

## BAYESIAN ESTIMATION FOR THE STABLE DISTRIBUTIONS IN THE PRESENCE OF COVARIATES WITH APPLICATIONS IN CLINICAL ISSUES

Jorge Alberto Achcar<sup>1</sup>, Roberto Molina de Souza<sup>2</sup>,  
Daiane Bussola<sup>3</sup> and Fernando A. Moala<sup>4\*</sup>

Received July 22, 2021 / Accepted March 14, 2022

**ABSTRACT.** In this paper we explore a Bayesian approach for stable distributions in presence of covariates. This class of distribution has great flexibility for fitting asymmetric and heavy-tailed empirical data. These models are commonly used for data sets in finance and insurance. In this paper we show that these distributions can also be used to fit clinical data. Since there is not an analytical form for the density probability function which implies in serious difficulties to obtain the maximum likelihood estimators for the parameters, we use Bayesian methods with data augmentation techniques to get the inferences of interest. In this study we also discuss the choice of different prior distributions for the parameters considering regression models for the location and scale parameters of the stable distribution. We use MCMC (Markov Chain Monte Carlo) algorithms to generate samples from the posterior distributions in order to evaluate the point and interval estimators. A great simplification is obtained using the OpenBugs software. Two real data examples illustrate the applicability of the proposed modeling approach.

**Keywords:** table distributions, Bayesian approach, regression models, prior distributions, MCMC methods.

### 1 INTRODUCTION

In many areas of applications, the usual assumption of normality of the data needed in many statistical models is not satisfied in practice, even using data transformations to improve the symmetry of the data. The assumption of normality is essential for the use of traditional techniques such as ANOVA (analysis of variance) models, linear regression models, hypothesis tests for comparison of means, medians or variances, especially with small sample sizes. In this situation, usually

---

\*Corresponding author

<sup>1</sup>University of São Paulo, Department of Public Health, Av. Bandeirantes, 3900, Monte Alegre, 14049-900 Ribeirão Preto, SP, Brazil – E-mail: achcar@fmrp.usp.br – <http://orcid.org/0000-0002-9868-9453>

<sup>2</sup>Federal Technological University of Paraná, Department of Mathematics, Av. Alberto Carazzai, 1640, Centro, 86300-000 Cornélio Procopio, PR, Brazil – E-mail: rmolinasouza@utfpr.edu.br – <http://orcid.org/0000-0002-0638-3650>

<sup>3</sup>University of São Paulo, Department of Public Health, Av. Bandeirantes, 3900, Monte Alegre, 14049-900 Ribeirão Preto, SP, Brazil – E-mail: dpsampaio@gmail.com – <http://orcid.org/0000-0002-0362-2871>

<sup>4</sup>São Paulo State University, Department of Statistics, R. Roberto Símonsens, 305, Centro Educacional, 19060-900 Presidente Prudente, SP, Brazil – E-mail: f.moala@unesp.br – <http://orcid.org/0000-0002-2445-0407>

the data analyst uses existing non-parametrical methods, but this approach usually gives poor statistical inference results, and in general it is not possible to get predictions, among many other problems. A recent alternative that has been explored in the literature given the advancement of computational methods and better computers is the use of super models that generalize the normal distribution allowing the adjustment to asymmetric data, multimodal data, among others. In this case we can cite the family of stable distributions that have already been used in areas such as economics and finance. The family of stable distributions arises from a generalized version of the central limit theorem (Gnedenko & Kolmogorov, 1954) where the Gaussian distribution is a particular case of this family of distributions. The stable distribution has great flexibility to fit data with very different shapes, allowing to fit asymmetric and heavy-tailed data, which makes it suitable for modeling in different research areas such as complex systems in physics, engineering, biology, sociology and other fields. Several analytical properties of this class of stable distributions and several practical applications can be seen in Zolotarev (1986).

A major difficulty associated with stable distributions is that, in general, there is no simple and closed formula for the probability density functions (pdf), making it difficult to apply well-known estimation methods, such as parameter estimation using maximum likelihood methods. In theory, the pdf of a stable distribution could be obtained numerically from an integration based on the inversion formula or using a characteristic function-based estimation method, but these techniques are often complex and difficult to be used in practice.

Borak et al. (2005) and Kateregga et al. (2017) reviewed some classical approaches proposed to estimate the parameters of stable distributions and investigated their performance from simulation studies. They observed that the inferences obtained by these approaches can present several computational difficulties and often inaccurate estimates of model parameters.

In the last decades, Bayesian inference methods have been considered as an efficient and robust alternative approach to data analysis, especially using more complex models in general with several parameters and with great flexibility to be fitted by the data. The literature presents several studies considering the use of Bayesian estimation methods to get the estimators for the parameters of stable distributions. In this direction, one of the first papers written to explore the use of Bayesian methods for stable distribution was introduced by Buckle (1995) proposing a Bayesian analysis using Markov Chain Monte Carlo (MCMC) methods and data augmentation techniques (introduction of unobserved latent variables) that made it possible to get precise inferences of interest to the stable distribution. Lombardi (2004) introduced a novel approach for Bayesian inference in the setting of stable distributions that resorts to a Fast Fourier Transform (FFT) of the characteristic function in order to approximate the likelihood function; the posterior distributions of the parameters are then produced via a random walk MCMC method. Oral and Erdemir (2012) compared the FFT with their Bayesian approach using Metropolis random walk chain and direct numerical integration.

Achcar et al. (2013a) studied computational aspects for the Bayesian analysis involving stable distributions. Following in the same direction, Achcar et al. (2013b) presented some robustness aspects of linear regression models in the presence of outliers or discordant observations consid-

ering the use of stable distributions for the response in place of the usual normality assumption—where the main aspects of the methodology, as well as two examples of implementation, were presented: one using data from the industrial area and other from simulated data. The authors elaborated the analysis and discussion including measures of goodness of fit of the models. The influence of outliers in the model fit was also discussed in this publication.

Lemke et al. (2015) considered a fully Bayesian inference for  $\alpha$ -stable distributions using a Poisson series representation, and more recently Karling et al. (2021) provided a Bayesian algorithm that makes use only of the power series representation.

The main goal of this paper is to carry out a Bayesian inference with data augmentation techniques and MCMC (Markov Chain Monte Carlo) methods proposed by Buckle (1995) to analyse data under stable distributions in presence of covariates. We provide major applications of the stable distributions in the medical field involving covariates in order to determine important prognostic factors.

We also discuss the choice of appropriate prior distributions for the parameters of the model and some computational issues to obtain accurate Bayesian inferences. The popular existing free Openbugs software (Spiegelhalter et al., 2003) is used to obtain the posterior summaries of interest.

The paper is organized as follows: in Section 2, we review the basic concepts of the stable distribution family and some basic properties of the model. Section 3 presents a Bayesian analysis for the stable distributions not considering the presence of covariates. Section 4 presents a Bayesian analysis of the model assuming uniform prior distributions for the parameters of the model and also considering regression models for the location and scale parameters. Section 5 presents two applications with real medical data sets. Finally, section 6 provides some concluding remarks on the obtained results.

## 2 DEFINITION AND BASIC PROPERTIES

The stable distribution does not have an analytic closed form for the probability density and distribution functions, but can be expressed by its characteristic function given by

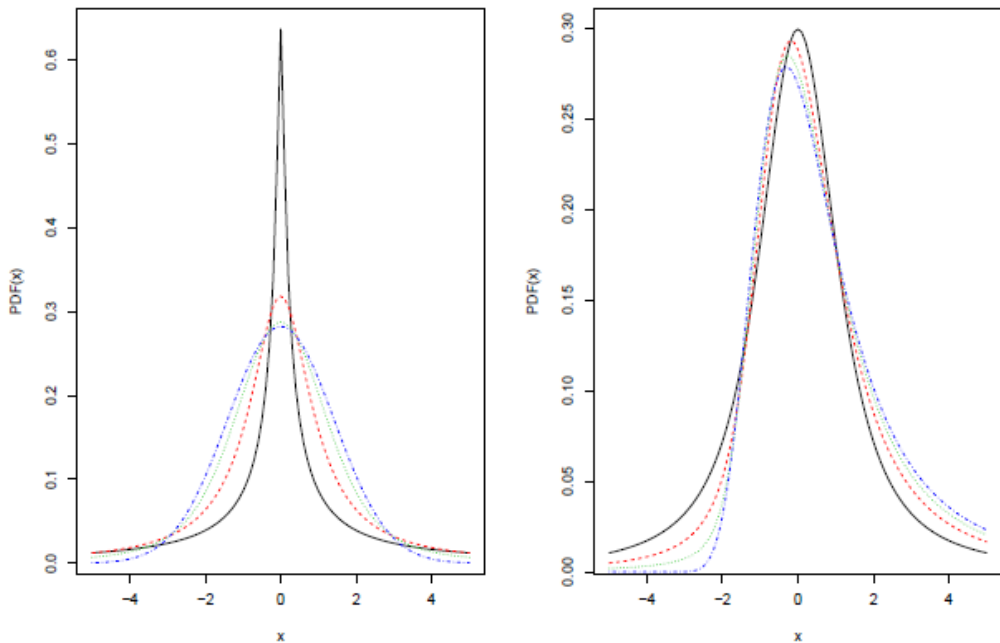
$$\phi(t; \alpha, \beta, \mu, \sigma) = E[e^{itX}] = \begin{cases} \exp\{i\mu t - |\sigma t|^\alpha (1 - i\beta(\text{sign}t) \tan(\frac{\pi\alpha}{2}))\}, & \alpha \neq 1 \\ \exp\{i\mu t - \alpha|t|(1 + i\beta\frac{2}{\pi}(\text{sign}t)\ln|t|)\}, & \alpha = 1 \end{cases}. \quad (1)$$

where  $\text{sign}t = 1$  if  $t > 0$ ;  $\text{sign}t = 0$  if  $t = 0$  and  $\text{sign}t = -1$ , if  $t < 0$ .

The stable distribution, denoted by  $S_\alpha(\beta, \mu, \sigma)$ , requires four parameters  $\alpha, \beta, \mu$  and  $\sigma$  to complete description. The parameter  $\alpha \in (0, 2]$  defines the index of stability or the shape parameter (heaviness of the tails), and when  $\alpha = 2$  this class reduces to Gaussian distributions. The parameter  $\beta \in [-1, 1]$  is the skewness parameter, where for  $\beta = 0$  we have symmetric distributions. The location and scale parameters are, respectively,  $\mu \in (-\infty, \infty)$  and  $\sigma \in (0, \infty)$ . An important property of the stable distribution is that if a random variable has a stable distribution, that is,  $X \sim S_\alpha(\beta, \mu, \sigma)$  then  $Z = (X - \mu) / \sigma \sim S_\alpha(\beta, 0, 1)$ .

