A Bayesian Approach for Stable Distributions: Some Computational Aspects

J.A. Achcar[a1]; S.R.C. Lopes[b]; J. Mazucheli[c] and R.R. Linhares[b]

[a]*Medical School - USP - Ribeirão Preto, SP, Brazil*

[b]*Mathematical Institute - UFRGS - Porto Alegre, RS, Brazil*

[c]*Departament of Statistics - UEM - Maringá, PR, Brazil*

June 5, 2013

Abstract

In this work we study some computational aspects for the Bayesian analysis involving stable distributions. It is well known that, in general, there is no closed form for the probability density function of stable distributions. However, the use of a latent or auxiliary random variable facilitates 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 daily price returns of Abbey National shares, considered in [1], and the other is the length distribution analysis of coding and non-coding regions in a Homo sapiens chromosome DNA sequence, considered in [2]. Posterior summaries of interest are obtained using the OpenBUGS software.

Keywords: Stable Laws; Bayesian Analysis; DNA Sequences; MCMC Methods; OpenBUGS Software.

1 Introduction

A wide class of distributions that encompasses the Gaussian one is given by the class of stable distributions. This larger 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, δ and σ. The $\alpha \in (0,2]$ parameter defines the "fatness of the tails", and 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, $\delta \in (-\infty,\infty)$ and $\sigma \in (0,\infty)$ (see [3]).

Stable distributions are usually denoted by $S_\alpha(\beta,\delta,\sigma)$. If a random variable $X \sim S_\alpha(\beta,\delta,\sigma)$, then $Z = \frac{X-\delta}{\sigma} \sim S_\alpha(\beta,0,1)$ (see [4] and [5]).

The difficulty associated to stable distributions $S_\alpha(\beta,\delta,\sigma)$, is that in general there is no simple closed form for their probability density functions. However, it is known the probability density functions of stable distributions are continuous ([6]; [7]) and unimodal ([8]; [9]). 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 [10]).

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

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

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

$$sign(x) = \begin{cases} -1, & \text{if } x < 0 \\ 0, & \text{if } x = 0 \\ 1, & \text{if } x > 0. \end{cases} \tag{1.2}$$

[1]Corresponding author. E-mail: achcar@fmrp.usp.br

Although a good class for data modeling in different areas, one has difficulties to obtain estimates under a classical inference approach due to the lack of closed form expressions for their probability density functions. An alternative is the use of Bayesian methods. However, the computational cost can be further exacerbated in assessing posterior summaries of interest.

A Bayesian analysis of stable distributions is introduced by Buckle (1995) using Markov Chain Monte Carlo (MCMC) methods. The use of Bayesian methods with MCMC simulation can have great flexibility by considering latent variables (see, for instance, [11] and [12]), where samples of latent variables are simulated in each step of the Gibbs or Metropolis-Hastings algorithms.

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

$$a_{\alpha,\beta} = -\frac{b_{\alpha,\beta}}{\alpha\,\pi}, \tag{1.3}$$

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

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

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

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

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

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

From the bivariate density (1.4), [1] shows 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 is 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 a 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 a special case of the stable distributions, namely, the Lévy distribution. In Section 3 we introduce a Bayesian analysis for stable distributions. Two applications are presented in Section 4. Section 5 is devoted to some concluding remarks.

2 A Special Case of Stable Distributions: Lévy Distribution

Some special cases of stable distributions are given for specified values of α and β. If $\alpha = 2$ and $\beta = 0$ one has the Gaussian distribution with δ mean and variance equals to $2\sigma^2$. If $\alpha = 0.5$ and $\beta = 1$ one has a Lévy distribution with probability density function given by

$$f(x \mid \delta, \sigma) = \left(\frac{\sigma}{2\pi}\right)^{\frac{1}{2}} (x-\delta)^{-\frac{3}{2}} \exp\left(-\frac{0.5\sigma}{x-\delta}\right), \tag{2.1}$$

for $\delta < x < \infty$. Figure 1 presents Lévy probability density functions for $\delta = 0$ and different values of the σ scale parameter.

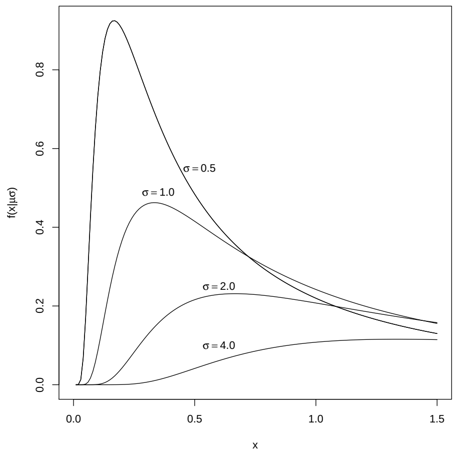


Figure 1: Lévy density function for $\delta = 0$ and different values for the σ scale parameter.

