Robust Linear Regression Models: Use of a Stable Distribution for the Response Data

Jorge A. Achcar[*], Angela Achcar, Edson Zangiacomi Martinez

Medical School, University of São Paulo, Ribeirão Preto, Brazil
Email: [*]achcar@fmrp.usp.br

ABSTRACT

In this paper, we study some robustness aspects of linear regression models of the presence of outliers or discordant observations considering the use of stable distributions for the response in place of the usual normality assumption. It is well known that, in general, there is no closed form for the probability density function of stable distributions. However, under a Bayesian approach, the use of a latent or auxiliary random variable gives some simplification to obtain any posterior distribution when related to stable distributions. To show the usefulness of the computational aspects, the methodology is applied to two examples: one is related to a standard linear regression model with an explanatory variable and the other is related to a simulated data set assuming a 2^3 factorial experiment. Posterior summaries of interest are obtained using MCMC (Markov Chain Monte Carlo) methods and the OpenBugs software.

Keywords: Stable Distribution; Bayesian Analysis; Linear Regression Models; MCMC Methods; OpenBugs Software

1. Introduction

A wide class of distributions that encompass the Gaussian distribution is given by the class of stable distributions. This large class defines location-scale families that are closed under convolution. The Gaussian distribution is a special case of this distribution family (see for instance, [1]), described by four parameters α, β, γ and δ. The $\alpha \in (0,2]$ parameter defines the "fatness of the tails". When $\alpha = 2$, this class reduces to Gaussian distributions. The $\beta \in [-1,1]$ is the skewness parameter and for $\beta = 0$ one has symmetric distributions. The location and scale parameters are, respectively, $\gamma \in (-\infty, \infty)$ and $\delta \in (0, \infty)$ (see [2]).

Stable distributions are usually denoted by $S_\alpha(\beta, \gamma, \delta)$. If a random variable $X \sim S_\alpha(\beta, \gamma, \delta)$, then

$$Z = \frac{X - \gamma}{\delta} \sim S_\alpha(\beta, 0, 1)$$

follows a standardized stable distribution (see [3] and [4]).

A great difficulty associated to stable distributions $S_\alpha(\beta, \gamma, \delta)$, is that in general there is no simple closed form for their probability density functions. However, it is known that the probability density functions of stable distributions are continuous ([5,6]) and unimodal ([7,8]). Also 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 $\beta = -1$ (see [9]).

The characteristic function $\Phi(.)$ of a stable distribution is given by,

$$\log[\Phi(t)]$$
$$= \begin{cases} i\gamma t - |\delta t|^\alpha \left[1 - i\beta \text{sign}(t) \tan\left(\frac{\pi\alpha}{2}\right)\right], \text{for } \alpha \neq 1 \\ i\gamma t - |\delta t|^\alpha \left[1 + i\beta \text{sign}(t)\left(\frac{2}{\pi}\right)\log(t)\right], \text{for } \alpha = 1 \end{cases} \quad (1.1)$$

where $i = \sqrt{-1}$ and sign (.) function is given by

$$\text{sign}(x) = \begin{cases} -1, \text{if } x < 0 \\ 0, \text{ if } x = 0 \\ 1, \text{ if } x > 0. \end{cases}$$

It is important to point out that if $\alpha < 1$, the variance is infinite and the mean of the stable distribution does not

exist.

Although this class of distributions is a good alternative for data modeling in different areas, we usually have difficulties to obtain estimates under a classical inference approach due to the lack of closed form expressions for their probability density functions. One possibility in applications is to get the probability density function from the inversion formula (see, for example [10]),

$$f(x) = 1/2\pi \int_{-\infty}^{\infty} e^{-itx} \Phi(t) dt, \qquad (1.2)$$

where $\Phi(t)$ is the characteristic function. In applications, we need use numerical methods to solve the integral in (1.2), usually taking a great computational time.

An alternative is the use of Bayesian methods. However, the computational cost can be further high to get the posterior summaries of interest. A good alternative is to use latent or artificial variables that could improve the simulation computation of samples of the joint posterior distributions of interest (see, for example [11,12]).

In this way, a Bayesian analysis of stable distributions was introduced by [1] using Markov Chain Monte Carlo (MCMC) methods and latent variables (see also, [13]). The use of Bayesian methods with MCMC simulation can have great flexibility by considering latent variables where samples of latent variables are simulated in each step of the Gibbs or Metropolis-Hastings algorithms(see for example [14,15]).