A great difficulty associated to stable distributions  $S_\alpha(\beta, \mu, \sigma)$  is that in general there is no simple closed form for their probability density functions except for the case of a Gaussian ( $\alpha = 2$ ), Cauchy ( $\alpha = 1, \beta = 0$ ) and Lévy ( $\alpha = 0.5, \beta = \pm 1$ ) distributions. However, it is known that the probability density functions of stable distributions are continuous and unimodal. The support of all stable distributions is given in  $(-\infty, \infty)$ , except for  $\alpha < 1$  and  $|\beta| = 1$  when the support is  $(-\infty, 0)$  for  $\beta = 1$  and  $(0, \infty)$  for  $\beta = -1$ . It is important to point out that, if  $\alpha < 1$ , the variance is infinite and the mean of the stable distribution does not exist. When  $\alpha > 1$ , the mean of the distribution exists and is equal to  $\mu$ . When the skewness parameter  $\beta$  is positive, the distribution is skewed to the right; when it is negative it is skewed to the left. When  $\beta = 0$ , the distribution is symmetric about  $\mu$ . As  $\alpha$  approaches 2,  $\beta$  loses its effect and the distribution approaches the Gaussian distribution regardless the values of  $\beta$ . If the scale parameter  $\sigma = 1$  and the location parameter  $\mu = 0$ , the distribution is called standard stable.

Figure 1 shows the density plots for different parameter values of the stable distribution.



**Figure 1** – Densities of the stable distribution  $S_\alpha(\beta, \mu, \sigma)$  with parameters  $\mu = 0, \sigma = 1$ .

Right panel: closed form for the known densities Gauss, Cauchy and Levy.

Left panel: stable density functions for  $\alpha = 1.2$  and  $\beta = 0, 0.5, 0.8$  and 1.

Some properties of the  $\alpha$ -stable distribution are:

- The tail of the density function decays like a power function, that is,  $P(|X| > x) \propto Cx^{-\alpha}$  when  $x \rightarrow \infty$  for some constant  $C$ .

- $E(|X|^p) < \infty, 0 < p < \alpha; E(|X|^p) = \infty, p \geq \alpha.$
- $E(|X|) = \mu, \text{ if } \alpha > 1; E(|X|) = \infty, \text{ if } \alpha \leq 1.$
- The stability property is preserved under linear transformation.

Although this class of distributions is a good alternative for data modeling in different areas, we usually have great difficulties to obtain estimates under classical approaches due to the lack of closed form expressions for their probability density functions. One possibility in applications is to compute the probability density function from the inversion formula,

$$f(x) = 1/2\pi \int_{-\infty}^{\infty} e^{-itx} \Phi(t) dt \tag{2}$$

where  $\Phi(t)$  is the characteristic function. In applications, we must use numerical methods to solve the integral in (2), usually taking a great computational time.

The literature presents different studies leading to different techniques to compute the density functions of the family of stable distributions  $S_{\alpha}(\beta, \mu, \sigma)$  (Oral & Erdemir, 2012).

### 3 BAYESIAN INFERENCES FOR THE PARAMETERS OF THE STABLE DISTRIBUTION

An alternative estimation method to obtain inferences for stable distributions is the use of Bayesian inference (Gelman et al., 2013; O’Hagan & Forster, 2004; Bernardo & Smith, 2007).

In Bayesian inference, a prior distribution is specified for the parameters of the distribution, which reflects existing knowledge (or lack of knowledge) regarding the probability of the possible values. The ability to specify a prior distribution is seen as an advantage of the Bayesian approach because the uncertainty surrounding each parameter is taken into consideration.

It is important to note that, in the Bayesian approach, the inference is based on the posterior density of the parameters, which is proportional to the product of likelihood and prior distribution, that is  $p(\theta|\mathbf{x}) \propto f(\mathbf{x}|\theta)\pi_0(\theta)$ . However, for stable distributions the likelihood function based on the the data is not possible to derive directly from the density (2) since this one is not given in a closed form. Therefore, an alternative to estimate the parameters of  $S_{\alpha}(\beta, \mu, \sigma)$  is to use latent or artificial variables.

If it is not possible to obtain the probability density function of  $x$  in closed form whereas the joint density function of  $x$  and  $y$  exists, then the posterior density is found by taking integration,

$$\pi(\theta|x) \propto \int f(x,y|\theta)\pi_0(\theta)dy \tag{3}$$

where  $\pi_0$  is the prior distribution for the parameter.

Considering a latent or an auxiliary variable, Buckle (1995) proved a theorem that is very useful to simulate samples of the joint posterior distribution for the parameters  $\alpha, \beta, \gamma$  and  $\delta$ . This theorem establishes that a stable distribution for a random variable  $Z$  defined in  $(-\infty, \infty)$  is obtained as

the marginal of a bivariate distribution for the random variable  $Z$  itself and an auxiliary or latent random variable  $Y$ . This non-observed random variable  $Y$  is defined in the interval  $(-0.5, a_{\alpha,\beta})$ , when  $Z \in (-\infty, 0)$ , and in  $(a_{\alpha,\beta}, 0.5)$ , when  $Z \in (0, \infty)$ .

The quantity  $a_{\alpha,\beta}$  is given by

$$a_{\alpha,\beta} = -b_{\alpha,\beta} \alpha \pi \quad (4)$$

where  $b_{\alpha,\beta} = (\beta \pi / 2) \min\{\alpha, 2 - \alpha\}$ .

The joint probability density function for the random variables  $Z$  and  $Y$  is given by

$$f(z, y | \alpha, \beta) = \{\alpha / |\alpha - 1|\} \{1/|z|\} \{|z/t_{\alpha,\beta}(y)|^\theta\} \exp\{-|z/t_{\alpha,\beta}(y)|^\theta\} / \sigma \quad (5)$$

where  $\theta = \alpha / (\alpha - 1)$ ,

$$t_{\alpha,\beta}(y) = \{\sin(\pi \alpha y + b_{\alpha,\beta}) / \cos(\pi y)\} \{\cos(\pi y) / [\cos(\pi(\alpha - 1)y) + b_{\alpha,\beta}]\}^{1/\theta} \quad (6)$$

and  $Z = (X - \mu) / \sigma$  for  $\sigma \neq 0$ .

Buckle (1995) showed that the marginal distribution for the random variable  $Z$  of the bivariate density (5), is a stable  $S_\alpha(\beta, 0, 1)$  distribution.

Let us assume that  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  is a random sample of size  $n$ , where  $X \sim S_\alpha(\beta, \mu, \sigma)$  and  $\pi_0(\alpha, \beta, \mu, \sigma)$  is a joint prior distribution for the parameters  $\alpha, \beta, \mu$  and  $\sigma$ . Then, from the bivariate density (5), the joint posterior distribution for  $\alpha, \beta, \mu$  and  $\sigma$  is given by:

$$\pi(\alpha, \beta, \mu, \sigma | \mathbf{x}) \propto \left( \frac{\alpha}{|\alpha - 1| \sigma} \right)^n \exp \left\{ - \sum_{i=1}^n \left| \frac{z_i}{t_{\alpha,\beta}(y_i)} \right|^\theta \right\} \prod_{i=1}^n \left| \frac{z_i}{t_{\alpha,\beta}(y_i)} \right|^\theta \frac{1}{|z_i|} \pi_0(\alpha, \beta, \mu, \sigma) dy. \quad (7)$$

where  $\theta = \alpha / (\alpha - 1)$ ,  $Z_i = (X_i - \mu) / \sigma$ , for  $i = 1, \dots, n$ ,  $\alpha \in (0, 2]$ ,  $\beta \in [-1, 1]$ ,  $\mu \in (-\infty, +\infty)$ ,  $\sigma \in (0, +\infty)$ .

We assume a uniform  $U(-0.5, 0.5)$  prior distribution  $h(y_i)$  for the latent random quantities  $Y_i$  for  $i = 1, 2, \dots, n$ . Then the joint posterior probability distribution for  $\alpha, \beta, \mu, \sigma$  and  $\mathbf{y}$  is given by

$$\pi(\alpha, \beta, \mu, \sigma, \mathbf{y} | \mathbf{x}) \propto \left( \frac{\alpha}{|\alpha - 1| \sigma} \right)^n \exp \left\{ - \sum_{i=1}^n \left| \frac{z_i}{t_{\alpha,\beta}(y_i)} \right|^\theta \right\} \prod_{i=1}^n \left| \frac{z_i}{t_{\alpha,\beta}(y_i)} \right|^\theta \frac{1}{|z_i|} \prod_{i=1}^n h(y_i) \pi_0(\alpha, \beta, \mu, \sigma). \quad (8)$$

The Gibbs sampler can be implemented through the Openbugs software by the following procedure. Firstly, start with the initial values  $\alpha^0, \beta^0, \mu^0, \sigma^0$ . For each observation  $x_i$ , then  $y_i$  is generated from  $\pi(y_i | \alpha, \beta, \mu, \sigma, x_i)$ , for  $i = 1, 2, \dots, n$ . After generating the vector  $\mathbf{y} = (y_1, y_2, \dots, y_n)$  then the parameters  $\alpha, \beta, \mu$  and  $\sigma$  are generated from the conditional distributions  $\pi(\alpha / \beta, \mu, \sigma, \mathbf{x}, \mathbf{y})$ ,  $\pi(\beta / \alpha, \mu, \sigma, \mathbf{x}, \mathbf{y})$ ,  $\pi(\mu / \alpha, \beta, \sigma, \mathbf{x}, \mathbf{y})$  and  $\pi(\sigma / \alpha, \beta, \mu, \mathbf{x}, \mathbf{y})$ , respectively.

In this study we assume uniform prior distributions on the domain of each parameter, that is,  $\alpha \sim U(0, 2)$ ,  $\beta \sim U(-1, 1)$ ,  $\sigma \sim U(0, a)$ ,  $\mu \sim U(-b, b)$ , where the symbol  $\sim$  denotes that the parameter has the distribution,  $a$  and  $b$  are known hyperparameters, usually very large to have approximately non-informative prior distributions. Other possibility in the data analysis, is to assume a gamma prior distribution for the scale parameter  $\sigma$  and a normal prior distribution for the

location parameter  $\mu$ , but it was observed in the examples considered in this study that, assuming uniform priors for all parameters, the obtained posterior inferences were better, especially in terms of convergence of the simulation algorithms.

With this choice of priors, the simulated Gibbs sampling for the joint posterior distribution is carried out to obtain the posterior summaries of interest.