The probability distribution function of the random variable X with a Lévy distribution defined in (2.1) is given by

$$F(x \mid \delta, \sigma) = \mathbb{P}(X \leq x) = \operatorname{erfc}\left(-\frac{0.5\sigma}{x-\delta}\right)^{\frac{1}{2}}, \tag{2.2}$$

where $\operatorname{erfc}(x) = 1 - \operatorname{erf}(x)$ is the complementary error function with the error function $\operatorname{erf}(\cdot)$ given by

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt. \tag{2.3}$$

The Lévy distribution with probability density function (2.1) has undefined mean and undefined variance but its median is given by

$$Median = \delta + \frac{\sigma}{2\left[\operatorname{erfc}^{-1}\left(\frac{1}{2}\right)\right]^2}, \tag{2.4}$$

where the inverse complementary error function is

$$\operatorname{erfc}^{-1}(1-x) = \operatorname{erf}^{-1}(x).$$

To obtain the probability density function or the median of a random variable X with a Lévy density function different approximations for the complementary error function are introduced in the literature (see [13]). Some special cases are presented below.

1.
$$\operatorname{erf}(x) \approx 1 - \frac{1}{(1 + a_1x + a_2x^2 + a_3x^3 + a_4x^4)^4}, \tag{2.5}$$
where the maximum error is 5×10^{-4} and $a_1 = 0.278393$, $a_2 = 0.230389$, $a_3 = 0.000972$ and $a_4 = 0.078108$;

2.
$$\operatorname{erf}(x) \approx 1 - \frac{1}{(1 + a_1x + \ldots + a_6x^6)^{16}}, \tag{2.6}$$
where the maximum error is 3×10^{-7}, $a_1 = 0.0705230784$, $a_2 = 0.0422820123$, $a_3 = 0.0092705272$, $a_4 = 0.0001520143$, $a_5 = 0.0002765672$ and $a_6 = 0.0000430638$;

3.
$$\operatorname{erf}(x) \approx sign(x)\sqrt{1-\exp\left(-x^2\frac{\frac{4}{\pi}+ax^2}{1+ax^2}\right)}, \tag{2.7}$$

where $a=\frac{8(\pi-3)}{3\pi(4-\pi)}\approx 0.140012$ and $sign(\cdot)$ is given by (1.2).

4. An approximation for the inverse error function is given by

$$\operatorname{erf}(x)^{-1} = sign(x)\sqrt{\sqrt{\left(\frac{2}{\pi a}-\frac{\log\left(1-x^2\right)}{2}\right)^2-\frac{\log\left(1-x^2\right)}{a}}-\left(\frac{2}{\pi a}-\frac{\log\left(1-x^2\right)}{2}\right)}, \tag{2.8}$$

where the constant a and the $sign(\cdot)$ are given in (2.7).

Assuming a random sample of size n with a Lévy distribution with probability density as (2.2), the likelihood function for δ and σ is given by

$$L(\delta,\sigma) = \left(\frac{\sigma}{2\pi}\right)^{\frac{n}{2}}\prod_{i=1}^{n}\left(x_i-\delta\right)^{-\frac{3}{2}}\exp\left[-\frac{\sigma}{2\left(x_i-\delta\right)}\right]\prod_{i=1}^{n} I\left(x_i>\delta\right), \tag{2.9}$$

where $I(A)$ denotes the indicator function of set A.

Inferences for δ and σ parameters in the case of Lévy distribution are obtained using standard Markov Chain Monte Carlo methods (see [14,15]).

3 A Bayesian Analysis for General Stable Distributions

Let us assume that x_i, for $i=1,\ldots,n$, is a random sample of size n, where $X_i \sim S_\alpha(\beta,\delta,\sigma)$, that is, $Z_i=\frac{X_i-\delta}{\sigma}\sim S_\alpha(\beta,0,1)$. Assuming a joint prior distribution for α,β,δ and σ, given by $\pi_0(\alpha,\beta,\delta,\sigma)$, [1] shows that the joint posterior distribution for parameters α,β,δ and σ is given by