Considering a latent or an auxiliary variable, [1] proved a theorem that is useful to simulate samples of the joint posterior distribution for the parameters α, β, δ and σ. 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 random variable Y. This 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

$$a_{\alpha,\beta} = -\frac{b_{\alpha,\beta}}{\alpha\pi}, \qquad (1.3)$$

where $b_{\alpha,\beta} = \beta \min\{\alpha, 2-\alpha\} \dfrac{\pi}{2}$.

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

$$f(z, y | \alpha, \beta)$$

$$= \frac{\alpha}{|\alpha-1|} \exp\left[-\left|\frac{z}{t_{\alpha,\beta}(y)}\right|^{\theta}\right] \left|\frac{z}{t_{\alpha,\beta}(y)}\right|^{\theta} \frac{1}{|z|}, \qquad (1.4)$$

where $\theta = \dfrac{\alpha}{\alpha-1}$,

$$t_{\alpha,\beta}(y)$$

$$= \left(\frac{\sin(\pi\alpha y + b_{\alpha,\beta})}{\cos(\pi y)}\right)\left(\frac{\cos(\pi y)}{\cos(\pi(\alpha-1)y + b_{\alpha,\beta})}\right)^{\frac{1}{\theta}} \qquad (1.5)$$

and $Z = \dfrac{X-\delta}{\sigma}$, for $\sigma \neq 0$.

From the bivariate density (1.4), [1] showed the marginal distribution for the random variable Z is stable $S_\alpha(\beta, 0, 1)$ distributed. Usually, the computational costs to obtain posterior summaries of interest using MCMC methods are high for this class of models, which could give some limitations for practical applications. One problem can be the simulation algorithm convergence. In this paper, we propose the use of the popular free available software to obtain the posterior summaries of interest: the OpenBUGS software.

The paper is organized as follows: in Section 2, we introduce linear regression models assuming a stable distribution; in Section 3, we introduce a Bayesian analysis for linear regression models assuming a stable distribution; in Section 4, we present two examples to illustrate the proposed methodology; finally, in Section 5, we present some concluding remarks.

2. Linear Regression Models Assuming a Stable Distribution

Consider a random variable X related to a controlled variable v given by the linear relationship,

$$x_i = d_0 + d_1 v_i + \varepsilon_i, \qquad (2.1)$$

for $i = 1, 2, \cdots, n$, where,

1) The random variable X_i represents the response for the i-th unit associated with an experimental value of the independent or explanatory variable v_i assumed as a fixed value. In this way, x_i is an observation of X_i.

2) The variables $\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_n$ are considered as components of unknown errors as unobserved random variables. Assume that these random variables ε_i are independent and identically distributed with normal distribution $N(0, \sigma^2)$.

3) The parameters d_0 and d_1 are unknown.

From the above assumptions, we have normality for the responses, that is,

$$X_i \sim N\{d_0 + d_1 v_i; \sigma^2\}. \qquad (2.2)$$

In this way X_i has a normal distribution with mean $d_0 + d_1 v_i$ and common variance σ^2.

Usually we get estimators for the regression parameters using minimum squares approach or standard maximum likelihood methods (see for example, [16,17]).

Standard generalization for the linear model (2.1) is given in presence of k independent or explanatory vari-

ables, that is, a multiple linear regression model, given by,

$$x_i = d_0 + d_1 v_{1i} + d_2 v_{2i} + \cdots + d_k v_{ki} + \varepsilon_i. \qquad (2.3)$$

From the normality assumption for the error ε_i in (2.3), the random variable X_i has a normal distribution with mean $d_0 + d_1 v_{1i} + d_2 v_{2i} + \cdots + d_k v_{ki}$ and variance σ^2.

In practical work, in all applications we need to check if the above assumptions are verified. In this way, we consider graphical approaches to verify if the residuals of the model satisfy the above assumptions.

In presence of outliers or discordant observations we could have great effects on the obtained estimators for the regression model given by (2.3) which could invalidate the obtained inferences. In this way, we could use non-parametric regression models or to assume more robust probability distributions for the data. One possibility is to assume that the random variable X in (2.3) or (2.1) have a stable distribution $S_\alpha(\beta, \gamma, \delta)$.