#### 4 A BAYESIAN ANALYSIS ASSUMING REGRESSION MODELS FOR STABLE DISTRIBUTIONS

In this section, let us assume the presence of a vector  $\mathbf{v} = (v_1, v_2, \dots, v_k)$  of covariates associated to the variable  $X_i, i = 1, \dots, n$ , under a stable distribution  $X_i \sim S_\alpha(\beta, \mu_i, \sigma_i)$ , which implies that  $Z_i = (X_i - \mu_i)/\sigma_i \sim S_\alpha(\beta, 0, 1)$ , where the location and the scale parameters  $\mu_i$  and  $\sigma_i$  are assumed to be dependent of the covariates as follows:

$$\begin{aligned} \mu_i &= d_0 + d_1 v_{1i} + d_2 v_{2i} + \dots + d_k v_{ki} \\ \text{and} \\ \sigma_i &= \exp(e_0 + e_1 v_{1i} + e_2 v_{2i} + \dots + e_k v_{ki}), \end{aligned} \tag{9}$$

$i=1, 2, \dots, n$ , with  $\mathbf{d} = (d_0, d_1, \dots, d_k)$  and  $\mathbf{e} = (e_0, e_1, \dots, e_k)$  as the regression coefficients.

Assuming a joint prior distribution  $\pi_0(\alpha, \beta, \mathbf{d}, \mathbf{e})$  for  $\alpha, \beta, \mathbf{d}$  and  $\mathbf{e}$ , Buckle (1995) shows that the joint posterior distribution for the parameters  $\alpha, \beta, \mathbf{d}$  and  $\mathbf{e}$ , is given by

$$\pi(\alpha, \beta, \mathbf{d}, \mathbf{e} | \mathbf{x}) \propto \int \left( \frac{\alpha}{|\alpha - 1| \sigma_i} \right)^n \exp \left\{ - \sum_{i=1}^n \left| \frac{z_i}{t_{\alpha, \beta}(Y_i)} \right|^\theta \right\} \prod_{i=1}^n \left| \frac{z_i}{t_{\alpha, \beta}(y_i)} \right| \frac{1}{|z_i|} \pi_0(\alpha, \beta, \mathbf{d}, \mathbf{e}) dy, \tag{10}$$

where  $\theta = \alpha/(\alpha-1)$ ,  $Z_i = (X_i - \mu_i)/\sigma_i$ , for  $i = 1, \dots, n$ ,  $\alpha \in (0, 2], \beta \in [-1, 1], \mathbf{x} = (x_1, x_2, \dots, x_n)$  and  $\mathbf{y} = (y_1, y_2, \dots, y_n)$  are, respectively, the observed and non-observed data vectors. Note that when  $\alpha = 2$ , we have  $\theta = 2$  and  $b_{\alpha, \beta} = 0$ . In this case, we have a Gaussian distribution with mean equal to  $\mu$  and variance equal to  $2\sigma^2$ .

For a Bayesian analysis of the proposed model, we assume uniform  $U(a, b)$  prior distributions for the parameters  $\alpha, \beta, d_0, d_1, d_2, \dots, d_k, e_0, e_1, e_2, \dots, e_k$  where the hyperparameters  $a$  and  $b$  are assumed to be known in each application following the restrictions  $\alpha \in (0, 2]$  and  $\beta \in [-1, 1]$ . We further assume prior independence among all parameters.

In the simulation algorithm to obtain a Gibbs sample for the parameters  $\alpha, \beta, \mathbf{d}$  and  $\mathbf{e}$ , having the joint posterior distribution (10), it is also assumed uniform distributions for the latent variable  $Y$  defined in the interval  $(-0.5, 0.5)$ . With this choice of prior distributions, it is used the standard free Openbugs software to get the posterior summaries of interest, which gives great simplification to obtain the simulated Gibbs samples for the joint posterior distribution of interest.

From the posterior distribution (10), the joint posterior probability distribution for  $\alpha, \beta, d, e$  and  $\mathbf{y} = (y_1, y_2, \dots, y_n)$  is given by,

$$\pi(\alpha, \beta, \mathbf{d}, \mathbf{e}, \mathbf{y} | \mathbf{x}) \propto \left( \frac{\alpha}{|\alpha - 1| \sigma_i} \right)^n \exp \left\{ \sum_{i=1}^n \left| \frac{z_i}{t_{\alpha, \beta}(y_i)} \right| \right\} \prod_{i=1}^n \left| \frac{z_i}{t_{\alpha, \beta}(y_i)} \right|^{-\theta} \frac{1}{|z_i|} \cdot \prod_{i=1}^n h(y_i) \pi_0(\alpha, \beta, \mathbf{d}, \mathbf{e}), \quad (11)$$

where  $\theta$  and  $t_{\alpha, \beta}(y)$  are respectively defined in (7) and (8) and the variable  $Y$  is defined in the interval  $(-0.5, 0.5)$ .

Since we are using the Openbugs software to simulate samples for the joint posterior distribution we do not present all full conditional distributions needed for the Gibbs sampling algorithm. This software only requires the data distribution and prior distributions of the interested random quantities. This gives great computational simplification to get the posterior summaries of interest as shown in the applications as follows.

## 5 APPLICATIONS WITH REAL DATA

In this section two examples with real data sets showing the usefulness of the proposed methodology are introduced.

### 5.1 Effects of smoking on health

The literature presents many studies on the effects of smoking on health. In this direction, Tager et.al (1983) reported analyses of a study aimed at assessing children's breathing in the absence or presence of smoking cigarettes, as well as exposure to passive smoke from at least one parent ( $n = 654$  individuals). These studies present some of the earliest attempts to document the obvious signs of reduced pulmonary function from smoking and from exposure to second-hand smoke. The data in this investigation comes from an observational study where the subjects self-select which group they think they belong, to the smoking or non-smoking group. Subjects also self-report smoking status. Using a spirometer, FEV (Forced Expiratory Volume) is recorded for each subject.

The following variables are reported associated to each individual: age (age in years), FEV (forced expiratory volume in litres), ht (height of subject in cm), gender (female (0), male (1)), smoker (non-smoker (0), smoker (1)). Figure 2 shows the histograms and normal plots of the response variable FEV in the original and logarithmic scales. From these plots, it is observed that the normality is improved in the logarithmic scale. Not considering the logarithm transformation, the distribution is approximately skewed (Figure 2). In this way, we assume a stable distribution to have better fit for the data in the original scale.

As a first statistical analysis, it is assumed a linear regression model with normal errors under a classical approach (least squares estimators-LSE) relating the response FEV in the logarithm scale with the covariates age, height, gender and smoker. The fitted regression model is given by

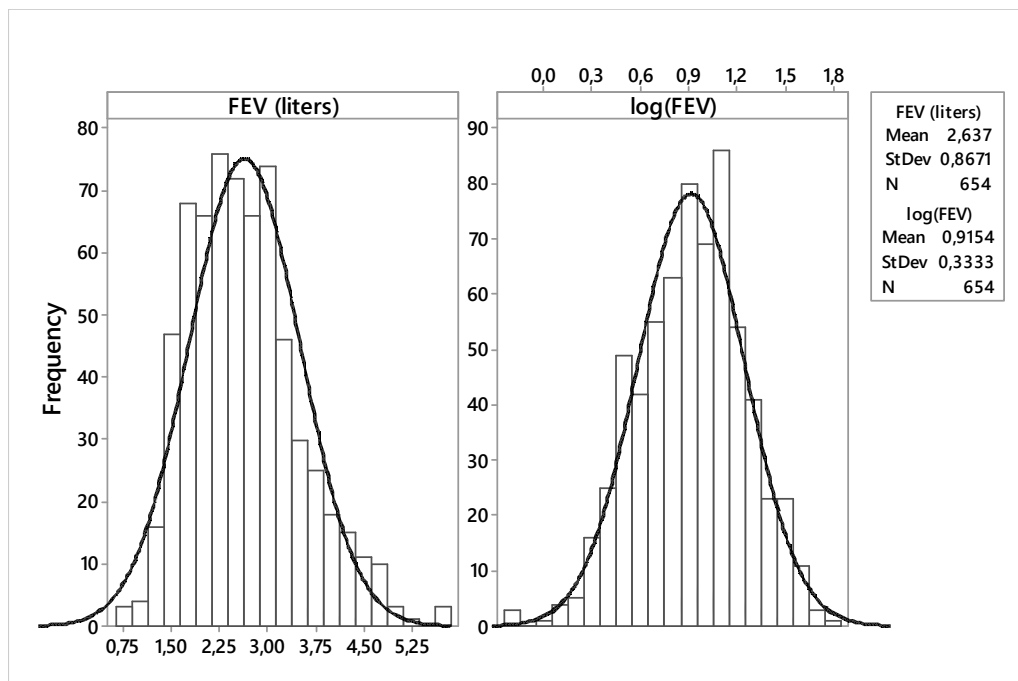
$$\log(\text{FEV}) = -1.9414 + 0.016819 \text{ height} + 0.02363 \text{ age} + 0.0288 \text{ gender} - 0.0471 \text{ smoker}. \quad (12)$$



Table 1 shows the summaries for LSE analysis ( $S = 0.145882$ ;  $R\text{-sq} = 80.96\%$ ). From the results presented in Table 1, it is observed that assuming a linear multiple regression model with normal errors (response in logarithmic scale) all covariates, height, age, gender and smoking, show significative effects on the response FEV since the p-values are smaller than 0.05 (assumed significance level).

**Table 1** – Linear normal regression model (FEV data).

	LSE	SE	T-Value	P-Value
Constant	-1.9414	0.0790	-24.58	< 0.001
Age	0.0236	0.0033	7.04	< 0.001
gendercoded	0.0288	0.0118	2.45	0.015
smokercoded	-0.0471	0.0210	-2.25	0.025
Height	0.0168	0.0006	25.34	< 0.001



**Figure 2** – Histograms (original and transformed scales).

Older people have larger FEV; males have larger FEV when compared to females; smokers have smaller FEV when compared to non-smokers and larger height implies in larger FEV. Figure 3 presents the residual analysis, from where it is observed that the needed assumptions (normality and constant variance of the residuals) assuming the logarithmic scale for the responses (a log-normal distribution for FEV) are only approximately verified.

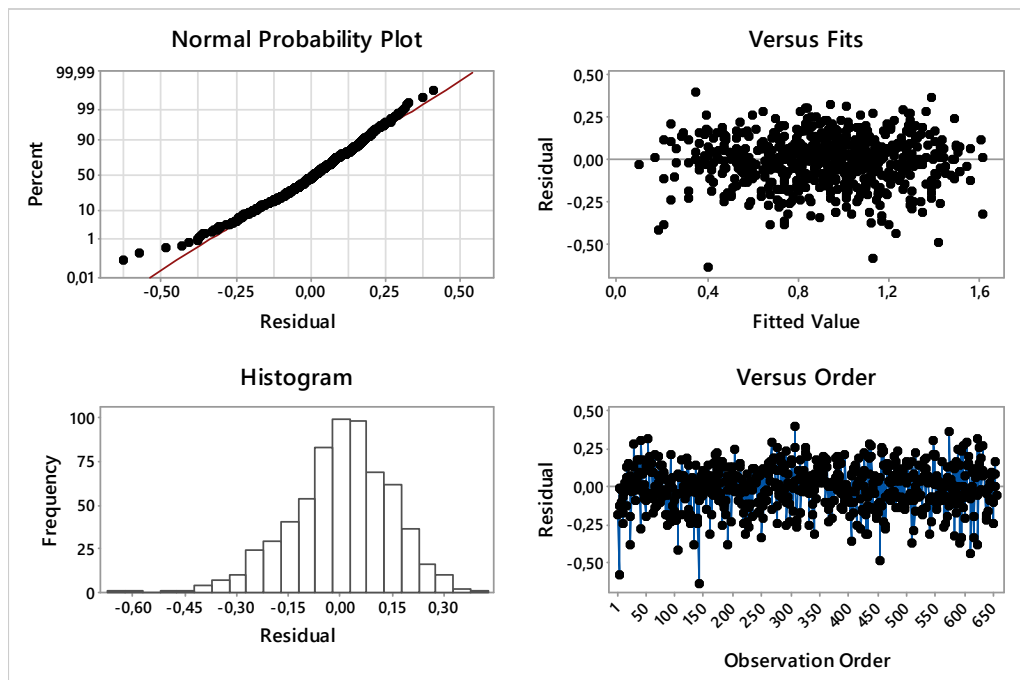


Figure 3 – Residual plots (normal linear regression model-FEV data).