$$\begin{aligned}\pi(\alpha,\beta,\delta,\sigma|\mathbf{x}) \;&\propto\; \int\left(\frac{\alpha}{|\alpha-1|\,\sigma}\right)^n\times\exp\left\{-\sum_{i=1}^{n}\left|\frac{z_i}{t_{\alpha,\beta}(y_i)}\right|^{\theta}\right\}\prod_{i=1}^{n}\left|\frac{z_i}{t_{\alpha,\beta}(y_i)}\right|^{\theta}\frac{1}{|z_i|}\\ &\times\;\pi_0(\alpha,\beta,\delta,\sigma)\,\mathbf{dy},\end{aligned} \tag{3.1}$$

where $\theta=\frac{\alpha}{\alpha-1}$, $z_i=\frac{x_i-\delta}{\sigma}$, for $i=1,\cdots,n$, $\alpha\in(0,2]$, $\beta\in[-1,1]$, $\delta\in(-\infty,\infty)$ and $\sigma\in(0,\infty)$; $\mathbf{x}=(x_1,x_2,\cdots,x_n)$ and $\mathbf{y}=(y_1,y_2,\cdots,y_n)$ are respectively, the observed and non-observed data vectors. Notice that the bivariate distribution in expression (3.1) 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 $\alpha=2$ one has $\theta=2$ and $b_{\alpha,\beta}=0$. In this case one has a Gaussian distribution with δ mean and $2\sigma^2$ variance.

For a Bayesian analysis of the proposed model, we assume uniform $\mathcal{U}(a,b)$ independent priors for α, β, δ and σ, where the hyperparameters a and b are assumed to be known in each application following the restrictions $\alpha\in(0,2]$, $\beta\in[-1,1]$, $\delta\in(-\infty,\infty)$ and $\sigma\in(0,\infty)$.

In the simulation algorithm to obtain a Gibbs sample for the random quantities α,β,δ and σ having the joint posterior distribution (3.1), we assume a uniform $\mathcal{U}(-0.5,0.5)$ 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, one has the possibility to use standard software package like OpenBus (see [16]) with great simplification to obtain the simulated Gibbs samples for the joint posterior distribution.

In this way, one has the following algorithm:

(i) Start with the initial values $\alpha^{(0)}, \beta^{(0)}, \delta^{(0)}, \sigma^{(0)}$;

(ii) Simulate a sample $\mathbf{y} = (y_1, \ldots, y_n)$ from the conditional distributions $\pi(y_i|\alpha^{(0)}, \beta^{(0)}, \delta^{(0)}, \sigma^{(0)}, \mathbf{x})$, for $i = 1, \cdots, n$;

(iii) Update $\alpha^{(0)}, \beta^{(0)}, \delta^{(0)}, \sigma^{(0)}$ by $\alpha^{(1)}, \beta^{(1)}, \delta^{(1)}, \sigma^{(1)}$ from the conditional distributions $\pi(\alpha|\beta^{(0)}, \delta^{(0)}, \sigma^{(0)}, \mathbf{x}, \mathbf{y})$, $\pi(\beta \mid \alpha^{(0)}, \delta^{(0)}, \sigma^{(0)}, \mathbf{x}, \mathbf{y})$, $\pi(\delta|\alpha^{(0)}, \beta^{(0)}, \sigma^{(0)}, \mathbf{x}, \mathbf{y})$ and $\pi(\sigma|\alpha^{(0)}, \beta^{(0)}, \delta^{(0)}, \mathbf{x}, \mathbf{y})$;

(iv) Repeat steps (i), (ii) and (iii) until convergence.

From expression (3.1), the joint posterior probability distribution for $\alpha, \beta, \delta, \sigma$ and $\mathbf{y} = (y_1, y_2, \cdots, y_n)$ is given by

$$\begin{aligned} \pi(\alpha, \beta, \delta, \sigma, \mathbf{y}|\mathbf{x}) \quad &\propto \quad \left(\frac{\alpha}{|\alpha-1|\,\sigma}\right)^n \exp\left\{-\sum_{i=1}^{n}\left|\frac{z_i}{t_{\alpha,\beta}(y_i)}\right|\right\} \\ &\times \quad \prod_{i=1}^{n}\left|\frac{z_i}{t_{\alpha,\beta}(y_i)}\right|^{\theta}\frac{1}{|z_i|} \times \prod_{i=1}^{n} h(y_i)\pi_0(\alpha, \beta, \delta, \sigma), \end{aligned} \tag{3.2}$$

where θ and $t_{\alpha,\beta}(\cdot)$ are respectively defined in (1.4) and (1.5) and $h(y_i)$ is a $\mathcal{U}(-0.5, 0.5)$ 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 below.

4 Some Applications

4.1 Buckle's Data

In Table 1, we have a data set introduced by [1]. This is the daily price return data of Abbey National shares in the period from July 31, 1991 to October 08, 1991.