3. A Bayesian Analysis for Linear Regression Models Assuming a Stable Distribution

In this section, let us assume that the response x_i in the linear regression model (2.3) for $i = 1, \cdots, n$, have a stable distribution $X_i \sim S_\alpha(\beta, \gamma_i, \delta)$, that is,

$$Z_i = \frac{X - \gamma_i}{\delta} \sim S_\alpha(\beta, 0, 1)$$

and where the location parameter γ_i of the stable distribution is related to the explanatory variables by a linear relation given by,

$$\gamma_i = \beta_0 + \beta_1 v_{1i} + \beta_2 v_{2i} + \cdots + \beta_k v_{ki}. \qquad (3.1)$$

Assuming a joint prior distribution for α, β, d and δ, where $d = (d_0, d_1, d_2, \cdots, d_k)$ given by $\pi_0(\alpha, \beta, d, \delta)$, [1] shows that the joint posterior distribution for parameters α, β, d and δ, is given by,

$$\pi_0(\alpha, \beta, d, \delta/x)$$

$$\propto \int \left(\frac{\alpha}{|\alpha - 1|\delta} \right)^n \exp\left\{ -\sum_{i=1}^n \left| \frac{z_i}{t_{\alpha,\beta}(Y_i)} \right|^\theta \right\} \qquad (3.2)$$

$$\prod_{i=1}^n \left| \frac{z_i}{t_{\alpha,\beta}(y_i)} \right|^\theta \cdot \frac{1}{|z_i|} \pi_0(\alpha, \beta, d, \delta) \mathrm{d}y$$

where $\theta = \dfrac{\alpha}{\alpha - 1}$, $z_i = \dfrac{x_i - \gamma_i}{\delta}$, for $i = 1, \cdots, n$, $\alpha \in (0, 2]$, $\beta \in [-1, 1]$ and $\delta \in (0, \infty)$; $x = (x_1, x_2, \cdots, x_n)$ and $y = (y_1, y_2, \cdots, y_n)$ are respectively, the observed and non-observed data vectors. Observe that the bivariate distribution in expression (3.2) is given in terms of x_i

and the latent variables y_i, and not in terms of z_i and y_i (there is the Jacobian σ^{-1} multiplied by the right-hand-side of expression (1.4)).

Observe that when $\alpha = 2$ we have $\theta = 2$ and $b_{\alpha,\beta} = 0$. In this case we have a Gaussian distribution with mean equals to δ and variance equals to $2\sigma^2$.

For a Bayesian analysis of the proposed model, we assume uniform $U(a, b)$ priors for α, β and δ where the hyperparameters a and b are assumed to be known in each application following the restrictions $\alpha \in (0, 2]$, $\beta \in [-1, 1]$ and $\delta \in (0, \infty)$. We also assume Normal $N(a, b^2)$ prior distributions for the regression parameters $d_0, d_1, d_2, \cdots, d_k$ considering known hyperparameter values a and b^2. We further assume independence among all parameters.

In the simulation algorithm to obtain a Gibbs sample for the random quantities α, β, d and δ, having the joint posterior distribution (3.2), we assume a uniform $U(-0.5, 0.5)$ prior distribution for the latent random quantities Y_i for $i = 1, \cdots, n$. Observe that, in this case, we are assuming $a_{\alpha,\beta} = 0 (b_{\alpha,\beta} = 0)$. With this choice of priors, we use standard available software packages like OpenBugs (see [18]) which gives great simplification to obtain the simulated Gibbs samples for the joint posterior distribution of interest.

From expression (3.2), the joint posterior probability distribution for α, β, d, δ and $y = (y_1, y_2, \cdots, y_n)$ is given by,

$$\pi(\alpha, \beta, d, \delta, y | x)$$

$$\propto \left(\frac{\alpha}{|\alpha - 1|\delta} \right)^n \exp\left\{ \sum_{i=1}^n \left| \frac{z_i}{t_{\alpha,\beta}(y_i)} \right| \right\} \prod_{i=1}^n \left| \frac{z_i}{t_{\alpha,\beta}(y_i)} \right|^\theta \qquad (3.3)$$

$$\frac{1}{|z_i|} \prod_{i=1}^n h(y_i) \pi_0(\alpha, \beta, d, \delta),$$

where θ and $t_{\alpha,\beta}(.)$ are respectively defined in (1.4) and (1.5) and $h(y_i)$ is a $U(-0.5, 0.5)$ density function, for $i = 1, \cdots, n$.

Since we are using the OpenBugs software to simulate samples for the joint posterior distribution we do not present here 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 for determining posterior summaries of interest as shown in the applications as follows.

In these applications we illustrate the proposed methodology to real and simulated data sets, in special, showing the robustness of the model in presence of outliers.

4. Some Applications

4.1. An Example with a Simple Linear Regression Model

In **Table 1**, we have a data set related to an industrial experiment, where x denotes the response and v denotes an explanatory variable associated to each response ($n = 15$ observations).