### 5.1.1 A Bayesian approach using a stable distribution not considering the presence of covariates

Now, let us assume a stable distribution for the dataset FEV in the original scale not considering the presence of covariates, assuming the following prior distributions for the parameters  $\alpha$ ,  $\beta$ ,  $\mu$  and  $\sigma$  of the stable distribution:  $\alpha \sim U(1, 2)$ ,  $\beta \sim U(-1, 1)$ ,  $\mu \sim U(-2, 2)$  and  $\sigma \sim U(0,2)$ , where  $U(a, b)$  denotes an uniform distribution in the interval  $(a, b)$ . Using informative priors for the parameters of the model, some information on the parameters of the stable distribution could be obtained from the histograms presented in Figure 2, as close symmetry leading to prior distributions that should be concentrated in  $\alpha = 2$ . In this way, it is considered a uniform prior for the parameter  $\alpha$  concentrated in the interval  $(1, 2)$ . Table 2 shows the posterior summaries of interest (burn-in sample = 100,000 and other 400,000 Gibbs samples taking every 100<sup>th</sup> sample) obtained using the Openbugs software. Convergence of the MCMC algorithm was verified from standard traceplots of the simulated samples. The Openbugs code is presented in Appendix 1.

From the results of Table 2, it is possible to see that the posterior mean for the parameter  $\alpha$  has a Monte Carlo estimator based on the simulated Gibbs samples close to the value 2 (Monte Carlo estimator of the posterior mean equal to 1.912) indicating that the fitted stable distribution is close to the usual normal distribution. The convergence of the simulated Gibbs samples was verified from trace plots of the generated samples for each parameter.

**Table 2** – Bayesian analysis - FEV data no covariates.

	mean	s.d.	Lower 95%	Upper 95%
$\alpha$	1.912	0.0467	1.83	1.993
$\beta$	-0.0755	0.4961	-0.9024	0.9151
$\mu$	1.997	0.0024	1.991	2.0
$\sigma$	0.7389	0.0261	0.6908	0.7926

**5.1.2 A Bayesian approach assuming a stable distribution in presence of covariates**

For a second statistical analysis now considering the presence of a vector of covariates, it is assumed the regression model for the location and scale parameters of the stable distribution defined by (9) under a Bayesian approach (use of MCMC methods) relating the location and scale parameters  $\mu$  and  $\sigma$  of the stable distribution for the response FEV with the covariates age, height, gender and smoke, that is,

$$\begin{aligned} \mu_i &= d_0 + d_1 \text{height}_i + d_2 \text{age}_i + d_3 \text{gender}_i + d_4 \text{smoker}_i \\ \sigma_i &= \exp(e_0 + e_1 \text{height}_i + e_2 \text{age}_i + e_3 \text{gender}_i + e_4 \text{smoker}_i). \end{aligned} \tag{13}$$

For a Bayesian analysis, we assume an uniform U(1.5, 2) prior distribution for  $\alpha$ , an uniform U(-0.5, 0.5) prior distribution for  $\beta$  and uniform prior U(-1, 1) distributions for the regression parameters  $d_0, d_1, d_2, d_3, d_4, e_0, e_1, e_2, e_3, e_4$ . Table 3 shows the posterior summaries of interest.

**Table 3** – Bayesian analysis – (FEV data presence of covariates).

	Mean	s.d.	Lower 95%	Upper 95%
$\alpha$	1.975	0.0215	1.923	1.999
$\beta$	-0.0138	0.2831	-0.4783	0.4713
$d_0$	2.015	0.0157	2.0	2.061
$d_1$	-0.0119	3,11E-01	-0.0124	-0.0112
$d_2$	0.2394	0.0049	0.2292	0.2479
$d_3$	0.2355	0.0426	0.1491	0.3178
$d_4$	0.0499	0.1060	-0.1615	0.2431
$e_0$	-0.8975	0.1053	-0.9978	-0.613
$e_1$	-0.0066	0.0013	-0.0095	-0.0044
$e_2$	0.0854	0.0156	0.0549	0.1167
$e_3$	0.3526	0.0581	0.2427	0.4694
$e_4$	0.1683	0.1060	-0.0281	0.3937

From the results presented in Table 3, it is observed that assuming the stable distribution in presence of the covariates, the covariate smoker do not show significative effect on the location and scale parameters of the stable distribution assumed for the response FEV (Forced Expiratory

Volume) since zero is inside the 95% credible intervals for the corresponding regression parameters  $d_4$  and  $e_4$ . All other covariates (age, height and gender) show significant effects on the location and scale parameters of the stable distribution since the zero value is not included in the respective 95% credible intervals for the regression parameters  $d_1$ ,  $d_2$ ,  $d_3$ ,  $e_1$ ,  $e_2$  and  $e_3$ . The convergence of the simulated Gibbs samples was verified from trace plots of the generated samples for each parameter.

Figure 4 shows the scatter plots of the response FEV versus each covariate, from where it is possible to see the great linear dependence between the response FEV and the covariates age and height.

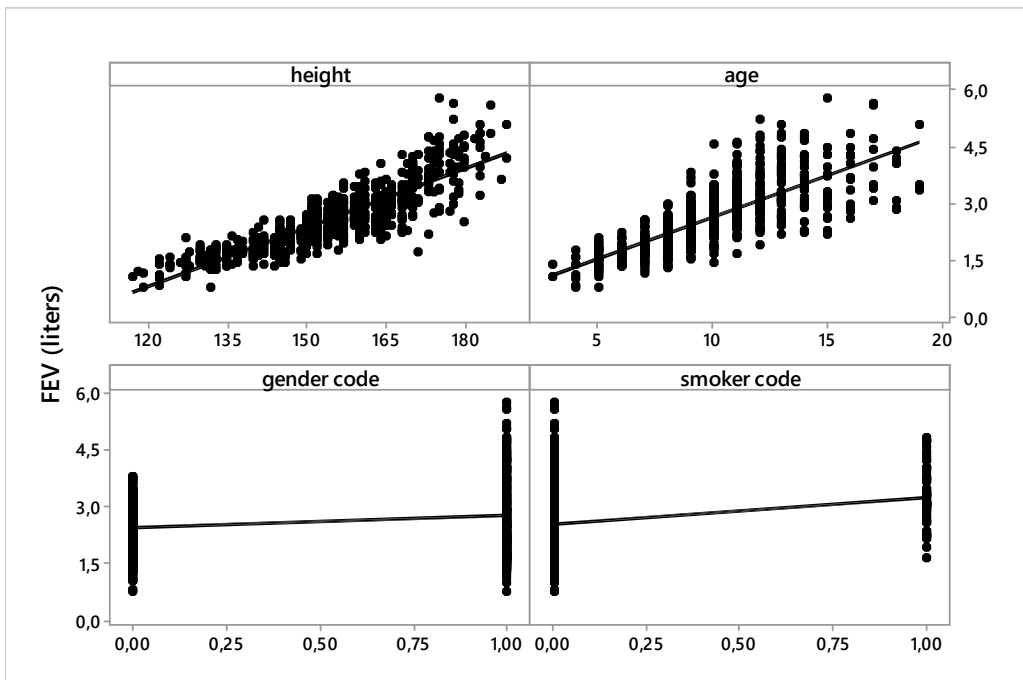


Figure 4 – Scatter plots of the response FEV versus each covariate.

## 5.2 Association between hemoglobin (HbA1c) levels and some covariates for type 2 diabetes mellitus (T2DM)

As another illustrative example, we assume in this example, a data set related to the association between some covariates for type 2 diabetes mellitus (T2DM) patients with different glycated hemoglobin (HbA1c) levels (data set introduced by Shu et al., 2017). The sample size consists of  $n = 154$  patients. In this application, we consider some covariates related to the response glycated hemoglobin (HbA1c) for each patient: age, education in years, duration diagnosis in months, medication compliance, barrier score, dietary knowledge score, BMI (body mass score),

gender and total HEI (healthy eating index) score. Figure 5 shows the scatter plots of the response HbA1c average versus each covariate.

A first statistical analysis is considered assuming a classical multiple linear regression model in presence of the covariates (age, education in years, duration diagnosis in months, medication compliance, barrier score, dietary knowledge score, BMI (body mass score), gender and total HEI (healthy eating index) score. Table 4 shows the summaries for the LSE analysis. The fitted regression model (use of Minitab® software) is given by

$$\begin{aligned} \text{HBA1c.average} = & 12.37 + 0.0188 \text{ age} + 0.0010 \text{ educ.years} + 0.00221 \text{ duration.diagnosis} - \\ & 0.0263 \text{ medication.compliance} - 0.0100 \text{ barrier.score} - 0.1267 \text{ dietary.knowledge.score} + \\ & 0.0581 \text{ BMI} + 0.057 \text{ gender} - 0.0221 \text{ total.HEI.score} \end{aligned} \quad (14)$$

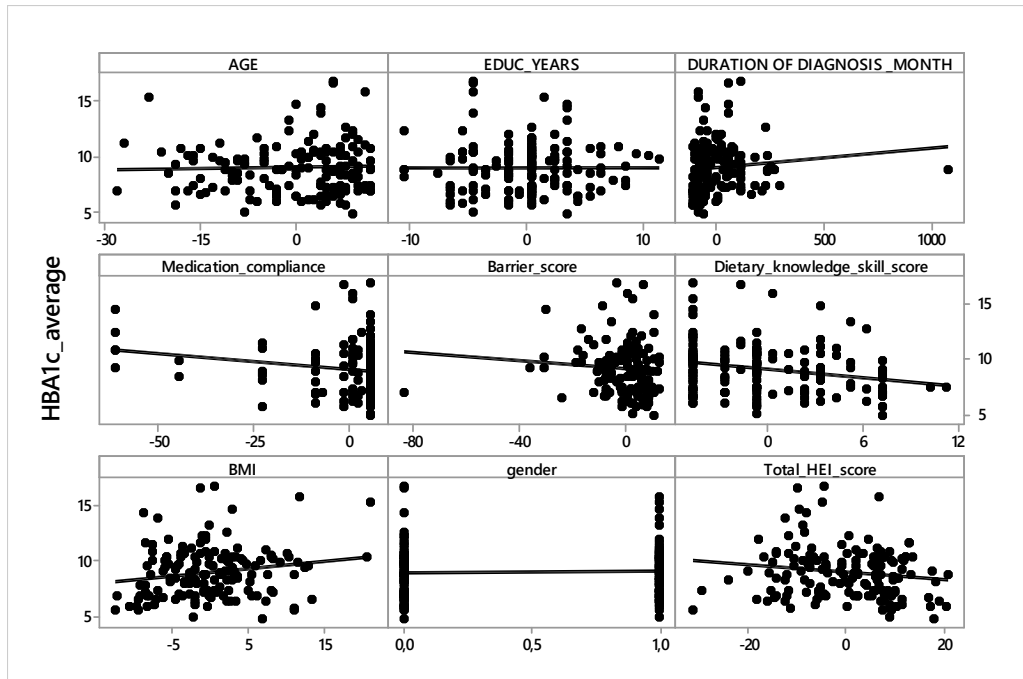


Figure 5 – Scatter plots of the response HbA1c versus each covariate.

