

# GOMPERTZ GROWTH CURVES ASSUMING STABLE DISTRIBUTIONS: AN APPLICATION TO INTRAUTERINE GROWTH FOR PRETERM INFANTS

Jorge Alberto ACHCAR<sup>1</sup>  
Breno Raphael Gomes de OLIVEIRA<sup>1</sup>

- **ABSTRACT:** *In this paper we consider the use of Bayesian methods to analyze some standard existing growth models, a class of nonlinear regression models. For the nonlinear modeling we assume usual normal errors and also stable distributions for the response variable. We also study some robustness aspects of nonlinear regression models to 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 an example related to the intrauterine growth curves for preterm infants. Posterior summaries of interest are obtained using MCMC (Markov Chain Monte Carlo) methods and the OpenBugs software.*
- **KEYWORDS:** *Stable distribution; Bayesian analysis; nonlinear regression models; MCMC methods; OpenBugs software.*

## 1 Introduction

Growth curves are included in a class of nonlinear models widely used in biology to model different problems as the population size or biomass (in population ecology and demography, for population growth analysis) or the individual body height or biomass (in physiology, for growth analysis of individuals). A growth curve is an empirical model of the evolution of a quantity over time. Growth curves are employed in many disciplines besides biology, particularly in statistics, where there is a great literature on this subject related to nonlinear models. Under a more probabilistic and mathematical statistics approach, growth curves are often modeled as being a continuous stochastic processes.

Standard classical inference methods to get point or interval estimates for the parameters of growth curves are presented within the nonlinear modeling methodology.

Nonlinear regression methodology is similar to the linear regression methodology, that is, a modeling approach to relate a response  $X$  to a vector of covariates,

---

<sup>1</sup>Universidade de São Paulo - USP, Faculdade de Medicina de Ribeirão Preto, Departamento de Medicina Social, CEP: 14049-900, Ribeirão Preto, SP, Brasil. E-mail: [achcar@fmrp.usp.br](mailto:achcar@fmrp.usp.br); [brenooliveira@usp.br](mailto:brenooliveira@usp.br)

$\mathbf{v} = (v_1, \dots, v_k)^T$ . Different of linear models, nonlinear regression is characterized by the fact that the prediction equation depends nonlinearly on one or more unknown parameters.

Different of linear regression methodology often used for building a purely empirical model, nonlinear regression methodology usually arises when there is some physical reason which implies that the relationship between the response and the predictors follows a particular functional form.

A nonlinear regression model has the general form,

$$X_i = f(\mathbf{v}_i, \boldsymbol{\theta}) + \varepsilon_i \quad (1)$$

where  $X_i$  are the responses,  $i = 1, \dots, n$ ;  $f$  is a known function of the covariate vector  $\mathbf{v}_i = (v_{i1}, \dots, v_{ik})^T$  is a vector of  $k$  covariates or independent variables;  $\boldsymbol{\theta} = (\theta_1, \dots, \theta_p)^T$  is the vector of  $p$  parameters and  $\varepsilon_i$  are random errors. The errors  $\varepsilon_i$  are usually assumed to be uncorrelated with a normal distribution with mean zero and constant variance.

The most popular criterion to estimate the  $p$  parameter vector  $\boldsymbol{\theta}$  in the nonlinear model (1) is to find estimates for the parameters (nonlinear least squares) which minimizes the sum of squared errors, given by,

$$\sum_{i=1}^n [x_i - f(\mathbf{v}_i, \boldsymbol{\theta})]^2 \quad (2)$$

**Remark:** If the errors  $\varepsilon_i$  follow a normal distribution, then the least squares estimator for  $\boldsymbol{\theta}$  is also the maximum likelihood estimator.

Usually, nonlinear regression estimates must be computed by iterative procedures using optimization methods to minimize the sum of squared errors expression (2). It is important to point out that the definition of nonlinearity is related to the unknown parameters and not to the relationship between the covariates and the response. As an example,  $X = \beta_0 + \beta_1 v + \beta_2 v^2 + \varepsilon$ , is considered as a linear model (see for example, Bates and Watts, 1988; Ratkowsky, 1983; Seber and Wild, 1989).

A popular iterative technique to find the least squares estimator of nonlinear models is the Gauss–Newton algorithm. The Gauss–Newton algorithm increments the working estimate  $\boldsymbol{\theta}$  at each iteration by an amount equal to the coefficients from the linear regression of the current residuals  $\mathbf{e}$  on the current gradient matrix  $\mathbf{X}$ .