From a preliminary data analysis, we see that a linear regression model (2.1) is suitable for the data set. The estimated regression straight line obtained by minimum squares estimates and using the software MINITAB version 16 is given by $x_i = 0.603 + 0.0444v_i$ where the regression parameter d_1 is statistically different of zero (p-value < 0.05). From standard residuals plots we verify that the required assumptions (normality of the residuals and constancy of the variance are verified).

Under a Bayesian approach, we have in **Table 2**, the posterior summaries of interest assuming the linear regression model defined by (3.1) with a stable distribution for the response and the OpenBugs software assuming the following prior distributions: $\alpha \sim U(0,2)$, $\beta \sim U(-1,0)$, $\delta \sim U(0,3)$, $d_0 \sim N(0,1)$ and $d_1 \sim N(0,1)$. We also assume a uniform $U(-0.5, 0.5)$ distribution for the latent variable Y_i, $i = 1, 2, \cdots, 15$. We simulated 800,000 Gibbs samples, with a "burn-in-sample" of 300,000 samples discarded to eliminate the effects of the initial values in the iterative simulation process and taking a final sample of size 1,000 (every 500th sample chosen from the

Table 1. An industrial experiment data set.

Row	x	v
1	1.2	19
2	1.5	15
3	1.5	35
4	3.3	52
5	2.5	35
6	2.1	33
7	2.5	30
8	3.2	57
9	2.8	49
10	1.5	26
11	2.2	45
12	2.2	39
13	1.9	25
14	1.8	40
15	2.8	40

Table 2. Posterior summaries (linear regression).

Stable distribution (SAV = 4.496)			
Parameter	Mean	Standard Deviation	95% Credible Interval
α	1,716	0.2123	(1.2380, 1.9890)
β	−0.6183	0.2391	(−0.9881, −0.07191)
δ	0.2931	0.06375	(0.1960, 0.446)
d_0	0.5431	0.3559	(−0.1981, 1.2270)
d_1	0.04602	0.009251	(0.02831, 0.06521)
Normal distribution (SAV = 4.504)			
Parameter	Mean	Standard Deviation	95% Credible Interval
d_0	0.4802	0.4927	(−0.56587, 1.4230)
d_1	0.04747	0.01312	(0.02132, 0.07369)
ζ	0.4949	0.1709	(0.2119, 0.8834)

500,000 samples). Convergence of the Gibbs sampling algorithm was monitored from standard trace plots of the simulated samples.

In **Table 2**, we also have the sum of absolute values between (SAV) the observed and fitted values, given by,

$$SAV = \sum_{i=1}^{n} abs\left[\text{observed}(i) - \text{fitted mean}(i) \right] \quad (4.1)$$

In **Table 2**, we also have the posterior summaries of the regression model (2.1) assuming a normal $N(0, \sigma^2)$ distribution for the error and the following priors for the parameters of the model: $\zeta = 1/\sigma^2 \sim U(0,3)$, $d_0 \sim N(0,1)$ and $d_1 \sim N(0,1)$. In this case, we simulated 55,000 Gibbs samples taking a "burn-in-sample" of size 5,000 using the OpenBugs software. From the results of **Table 2**, we observe similar results assuming normality or a stable distribution for the data. In this case, we conclude that we do not need to assume a stable distribution for the data, since the results are very similar assuming usual normality for the errors and the computational cost using stable distributions is very high.

In **Figure 1**, we have the plots of observed, fitted means considering both models versus samples. From the plots of **Figure 1**, we observe similar fit of both models (linear regression model assuming normality or a stable distribution). Observe that we have SAV = 4.496 assuming a stable distribution and SAV = 4.504 assuming a normal distribution, that is, very close results.

Now let us consider the presence of an outlier or discordant response (considered as a measure error) replacing the 15th response (2.8) in **Table 1** by the value 8.0 (denoted as outlier 1). In **Table 3**, we have the obtained posterior summaries assuming the same priors and simulation procedure assumed for the results of **Table 2**. In **Figure 2**, we have the plots of observed, fitted means
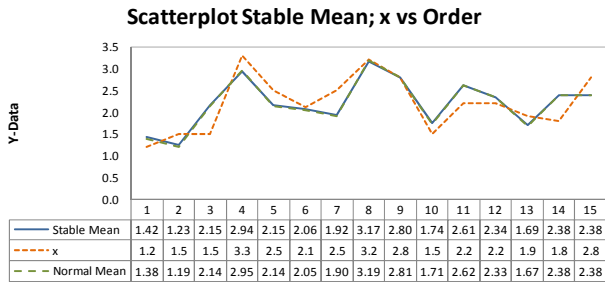
Scatterplot Stable Mean; x vs Order

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Stable Mean	1.42	1.23	2.15	2.94	2.15	2.06	1.92	3.17	2.80	1.74	2.61	2.34	1.69	2.38	2.38
x	1.2	1.5	1.5	3.3	2.5	2.1	2.5	3.2	2.8	1.5	2.2	2.2	1.9	1.8	2.8
Normal Mean	1.38	1.19	2.14	2.95	2.14	2.05	1.90	3.19	2.81	1.71	2.62	2.33	1.67	2.38	2.38

Figure 1. Plots of observed, fitted means considering both models versus samples.