From the results of Table 4, it is observed that assuming a multiple linear regression model with normal errors, the covariates medication compliance, dietary knowledge skill score and BMI show significant effects on the response FEV since the respective p-values are smaller than 0.05 (assumed significance level). Figure 6 shows the residual analysis, from where it is observed that the needed assumptions (normality and constant variance of the residuals) assuming the logarithmic scale for the responses are not well verified, especially the normality of the residuals.

**Table 4** – Linear normal regression model (HbA1c data).

	Coef	SE	T-value	P-Value
Constant	12.37	2.39	5.18	< 0.001
Age	0.0188	0.0207	0.91	0.366
Educ years	0.0010	0.0459	0.02	0.983
Durationdiagnosis in months	0.0022	0.0014	1.52	0.131
Medication_compliance	-0.0263	0.0127	-2.07	0.040
Barrier_score	-0.0100	0.0166	-0.60	0.549
Dietary_knowledge_skill_score	-0.1267	0.0446	-2.84	0.005
BMI	0.0581	0.0289	2.01	0.046
Gender	0.057	0.387	0.15	0.883
total_HEI_score	-0.0221	0.0179	-1.23	0.220

### 5.2.1 A Bayesian approach using a stable distribution not considering the presence of covariates

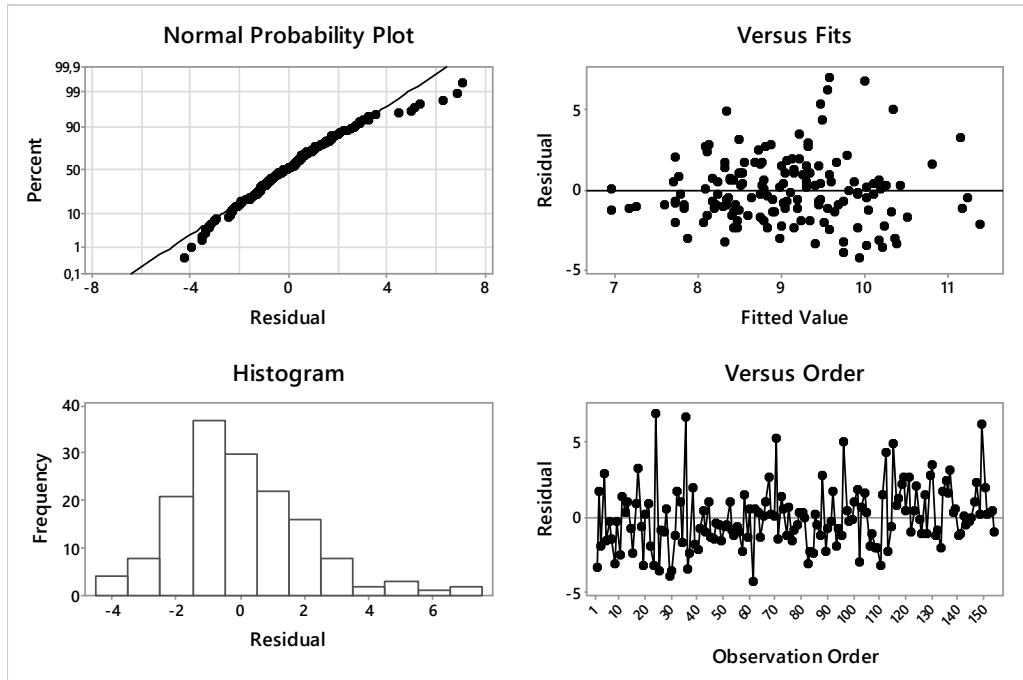
In this subsection, let us assume a stable distribution for the dataset, first not considering the presence of covariates assuming the same uniform prior distributions for the parameters  $\alpha$ ,  $\beta$ ,  $\mu$  and  $\sigma$  assumed in section 5.1.1. Table 5 presents the posterior summaries of interest (burn-in sample = 300,000 and other 511,000 Gibbs samples taking every 100th sample) obtained using the Openbugs software. Convergence of the MCMC algorithm was verified from standard traceplots of the simulated samples. The convergence of the simulated Gibbs samples was verified from trace plots of the generated samples for each parameter.

**Table 5** – Bayesian analysis - stable distribution not considering the presence of covariates (HbA1c data).

	mean	s.d.	Lower 95%	Upper 95%
$\alpha$	1.65	0.1424	1.389	1.909
$\beta$	-0.2276	0.7033	-0.9878	0.975
$\mu$	8.791	0.1734	8.484	9.153
$\sigma$	1.393	0.103	1.208	1.616

### 5.2.2 A Bayesian approach assuming a stable distribution in presence of covariates

Now, considering the presence of a vector of covariates (age, education in years, duration diagnosis in months, medication compliance, barrier score, dietary knowledge score, BMI, gender



**Figure 6** – Residual plots (normal linear regression model- HbA1c data).

and total HEI, it is assumed the regression models for the location and scale parameters of the stable distribution defined by (9) under a Bayesian approach. That is,

$$\begin{aligned} \mu_i = & d_0 + d_1 age_i + d_2 educ.years_i + d_3 duration.diagnosis_i + \\ & d_4 medication.compliance_i + d_5 barrier.score_i + d_6 dietary.knowledge.score_i + \\ & d_7 BMI_i + d_8 total.HEI.score_i + d_9 gender_i \\ \sigma_i = & \exp(e_0 + e_1 age_i + e_2 educ.years_i + e_3 duration.diagnosis_i + \\ & e_4 medication.compliance_i + e_5 barrier.score_i + e_6 dietary.knowledge.score_i + \\ & e_7 BMI_i + e_8 total.HEI.score_i + e_9 gender_i) \end{aligned} \quad (15)$$

For a Bayesian analysis, it is assumed an uniform  $U(0.5,2)$  prior for  $\alpha$ , an uniform  $U(-1, 1)$  prior for  $\beta$ , an uniform  $U(-10,10)$  prior for the regression parameter  $d_0$ , an uniform prior  $U(-1, 1)$  for the regression parameters  $d_9$  and  $e_0$  and uniform prior distributions  $U(-0.5, 0.5)$  for the other regression parameters. Table 6 presents the posterior summaries of interest.

From the results presented in Table 6, it is observed that assuming the stable distribution in presence of the covariates, the covariate duration of diagnosis, dietary knowledge score and BMI show significative effect on the location parameter of the stable distribution assumed for the

**Table 6** – Bayesian analysis – stable regression model (HbA1c data).

	mean	s.d.	Lower 95%	Upper 95%
$\alpha$	1.82	0.1479	1.47	1.994
$\beta$	0.0196	0.6218	-0.9489	0.9678
$d_0$	8.92	0.2138	8.506	9.3380
$d_1$	0.0267	0.0186	-0.0098	0.0629
$d_2$	-4,75E-01	0.0413	-0.0804	0.0879
$d_3$	0.0021	0.0014	4.09E-4	0.0059
$d_4$	-0.0164	0.0129	-0.0404	0.0121
$d_5$	-0.0428	0.0228	-0.0897	0.0014
$d_6$	-0.1468	0.0404	-0.2229	-0.0646
$d_7$	0.0734	0.0309	0.0100	0.1313
$d_8$	-0.0251	0.0179	-0.0603	0.0093
$d_9$	0.2887	0.3430	-0.4444	0.91
$e_0$	0.192	0.0995	0.0026	0.3804
$e_1$	0.0157	0.0080	0.0031	0.0316
$e_2$	-0.0213	0.0193	-0.0592	0.0163
$e_3$	-6,85E-01	0.0010	-0.0027	0.0012
$e_4$	0.0014	0.0058	-0.0113	0.0117
$e_5$	-0.0116	0.0067	-0.0246	0.0013
$e_6$	-0.0384	0.0206	-0.0813	0.0020
$e_7$	0.0155	0.0126	-0.0105	0.0393
$e_8$	-0.0079	0.0081	-0.0233	0.0098
$e_9$	0.2315	0.1399	-0.0732	0.4706

response HbA1c data since zero is not inside the 95% credible intervals for the corresponding regression parameters  $d_3$ ,  $d_6$  and  $d_7$ . Also the covariate age shows significant effect in the scale parameter since zero is not inside the 95% credible interval for the corresponding regression parameter  $e_1$ . The convergence of the simulated Gibbs samples was also verified from trace plots of the generated samples for each parameter.

## 6 CONCLUDING REMARKS

The use of stable distributions can be a good alternative to analyze data from different application areas when the data indicate in some way (descriptive analyses, theoretical interpretations) that the normality assumption usually necessary for several traditional statistical techniques as ANOVA models, multiple linear regression models, among many others, is not satisfied. The stable distribution presents great fit flexibility for asymmetric or heavy tail data sets and good robustness to the presence of outliers, as pointed out by Achcar et al. (2013a,b). For the situation of non-normality of the data, usually data analysts (especially non statisticians) use non-parametric



techniques that often do not respond adequately to the objectives of a research, especially in the medical area. The literature presents several alternatives to analyze data when there is no indication of normality, such as the use of appropriate distributions for heavy tails or finite mixtures of parametric distributions, among several other possibilities, but the use of stable distributions could be a good alternative in the data analysis. The two examples with medical data presented in the study show that stable distributions can be very useful in clinical studies, in addition to what is already considered in other areas of application, such as in particular, the area of finance where the use of stable distributions is already very popular.