Table 1: Daily price returns with $n = 50$.

296	296	300	302	300	304	303	299	293	294
294	293	295	287	288	297	305	307	307	304
303	304	304	309	309	309	307	306	304	300
296	301	298	295	295	293	292	297	294	293
306	303	301	303	308	305	302	301	297	299

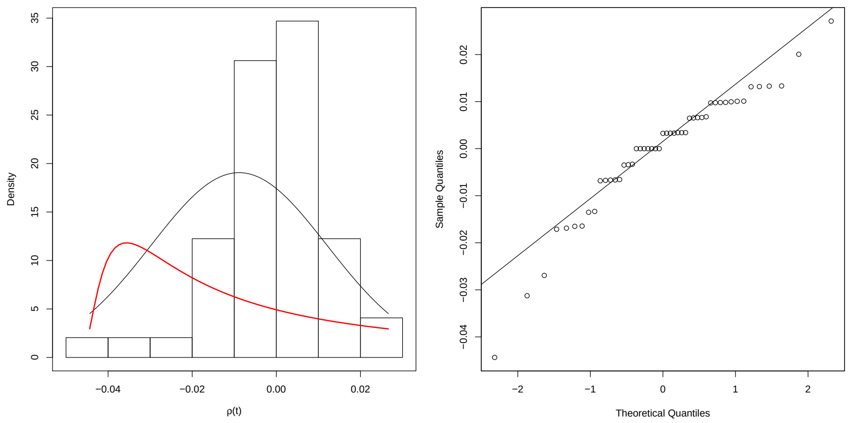
Table 2: Returns $\rho(t)$, at time t, for $n = 49$.

0.0000	0.0135	0.0067	-0.0066	0.0133	-0.0033	-0.0132	-0.0201	0.0034	0.0000
-0.0034	0.0068	-0.0271	0.0035	0.0312	0.0269	0.0066	0.0000	-0.0098	-0.0033
0.0033	0.0000	0.0164	0.0000	0.0000	-0.0065	-0.0033	-0.0065	-0.0132	-0.0133
0.0169	-0.0100	-0.0101	0.0000	-0.0068	-0.0034	0.0171	-0.0101	-0.0034	0.0444
-0.0098	-0.0066	0.0066	0.0165	-0.0097	-0.0098	-0.0033	-0.0133	0.0067	-

Table 3: Posterior summaries for the Lévy distribution.

Parameter	Mean	Standard Deviation	95% Credible Interval
δ	-0.04868	0.001669	$(-0.05284, -0.04628)$
σ	0.03901	0.008916	$(0.02343, 0.05948)$

In Figure 2(a) we present the histogram of the returns $\rho(\cdot)$ time series while in Figure 2(b) we have the Gaussian probability plot for the same data. From these figures, one observes that the Gaussian distribution does not fit well the data.



(a) (b)

Figure 2: (a) Empirical return distribution; (b) Normal probability plot.

Assuming a Lévy distribution with probability density function given in (2.1) for a Bayesian analysis we consider the following prior distributions for δ and σ, $\delta \sim \mathcal{U}(-1, -0.0271)$ and $\sigma \sim \mathcal{U}(0, 1)$, where $\mathcal{U}(a, b)$ denotes a uniform distribution on the interval (a, b). Observe that the minimum value for the $\rho(\cdot)$ data is given by -0.0271 and $x_i \geq \delta$, that is, $\min\{x_1, \cdots, x_n\} \geq \delta$. To simulate samples for the joint posterior distribution for δ and σ, using standard MCMC methods, we have used OpenBUGS software which only requires the log-likelihood function and prior distributions for model parameters. In Table 3 we present the posterior summaries of interest considering a burn-in-sample of size $5,000$ discarded to eliminate the initial value effect. After this burn-in-sample period we simulate another $200,000$ Gibbs samples taking every 10-th sample. This gives a final sample of size $20,000$ to be used for finding the posterior summaries of interest. Convergence of the Gibbs sample algorithm was verified by trace-plots of the simulated Gibbs samples. From OpenBUGS output we obtain a Deviance Information Criterium (DIC) value equals to -151.7. In Figure 2 (red line), we have the plot of the fitted Lévy density with $\delta = -0.0487$ mean and $\sigma = 0.0391$ as the scale parameter and the histogram of the $\rho(\cdot)$ returns.