Table 3. Posterior summaries (example 1, with outlier 1).

Stable distribution (SAV = 4.623)			
Parameter	Mean	Standard Deviation	95% Credible Interval
α	1,373	0.2195	(1.0580, 1.8430)
β	−0.5726	0.2703	(−0.9821, −0.04255)
δ	0.3221	0.08016	(0.2004, 0.4856)
d_0	0.4797	0.3461	(−0.3483, 1.0400)
d_1	0.0465	0.009595	(0.02922, 0.07010)
Normal distribution (SAV = 6.394)			
Parameter	Mean	Standard Deviation	95% Credible Interval
d_0	0.2408	0.7866	(−1.3100, 1.8200)
d_1	0.06316	0.02246	(0.0201, 0.1061)
ζ	0.4949	0.1709	(0.2119, 0.8834)

Scatterplot of Stable Mean outlier1; x vs Order

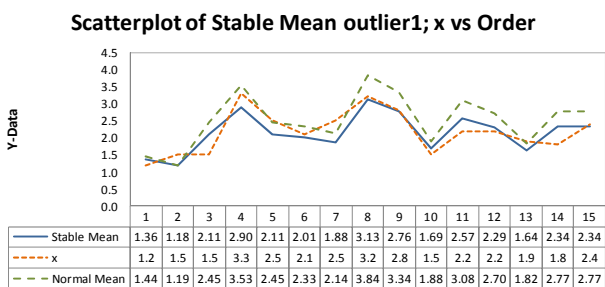|	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Stable Mean	1.36	1.18	2.11	2.90	2.11	2.01	1.88	3.13	2.76	1.69	2.57	2.29	1.64	2.34	2.34
x	1.2	1.5	1.5	3.3	2.5	2.1	2.5	3.2	2.8	1.5	2.2	2.2	1.9	1.8	2.4
Normal Mean	1.44	1.19	2.45	3.53	2.45	2.33	2.14	3.84	3.34	1.88	3.08	2.70	1.82	2.77	2.77

Figure 2. Plots of observed, fitted means considering both models (presence of outlier 1).

considering both models versus samples. From the plots of **Figure 2**, we observe that model with a stable distribution is very robust to the presence of the outlier given similar inference results as obtained without the presence of the outlier (see results in **Table 2**). We also observe in **Table 3**, that the estimated regression parameters with normal error are strongly affected by the presence of the outlier 1. Observe that we have SAV = 4.623 assuming a stable distribution (a value very close to the SAV values given in **Table 2**, without the presence of an outlier) and SAV = 6.394 assuming a normal distribution.

Similarly, we have in **Table 4** the posterior summaries assuming another outlier replacing the 15th response (2.8) in **Table 1** by the value 50.0 (denoted as outlier 2). In **Figure 3**, we have the plots of observed, fitted means considering both models versus samples for this case.

From the plots of **Figure 3**, we observe that model with a stable distribution is very robust to the presence of outliers even considering a large discordant observation (outlier 2). We also observe in **Table 4**, that the estimated regression parameters of regression model with normal error are strongly affected by the presence of the outlier 2. Observe that we have SAV = 4.632 assuming a stable distribution (a value very close to the values given in **Table 2**, without the presence of an outlier) and SAV = 46.083 assuming a normal distribution, that is, with very large values of outliers the obtained inferences are greatly affected assuming normality for the data.

4.2. An example with Simulated Data (5 Replicates of a Factorial 2^3 Experiment)

In **Table 5**, we have 40 responses simulated from a factorial 2^3 experiment with 5 replicates assuming the

Table 4. Posterior summaries (example 1, with outlier 2).