The results of this study show, especially for regression models, that the assumption of errors with stable distribution could be considered in medical studies, where the computational costs for the determination of Bayesian estimators are not very high using MCMC simulation methods and data augmentation techniques, especially using existing Bayesian software to simulate samples of the joint posterior distribution of interest like the Openbugs software used in this study.

We also point out that the use of the proposed methodology could be useful to check the normality of the errors in multiple regression parameters since the parameter  $\alpha$  of the stable distribution indicates a Gaussian distribution if  $\alpha = 2$ , that is, if the value 2 is inside a 95% credible interval for  $\alpha$ , there is an indication of normality for the residuals.

Although the stable distribution probability density function does not have a closed form, we obtained satisfactory results for parameter estimation in the presence of a vector of covariates by applying the Bayesian inference method based on data augmentation techniques.

In both examples presented in Section 5, it was shown that using joint regression models for the location and scale parameters of the stable distribution under a Bayesian approach and using existing MCMC methods assuming uniform priors for all parameters, good inferences were obtained for the model parameters. It is important to note that using a simple linear regression model for the location parameter of the stable distribution, assuming normal distributions as prior distributions for the regression parameters, it was not possible to obtain good convergence results from the MCMC simulation algorithm even after a large number of iterations.

We also showed that the MCMC methods for a Bayesian analysis of these models using the Openbugs software give a great simplification to obtain very accurate posterior summaries of interest, not requiring large computational time to perform, even when the simulation of a large number of Gibbs samples is necessary for the convergence of the algorithm.

Other important issue to use Bayesian methods is the choice of prior distributions for the parameters of the proposed model. In the two applications with medical data, we have used non-informative prior distributions since we did not have prior opinion of medical experts for the elicitation of priors. The data sets were obtained from the literature to illustrate the proposed methodology. In practical work, usually we can use informative priors based on the information of medical experts leading to more accurate inference results. This is a great advantage of the use of Bayesian methods in applications in medical studies and applied work in general.

## References

- ACHCAR JA, LOPES SRC, MAZUCHELI J & LINHARES R. 2013a. A Bayesian Approach for Stable Distributions: Some Computational Aspects. *Open Journal of Statistics*, **3**: 268-277.
- ACHCAR JA, ACHCAR A & MARTINEZEZ. 2013b. Robust Linear Regression Models: Use of a Stable Distribution for the Response Data. *Open Journal of Statistics*, **3**(6): 409-416.
- BERNARDO JM & SMITH AFM. 2007. *Bayesian Theory*, 2nd edition. Wiley Series in Probability and Statistics. Wiley, New York.
- BORAK S, HARDLE WK & WERON R. 2005. *Stable Distributions. Statistical Tools for Finance and Insurance*. 10.1007/3-540-27395-6\_1.
- BUCKLE DJ. 1995. Bayesian inference for stable distributions. *Journal of the American Statistical Association*, **90**(430): 605-613.
- GELMAN A, CARLIN JB, STERN HS, DUNSON DB, VEHTARI A & RUBIN DB. 2013. *Bayesian data analysis*. London, United Kingdom: Chapman and Hall/CRC.
- GNEDENKO BV & KOLMOGOROV AN. 1954. *Limit Distributions for Sums of Independent Random Variables*. Cambridge, MA: Addison-Wesley.
- KARLING MJ, LOPES SRC & SOUZA RM. 2021. A Bayesian approach for estimating the parameters of an  $\alpha$ -stable distribution. *Journal of Statistical Computation and Simulation*, **91**(9):1713–1748.
- KATEREGGA M. 2017. Parameter estimation for stable distributions with application to commodity futures log-returns. *Cogent Economics & Finance*, **5**(1):1318813.
- LEMKE T, RIABIZ M & GODSILL SJ. 2015. Fully Bayesian inference for  $\alpha$ -stable distributions using a Poisson series representation, *Digital Signal Processing*, **47**(C), 96-115.
- LOMBARDI MJ. 2004. *Bayesian inference for alpha-stable distributions: a random walk MCMC approach*, Econometrics Working Papers Archive wp2004\_11, Universita' degli Studi di Firenze, Dipartimento di Statistica, Informatica, Applicazioni "G. Parenti". <https://ideas.repec.org/s/fir/econom.html>
- O'HAGAN A & FORSTER JJ. 2004. *Bayesian Inference*, 2nd edition, v. 2B of Kendall's Advanced Theory of Statistics. Arnold, London.
- ORAL E & ERDEMIR C. 2012. A Bayesian Estimation of Stable Distributions, *Journal of Statistical and Econometric Methods*, **1**(3):39-52.
- SHU PS, CHAN YM & HUANG SL. 2017. Higher body mass index and lower intake of dairy products predict poor glycaemic control among type2 diabetes patients in Malaysia. *PLoS One*, **12**:e0172231.

SPIEGELHALTER DJ, THOMAS A, BEST NG & LUNN D. 2003. *WinBUGS User's Manual*, MRC Biostatistics Unit, Cambridge.

TAGER I, WEISS SAM, ROSNER B & SPEIZER F. 1983. Longitudinal study of the effects of maternal smoking on pulmonary function. *New England Journal of Medicine*, **309**(12):699-703.

ZOLOTAREV VM. 1986. *One-Dimensional Stable Distributions*, Vol 65 of Translations of Mathematical Monographs, American Mathematical Society, Providence, Rhode Island.

### **How to cite**

ACHCAR JA, SOUZA RM, BUSSOLA D & MOALA FA. 2022. Bayesian estimation for the stable distributions in the presence of covariates with applications in clinical issues. *Pesquisa Operacional*, **42**: e254533. doi: 10.1590/0101-7438.2022.042.00254533.

## APPENDIX 1 (OPENBUGS-APPLICATION FEV DATA OF APPLICATION 4.1)

## • Stable distribution for the data (no presence of covariates)

```

model {
for (i in 1:n) {
dummy[i] <- 0
x[i] <- FEV[i]
dummy[i] ~ dloglik(logLike[i])
z[i] <- (x[i]-mu)/sigma
w[i] <- abs(z[i]/t[i])
a1[i] <- (sin(3.14159*alpha*y[i]+b1))/cos(3.14159*y[i])
a2[i] <- (cos(3.14159 *y[i]))/cos(3.14159*(alpha-1)*y[i]+b1)
a3[i] <- pow(a2[i],theta2)
t[i] <- a1[i]*a3[i]
logLike[i] <- log(alpha)-log(abs(alpha-1))-log(sigma)-pow(w[i],theta1)+
theta1*log(w[i])-log(abs(z[i]))
y[i] ~ dunif(-0.5,0.5)
}
theta1<-alpha/(alpha-1)
theta2<-(alpha-1)/alpha
b1<-(beta*min(alpha,2-alpha)*3.14159)/2
alpha ~ dunif(1, 2)
beta ~ dunif(-1, 1)
sigma~ dunif(0,2)
mu~ dunif(-2,2)
}

list (FEV=c(1.708,1.724,1.720,1.558,1.895,2.336,1.919,1.415,1.987,1.942,1.602
,1.735,2.193,2.118,2.258,1.932,1.472,1.878,2.352,2.604,1.400,1.256
,0.839,2.578,2.988,1.404,2.348,1.755,2.980,2.100,1.282,3.000,2.673
,2.093,1.612,2.175,2.725,2.071,1.547,2.004,3.135,2.420,1.776,1.931
,1.343,2.076,1.624,1.344,1.650,2.732,2.017,2.797,3.556,1.703,1.634
,2.570,3.016,2.419,1.569,1.698,2.123,2.481,1.481,1.577,1.940,1.747
,2.069,1.631,1.536,2.560,1.962,2.531,2.715,2.457,2.090,1.789,1.858
,1.452,3.842,1.719,2.111,1.695,2.211,1.794,1.917,2.144,1.253,2.659
,1.580,2.126,3.029,2.964,1.611,2.215,2.388,2.196,1.751,2.165,1.682
,1.523,1.292,1.649,2.588,0.796,2.574,1.979,2.354,1.718,1.742,1.603
,2.639,1.829,2.084,2.220,1.473,2.341,1.698,1.196,1.872,2.219,2.420
,1.827,1.461,1.338,2.090,1.697,1.562,2.040,1.609,2.458,2.650,1.429
,1.675,1.947,2.069,1.572,1.348,2.288,1.773,0.791,1.905,2.463,1.431
,2.631,3.114,2.135,1.527,2.293,3.042,2.927,2.665,2.301,2.460,2.592
,1.750,1.759,1.536,2.259,2.048,2.571,2.046,1.780,1.552,1.953,2.893
,1.713,2.851,1.624,2.631,1.819,1.658,2.158,1.789,3.004,2.503,1.933
,2.091,2.316,1.704,1.606,1.165,2.102,2.320,2.230,1.716,1.790,1.146
,2.187,2.717,1.796,1.335,2.119,1.666,1.826,2.709,2.871,1.092,2.262
,2.104,2.166,1.690,2.973,2.145,1.971,2.095,1.697,2.455,1.920,2.164
,2.130,2.993,2.529,1.726,2.442,1.102,2.056,1.808,2.305,1.969,1.556
,1.072,2.042,1.512,1.423,3.681,1.991,1.897,1.370,1.338,2.016,2.639
,1.389,1.612,2.135,2.681,3.223,1.796,2.010,1.523,1.744,2.485,2.335
,1.415,2.076,2.435,1.728,2.850,1.844,1.754,1.343,2.303,2.246,2.476
,3.239,2.457,2.382,1.640,1.589,2.056,2.226,1.886,2.833,1.715,2.631