Assuming a general stable distribution, we present in Table 4 the posterior summaries of interest obtained using OpenBUGS software considering the following priors: $\alpha \sim \mathcal{U}(1, 2)$, $\beta \sim \mathcal{U}(-1, 0)$, $\delta \sim \mathcal{U}(-0.5, 0.5)$ and $\sigma \sim \mathcal{U}(0, 0.5)$. In the simulation procedure we have used a burn-in-sample of size $10,000$ and another $490,000$ Gibbs samples taking every 100-th sample. This gives a final sample of size $4,900$ to be used for finding the posterior summaries of interest.

In Figure 3 we have the trace-plots of the simulated Gibbs samples. In Figure 2 we also have the plot of the fitted stable distribution with $\alpha = 1.653$, $\beta = -0.3455$, $\delta = 0.00782$ and $\sigma = 0.001132$. We observe good fit of the stable distribution (black line). The obtained DIC value is equal to -70480. From this value we conclude that the data is better fitted by the general stable distribution in contrast to the Lévy distribution (since it has smaller DIC value).

Table 4: Posterior summaries for the general stable distribution.

Parameter	Mean	Standard Deviation	95% Credible Interval
α	1.653	0.01639	(1.29,1.965)
β	-0.3455	0.02556	(-0.9188,-0.01257)
δ	0.00782	0.0702	(0.00549,0.01048)
σ	0.001132	1.35e-4	(-0.002478,0.004601)

4.2 Coding and Non-Coding Regions in DNA Sequences

Crato et al. [2] introduce the length distribution of coding and non-coding regions for all Homo Sapiens chromosomes available from the European Bioinformatics Institute. In this way they con-

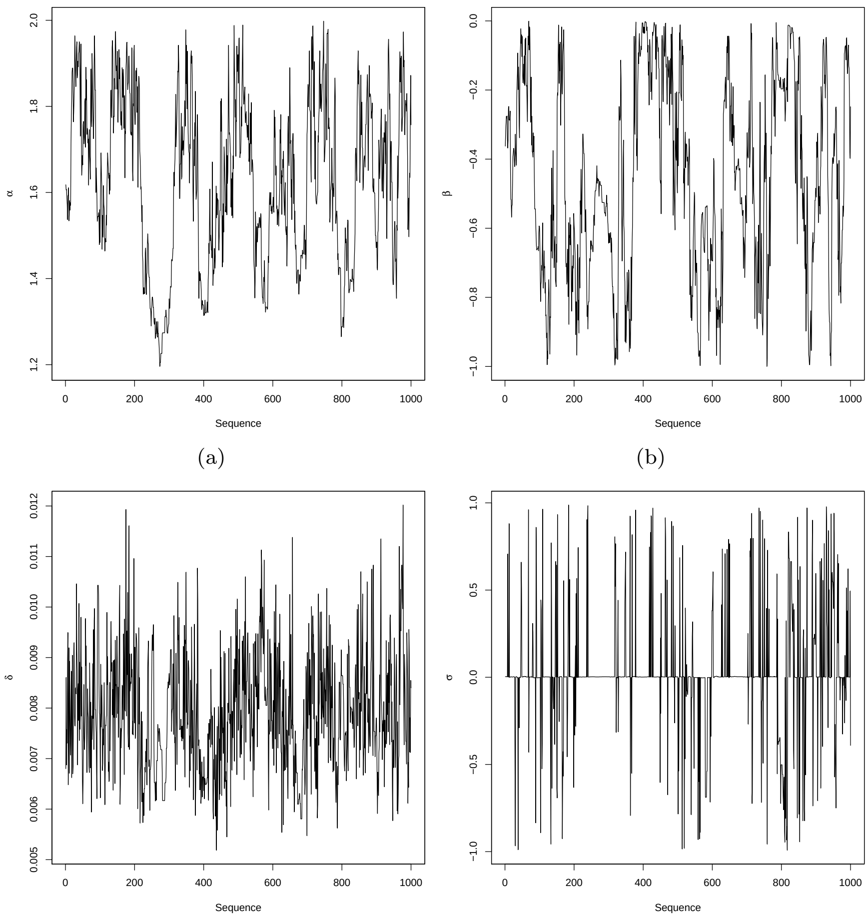
Table 5: Coding sequence CM000275.

