3.268;74.64]	7.454	5.173	[0.546;17.708]
	$\gamma = 0$	12.738	13.26	[0.552;41.335]	12.723	11.164	[0.645;36.229]	10.611	9.426	[0.98;28.234]	17.357	13.282	[2.269;37.765]
	$\gamma = .9$	16.005	16.694	[0.551;51.512]	13.219	11.183	[0.742;34.556]	15.716	13.422	[0.828;41.353]	8.961	11.235	[0.882;35.335]

Table 70 – Results of the simulation study for the skew contaminated normal model with $\nu_1 = 0.1$ and $\nu_2 = 0.9$.

		50			250			500			1000		
		Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_0	$\gamma = -0.9$	2.527	2.214	[0.34;7.227]	1.622	.953	[0.496;3.646]	2.823	1.57	[0.858;5.989]	.994	.206	[0.773;1.33]
	$\gamma = 0$	2.793	2.306	[0.516;8.057]	2.006	1.43	[0.35;4.915]	1.318	.376	[0.744;2.056]	2.467	1.636	[0.808;5.776]
	$\gamma = .9$	2.981	2.193	[-0.075;7.062]	3.398	1.904	[0.649;6.94]	1.29	.667	[0.562;2.753]	1.623	.764	[0.927;2.848]
β_1	$\gamma = -0.9$	4.249	3.68	[0.755;13.274]	5.304	3.291	[1.561;12.134]	5.864	3.183	[1.711;12.109]	1.872	.403	[1.427;2.472]
	$\gamma = 0$	4.12	3.47	[0.641;11.945]	4.456	3.045	[1.113;10.545]	2.842	.908	[1.586;4.813]	4.76	3.084	[1.585;10.826]
	$\gamma = .9$	6.727	4.456	[0.914;15.429]	6.568	3.773	[1.367;13.306]	3.122	1.719	[1.281;6.822]	2.976	1.34	[1.801;5.255]
γ	$\gamma = -0.9$	.857	.602	[-0.859;0.994]	-0.613	.511	[-0.994;0.737]	-0.524	.383	[-0.925;0.43]	-0.945	.193	[-0.995;-0.394]
	$\gamma = 0$	-0.907	.477	[-0.995;0.609]	.732	.621	[-0.995;0.938]	-0.659	.609	[-0.878;0.995]	.447	.322	[-0.337;0.872]
	$\gamma = .9$	.883	.585	[-0.891;0.995]	.692	.594	[-0.928;0.99]	.144	.504	[-0.993;0.753]	.884	.316	[0.015;0.995]
ν_1	$\gamma = -0.9$	.605	.267	[0.12;0.996]	.522	.235	[0.13;0.976]	.549	.203	[0.079;0.892]	.211	.237	[0.005;0.751]
	$\gamma = 0$	.581	.268	[0.122;1]	.656	.234	[0.165;0.997]	.228	.137	[0.041;0.509]	.611	.252	[0.116;1]
	$\gamma = .9$	.509	.24	[0.107;0.956]	.495	.199	[0.086;0.856]	.465	.246	[0.081;0.932]	.313	.177	[0.07;0.701]
ν_2	$\gamma = -0.9$	.266	.278	[0.002;0.868]	.19	.212	[0.005;0.695]	.146	.18	[0.007;0.608]	.267	.19	[0.004;0.713]
	$\gamma = 0$	.23	.273	[0.001;0.831]	.261	.28	[0.008;0.851]	.048	.046	[0.002;0.13]	.211	.219	[0.009;0.727]
	$\gamma = .9$	.131	.223	[0.001;0.729]	.098	.169	[0.003;0.525]	.166	.174	[0.003;0.569]	.12	.069	[0.011;0.245]

Table 71 – Results of the simulation study for the skew contaminated normal model with $\nu_1 = 0.9$ and $\nu_2 = 0.1$.

		50			250			500			1000		
		Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_0	$\gamma = -0.9$	-1.659	1.424	[-4.6;0.148]	1.074	1.101	[0.059;3.569]	.925	.943	[0.152;3.301]	.661	.613	[0.177;2.121]
	$\gamma = 0$	1.228	1.442	[-0.244;4.398]	.33	.472	[-0.133;1.438]	.882	.8	[0.106;2.725]	1.16	.829	[0.252;3.043]
	$\gamma = .9$	.915	1.542	[-0.584;4.426]	.941	.914	[0.054;2.849]	1.463	.99	[0.37;3.705]	.984	.979	[0.285;3.446]
β_1	$\gamma = -0.9$	4.734	3.279	[0.629;11.228]	1.949	1.844	[0.467;6.332]	1.531	1.484	[0.505;5.106]	1.376	1.234	[0.557;4.544]
	$\gamma = 0$	2.375	2.301	[0.372;7.867]	1.563	1.529	[0.494;4.911]	1.637	1.377	[0.454;4.497]	2.395	1.561	[0.7;5.532]
	$\gamma = .9$	.263	1.032	[-1.863;2.603]	1.535	1.352	[0.346;4.378]	2.184	1.6	[0.545;5.77]	1.55	1.493	[0.531;5.473]
γ	$\gamma = -0.9$	-0.826	.61	[-0.994;0.916]	.799	.627	[-0.92;0.995]	-0.86	.535	[-0.995;0.731]	-0.883	.301	[-0.994;-0.063]
	$\gamma = 0$	-0.817	.687	[-0.968;0.995]	-0.807	.566	[-0.995;0.807]	-0.858	.615	[-0.994;0.9]	.13	.499	[-0.982;0.703]
	$\gamma = .9$	.824	.678	[-0.995;0.972]	-0.856	.665	[-0.995;0.961]	.84	.617	[-0.911;0.995]	.859	.522	[-0.631;0.995]
ν_1	$\gamma = -0.9$	.593	.24	[0.189;0.999]	.653	.275	[0.093;0.998]	.629	.303	[0.087;0.996]	.635	.313	[0.078;1]
	$\gamma = 0$	.67	.272	[0.127;0.999]	.619	.286	[0.082;1]	.676	.287	[0.086;0.998]	.722	.204	[0.274;0.991]
	$\gamma = .9$	.655	.272	[0.111;1]	.64	.297	[0.074;0.996]	.634	.219	[0.28;0.995]	.646	.3	[0.083;0.999]
ν_2	$\gamma = -0.9$	.199	.253	[0.001;0.813]	.297	.288	[0.003;0.885]	.418	.301	[0.007;0.947]	.446	.293	[0.008;0.955]
	$\gamma = 0$	.336	.291	[0.002;0.899]	.365	.276	[0.003;0.869]	.335	.287	[0.006;0.881]	.176	.207	[0.006;0.638]
	$\gamma = .9$	.294	.289	[0.001;0.864]	.314	.291	[0.005;0.895]	.157	.221	[0.005;0.752]	.403	.306	[0.006;0.951]

APPENDIX G – Results of the simulations study: parameter recovery for section 3.4.2

Here we presented the tables with the results for the simulation study in Section 3.4.2 containing the scenarios not presented in this section.