```

, 2.550, 1.912, 1.877, 1.935, 1.539, 2.803, 2.923, 2.358, 2.094, 1.855, 1.535  
 , 2.135, 1.930, 2.182, 1.359, 2.002, 1.699, 2.500, 2.366, 2.069, 1.418, 2.333  
 , 1.514, 1.758, 2.535, 2.564, 2.487, 1.591, 1.624, 2.798, 1.691, 1.999, 1.869  
 , 1.004, 1.427, 1.826, 2.688, 1.657, 1.672, 2.015, 2.371, 2.115, 2.328, 1.495  
 , 2.884, 2.328, 3.381, 2.170, 3.470, 3.058, 1.811, 2.524, 2.642, 3.741, 4.336  
 , 4.842, 4.550, 2.841, 3.166, 3.816, 2.561, 3.654, 2.481, 2.665, 3.203, 3.549  
 , 3.222, 3.111, 3.490, 3.147, 2.520, 2.292, 2.889, 2.246, 1.937, 2.646, 2.957  
 , 4.007, 2.386, 3.251, 2.762, 3.011, 4.305, 3.906, 3.583, 3.236, 3.436, 3.058  
 , 3.007, 3.489, 2.864, 2.819, 2.250, 4.683, 2.352, 3.108, 3.994, 4.393, 2.592  
 , 3.193, 2.346, 3.515, 2.754, 2.720, 2.463, 2.633, 3.048, 3.111, 3.745, 2.094  
 , 3.183, 3.977, 3.354, 3.411, 3.171, 3.887, 2.646, 2.504, 3.587, 3.845, 2.971  
 , 2.891, 1.823, 2.417, 2.175, 2.735, 4.273, 2.976, 4.065, 2.318, 3.596, 3.395  
 , 2.751, 2.673, 2.556, 2.542, 2.608, 2.354, 1.458, 3.795, 2.491, 3.060, 2.545  
 , 2.993, 3.305, 3.774, 2.855, 2.988, 2.498, 3.169, 2.887, 2.704, 3.515, 3.425  
 , 2.287, 2.434, 2.365, 2.696, 2.868, 2.813, 3.255, 4.593, 4.111, 1.916, 1.858  
 , 3.350, 2.901, 2.241, 4.225, 3.223, 5.224, 4.073, 4.080, 2.606, 4.411, 3.791  
 , 3.089, 2.465, 3.200, 2.913, 4.877, 2.358, 3.279, 2.581, 2.347, 2.691, 2.827  
 , 1.873, 2.538, 2.758, 3.050, 3.079, 2.201, 1.858, 3.403, 3.501, 2.578, 1.665  
 , 2.081, 2.974, 4.073, 4.448, 3.984, 2.250, 2.752, 3.680, 2.862, 3.023, 3.681  
 , 3.255, 3.692, 2.356, 4.591, 3.082, 3.258, 2.216, 3.247, 4.324, 2.362, 2.563  
 , 3.206, 3.585, 4.720, 3.331, 5.083, 2.417, 2.364, 2.341, 3.231, 3.078, 3.369  
 , 3.529, 2.866, 2.891, 3.022, 3.127, 2.866, 2.605, 3.056, 2.569, 2.501, 3.320  
 , 2.123, 3.780, 3.847, 3.924, 2.132, 2.752, 2.449, 3.456, 3.073, 2.688, 3.329  
 , 4.271, 3.530, 2.928, 2.689, 2.332, 2.934, 3.110, 2.894, 2.435, 2.838, 3.035  
 , 4.831, 2.812, 2.714, 3.086, 3.519, 4.232, 2.770, 3.341, 3.090, 2.531, 2.822  
 , 2.935, 2.568, 2.387, 2.499, 4.130, 3.001, 3.132, 3.577, 3.222, 3.280, 2.659  
 , 2.822, 2.140, 4.203, 2.997, 2.562, 3.082, 3.806, 2.458, 2.391, 3.141, 2.579  
 , 2.100, 2.785, 4.284, 2.906, 5.102, 4.429, 4.279, 4.500, 2.635, 3.082, 3.387  
 , 5.793, 3.985, 4.220, 4.724, 3.731, 3.500, 3.674, 5.633, 3.645, 2.887, 3.960  
 , 4.299, 2.981, 4.504, 5.638, 2.853, 3.211, 1.953, 2.236, 3.428, 3.208, 1.694  
 , 3.957, 4.789, 2.384, 3.074, 2.387, 3.835, 2.599, 4.756, 3.086, 4.309, 3.413  
 , 2.975, 3.169, 3.343, 3.751, 2.216, 3.078, 3.186, 3.297, 2.304, 3.102, 2.677  
 , 3.297, 3.498, 2.759, 2.953, 3.785, 2.276, 4.637, 3.038, 3.120, 3.339, 3.152  
 , 3.104, 4.045, 4.763, 3.069, 4.506, 3.519, 3.688, 2.679, 2.198, 3.345, 3.082  
 , 2.903, 3.004, 3.406, 3.122, 3.330, 2.608, 3.799, 4.086, 4.070, 2.264, 4.404  
 , 2.278, 4.872, 4.270, 3.727, 2.795), n = 654)

• **Stable distribution (regression in the location parameter and scale parameters)**

```

model {
for (i in 1:n) {
dummy[i] <- 0
x[i] <- FEV[i]
dummy[i] ~ dloglik(logLike[i])
z[i] <- (x[i]-mu[i])/sigma[i]
mu[i] <- d0+d1*height[i]+d2*age[i]+d3*gender[i]+d4*smoker[i]
sigma[i] <- exp(e0+e1*height[i]+e2*age[i]+e3*gender[i]+e4*smoker[i])
w[i] <- abs(z[i]/t[i])
a1[i] <- (sin(3.14159*alpha*y[i]+b1))/cos(3.14159*y[i])
a2[i] <- (cos(3.14159 *y[i]))/cos(3.14159*(alpha-1)*y[i]+b1)
a3[i] <- pow(a2[i],theta2)
    
```

```

      t[i] <- a1[i]*a3[i]
logLike[i] <- log(alpha)-log(abs(alpha-1))-log(sigma[i])-pow(w[i],theta1)+
theta1*log(w[i])-log(abs(z[i]))
      y[i] ~ dunif(-0.5,0.5)
}
theta1<-alpha/(alpha-1)
theta2<-(alpha-1)/alpha
b1<-(beta*min(alpha,2-alpha)*3.14159)/2
alpha ~ dunif(1.5, 2)
beta ~ dunif(-0.5, 0.5)
d0 ~ dunif(2,5)
d1~ dunif(-0.5,0.5)
d2~ dunif(-1,1)
d3~ dunif(-1,1)
d4~ dunif(-1,1)
e0 ~ dunif(-1,1)
e1~ dunif(-1,1)
e2~ dunif(-1,1)
e3~ dunif(-1,1)
e4~ dunif(-1,1)
}

```

```

list(height=c(145,171,138,135,145,155,147,142,149,152,135,137,149,154,147
,135,127,135,150,156,124,133,122,159,165,131,152,132,152,152
,124,166,152,146,132,150,150,140,145,145,152,150,130,145,127
,145,137,133,140,154,138,156,157,138,137,145,159,152,127,146
,152,152,130,124,150,146,147,141,132,154,145,147,152,150,151
,142,135,130,175,135,145,135,160,138,147,160,132,156,133,157
,156,164,146,152,152,155,147,156,140,140,132,137,160,119,154
,142,149,140,149,130,151,137,147,147,133,154,138,118,144,140
,145,138,137,135,145,150,140,141,131,155,161,146,135,144,137
,132,135,156,149,132,147,155,130,157,164,149,133,147,168,161
,163,149,163,154,140,135,122,149,164,154,142,149,137,147,164
,128,152,131,150,135,135,136,132,163,160,147,149,151,130,146
,119,141,145,155,141,136,127,156,156,140,144,145,132,133,159
,165,127,146,144,146,137,151,151,147,145,140,152,144,152,150
,160,150,135,156,122,160,141,164,150,149,117,157,135,126,173
,151,141,140,131,142,160,122,144,150,154,165,140,140,130,133
,163,150,136,154,151,144,160,144,156,132,145,161,160,165,156
,157,140,130,137,145,142,156,135,150,142,150,133,133,127,151
,163,155,146,152,140,142,130,151,128,146,137,145,147,152,124
,145,132,132,151,147,163,145,135,157,135,144,145,122,126,130
,151,142,137,146,141,127,152,145,175,163,160,147,169,154,145
,163,155,174,177,183,180,160,156,161,157,165,155,160,168,173
,183,168,170,163,154,160,163,154,157,152,164,170,156,168,152
,163,174,170,170,168,159,155,157,169,152,157,147,174,156,164
,170,174,165,178,150,171,166,166,164,157,166,171,173,149,166
,179,160,161,160,171,156,152,164,174,164,155,145,159,147,159
,184,166,169,150,173,170,160,164,157,157,168,157,145,174,150
,156,165,169,165,170,164,178,152,163,159,155,163,166,155,166
,161,168,157,156,168,175,180,154,147,175,151,163,188,164,178
,170,164,165,173,174,171,152,165,163,185,150,179,168,156,170

```

, 159, 133, 180, 166, 152, 152, 154, 150, 157, 164, 160, 145, 160, 157, 174  
, 175, 180, 147, 161, 170, 155, 171, 173, 169, 170, 154, 178, 161, 160, 155  
, 166, 171, 155, 160, 161, 178, 182, 166, 188, 155, 155, 155, 160, 171, 179  
, 179, 157, 157, 156, 157, 154, 159, 160, 160, 157, 166, 165, 178, 168, 173  
, 150, 174, 160, 160, 168, 157, 173, 184, 163, 166, 156, 145, 163, 164, 170  
, 165, 160, 157, 180, 155, 166, 157, 166, 179, 157, 166, 165, 155, 177, 166  
, 161, 154, 165, 170, 161, 151, 161, 155, 168, 163, 157, 154, 180, 164, 159  
, 164, 173, 152, 151, 155, 160, 147, 175, 178, 168, 183, 178, 171, 178, 163  
, 164, 169, 175, 180, 173, 179, 170, 157, 171, 185, 187, 160, 178, 168, 168  
, 183, 178, 152, 169, 147, 168, 163, 155, 152, 183, 175, 161, 165, 168, 177  
, 159, 173, 171, 175, 168, 160, 159, 173, 183, 173, 168, 170, 165, 169, 163  
, 170, 165, 173, 156, 170, 160, 168, 183, 165, 155, 174, 157, 171, 175, 173  
, 165, 180, 168, 173, 168, 157, 166, 170, 160, 163, 175, 163, 174, 157, 169  
, 170, 177, 160, 179, 152, 183, 170, 173, 160),