Stable distribution, outlier 2 (SAV = 4.632)			
Parameter	Mean	Standard Deviation	95% Credible Interval
α	1,351	0.1920	(1.0790, 1.7820)
β	−0.5289	0.2733	(−0.9705, −0.0344)
δ	0.3380	0.09538	(0.1943, 0.5680)
d_0	0.5373	0.3477	(−0.2225, 1.2040)
d_1	0.04513	0.009471	(0.02728, 0.06528)
Normal distribution, outlier 2 (SAV = 46.083)			
Parameter	Mean	Standard Deviation	95% Credible Interval
d_0	0.00849	1.0020	(−1.9710, 1.8500)
d_1	0.1462	0.08654	(−0.02591, 0.3275)
ζ	0.007598	0.0026	(0.003284, 0.01354)

Scatterplot of Stable Mean outlier2; x vs r

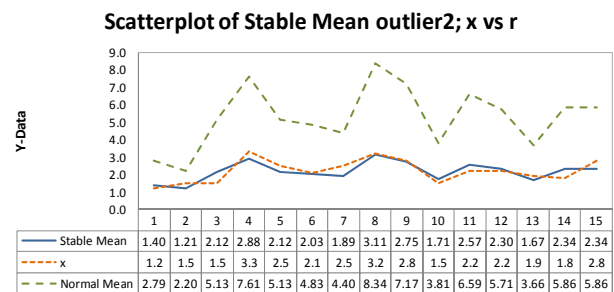|	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Stable Mean	1.40	1.21	2.12	2.88	2.12	2.03	1.89	3.11	2.75	1.71	2.57	2.30	1.67	2.34	2.34
x	1.2	1.5	1.5	3.3	2.5	2.1	2.5	3.2	2.8	1.5	2.2	2.2	1.9	1.8	2.8
Normal Mean	2.79	2.20	5.13	7.61	5.13	4.83	4.40	8.34	7.17	3.81	6.59	5.71	3.66	5.86	5.86

Figure 3. Plots of observed, fitted means considering both models (presence of outlier 2).

following model,

$$x_i = d_0 + d_1 v_{1i} + d_2 * v_{2i} + d_3 * v_{3i} + \varepsilon_i, , \qquad (4.2)$$

where $d_0 = 2$, $d_1 = 1$, $d_2 = 0.5$, $d_3 = 1.8$ and ε_i are independent and identically distributed errors with a normal distribution N (0, 1).

The estimated regression obtained by minimum squares estimates and using the software MINITAB version 16 is given by $x_i = 2.19 + 0.964 v_{1i} + 0.828 v_{2i} + 2.08 v_{3i}$ where the regression parameters d_1, d_2, and d_3 are statistically different of zero (p-value < 0.05 for all regression parameters). From standard residuals plots we verify that the required assumptions (normality of the residuals and constancy of the variance are verified).

Under a Bayesian approach, we have in **Table 6**, the posterior summaries of interest assuming the linear regression model defined by (3.1) with a stable distribution for the response and the OpenBugs software assuming the following prior distributions: $\alpha \sim U(0,2)$, $\beta \sim U(-1,0)$, $\delta \sim U(0,3)$, $d_0 \sim N(0,1)$ and $d_0 \sim N(2,10)$, $d_1 \sim N(1,10)$, $d_2 \sim N(0.5,10)$ and $d_3 \sim N(1.5,10)$ .We also assume a uniform $U(-0.5,0.5)$ distribution for the latent variable Y_i, $i=1,2,\cdots,15$. We simulated 400,000 Gibbs samples, with a "burn-in-sample" of 100,000 samples discarded to eliminate the effects of the initial values in the iterative simulation process and taking a final sample of size 1,500 (every 200$^{\text{th}}$ sample chosen from the 300,000 samples). Convergence of the Gibbs sampling algorithm was monitored from standard trace plots of the simulated samples**.** From the results of **Table 6**, we observe that the parameter α has a Monte Carlo estimate for the posterior mean for α given by 0.9089, that is, the mean of the stable distribution is not defined in this case ($\alpha < 1$).

In **Table 6**, we also have the posterior summaries of the regression model (4.2) assuming a normal $N(0,\sigma^2)$ distribution for the error and the following priors for the parameters of the model: $\zeta = 1/\sigma^2 \sim \text{Gamma}(0.01,0.01)$, where Gamma (a,b) denotes a gamma distribution with mean a/b and variance a/b^2 , $d_0 \sim N(2,1000)$, $d_1 \sim N(0,100)$, $d_2 \sim N(0,100)$ and $d_3 \sim N(0,100)$. In this case, we simulated 55,000 Gibbs samples taking a "burn-in-sample" of size 5,000 using the OpenBugs software. From the results of **Table 6**, we observe similar results for the Bayesian estimates of the regression parameters assuming normality or a stable distribution for the data.