If the function  $f$  in (1) is continuously differentiable in  $\boldsymbol{\theta}$ , then it can be linearized locally as,

$$f(\mathbf{v}, \boldsymbol{\theta}) = f(\mathbf{v}, \boldsymbol{\theta}_0) + V_0(\boldsymbol{\theta} - \boldsymbol{\theta}_0), \quad (3)$$

where  $V_0$  is the  $n \times p$  gradient matrix with elements  $\partial f(\mathbf{v}_i, \boldsymbol{\theta}_0) / \partial \theta_j$  and  $\boldsymbol{\theta}_0$  is a vector of initial values for the iterative procedure. This leads to the Gauss–Newton algorithm for estimating  $\boldsymbol{\theta}$ ,

$$\boldsymbol{\theta}_1 = \boldsymbol{\theta}_0 + (V_0^T V_0)^{-1} V_0^T \mathbf{e}, \quad (4)$$

Where  $\mathbf{e}$  is the vector of working residuals  $x_i - f(\mathbf{v}_i, \boldsymbol{\theta}_0)$ .

If the errors  $\varepsilon_i$  are independent and normally distributed  $N(0, \sigma^2)$ , then the Gauss–Newton algorithm is an application of Fisher’s method of scoring. This algorithm is implemented in many existing statistical softwares as for example, R, Minitab version 16 or SAS.

If  $\mathbf{X}$  is of full column rank in a neighborhood of the least squares solution, then it can be shown that the Gauss–Newton algorithm will converge to the solution from a sufficiently good starting value. In practical work, there is no guarantee, though, that the algorithm will converge from values further from the solution. Some improvement of Gauss–Newton algorithm are given in the literature as the Levenberg–Marquart damping algorithm (see for example, Seber and Wild, 1989).

Standard inferences for the parameters of nonlinear models are obtained from the asymptotical normality of the least squares estimators  $\hat{\boldsymbol{\theta}}$  with mean  $\boldsymbol{\theta}$  and variance-covariance matrix  $\sigma^2(\mathbf{V}^T\mathbf{V})^{-1}$  where the variance  $\sigma^2$  is usually estimated by,

$$s^2 = \{ \sum_{i=1}^n [x_i - f(v_i, \hat{\boldsymbol{\theta}})]^2 \} / (n - p). \quad (5)$$

It is important to point out that since most asymptotic inference for nonlinear regression models is based on analogy with linear models, and since this inference is only approximated as the actual model differs from a linear model, various measures of nonlinearity have been proposed in the literature to verify how good linear approximations are likely to be in each case. One class of measures focuses on curvature (intrinsic curvatures) of the function  $f$  and it is based on the sizes of the second derivatives of  $f$  (see for example, Bates and Watts, 1980). Other class of measures: intrinsic curvature defined by the residuals  $u_i = x_i - f(v_i, \hat{\boldsymbol{\theta}})$  or parameter effect curvatures (see for example, Bates and Watts, 1988).

A good alternative to get accurate inferences and predictions for non-linear models is the use of Bayesian methods especially using MCMC (Markov Chain Monte Carlo) methods as the popular Gibbs sampling algorithm (see for example, Gelfand and Smith, 1990; or Casella and George, 1992) or the Metropolis-Hastings algorithm (see for example, Chib and Greenberg, 1995).

The paper is organized as follows: in section 2, we introduce some special growth functions: the use of stable distribution for the response; in section 3, we introduce a Bayesian approach for nonlinear models assuming a stable distribution; in section 4, we present an example with an intrauterine growth data for preterm infants; in section 5, we present the intrauterine growth curves for preterm infants in presence of outliers; finally in section 6, we present some concluding remarks.

## 2 Some special growth functions: the use of stable distribution for the response

In many applications, the systematic part of the response is known to be monotonic increasing in  $v$ , where  $v$  might represent time or dosage. Nonlinear regression models with this property are called growth models. The simplest growth model is the exponential growth model,

$$f(v, \boldsymbol{\theta}) = \theta_1 \exp(-\theta_2 v), \quad (6)$$

Where  $\boldsymbol{\theta} = (\theta_1, \theta_2)$ .

Many other growth models are introduced in the literature. Special cases are given by:

(i)  $f(v, \boldsymbol{\theta}) = \theta_1 + \theta_2 v^{\theta_3}$ , where  $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$ .

- (ii)  $f(v, \theta) = \theta_1 + \theta_2 \exp[\theta_3 v]$  (exponential growth model), where  $\theta = (\theta_1, \theta_2, \theta_3)$ .
- (iii)  $f(v, \theta) = \theta_1 \exp[-\theta_2 v] + \theta_3 \exp[-\theta_4 v]$ , where  $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$ .
- (iv)  $f(v, \theta) = \theta_1 v / [\theta_2 + v]$  (Michaelis-Menten), where  $\theta = (\theta_1, \theta_2)$ .
- (v)  $f(v, \theta) = \theta_1 \exp[-\exp(\theta_2 - \theta_3 v)]$  (Gompertz), where  $\theta = (\theta_1, \theta_2, \theta_3)$ .
- (vi)  $f(v, \theta) = \theta_1 + (\theta_2 - \theta_1) / [1 + \exp((v - \theta_3) / \theta_4)]$  (logistic), where  $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$ .
- (vii)  $f(v, \theta) = \theta_1 + (\theta_2 - \theta_1) \exp[-\theta_3 v^{\theta_4}]$  (Weibull), where  $\theta = (\theta_1, \theta_2, \theta_3, \theta_4)$ .
- (viii)  $f(v, \theta) = 1 - \exp[-\exp(\theta_1 - \theta_2 v)]$  (sigmoid), where  $\theta = (\theta_1, \theta_2)$ .