108	103	68	55	97	73	87	110	320	111	152	177	11	297	61
42	88	68	64	78	272	190	39	254	18	95	119	263	168	165
20	101	165	127	74	121	60	97	63	141	132	252	145	57	53
47	44	425	5	379	246	87	97	179	102	74	161	34	11	116
431	101	104	58	74	38	9	54	76	111	110	95	124	80	77
353	215	34	111	77	152	77	60	394	77	111	144	51	353	77
111	144	51	128	94	110	113	146	174	11	155	254	121	117	212
48	57	156	183	76	353	54	91	781	69	149	77	122	170	134
129	145	158	119	158	181	162	119	194	181	124	147	96	358	138
179	137	599	69	199	350	149	77	122	134	129	145	158	119	158
181	162	119	194	181	124	147	96	358	138	179	137	599	69	199
350	323	95	92	32	20	91	282	112	282	1659	554	161	263	46
90	346	11	139	46	33	183	212	341	512	98	512	109	21	512
692	101	107	84	151	185	20	15	83	103	50	81	91	5	29
103	147	64	3	26	180	97	171	157	101	26	180	97	171	4
479	105	33	74	159	64	94	364	56	31	143	88	78	18	81
300	103	108	144	458	104	145	200	342	353	77	111	148	125	56
160	74	37	201	86	131	127	114	278	258	115	68	30	115	68
57	137	98	91	57	137	91	18	114	152	177	103	108	272	103
108	227	108	103	177	152	114	110	87	98	50	195	90	53	66
28	139	159	118	136	141	139	178	191	159	122	89	80	370	159
31	150	86	83	122	467	91	51	63	139	71	71	37	96	72
1591	1622	767	122	29	188	88	18	248	88	74	57	8	273	379
260	44	59	257	260	44	233	33	211	173	77	117	105	18	139
390	-	-	-	-	-	-	-	-	-	-	-	-	-	-

sider a transformation of the genomes in numerical sequences. As an illustration, we have, respectively, in Tables 5 and 6, the data for coding and non-coding length sequences for H. Sapiens chromosomes transformed in a logarithm scale (sequence CM000275 extracted from Table 2, in [2]).

Figure 4 presents the histograms of the data given in Tables 5 and 6, assuming a logarithm transformation. From these plots, we observe that a Gaussian distribution could not be a reasonable model for fitting the data. Assuming a Lévy distribution with probability density function (2.1) for a Bayesian analysis we consider the following prior distributions for $\delta \sim \mathcal{U}(-1000, 1.0986)$, where 1.0986 is the minimum of the observations in logarithm scale, and for $\sigma \sim \mathcal{U}(0, 10000)$. In Table 7 we have the posterior summaries of interest considering the transformed coding and non-coding data using OpenBUGS software.

Figure 4 shows the fitted Lévy density with $\delta = 0.9693$ and $\sigma = 3.167$ (for coding data) and with $\delta = 4.182$ and $\sigma = 1.633$ (for non-coding data). From this figure we observe that the data is



(a) (b) (c) (d)

Figure 3: Trace-plots for α, β,δ and σ — Buckle's data.

Table 6: Non-coding sequence CM000275.

473	3014	46804	1610	596	315546	82	438	122	1995	1886	2686
117891	4797	507006	439	3003	254680	132	208227	33316	27057	77385	299389
11126	14901	14946	21801	29460	182825	1282	2657	850	318	5045	6112
2730	110	699	392	6047	20772	17837	151831	2502	1230	270	1959
331299	126	52160	478	19056	94	689	2308	942	1034	936	42731
2590	2357	42188	4426	686	3091	2781	1250	14946	398	808	2001
70798	129	607	3731	155639	98	129	358	129	250729	99	608
129	102	107	1682	608	129	102	107	60835	402	86	858
2556	1917	3300	515979	894268	38404	20693	2501	149896	15176	12545	27795
12119	27552	28108	441	226144	310	380715	1146	1305	247	1656	3778
124	2610	419	116	10458	96	4157	123	275	84	116	4880
567	438	452	173	188	284	154	96	38536	1146	1305	2073
3778	124	2610	419	116	10458	96	4157	123	275	84	116
4880	567	438	452	173	188	284	154	96	101875	1486	2385
2074	7530	471	2014	1257	863	102886	426	1369	5266	496	103
6531	398292	19811	3485	547	1316	52053	114313	104170	41236	41846	2814
560	4016	479	33776	5064	67998	441	386840	1522	107156	3721	2899
1781	3500	1379	2766	135123	1931	392	699	110	25884	3013	1236
7424	10533	1265	24203	3013	1236	7424	11955	27994	142294	6042	349782
290	164	64038	152	117	2263	8885	409	83	489	882	2368
836	451	4275	39510	2967	1082	878	1046	15536	610	129	100
100912	1610	12498	4740	197823	433	11102	1320	1133	654	2979	157170
9491	2576	11103	8314	2576	60172	219	111	949	1662	219	1158
119353	1466	2036	1925	2995	423	3901	2995	423	13114	157134	454
2709	1936	2008	870	83	425	19373	45480	21308	11169	16584	43630
48317	5045	9466	5710	233467	12543	13912	198	651	6242	617	985
2978	122	1039	2194	2636	1138	2212	10411	94951	561	437	90
1401	2373	382432	2382	167632	421	1622	145848	1639	18367	90	256
421	1031	256	421	1357	297	66358	271	78	87	924	139
78	87	233	51764	138	275166	19259	2060	2855	693	52166	189

not well fitted by the Lévy distribution.