Table 72 – Results of the simulation study under $\gamma_c = 0.9$, $\gamma_w = 0$ and $\gamma_b = 0$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	1.133	.429	[.392; 1.994]	.822	.131	[.558; 1.059]	.945	.113	[.705; 1.156]	.940	.081	[.786; 1.100]
β_{b1}	2.000	2.277	.574	[1.208; 3.337]	1.599	.193	[1.220; 1.975]	1.962	.173	[1.628; 2.319]	1.820	.128	[1.533; 2.044]
β_{c0}	1.000	1.399	.233	[.971; 1.873]	1.129	.083	[.980; 1.299]	.980	.065	[.865; 1.117]	1.022	.046	[.934; 1.111]
β_{c1}	2.000	1.977	.229	[1.564; 2.483]	2.109	.076	[1.953; 2.249]	1.994	.061	[1.887; 2.113]	2.044	.039	[1.959; 2.114]
γ_b	.000	.022	.658	[-0.961; .995]	-0.718	.313	[-0.995; -0.067]	-0.340	.463	[-0.995; .504]	.043	.515	[-0.995; .830]
γ_c	.900	.171	.641	[-0.934; .990]	.840	.186	[.465; .993]	.854	.170	[.512; .992]	.850	.127	[.590; .988]
γ_w	.000	-0.461	.532	[-0.994; .658]	-0.801	.213	[-0.988; -0.330]	.104	.272	[-0.446; .701]	.115	.193	[-0.238; .502]
σ_c^2	1.000	1.088	.550	[.177; 2.231]	1.065	.160	[.764; 1.367]	1.145	.176	[.825; 1.511]	1.106	.119	[.893; 1.339]
σ_w^2	1.000	1.432	.583	[.431; 2.636]	.649	.135	[.399; .937]	1.044	.181	[.704; 1.386]	.991	.120	[.773; 1.207]

Table 73 – Results of the simulation study under $\gamma_c = -0.9$, $\gamma_w = 0$ and $\gamma_b = 0$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	.997	.436	[.156; 1.797]	.965	.158	[.668; 1.285]	.961	.115	[.730; 1.174]	.926	.075	[.774; 1.063]
β_{b1}	2.000	2.265	.613	[1.178; 3.449]	1.781	.194	[1.386; 2.148]	1.943	.184	[1.544; 2.275]	1.773	.127	[1.549; 2.042]
β_{c0}	1.000	.734	.245	[.284; 1.181]	1.031	.095	[.849; 1.204]	.988	.065	[.870; 1.116]	.996	.043	[.920; 1.091]
β_{c1}	2.000	1.835	.225	[1.439; 2.267]	2.009	.085	[1.823; 2.166]	2.014	.068	[1.891; 2.148]	1.972	.039	[1.891; 2.039]
γ_b	.000	.397	.589	[-0.870; .994]	-0.255	.625	[-0.995; .854]	-0.131	.562	[-0.994; .858]	-0.342	.415	[-0.995; .414]
γ_c	-0.900	-0.070	.657	[-0.990; .963]	-0.885	.145	[-0.993; -0.623]	-0.411	.305	[-0.942; .012]	-0.909	.086	[-0.991; -0.715]
γ_w	.000	-0.408	.551	[-0.993; .762]	.024	.314	[-0.703; .615]	.203	.342	[-0.225; .968]	.299	.189	[-0.037; .659]
σ_c^2	1.000	1.245	.731	[.016; 2.704]	1.198	.249	[.751; 1.691]	1.173	.170	[.806; 1.470]	.920	.097	[.743; 1.121]
σ_w^2	1.000	1.550	.767	[.169; 3.016]	1.114	.232	[.684; 1.561]	.911	.178	[.555; 1.248]	1.039	.096	[.864; 1.231]

Table 74 – Results of the simulation study under $\gamma_c = 0$, $\gamma_w = 0.9$ and $\gamma_b = 0$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	.921	.407	[.133; 1.669]	.917	.168	[.604; 1.257]	.871	.114	[.635; 1.091]	1.005	.083	[.845; 1.154]
β_{b1}	2.000	1.947	.503	[1.060; 2.926]	2.192	.254	[1.725; 2.650]	2.053	.187	[1.698; 2.406]	2.012	.125	[1.771; 2.252]
β_{c0}	1.000	1.259	.208	[.882; 1.666]	.988	.092	[.802; 1.150]	1.035	.063	[.911; 1.161]	.996	.043	[.917; 1.077]
β_{c1}	2.000	1.903	.190	[1.541; 2.275]	2.075	.080	[1.917; 2.234]	1.933	.063	[1.810; 2.048]	1.898	.040	[1.822; 1.981]
γ_b	.000	-0.098	.658	[-0.993; .960]	-0.386	.574	[-0.995; .852]	.332	.468	[-0.565; .995]	.162	.413	[-0.523; .985]
γ_c	.000	.378	.611	[-0.847; .993]	.212	.420	[-0.502; .970]	.207	.238	[-0.106; .743]	-0.071	.158	[-0.408; .268]
γ_w	.900	.646	.490	[-0.521; .994]	.175	.392	[-0.509; .965]	.859	.195	[.445; .992]	.863	.117	[.621; .991]
σ_c^2	1.000	.769	.614	[.012; 1.940]	1.021	.247	[.542; 1.476]	1.192	.205	[.813; 1.566]	1.035	.108	[.836; 1.236]
σ_w^2	1.000	1.370	.691	[.149; 2.596]	1.076	.258	[.574; 1.562]	.911	.207	[.520; 1.302]	1.037	.120	[.816; 1.279]