Now let us assume the presence of an outlier or discordant response (measure error) replacing the 30$^{\text{th}}$ response (5.04875) in **Table 5**, by the value 25.250. In **Table 7**, we have the obtained posterior summaries assuming the same priors and simulation procedure assumed for the results of **Table 6**. We observe in **Table 7**, that the estimated regression parameters assuming normal error are strongly affected by the presence of the outlier.

Table 5. A simulated data set.

Row	response	v_1	v_2	v_3
1	−1.39519	−1	−1	−1
2	−0.39840	−1	−1	−1
3	−1.51516	−1	−1	−1
4	−1.63693	−1	−1	−1
5	−2.52282	−1	−1	−1
6	1.04163	1	−1	−1
7	−1.20421	1	−1	−1
8	1.63848	1	−1	−1
9	−0.79469	1	−1	−1
10	−0.66038	1	−1	−1
11	0.15325	−1	1	−1
12	−1.44337	−1	1	−1
13	−0.02142	−1	1	−1
14	−0.84649	−1	1	−1
15	−0.02497	−1	1	−1
16	3.25672	1	1	−1
17	2.39574	1	1	−1
18	1.26225	1	1	−1
19	1.61358	1	1	−1
20	3.37443	1	1	−1
21	3.26868	−1	−1	1
22	4.09273	−1	−1	1
23	1.14569	−1	−1	1
24	2.84898	−1	−1	1
25	2.57364	−1	−1	1
26	4.43130	1	−1	1
27	3.45746	1	−1	1
28	4.88136	1	−1	1
29	2.98939	1	−1	1
30	5.04875	1	−1	1
31	4.54937	−1	1	1
32	2.58553	−1	1	1
33	4.18825	−1	1	1
34	5.12804	−1	1	1
35	3.82664	−1	1	1
36	5.12793	1	1	1
37	7.41328	1	1	1
38	4.60123	1	1	1
39	6.52666	1	1	1
40	6.73205	1	1	1

Table 6. Posterior summaries (simulated data).

Stable distribution			
Parameter	Mean	Standard Deviation	95% Credible Interval
α	0.9089	0.06488	(0.7590, 0.9767)
β	−0.7971	0.08412	(−0.9629, −0.6583)
δ	0.8275	0.07882	(0.7016, 0.9702)
d_0	1.8310	0.0984	(1.6650, 2.053)
d_1	1.1000	0.09937	(0.8604, 1.2410)
d_2	0.8113	0.05056	(0.7089, 0.9084)
d_3	2.0210	0.07688	(1.8700, 2.1790)
Normal distribution			
Parameter	Mean	Standard Deviation	95% Credible Interval
d_0	2.0460	0.1724	(1.7090, 2.4090)
d_1	0.8094	0.1804	(0.4643, 1.1670)
d_2	0.6862	0.1810	(0.3318, 1.0560)
d_3	2.0220	0.1804	(1.6680, 2.3850)
ζ	0.8654	0.2121	(0.5129, 1.3540)

Table 7. Posterior summaries (simulated data in the presence of an outlier).

Stable distribution			
Parameter	Mean	Standard Deviation	95% Credible Interval
α	0.9020	0.05369	(0.7718, 0.9772)
β	−0.8323	0.06828	(−0.9630, -0.6798)
δ	0.8867	0.07621	(0.7061, 0.9941)
d_0	1.8130	0.08077	(1.6720, 2.006)
d_1	1.1130	0.08735	(0.9105, 1.2640)
d_2	0.8002	0.07881	(0.6454, 0.9422)
d_3	2.0430	0.08342	(1.8770, 2.2080)
Normal distribution			
Parameter	Mean	Standard Deviation	95% Credible Interval
d_0	2.5170	0.5793	(1.3590, 3.6580)
d_1	1.2580	0.5634	(0.2024, 2.322)
d_2	0.1234	0.5635	(−1.040, 1.247)
d_3	2.4880	0.5664	(1.375, 3.605)
ζ	0.0811	0.0184	(0.0478, 0.1185)

5. Concluding Remarks

The presence of outliers or discordant observations is due to the measure errors many times. This situation is very common in applications of the regression analysis. In the presence of these discordant observations, the usual obtained inferences on the regression parameters or in the predictions based on minimum squares approach or maximum likelihood approach under the usual assumption of normality for the errors and constant variance could be greatly affected, which could imply in wrong inference results. The use of stable distributions could be a good alternative for many applications in the data analysis to have robust inference results, since this distribution has a great flexibility to fit for the data. With the use of Bayesian methods and MCMC simulation algorithms, it is possible to get inferences for the model despite of the nonexistence of an analytical form for the density function as it was showed in this paper. It is important to point out that the computational work in the sample simulations for the joint posterior distribution of interest can be greatly simplified using standard free softwares like the Open-Bugs software.