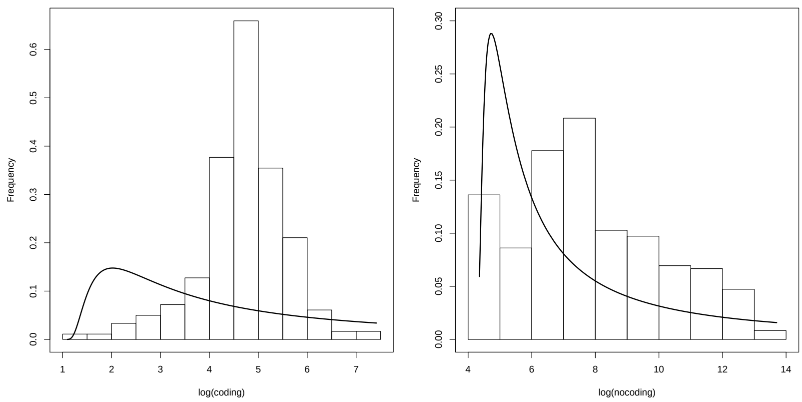
For a Bayesian analysis of the data assuming a general stable distribution we consider the following prior distributions: $\alpha \sim \mathcal{U}(0,2)$, $\beta \sim \mathcal{U}(-1,0)$, $\delta \sim \mathcal{U}(0,3)$ and $\sigma \sim \mathcal{U}(0,10)$. Using the OpenBugs software, we simulated $600,000$ Gibbs samples. From these $600,000$ samples, we discarded the first $100,000$ as a "burn-in-sample" to eliminate the initial value effects. After this "burn-in-sample" period, we took every 500-th sample, which gives a final Gibbs sample of size $1,000$ to be used for Monte Carlo of the interested random quantities. Convergence of the Gibbs sampling algorithm was verified from trace plots of the simulated samples for each parameter. Table 8 presents the posterior summaries of interest. Figure 5 shows the fitted stable distributions for coding and non-coding data. We observe good fit of the stable distributions in both cases.

Table 7: Posterior summaries, in the case of the Lévy distribution, for coding and non-coding regions of CM000275 sequence.

Parameter	Mean	Standard Deviation	95% Confidence Interval
		Coding	
δ	0.9693	0.0260	$(0.9106, 1.0160)$
σ	3.1670	0.2437	$(2.7060, 3.6490)$
		Non-coding	
δ	4.182	0.0325	$(4.114, 4.239)$
σ	1.633	0.1496	$(1.345, 1.628)$

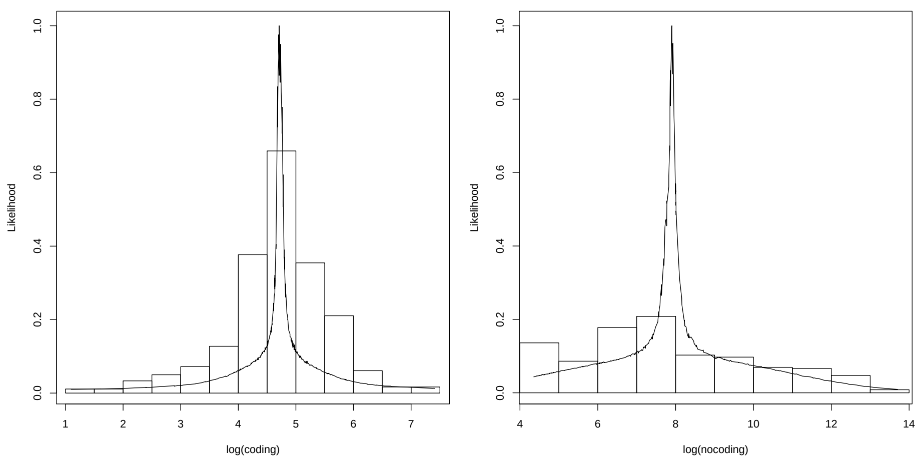
Table 8: Posterior summaries, in the case of general stable distributions, for coding and non-coding regions of CM000275 sequence.