Table 75 – Results of the simulation study under $\gamma_c = 0$, $\gamma_w = -0.9$ and $\gamma_b = 0$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	1.221	.440	[.335; 1.983]	.956	.133	[.699; 1.212]	.982	.129	[.722; 1.215]	.943	.083	[.778; 1.106]
β_{b1}	2.000	2.211	.634	[1.131; 3.549]	1.644	.225	[1.245; 2.129]	2.034	.189	[1.652; 2.394]	1.931	.128	[1.708; 2.212]
β_{c0}	1.000	1.016	.224	[.605; 1.445]	.944	.081	[.796; 1.107]	1.021	.066	[.890; 1.145]	1.026	.042	[.945; 1.111]
β_{c1}	2.000	1.803	.216	[1.312; 2.179]	1.929	.073	[1.799; 2.072]	2.006	.067	[1.874; 2.132]	1.974	.037	[1.907; 2.053]
γ_b	.000	.355	.605	[-0.841; .994]	-0.568	.391	[-0.995; .278]	.433	.464	[-0.464; .994]	-0.312	.436	[-0.994; .440]
γ_c	.000	-0.571	.533	[-0.993; .782]	-0.507	.276	[-0.926; .015]	-0.000	.235	[-0.505; .532]	.107	.173	[-0.129; .519]
γ_w	-0.900	-0.443	.584	[-0.993; .881]	-0.543	.450	[-0.990; .379]	-0.659	.323	[-0.989; .001]	-0.838	.154	[-0.990; -0.531]
σ_c^2	1.000	1.507	.765	[.232; 2.910]	1.323	.216	[.886; 1.731]	1.198	.176	[.884; 1.542]	.901	.101	[.689; 1.090]
σ_w^2	1.000	.988	.644	[.053; 2.215]	.367	.163	[.085; .717]	.967	.183	[.609; 1.305]	.961	.111	[.732; 1.164]

Table 76 – Results of the simulation study under $\gamma_c = 0$, $\gamma_w = 0$ and $\gamma_b = 0.9$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	.961	.392	[.224; 1.764]	.951	.152	[.657; 1.230]	.968	.126	[.717; 1.199]	.970	.083	[.813; 1.138]
β_{b1}	2.000	1.978	.566	[.931; 3.013]	1.747	.211	[1.353; 2.182]	1.995	.181	[1.645; 2.359]	1.864	.120	[1.637; 2.091]
β_{c0}	1.000	1.178	.211	[.793; 1.579]	1.049	.086	[.874; 1.204]	1.002	.067	[.885; 1.145]	1.017	.047	[.937; 1.118]
β_{c1}	2.000	1.796	.207	[1.438; 2.288]	2.047	.080	[1.881; 2.203]	1.968	.067	[1.841; 2.104]	1.958	.041	[1.868; 2.032]
γ_b	.900	.137	.644	[-0.924; .994]	-0.487	.462	[-0.995; .398]	.526	.412	[-0.361; .995]	-0.023	.503	[-0.848; .959]
γ_c	.000	.143	.634	[-0.933; .991]	-0.152	.406	[-0.977; .560]	.312	.294	[-0.049; .915]	-0.276	.330	[-0.951; .149]
γ_w	.000	-0.187	.597	[-0.993; .923]	-0.310	.400	[-0.989; .260]	.521	.376	[-0.020; .984]	.239	.240	[-0.092; .782]
σ_c^2	1.000	.867	.515	[.012; 1.842]	.936	.222	[.575; 1.404]	1.364	.196	[1.008; 1.772]	.886	.118	[.647; 1.109]
σ_w^2	1.000	1.377	.570	[.291; 2.454]	1.013	.222	[.557; 1.418]	1.000	.201	[.632; 1.403]	1.156	.126	[.934; 1.425]

Table 77 – Results of the simulation study under $\gamma_c = 0$, $\gamma_w = 0$ and $\gamma_b = -0.9$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	.973	.392	[.213; 1.739]	.951	.144	[.657; 1.230]	.982	.127	[.750; 1.225]	.970	.081	[.824; 1.125]
β_{b1}	2.000	1.934	.542	[.823; 2.861]	1.746	.205	[1.353; 2.182]	2.000	.189	[1.626; 2.331]	1.882	.122	[1.602; 2.074]
β_{c0}	1.000	1.185	.211	[.782; 1.602]	1.044	.089	[.874; 1.204]	1.004	.071	[.859; 1.128]	1.015	.043	[.934; 1.096]
β_{c1}	2.000	1.808	.202	[1.415; 2.216]	2.043	.080	[1.881; 2.203]	1.967	.064	[1.837; 2.090]	1.963	.039	[1.892; 2.046]
γ_b	-0.900	.161	.665	[-0.952; .994]	-0.452	.511	[-0.995; .398]	.351	.537	[-0.696; .993]	.193	.482	[-0.724; .994]
γ_c	.000	.210	.609	[-0.904; .987]	-0.106	.462	[-0.977; .560]	.330	.309	[-0.075; .938]	-0.348	.365	[-0.974; .085]
γ_w	.000	-0.246	.587	[-0.991; .919]	-0.374	.397	[-0.989; .260]	.532	.372	[-0.037; .989]	.266	.242	[-0.105; .715]
σ_c^2	1.000	.893	.490	[.017; 1.845]	.918	.220	[.575; 1.404]	1.367	.201	[1.010; 1.776]	.874	.116	[.655; 1.115]
σ_w^2	1.000	1.369	.600	[.237; 2.561]	1.034	.232	[.557; 1.418]	1.004	.201	[.624; 1.403]	1.181	.135	[.935; 1.445]

Table 78 – Results of the simulation study under $\gamma_c = 0$, $\gamma_w = 0.9$ and $\gamma_b = 0.9$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	2.158	.648	[.983; 3.443]	.658	.126	[.415; .906]	1.040	.126	[.796; 1.289]	.998	.081	[.849; 1.152]
β_{b1}	2.000	4.098	1.058	[2.038; 6.049]	1.562	.209	[1.193; 2.009]	2.107	.185	[1.750; 2.472]	2.038	.129	[1.772; 2.268]
β_{c0}	1.000	.900	.226	[.463; 1.337]	1.044	.090	[.874; 1.212]	1.000	.060	[.892; 1.112]	.994	.045	[.904; 1.080]
β_{c1}	2.000	2.078	.227	[1.679; 2.569]	1.980	.089	[1.802; 2.143]	2.047	.060	[1.933; 2.169]	2.053	.042	[1.977; 2.140]
γ_b	.900	-0.190	.659	[-0.995; .949]	-0.583	.441	[-0.995; .439]	.575	.350	[-0.128; .995]	.143	.431	[-0.501; .987]
γ_c	.000	.034	.602	[-0.986; .941]	.822	.180	[.433; .992]	-0.278	.324	[-0.939; .103]	-0.340	.267	[-0.856; .044]
γ_w	.900	-0.226	.542	[-0.991; .959]	.232	.472	[-0.626; .985]	.836	.176	[.476; .991]	.849	.139	[.561; .990]
σ_c^2	1.000	.530	.462	[.003; 1.487]	1.365	.238	[.901; 1.809]	.945	.138	[.690; 1.214]	.905	.108	[.682; 1.097]
σ_w^2	1.000	1.990	.646	[.751; 3.240]	.685	.196	[.320; 1.077]	.836	.143	[.562; 1.105]	1.134	.125	[.903; 1.377]