Observe that in general, the appearance of outliers will absolutely affect the regression model under standard normality assumptions. The ideal results not affected by outliers could be obtained using the proposed methodology as observed in our applications. These results could be of great interest in applications.

In both examples introduced in Section 4, the use of data augmentation techniques (see, for instance, [11]) is the key to obtain a good performance for the MCMC simulation method for applications using stable distributions.

We emphasize that the use of OpenBugs software does not require large computational time to get the posterior summaries of interest, even when the simulation of a large number of Gibbs samples is needed for the algorithm convergence. These results could be of great interest for researchers and practitioners, when dealing with non-Gaussian data, as in the applications presented here.

6. Acknowledgements

REFERENCES

[1] D. J. Buckle, "Bayesian Inference for Stable Distributions," *Journal of the American Statistical Association*, Vol. 90, 1995, pp. 605-613.
http://dx.doi.org/10.1080/01621459.1995.10476553

[2] P. Lévy, "Théorie des Erreurs la loi de Gauss et les Lois Exceptionelles," *Bulletin Society Mathematical*, Vol. 52, 1924, pp. 49-85.

[3] E. Lukacs, "Characteristic Functions," Hafner Publishing, New York, 1970.

[4] J. P. Nolan, "Stable Distributions—Models for Heavy Tailed Data," Birkhäuser, Boston, 2009. In Progress, Chapter 1.

academic2.american.edu/~jpnolan

[5]   B. V. Gnedenko and A. N. Kolmogorov, "Limit Distributions for Sums of Independent Random Variables," Addison-Wesley, Massachussetts, 1968.

[6]   A. V. Skorohod, "On a Theorem Concerning Stable Distributions," In: *Selected Translations in Mathematical Statistics and Probability*, Vol. 1, Institute of Mathematical Statistics and American Mathematical Society, Providence, 1961, pp. 169-170.

[7]   I. A. Ibragimov and K. E. Černin, "On the Unimodality of Stable Laws," *Teoriya Veroyatnostei i ee Primeneniya*, Vol. 4, No. 1, 1959, pp. 453-456.

[8]   M. Kanter, "On the Unimodality of Stable Densities," *Annals of Probability*, Vol. 4, No. 6, 1976, pp. 1006-1008. http://dx.doi.org/10.1214/aop/1176995944

[9]   W. Feller, "An Introduction to Probability Theory and Its Applications," Vol. II, John Wiley, New York, 1971.

[10] G. G. Roussas, "Measure-Theoretic Probability," Elsevier, San Diego, 2005.

[11] P. Damien, J. Wakefield and S. Walker, "Gibbs Sampling for Bayesian Non-Conjugate and Hierarchical Models by Using Auxiliary Variables," *Journal of the Royal Statistical Society*, *Series B*, Vol. 61, No. 2, 1999, pp. 331-344. http://dx.doi.org/10.1111/1467-9868.00179

[12] M. A. Tanner and W. H. Wong, "The Calculation of Posterior Distributions by Data Augmentation," *Journal of American Statistical Association*, Vol. 82, No. 398, 1987, pp. 528-550. http://dx.doi.org/10.1080/01621459.1987.10478458

[13] J. A. Achcar, S. R. C. Lopes, J. Mazucheli and R. Linhares, "A Bayesian Approach for Stable Distributions: Some Computational Aspects," *Open Journal of Statistics*, Vol. 3, No. 4, 2013, pp. 268-277. http://dx.doi.org/10.4236/ojs.2013.34031

[14] S. Chib and E. Greenberger, "Understanding the Metropolis-Hastings Algorithm," *The American Statistician*, Vol. 49, No. 4, 1995, pp. 327-335.

[15] C. P. Robert and G. Casella, "Monte Carlo Statistical Methods," 2nd Edition, Springer-Verlag, New York, 2004. http://dx.doi.org/10.1007/978-1-4757-4145-2

[16] N. R. Draper and H. Smith, "Applied Regression Analysis," Wiley Series in Probability and Mathematical Statistics, 1981.

[17] G. A. F. Seber and A. J. Lee, "Linear Regression Analysis," 2nd Edition, Wiley Series in Probability and Mathematical Statistics, 2003.

[18] D. J. Spiegelhalter, A. Thomas, N. G. Best and D. Lunn, "WinBUGS User's Manual," MRC Biostatistics Unit, Cambridge, 2003.