FEV=c(1.708, 1.724, 1.720, 1.558, 1.895, 2.336, 1.919, 1.415, 1.987, 1.942, 1.602  
, 1.735, 2.193, 2.118, 2.258, 1.932, 1.472, 1.878, 2.352, 2.604, 1.400, 1.256  
, 0.839, 2.578, 2.988, 1.404, 2.348, 1.755, 2.980, 2.100, 1.282, 3.000, 2.673  
, 2.093, 1.612, 2.175, 2.725, 2.071, 1.547, 2.004, 3.135, 2.420, 1.776, 1.931  
, 1.343, 2.076, 1.624, 1.344, 1.650, 2.732, 2.017, 2.797, 3.556, 1.703, 1.634  
, 2.570, 3.016, 2.419, 1.569, 1.698, 2.123, 2.481, 1.481, 1.577, 1.940, 1.747  
, 2.069, 1.631, 1.536, 2.560, 1.962, 2.531, 2.715, 2.457, 2.090, 1.789, 1.858  
, 1.452, 3.842, 1.719, 2.111, 1.695, 2.211, 1.794, 1.917, 2.144, 1.253, 2.659  
, 1.580, 2.126, 3.029, 2.964, 1.611, 2.215, 2.388, 2.196, 1.751, 2.165, 1.682  
, 1.523, 1.292, 1.649, 2.588, 0.796, 2.574, 1.979, 2.354, 1.718, 1.742, 1.603  
, 2.639, 1.829, 2.084, 2.220, 1.473, 2.341, 1.698, 1.196, 1.872, 2.219, 2.420  
, 1.827, 1.461, 1.338, 2.090, 1.697, 1.562, 2.040, 1.609, 2.458, 2.650, 1.429  
, 1.675, 1.947, 2.069, 1.572, 1.348, 2.288, 1.773, 0.791, 1.905, 2.463, 1.431  
, 2.631, 3.114, 2.135, 1.527, 2.293, 3.042, 2.927, 2.665, 2.301, 2.460, 2.592  
, 1.750, 1.759, 1.536, 2.259, 2.048, 2.571, 2.046, 1.780, 1.552, 1.953, 2.893  
, 1.713, 2.851, 1.624, 2.631, 1.819, 1.658, 2.158, 1.789, 3.004, 2.503, 1.933  
, 2.091, 2.316, 1.704, 1.606, 1.165, 2.102, 2.320, 2.230, 1.716, 1.790, 1.146  
, 2.187, 2.717, 1.796, 1.335, 2.119, 1.666, 1.826, 2.709, 2.871, 1.092, 2.262  
, 2.104, 2.166, 1.690, 2.973, 2.145, 1.971, 2.095, 1.697, 2.455, 1.920, 2.164  
, 2.130, 2.993, 2.529, 1.726, 2.442, 1.102, 2.056, 1.808, 2.305, 1.969, 1.556  
, 1.072, 2.042, 1.512, 1.423, 3.681, 1.991, 1.897, 1.370, 1.338, 2.016, 2.639  
, 1.389, 1.612, 2.135, 2.681, 3.223, 1.796, 2.010, 1.523, 1.744, 2.485, 2.335  
, 1.415, 2.076, 2.435, 1.728, 2.850, 1.844, 1.754, 1.343, 2.303, 2.246, 2.476  
, 3.239, 2.457, 2.382, 1.640, 1.589, 2.056, 2.226, 1.886, 2.833, 1.715, 2.631  
, 2.550, 1.912, 1.877, 1.935, 1.539, 2.803, 2.923, 2.358, 2.094, 1.855, 1.535  
, 2.135, 1.930, 2.182, 1.359, 2.002, 1.699, 2.500, 2.366, 2.069, 1.418, 2.333  
, 1.514, 1.758, 2.535, 2.564, 2.487, 1.591, 1.624, 2.798, 1.691, 1.999, 1.869  
, 1.004, 1.427, 1.826, 2.688, 1.657, 1.672, 2.015, 2.371, 2.115, 2.328, 1.495  
, 2.884, 2.328, 3.381, 2.170, 3.470, 3.058, 1.811, 2.524, 2.642, 3.741, 4.336  
, 4.842, 4.550, 2.841, 3.166, 3.816, 2.561, 3.654, 2.481, 2.665, 3.203, 3.549  
, 3.222, 3.111, 3.490, 3.147, 2.520, 2.292, 2.889, 2.246, 1.937, 2.646, 2.957  
, 4.007, 2.386, 3.251, 2.762, 3.011, 4.305, 3.906, 3.583, 3.236, 3.436, 3.058  
, 3.007, 3.489, 2.864, 2.819, 2.250, 4.683, 2.352, 3.108, 3.994, 4.393, 2.592  
, 3.193, 2.346, 3.515, 2.754, 2.720, 2.463, 2.633, 3.048, 3.111, 3.745, 2.094  
, 3.183, 3.977, 3.354, 3.411, 3.171, 3.887, 2.646, 2.504, 3.587, 3.845, 2.971  
, 2.891, 1.823, 2.417, 2.175, 2.735, 4.273, 2.976, 4.065, 2.318, 3.596, 3.395  
, 2.751, 2.673, 2.556, 2.542, 2.608, 2.354, 1.458, 3.795, 2.491, 3.060, 2.545  
, 2.993, 3.305, 3.774, 2.855, 2.988, 2.498, 3.169, 2.887, 2.704, 3.515, 3.425

, 2.287, 2.434, 2.365, 2.696, 2.868, 2.813, 3.255, 4.593, 4.111, 1.916, 1.858  
, 3.350, 2.901, 2.241, 4.225, 3.223, 5.224, 4.073, 4.080, 2.606, 4.411, 3.791  
, 3.089, 2.465, 3.200, 2.913, 4.877, 2.358, 3.279, 2.581, 2.347, 2.691, 2.827  
, 1.873, 2.538, 2.758, 3.050, 3.079, 2.201, 1.858, 3.403, 3.501, 2.578, 1.665  
, 2.081, 2.974, 4.073, 4.448, 3.984, 2.250, 2.752, 3.680, 2.862, 3.023, 3.681  
, 3.255, 3.692, 2.356, 4.591, 3.082, 3.258, 2.216, 3.247, 4.324, 2.362, 2.563  
, 3.206, 3.585, 4.720, 3.331, 5.083, 2.417, 2.364, 2.341, 3.231, 3.078, 3.369  
, 3.529, 2.866, 2.891, 3.022, 3.127, 2.866, 2.605, 3.056, 2.569, 2.501, 3.320  
, 2.123, 3.780, 3.847, 3.924, 2.132, 2.752, 2.449, 3.456, 3.073, 2.688, 3.329  
, 4.271, 3.530, 2.928, 2.689, 2.332, 2.934, 3.110, 2.894, 2.435, 2.838, 3.035  
, 4.831, 2.812, 2.714, 3.086, 3.519, 4.232, 2.770, 3.341, 3.090, 2.531, 2.822  
, 2.935, 2.568, 2.387, 2.499, 4.130, 3.001, 3.132, 3.577, 3.222, 3.280, 2.659  
, 2.822, 2.140, 4.203, 2.997, 2.562, 3.082, 3.806, 2.458, 2.391, 3.141, 2.579  
, 2.100, 2.785, 4.284, 2.906, 5.102, 4.429, 4.279, 4.500, 2.635, 3.082, 3.387  
, 5.793, 3.985, 4.220, 4.724, 3.731, 3.500, 3.674, 5.633, 3.645, 2.887, 3.960  
, 4.299, 2.981, 4.504, 5.638, 2.853, 3.211, 1.953, 2.236, 3.428, 3.208, 1.694  
, 3.957, 4.789, 2.384, 3.074, 2.387, 3.835, 2.599, 4.756, 3.086, 4.309, 3.413  
, 2.975, 3.169, 3.343, 3.751, 2.216, 3.078, 3.186, 3.297, 2.304, 3.102, 2.677  
, 3.297, 3.498, 2.759, 2.953, 3.785, 2.276, 4.637, 3.038, 3.120, 3.339, 3.152  
, 3.104, 4.045, 4.763, 3.069, 4.506, 3.519, 3.688, 2.679, 2.198, 3.345, 3.082  
, 2.903, 3.004, 3.406, 3.122, 3.330, 2.608, 3.799, 4.086, 4.070, 2.264, 4.404  
, 2.278, 4.872, 4.270, 3.727, 2.795),  
age=c(9, 8, 7, 9, 9, 8, 6, 6, 8, 9, 6, 8, 8, 8, 8, 7, 5, 6  
, 9, 9, 5, 5, 4, 7, 9, 3, 9, 5, 8, 9, 5, 9, 8, 7, 5, 8  
, 9, 8, 8, 8, 9, 8, 5, 8, 5, 9, 7, 8, 6, 8, 5, 9, 9, 8  
, 6, 9, 9, 7, 4, 8, 8, 8, 6, 4, 8, 6, 9, 7, 5, 9, 8, 8  
, 9, 9, 9, 7, 5, 5, 9, 6, 7, 6, 8, 8, 7, 8, 7, 9, 5, 9  
, 9, 9, 7, 8, 8, 9, 9, 9, 7, 8, 8, 7, 9, 4, 9, 6, 8, 6  
, 7, 7, 8, 7, 7, 7, 7, 8, 7, 5, 8, 7, 9, 7, 7, 6, 8, 8  
, 8, 9, 7, 8, 9, 8, 8, 9, 8, 6, 6, 8, 9, 5, 7, 9, 6, 9  
, 9, 9, 6, 8, 9, 8, 8, 9, 9, 9, 7, 8, 6, 9, 9, 9, 7, 8  
, 5, 8, 9, 6, 9, 6, 8, 5, 7, 7, 4, 9, 8, 9, 9, 9, 5, 9  
, 7, 6, 9, 9, 9, 7, 5, 8, 9, 7, 8, 9, 6, 6, 8, 9, 5, 6  
, 6, 9, 7, 9, 8, 5, 7, 6, 9, 7, 9, 9, 8, 9, 7, 9, 4, 9  
, 5, 8, 9, 8, 3, 9, 8, 6, 9, 8, 8, 7, 6, 8, 9, 4, 7, 8  
, 8, 9, 6, 8, 6, 8, 9, 8, 7, 9, 8, 7, 9, 8, 9, 6, 8, 9  
, 8, 9, 9, 8, 7, 5, 7, 8, 9, 9, 6, 8, 7, 9, 7, 7, 5, 9  
, 9, 8, 8, 9, 6, 7, 5, 9, 5, 7, 6, 8, 7, 8, 4, 8, 5, 8  
, 7, 7, 9, 9, 8, 9, 6, 8, 9, 4, 6, 7, 9, 8, 6, 8, 7, 5  
, 8, 7, 11, 10, 14, 11, 11, 12, 10, 11, 10, 14, 13, 14, 12, 12, 10, 13  
, 10, 11, 10, 11, 10, 13, 11, 10, 11, 13, 10, 10, 12, 10, 10, 10, 11, 11  
, 11, 10, 11, 11, 13, 13, 11, 11, 14, 11, 10, 10, 10, 13, 10, 14, 10, 11  
, 13, 12, 10, 13, 11, 11, 11, 10, 11, 11, 10, 11, 13, 10, 10, 11, 10, 11  
, 11, 13, 13, 10, 11, 11, 12, 10, 10, 11, 10, 11, 14, 13, 11, 11, 11, 14  
, 12, 10, 12, 11, 10, 11, 10, 10, 11, 13, 10, 11, 10, 11, 10, 11, 11, 14  
, 11, 13, 11, 11, 10, 13, 10, 10, 12, 10, 12, 11, 14, 12, 10, 10, 10, 12  
, 13, 11, 12, 11, 12, 11, 12, 12, 13, 11, 10, 12, 13, 10, 12, 10, 12, 10  
, 11, 10, 14, 10, 10, 12, 10, 10, 12, 12, 11, 10, 11, 11, 12, 13, 13, 10  
, 12, 14, 10, 11, 11, 13, 12, 10, 10, 12, 11, 10, 11, 11, 11, 11, 11, 14  
, 12, 13, 13, 12, 10, 10, 12, 11, 11, 12, 12, 14, 11, 10, 11, 12, 13, 12  
, 11, 11, 11, 14, 11, 12, 10, 12, 13, 10, 10, 10, 10, 14, 12, 11, 11, 12  
, 14, 10, 11, 10, 10, 12, 12, 11, 12, 10, 12, 13, 10, 12, 10, 13, 12, 12