Table 79 – Results of the simulation study under $\gamma_c = 0$, $\gamma_w = 0.9$ and $\gamma_b = -0.9$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	.952	.455	[.155; 1.880]	.796	.147	[.510; 1.070]	1.040	.136	[.796; 1.289]	1.011	.085	[.850; 1.186]
β_{b1}	2.000	2.465	.597	[1.332; 3.663]	1.746	.211	[1.303; 2.121]	2.107	.193	[1.750; 2.472]	1.940	.129	[1.690; 2.190]
β_{c0}	1.000	1.150	.205	[.732; 1.545]	.996	.092	[.820; 1.176]	1.000	.065	[.892; 1.112]	1.005	.045	[.916; 1.086]
β_{c1}	2.000	1.995	.166	[1.652; 2.308]	2.038	.079	[1.891; 2.199]	2.047	.060	[1.933; 2.169]	1.998	.045	[1.918; 2.089]
γ_b	-0.900	-0.044	.696	[-0.995; .957]	.330	.601	[-0.900; .995]	.575	.296	[-0.128; .995]	-0.307	.511	[-0.995; .760]
γ_c	.000	.100	.569	[-0.881; .983]	.188	.593	[-0.918; .989]	-0.278	.226	[-0.939; .103]	-0.110	.169	[-0.475; .167]
γ_w	.900	.092	.477	[-0.991; .906]	.491	.356	[-0.040; .989]	.836	.152	[.476; .991]	.847	.135	[.565; .991]
σ_c^2	1.000	.301	.284	[.007; .895]	.745	.228	[.304; 1.194]	.945	.172	[.690; 1.214]	1.031	.111	[.813; 1.250]
σ_w^2	1.000	1.666	.494	[.782; 2.641]	1.409	.260	[.901; 1.898]	.836	.181	[.562; 1.105]	1.073	.123	[.820; 1.300]

Table 80 – Results of the simulation study under $\gamma_c = 0$, $\gamma_w = -0.9$ and $\gamma_b = 0.9$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	2.032	.661	[.825; 3.321]	1.382	.147	[1.027; 1.834]	.851	.136	[.644; 1.073]	.944	.087	[.783; 1.123]
β_{b1}	2.000	3.649	.967	[1.828; 5.553]	2.186	.211	[1.719; 2.699]	1.919	.193	[1.586; 2.231]	1.965	.126	[1.730; 2.234]
β_{c0}	1.000	1.183	.241	[.745; 1.684]	1.023	.092	[.862; 1.183]	1.048	.065	[.913; 1.161]	1.024	.044	[.940; 1.110]
β_{c1}	2.000	1.844	.216	[1.431; 2.281]	2.075	.079	[1.897; 2.238]	1.968	.060	[1.853; 2.093]	2.050	.043	[1.967; 2.136]
γ_b	.900	-0.058	.673	[-0.995; .953]	.649	.601	[-0.068; .995]	-0.136	.296	[-0.995; .664]	.363	.473	[-0.535; .995]
γ_c	.000	-0.154	.642	[-0.991; .935]	.329	.593	[-0.360; .978]	.286	.226	[-0.143; .814]	-0.020	.124	[-0.282; .286]
γ_w	-0.900	-0.710	.368	[-0.994; .007]	-0.800	.356	[-0.992; -0.332]	-0.830	.152	[-0.993; -0.437]	-0.868	.126	[-0.993; -0.606]
σ_c^2	1.000	.784	.629	[.021; 2.106]	.752	.228	[.390; 1.092]	1.112	.172	[.845; 1.412]	.950	.097	[.754; 1.135]
σ_w^2	1.000	2.165	.789	[.595; 3.629]	1.073	.260	[.621; 1.476]	1.004	.181	[.694; 1.310]	.983	.112	[.770; 1.209]

Table 81 – Results of the simulation study under $\gamma_c = 0$, $\gamma_w = -0.9$ and $\gamma_b = -0.9$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	1.372	.452	[.480; 2.256]	.798	.163	[.482; 1.100]	1.118	.128	[.644; 1.073]	1.120	.090	[.950; 1.297]
β_{b1}	2.000	1.959	.510	[.960; 2.985]	1.622	.175	[1.278; 1.961]	1.984	.189	[1.586; 2.231]	2.068	.113	[1.868; 2.311]
β_{c0}	1.000	.870	.231	[.430; 1.298]	.946	.093	[.765; 1.116]	1.099	.065	[.913; 1.161]	1.008	.044	[.929; 1.108]
β_{c1}	2.000	2.024	.226	[1.611; 2.471]	1.954	.080	[1.796; 2.107]	2.071	.059	[1.853; 2.093]	2.021	.040	[1.948; 2.101]
γ_b	-0.900	-0.089	.675	[-0.995; .968]	.444	.530	[-0.747; .995]	.300	.484	[-0.995; .664]	.845	.169	[.507; .995]
γ_c	.000	.132	.634	[-0.939; .990]	.632	.350	[-0.003; .992]	.113	.243	[-0.143; .814]	.043	.182	[-0.335; .483]
γ_w	-0.900	-0.787	.290	[-0.993; -0.033]	-0.861	.157	[-0.993; -0.520]	-0.818	.202	[-0.993; -0.437]	-0.810	.145	[-0.987; -0.539]
σ_c^2	1.000	.771	.405	[.094; 1.545]	.772	.173	[.457; 1.151]	1.082	.156	[.845; 1.412]	.807	.107	[.592; 1.006]
σ_w^2	1.000	1.841	.627	[.715; 3.013]	1.319	.228	[.891; 1.754]	.957	.173	[.694; 1.310]	1.241	.131	[.968; 1.483]