Parameter	Mean	Standard Deviation	95% Confidence Interval
		Coding	
α	1.583	0.09803	$(1.402, 1.783)$
β	-0.08868	0.06195	$(-0.246, -0.0029)$
δ	4.722	0.03743	$(4.661, 4.802)$
σ	0.4785	0.03409	$(0.4098, 0.5424)$
		Non-coding	
α	1.974	0.02802	$(1.909, 1.999)$
α	-0.5291	0.3111	$(-0.989, -0.0242)$
δ	7.882	0.1225	$(7.601, 8.100)$
σ	1.638	0.06139	$(1.520, 1.767)$



(a) (b)

Figure 4: Histograms for log(coding) and log(non-coding) and fitted Lévy distributions.



(a) (b)

Figure 5: Histograms for log(coding) and log(non-coding) and fitted stable distribution.

5 Concluding Remarks

The use of stable distributions could be a good alternative for many applications in data analysis, since this model has a great flexibility for fitting the data. With the use of Bayesian methods and MCMC simulation methods it is possible to get inferences for the model despite the nonexistence of an analytical form for the density function. 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 OpenBugs software.

In the simulation study considered 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. Observe that MCMC methods are a class of algorithms for sampling from probability distributions based on constructing a Markov Chain that has the desired distribution as its equilibrium distribution. The state of the chain after a large number of steps is then used as a sample of the desired distribution. The quality of the sample improves as a function of the number of steps. The obtained simulation results for the applications in Section 4, could be easily replicated using the same auxiliary random variable Y defined in Section 1 and the non-informative prior distributions defined in Section 3 for the parameters of the model. More accurate posterior summaries results could be obtained using informative prior distributions for the parameters of the model based on prior opinion of experts rather than using non-informative priors as it was assumed in this paper. Observe that although the nonexistence of an analytical form for the density function for stable distributions, the moments could be obtained from the characteristic function defined in (1.1).

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 are 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.

Acknowledgements

S.R.C. Lopes research was partially supported by CNPq-Brazil, by CAPES-Brazil, by Pronex *Probabilidade e Processos Estocásticos* - E-26/170.008/2008 -APQ1 and also by INCT *em Matemática.* J. Mazucheli gratefully acknowledge the financial support from the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

References

[1] D. J. Buckle, "Bayesian inference for stable distributions", *Journal of the American Statistical Association*, Vol. 90, 1995, pp. 605-613.

[2] N. Crato, R. R. Linhares and S. R. C. Lopes, "α-stable laws for noncoding regions in DNA sequences", *Journal of Applied Statistics*, Vol. 38, 2011, pp. 267-271.

[3] P. Lévy, "Théorie des erreurs la loi de Gauss et les lois exceptionelles", *Bulletin Society Mathematical*, Vol. 52, 1924, pp. 49-85.

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

[5] J. P. Nolan, "Stable Distributions - Models for Heavy Tailed Data", Birkhäuser, Boston, 2009. In progress, Chapter 1 online at academic2.american.edu/$\sim$jpnolan.

[6] B. V. Gnedenko and A. N. Kolmogorov, "Limit distributions for sums of independent random variables", Addison-Wesley, Massachussetts, 1968.

[7] A. V. Skorohod, "On a theorem concerning stable distributions". In: *Select. Transl. Math. Statist. and Probability*, Vol. 1, Inst. Math. Statist. and Amer. Math. Soc., Providence, R.I., pp. 169-170, 1961.

[8] I. A. Ibragimov and K. E. Černin, "On the unimodality of stable laws", *Teor. Veroyatnost. i Primenen.*, Vol. 4, 1959, pp. 453-456.

[9] M. Kanter, "On the unimodality of stable densities", *Annals of Probability*, Vol. 4, 1976, pp. 1006-1008.

[10] W. Feller, "An introduction to probability theory and its applications", Vol. II. John Wiley, New York, 1971.

[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, 1999, pp. 331-344.

[12] M. A. Tanner and W. H. Wong, "The calculation of posterior distributions by data augmentation", *Journal of American Statistical Association*, Vol. 82, 1987, pp. 528-550.

[13] M. Abramowitz and I. A. Stegun, "Handbook of mathematical functions with formulas, graphs, and mathematical tables", National Bureau of Standards Applied Mathematics Series, Vol. 65, Washington, D.C., 1964.

[14] S. Chib and E. Greenberger, "Understanding the Metropolis-Hastings algorithm", *The American Statistician*, Vol. 49, 1995, pp. 327-335.

[15] C. P. Robert and G. Casella, "Monte Carlo statistical methods", 2nd ed., Springer-Verlag, New York, 2004.

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