Table 82 – Results of the simulation study under $\gamma_c = 0.9$, $\gamma_w = 0$ and $\gamma_b = 0.9$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	.803	.357	[.480; 2.256]	1.017	.169	[.666; 1.306]	1.108	.140	[.853; 1.402]	.975	.087	[.807; 1.128]
β_{b1}	2.000	2.341	.580	[.960; 2.985]	1.944	.266	[1.463; 2.507]	2.220	.178	[1.889; 2.569]	1.777	.114	[1.586; 2.015]
β_{c0}	1.000	.775	.202	[.430; 1.298]	.966	.094	[.784; 1.152]	1.087	.070	[.952; 1.215]	1.019	.046	[.931; 1.113]
β_{c1}	2.000	2.262	.192	[1.611; 2.471]	2.052	.084	[1.900; 2.226]	1.987	.056	[1.878; 2.086]	2.050	.041	[1.964; 2.119]
γ_b	.900	.473	.549	[-0.995; .968]	-0.054	.620	[-0.994; .935]	.737	.283	[.157; .995]	.786	.211	[.369; .995]
γ_c	.900	.366	.594	[-0.939; .990]	.797	.243	[.218; .990]	.709	.252	[.107; .991]	.835	.129	[.591; .992]
γ_w	.000	.348	.613	[-0.993; -0.033]	-0.146	.416	[-0.955; .566]	-0.059	.287	[-0.737; .524]	.081	.186	[-0.259; .491]
σ_c^2	1.000	1.056	.513	[.094; 1.545]	1.158	.256	[.684; 1.656]	1.257	.204	[.853; 1.665]	1.201	.121	[.920; 1.401]
σ_w^2	1.000	.882	.495	[.715; 3.013]	.767	.227	[.306; 1.176]	1.045	.197	[.693; 1.411]	.909	.117	[.675; 1.125]

Table 83 – Results of the simulation study under $\gamma_c = 0.9$, $\gamma_w = 0$ and $\gamma_b = -0.9$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	1.485	.423	[.674; 2.274]	1.262	.175	[.933; 1.601]	1.183	.127	[.853; 1.402]	1.065	.081	[.894; 1.205]
β_{b1}	2.000	2.000	.621	[.878; 3.231]	1.774	.224	[1.361; 2.220]	2.181	.202	[1.889; 2.569]	2.136	.124	[1.897; 2.388]
β_{c0}	1.000	1.632	.233	[1.147; 2.051]	.895	.092	[.727; 1.081]	1.091	.067	[.952; 1.215]	1.022	.045	[.926; 1.103]
β_{c1}	2.000	1.907	.210	[1.508; 2.320]	2.021	.089	[1.861; 2.204]	1.994	.059	[1.878; 2.086]	2.025	.041	[1.949; 2.105]
γ_b	-0.900	.124	.652	[-0.968; .992]	.069	.611	[-0.930; .994]	-0.695	.313	[.157; .995]	-0.701	.346	[-0.995; .101]
γ_c	.900	.277	.630	[-0.916; .993]	.459	.363	[-0.030; .984]	.797	.186	[.107; .991]	.928	.069	[.789; .991]
γ_w	.000	.154	.639	[-0.958; .992]	.292	.502	[-0.687; .986]	-0.333	.355	[-0.737; .524]	.073	.148	[-0.158; .436]
σ_c^2	1.000	1.028	.551	[.050; 1.997]	1.348	.268	[.848; 1.863]	1.439	.186	[.853; 1.665]	1.073	.118	[.843; 1.290]
σ_w^2	1.000	1.471	.616	[.378; 2.712]	.798	.257	[.354; 1.322]	.895	.175	[.693; 1.411]	1.027	.111	[.835; 1.258]

Table 84 – Results of the simulation study under $\gamma_c = 0.9$, $\gamma_w = 0$ and $\gamma_b = 0.9$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	.849	.348	[.095; 1.449]	1.075	.183	[.757; 1.478]	1.063	.132	[.799; 1.309]	.962	.086	[.813; 1.147]
β_{b1}	2.000	2.364	.599	[1.265; 3.592]	2.205	.261	[1.693; 2.685]	2.163	.180	[1.806; 2.484]	1.756	.112	[1.536; 1.963]
β_{c0}	1.000	1.127	.197	[.777; 1.531]	1.067	.082	[.908; 1.225]	.988	.064	[.872; 1.119]	1.000	.047	[.919; 1.102]
β_{c1}	2.000	2.295	.197	[1.905; 2.698]	1.982	.086	[1.801; 2.138]	1.894	.053	[1.789; 1.989]	1.978	.041	[1.898; 2.053]
γ_b	.900	.560	.467	[-0.460; .995]	.600	.409	[-0.295; .994]	.617	.381	[-0.205; .995]	.788	.215	[.321; .995]
γ_c	-0.900	-0.227	.605	[-0.992; .931]	-0.662	.329	[-0.989; .005]	-0.911	.097	[-0.993; -0.707]	-0.888	.101	[-0.991; -0.671]
γ_w	.000	-0.131	.619	[-0.990; .924]	-0.032	.370	[-0.691; .964]	-0.050	.225	[-0.554; .362]	.414	.217	[-0.004; .821]
σ_c^2	1.000	1.087	.481	[.218; 2.140]	.865	.217	[.454; 1.282]	1.059	.161	[.744; 1.368]	1.007	.101	[.842; 1.245]
σ_w^2	1.000	.792	.447	[.022; 1.626]	.997	.210	[.577; 1.424]	.980	.151	[.684; 1.247]	.957	.102	[.754; 1.149]

Table 85 – Results of the simulation study under $\gamma_c = -0.9$, $\gamma_w = 0$ and $\gamma_b = -0.9$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	1.511	.421	[.735; 2.351]	1.351	.172	[1.013; 1.668]	1.156	.119	[.799; 1.309]	1.010	.079	[.869; 1.182]
β_{b1}	2.000	2.097	.685	[.857; 3.431]	1.966	.244	[1.529; 2.453]	2.148	.188	[1.806; 2.484]	2.051	.125	[1.847; 2.338]
β_{c0}	1.000	1.221	.236	[.759; 1.653]	1.159	.084	[1.009; 1.322]	.998	.065	[.872; 1.119]	1.002	.046	[.916; 1.091]
β_{c1}	2.000	2.201	.200	[1.795; 2.563]	1.891	.083	[1.730; 2.037]	1.914	.055	[1.789; 1.989]	1.955	.039	[1.878; 2.025]
γ_b	-0.900	.080	.705	[-0.966; .995]	-0.413	.565	[-0.995; .796]	-0.561	.403	[-0.205; .995]	-0.675	.309	[-0.995; -0.088]
γ_c	-0.900	-0.536	.465	[-0.992; .294]	-0.604	.413	[-0.990; .150]	-0.797	.214	[-0.993; -0.707]	-0.749	.200	[-0.990; -0.393]
γ_w	.000	.578	.488	[-0.475; .993]	-0.013	.366	[-0.942; .662]	-0.152	.279	[-0.554; .362]	.212	.194	[-0.079; .616]
σ_c^2	1.000	1.376	.563	[.391; 2.505]	.753	.184	[.389; 1.106]	1.096	.174	[.744; 1.368]	.996	.107	[.802; 1.211]
σ_w^2	1.000	1.153	.565	[.231; 2.308]	.895	.193	[.531; 1.273]	.966	.168	[.684; 1.247]	.968	.106	[.766; 1.173]

Table 86 – Results of the simulation study under $\gamma_c = 0.9$, $\gamma_w = 0.9$ and $\gamma_b = 0$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	.835	.373	[.209; 1.671]	1.351	.167	[.643; 1.275]	.834	.123	[.600; 1.085]	.949	.079	[.794; 1.113]
β_{b1}	2.000	2.196	.551	[1.243; 3.208]	1.966	.261	[1.772; 2.787]	1.995	.172	[1.641; 2.308]	1.961	.125	[1.736; 2.205]
β_{c0}	1.000	1.446	.233	[1.006; 1.924]	1.159	.090	[.909; 1.243]	1.010	.061	[.886; 1.123]	.996	.046	[.918; 1.080]
β_{c1}	2.000	2.054	.228	[1.633; 2.522]	1.891	.069	[1.987; 2.254]	1.992	.055	[1.879; 2.101]	1.982	.039	[1.900; 2.054]
γ_b	.000	.014	.636	[-0.962; .992]	-0.413	.357	[-0.995; .032]	-0.096	.561	[-0.990; .894]	.422	.309	[-0.214; .995]
γ_c	.900	.461	.559	[-0.775; .994]	-0.604	.132	[.663; .992]	.877	.103	[.670; .990]	.848	.200	[.637; .992]
γ_w	.900	.137	.667	[-0.953; .993]	-0.013	.102	[.766; .992]	.814	.200	[.415; .991]	.756	.194	[.463; .983]
σ_c^2	1.000	1.806	.712	[.647; 3.338]	.753	.231	[.360; 1.253]	1.113	.209	[.649; 1.467]	1.067	.107	[.854; 1.254]
σ_w^2	1.000	.953	.593	[.018; 2.160]	.895	.279	[.632; 1.662]	.874	.200	[.521; 1.281]	.822	.106	[.620; 1.042]

Table 87 – Results of the simulation study under $\gamma_c = 0.9$, $\gamma_w = -0.9$ and $\gamma_b = 0$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	1.443	.451	[.209; 1.671]	1.000	.145	[.643; 1.275]	1.004	.123	[.790; 1.248]	.951	.081	[.797; 1.108]
β_{b1}	2.000	2.506	.674	[1.243; 3.208]	1.718	.229	[1.772; 2.787]	2.126	.184	[1.795; 2.501]	1.959	.117	[1.729; 2.169]
β_{c0}	1.000	1.324	.235	[1.006; 1.924]	1.034	.081	[.909; 1.243]	.995	.063	[.871; 1.128]	1.025	.046	[.938; 1.111]
β_{c1}	2.000	1.975	.228	[1.633; 2.522]	2.020	.077	[1.987; 2.254]	1.991	.060	[1.874; 2.110]	2.077	.039	[2.006; 2.148]
γ_b	.000	.336	.610	[-0.962; .992]	-0.696	.273	[-0.995; .032]	.609	.388	[-0.222; .995]	-0.104	.474	[-0.936; .691]
γ_c	.900	.275	.663	[-0.775; .994]	.057	.205	[.663; .992]	-0.130	.402	[-0.960; .574]	.936	.062	[.820; .993]
γ_w	-0.900	-0.198	.618	[-0.953; .993]	-0.576	.408	[.766; .992]	.281	.336	[-0.208; .946]	-0.848	.153	[-0.989; -0.508]
σ_c^2	1.000	1.095	.590	[.647; 3.338]	1.282	.205	[.360; 1.253]	.817	.155	[.552; 1.134]	1.082	.105	[.872; 1.290]
σ_w^2	1.000	1.498	.595	[.018; 2.160]	.482	.173	[.632; 1.662]	1.048	.164	[.750; 1.368]	1.021	.100	[.845; 1.238]

Table 88 – Results of the simulation study under $\gamma_c = -0.9$, $\gamma_w = 0.9$ and $\gamma_b = 0$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	1.443	.451	[.209; 1.671]	1.000	.145	[.643; 1.275]	1.004	.123	[.790; 1.248]	.951	.081	[.797; 1.108]
β_{b1}	2.000	2.506	.674	[1.243; 3.208]	1.718	.229	[1.772; 2.787]	2.126	.184	[1.795; 2.501]	1.959	.117	[1.729; 2.169]
β_{c0}	1.000	1.324	.235	[1.006; 1.924]	1.034	.081	[.909; 1.243]	.995	.063	[.871; 1.128]	1.025	.046	[.938; 1.111]
β_{c1}	2.000	1.975	.228	[1.633; 2.522]	2.020	.077	[1.987; 2.254]	1.991	.060	[1.874; 2.110]	2.077	.039	[2.006; 2.148]
γ_b	.000	.336	.610	[-0.962; .992]	-0.696	.273	[-0.995; .032]	.609	.388	[-0.222; .995]	-0.104	.474	[-0.936; .691]
γ_c	-0.900	.275	.663	[-0.775; .994]	.057	.205	[.663; .992]	-0.130	.402	[-0.960; .574]	.936	.062	[.820; .993]
γ_w	.900	-0.198	.618	[-0.953; .993]	-0.576	.408	[.766; .992]	.281	.336	[-0.208; .946]	-0.848	.153	[-0.989; -0.508]
σ_c^2	1.000	1.095	.590	[.647; 3.338]	1.282	.205	[.360; 1.253]	.817	.155	[.552; 1.134]	1.082	.105	[.872; 1.290]
σ_w^2	1.000	1.498	.595	[.018; 2.160]	.482	.173	[.632; 1.662]	1.048	.164	[.750; 1.368]	1.021	.100	[.845; 1.238]

Table 89 – Results of the simulation study under $\gamma_c = -0.9$, $\gamma_w = -0.9$ and $\gamma_b = 0$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	1.331	.431	[.563; 2.181]	.973	.140	[.721; 1.251]	.932	.119	[.707; 1.164]	.934	.078	[.798; 1.101]
β_{b1}	2.000	2.513	.663	[1.448; 3.899]	1.700	.218	[1.248; 2.109]	1.900	.184	[1.553; 2.257]	1.928	.122	[1.680; 2.140]
β_{c0}	1.000	.608	.223	[.222; 1.089]	.932	.091	[.746; 1.094]	.995	.061	[.861; 1.098]	1.012	.042	[.928; 1.089]
β_{c1}	2.000	1.825	.216	[1.399; 2.225]	1.922	.075	[1.794; 2.092]	2.016	.051	[1.921; 2.119]	1.986	.036	[1.917; 2.057]
γ_b	.000	.081	.680	[-0.976; .993]	-0.501	.411	[-0.995; .318]	-0.241	.507	[-0.990; .742]	-0.325	.375	[-0.985; .397]
γ_c	-0.900	-0.646	.458	[-0.992; .465]	-0.913	.079	[-0.992; -0.754]	-0.839	.116	[-0.991; -0.625]	-0.837	.132	[-0.989; -0.563]
γ_w	-0.900	-0.614	.475	[-0.993; .455]	-0.504	.455	[-0.984; .311]	-0.819	.177	[-0.989; -0.455]	-0.870	.099	[-0.991; -0.675]
σ_c^2	1.000	1.444	.709	[.230; 2.872]	1.593	.219	[1.224; 2.057]	1.283	.167	[.930; 1.603]	.865	.100	[.694; 1.074]
σ_w^2	1.000	1.142	.661	[.065; 2.405]	.397	.146	[.130; .683]	.728	.150	[.487; 1.084]	.980	.104	[.772; 1.167]

Table 90 – Results of the simulation study under $\gamma_c = 0.9$, $\gamma_w = 0.9$ and $\gamma_b = -0.9$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	1.323	.385	[.576; 2.019]	1.221	.174	[.828; 1.516]	.910	.111	[.701; 1.130]	.854	.071	[.723; 1.000]
β_{b1}	2.000	2.075	.585	[.989; 3.282]	2.053	.259	[1.559; 2.530]	2.077	.182	[1.776; 2.468]	1.915	.127	[1.692; 2.174]
β_{c0}	1.000	1.541	.201	[1.187; 1.936]	.779	.083	[.619; .951]	1.026	.065	[.886; 1.146]	.997	.045	[.907; 1.076]
β_{c1}	2.000	2.016	.212	[1.619; 2.416]	2.080	.077	[1.935; 2.247]	2.034	.048	[1.944; 2.132]	1.989	.038	[1.915; 2.065]
γ_b	-0.900	.079	.676	[-0.963; .995]	-0.253	.582	[-0.995; .834]	-0.622	.300	[-0.994; -0.016]	-0.659	.344	[-0.995; .136]
γ_c	.900	.411	.602	[-0.925; .993]	.730	.289	[.010; .991]	.882	.099	[.680; .992]	.803	.120	[.590; .990]
γ_w	.900	.280	.622	[-0.894; .992]	.519	.420	[-0.147; .984]	.865	.151	[.573; .991]	.814	.142	[.539; .980]
σ_c^2	1.000	1.369	.567	[.367; 2.502]	.931	.244	[.503; 1.419]	1.394	.197	[1.006; 1.796]	1.017	.107	[.821; 1.227]
σ_w^2	1.000	.982	.596	[.051; 2.159]	.885	.240	[.481; 1.386]	.687	.170	[.350; 1.007]	.872	.110	[.669; 1.090]

Table 91 – Results of the simulation study under $\gamma_c = 0.9$, $\gamma_w = -0.9$ and $\gamma_b = 0.9$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	.731	.319	[.084; 1.327]	1.015	.167	[.700; 1.355]	1.033	.126	[.769; 1.255]	.891	.082	[.726; 1.051]
β_{b1}	2.000	1.383	.343	[.736; 2.004]	1.884	.230	[1.417; 2.285]	1.966	.160	[1.676; 2.268]	1.945	.108	[1.755; 2.158]
β_{c0}	1.000	.816	.207	[.397; 1.192]	.971	.084	[.787; 1.120]	1.076	.069	[.948; 1.203]	1.027	.045	[.946; 1.119]
β_{c1}	2.000	1.952	.196	[1.610; 2.328]	2.100	.087	[1.936; 2.266]	2.036	.060	[1.915; 2.145]	2.095	.040	[2.014; 2.175]
γ_b	.900	-0.225	.608	[-0.995; .899]	.348	.589	[-0.791; .995]	.489	.454	[-0.472; .995]	.840	.178	[.432; .995]
γ_c	.900	.049	.644	[-0.939; .986]	.726	.312	[-0.002; .990]	.757	.240	[.273; .991]	.933	.067	[.786; .992]
γ_w	-0.900	.105	.544	[-0.905; .993]	-0.531	.367	[-0.984; .048]	-0.682	.286	[-0.989; -0.051]	-0.837	.161	[-0.992; -0.509]
σ_c^2	1.000	.559	.410	[.007; 1.369]	.921	.202	[.508; 1.320]	1.265	.180	[.888; 1.588]	1.108	.099	[.905; 1.284]
σ_w^2	1.000	1.666	.582	[.540; 2.839]	.836	.198	[.433; 1.212]	1.028	.171	[.727; 1.381]	.992	.096	[.817; 1.187]

Table 92 – Results of the simulation study under $\gamma_c = -0.9$, $\gamma_w = 0.9$ and $\gamma_b = 0.9$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	1.468	.456	[.715; 2.443]	.945	.168	[.700; 1.355]	.931	.116	[.719; 1.176]	1.029	.082	[.870; 1.188]
β_{b1}	2.000	1.886	.555	[.835; 2.966]	2.158	.277	[1.417; 2.285]	1.963	.165	[1.650; 2.314]	2.047	.108	[1.784; 2.273]
β_{c0}	1.000	.993	.203	[.580; 1.399]	1.046	.077	[.787; 1.120]	.933	.068	[.792; 1.058]	.982	.045	[.885; 1.062]
β_{c1}	2.000	1.916	.205	[1.531; 2.324]	1.933	.082	[1.936; 2.266]	1.965	.055	[1.868; 2.075]	1.905	.040	[1.826; 1.980]
γ_b	.900	.493	.529	[-0.751; .995]	.662	.366	[-0.791; .995]	.270	.493	[-0.589; .995]	.761	.178	[.353; .995]
γ_c	-0.900	-0.423	.505	[-0.991; .524]	-0.271	.385	[-0.002; .990]	-0.844	.176	[-0.991; -0.501]	-0.912	.067	[-0.992; -0.758]
γ_w	.900	.048	.614	[-0.941; .987]	.311	.489	[-0.984; .048]	.764	.232	[.254; .991]	.867	.161	[.635; .990]
σ_c^2	1.000	1.077	.483	[.082; 1.968]	.879	.202	[.508; 1.320]	1.080	.144	[.814; 1.363]	1.046	.099	[.867; 1.236]
σ_w^2	1.000	.864	.483	[.039; 1.790]	.661	.192	[.433; 1.212]	1.075	.156	[.805; 1.401]	1.087	.096	[.890; 1.266]

Table 93 – Results of the simulation study under $\gamma_c = 0.9$, $\gamma_w = -0.9$ and $\gamma_b = -0.9$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	1.339	.423	[.611; 2.254]	1.222	.165	[.919; 1.534]	1.159	.129	[.923; 1.411]	1.037	.082	[.884; 1.206]
β_{b1}	2.000	2.277	.681	[1.136; 3.753]	1.824	.228	[1.375; 2.262]	2.177	.188	[1.826; 2.533]	2.215	.132	[1.949; 2.442]
β_{c0}	1.000	1.183	.227	[.764; 1.647]	.971	.083	[.824; 1.146]	1.084	.063	[.963; 1.200]	1.032	.045	[.930; 1.108]
β_{c1}	2.000	1.719	.190	[1.347; 2.060]	2.080	.083	[1.931; 2.256]	2.040	.059	[1.935; 2.158]	2.050	.040	[1.981; 2.137]
γ_b	-0.900	-0.182	.666	[-0.995; .959]	-0.505	.510	[-0.995; .574;	-0.151	.572	[-0.993; .915]	-0.788	.189	[-0.995; -0.405]
γ_c	.900	.058	.602	[-0.942; .981]	.440	.430	[-0.219; .989]	.794	.224	[.297; .993]	.887	.103	[.670; .992]
γ_w	-0.900	.105	.509	[-0.992; .927]	-0.263	.392	[-0.970; .303]	-0.584	.296	[-0.978; .000]	-0.636	.283	[-0.986; -0.110]
σ_c^2	1.000	.445	.383	[.009; 1.263]	.884	.218	[.486; 1.302]	1.136	.164	[.827; 1.437]	.970	.119	[.741; 1.195]
σ_w^2	1.000	1.917	.574	[.931; 3.082]	.881	.224	[.456; 1.304]	1.158	.175	[.806; 1.479]	1.122	.123	[.903; 1.370]

Table 94 – Results of the simulation study under $\gamma_c = -0.9$, $\gamma_w = 0.9$ and $\gamma_b = -0.9$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	1.339	.423	[.611; 2.254]	1.222	.165	[.919; 1.534]	1.159	.129	[.923; 1.411]	1.037	.082	[.884; 1.206]
β_{b1}	2.000	2.277	.681	[1.136; 3.753]	1.824	.228	[1.375; 2.262]	2.177	.188	[1.826; 2.533]	2.215	.132	[1.949; 2.442]
β_{c0}	1.000	1.183	.227	[.764; 1.647]	.971	.083	[.824; 1.146]	1.084	.063	[.963; 1.200]	1.032	.045	[.930; 1.108]
β_{c1}	2.000	1.719	.190	[1.347; 2.060]	2.080	.083	[1.931; 2.256]	2.040	.059	[1.935; 2.158]	2.050	.040	[1.981; 2.137]
γ_b	-0.900	-0.182	.666	[-0.995; .959]	-0.505	.510	[-0.995; .574]	-0.151	.572	[-0.993; .915]	-0.788	.189	[-0.995; -0.405]
γ_c	-0.900	.058	.602	[-0.942; .981]	.440	.430	[-0.219; .989]	.794	.224	[.297; .993]	.887	.103	[.670; .992]
γ_w	.900	.105	.509	[-0.992; .927]	-0.263	.392	[-0.970; .303]	-0.584	.296	[-0.978; .000]	-0.636	.283	[-0.986; -0.110]
σ_c^2	1.000	.445	.383	[.009; 1.263]	.884	.218	[.486; 1.302]	1.136	.164	[.827; 1.437]	.970	.119	[.741; 1.195]
σ_w^2	1.000	1.917	.574	[.931; 3.082]	.881	.224	[.456; 1.304]	1.158	.175	[.806; 1.479]	1.122	.123	[.903; 1.370]

Table 95 – Results of the simulation study under $\gamma_c = -0.9$, $\gamma_w = -0.9$ and $\gamma_b = 0.9$.

		50			250			500			1000		
Par	Real	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD	Est	SD	95% HPD
β_{b0}	1.000	.514	.269	[-0.049; .987]	1.048	.169	[.765; 1.392]	.957	.112	[.718; 1.139]	.900	.083	[.737; 1.063]
β_{b1}	2.000	1.294	.353	[.566; 1.945]	2.033	.244	[1.518; 2.451]	1.822	.172	[1.520; 2.170]	1.930	.109	[1.735; 2.152]
β_{c0}	1.000	.658	.235	[.202; 1.099]	1.236	.077	[1.089; 1.389]	.975	.061	[.865; 1.101]	1.014	.045	[.933; 1.105]
β_{c1}	2.000	1.900	.183	[1.544; 2.247]	1.918	.065	[1.795; 2.043]	1.940	.046	[1.850; 2.026]	1.989	.035	[1.920; 2.052]
γ_b	.900	.298	.603	[-0.904; .995]	.439	.515	[-0.594; .995]	-0.328	.502	[-0.994; .763]	.636	.244	[.176; .994]
γ_c	-0.900	-0.832	.273	[-0.994; -0.126]	-0.832	.203	[-0.991; -0.404]	-0.952	.038	[-0.993; -0.871]	-0.809	.125	[-0.987; -0.553]
γ_w	-0.900	-0.145	.629	[-0.987; .935]	-0.856	.165	[-0.991; -0.532]	-0.899	.094	[-0.989; -0.705]	-0.885	.094	[-0.991; -0.706]
σ_c^2	1.000	2.235	.678	[1.150; 3.791]	.739	.177	[.389; 1.044]	1.178	.163	[.817; 1.467]	.826	.096	[.603; .997]
σ_w^2	1.000	.540	.404	[.020; 1.315]	.857	.193	[.500; 1.224]	.817	.149	[.547; 1.097]	1.017	.110	[.814; 1.237]