For all cases,  $v$  is a covariate. Observe that these different models could be considered to model growth curves. In many cases we have physical interpretations for the use of a particular model. In other situations we try different growth models to decide by the best one to be fitted by the data.

In this paper we explore a particular growth model: the Gompertz model.

In many cases the usual assumption of normality for the errors in (1) could not be appropriated. This situation occurs for example, when we have discordant observations which could greatly affect the obtained inferences. In this way, we could assume more robust distributions for the data, as the stable distribution.

Other point of interest: the use of asymptotical inference results could not be accurate depending on the sample sizes and the intrinsic curvature of the function  $f$  in (1).

In this way, we propose the use of stable distributions for the response  $X_i$  in (1).

A wide class of distributions that encompasses 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, Nolan, 2009), described by four parameters  $\alpha$ ,  $\beta$ ,  $\gamma$  and  $\delta$ . 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,  $\gamma \in (-\infty, \infty)$  and  $\delta \in (0, \infty)$  (see Levy, 1924).

Stable distributions are usually denoted by  $S_\alpha(\beta, \gamma, \delta)$ . If a random variable  $X \sim S_\alpha(\beta, \gamma, \delta)$ , then  $Z = \frac{X-\gamma}{\delta} \sim S_\alpha(\beta, 0, 1)$  (see Lukacs, 1970 and Nolan, 2009).

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 (Gnedenko and Kolmogorov, 1968; Skorohod, 1961) and unimodal (Ibragimov and Černin, 1959); Kanter, 1976). 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 Feller, 1971).

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

$$\log[\Phi(t)] = \begin{cases} i\gamma t - |\delta t|^\alpha [1 - i\beta \text{sign}(t) \tan(\frac{\pi\alpha}{2})], & \text{for } \alpha \neq 1 \\ i\gamma t - |\delta t|^\alpha [1 + i\beta \text{sign}(t) (\frac{2}{\pi}) \log(t)], & \text{for } \alpha = 1 \end{cases} \quad (7)$$

where  $i = \sqrt{-1}$  and  $\text{sign}(\cdot)$  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 Roussas, 2005),

$$f(x) = 1/2\pi \int_{-\infty}^{\infty} e^{-itx} \Phi(t) dt, \quad (8)$$

where  $\Phi(t)$  is the characteristic function. In applications, we need use numerical methods to solve the integral in (8), 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. An 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 Damien, et al 1999; Tanner and Wong, 1987).

In this way, a Bayesian analysis of stable distributions was introduced by Buckle(1995) using Markov Chain Monte Carlo (MCMC) methods and latent variables (see also, Achcar et al, 2013). 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.

Considering a latent or an auxiliary variable, Buckle (1995) proved a theorem that is 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 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,

$$a_{\alpha,\beta} = -\frac{b_{\alpha,\beta}}{\alpha\pi}, \quad (9)$$

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

$$f(z, y | \alpha, \beta) = \frac{\alpha}{|\alpha-1|} \exp \left\{ -\left| \frac{z}{t_{\alpha,\beta}(y)} \right|^\alpha \right\} \left| \frac{z}{t_{\alpha,\beta}(y)} \right|^{\frac{1}{|\alpha-1|}} \quad (10)$$

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

$$t_{\alpha,\beta}(y) = \left( \frac{\sin(\pi\alpha y + b_{\alpha,\beta})}{\cos(\pi y)} \right) \left( \frac{\cos(\pi y)}{\cos(\pi(\alpha-1)y + b_{\alpha,\beta})} \right)^{\frac{1}{\theta}}, \quad (11)$$

and

$$Z = \frac{X-\gamma}{\delta}, \text{ for } \delta \neq 0.$$

From the bivariate density (10), Buckle (1995) shows the marginal distribution for the random variable  $Z$  is a stable  $S_\alpha(\beta, 0, 1)$  distribution. 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 (Spiegelhalter et al, 2003).

### 3 A Bayesian approach assuming a stable distribution

In this section, let us assume that the response  $x_i$  in the nonlinear regression model (1) for  $i = 1, \dots, n$ , have a stable distribution  $X_i \sim S_\alpha(\beta, \gamma_i, \delta)$ , that is,  $Z_i = \frac{x_i - \gamma_i}{\delta} \sim S_\alpha(\beta, 0, 1)$  and where the location parameter  $\gamma_i$  of the stable distribution is related to  $k$  explanatory variables  $v_i = (v_{1i}, v_{2i}, \dots, v_{ki})$  by a nonlinear relation given by,

$$\gamma_i = f(v_i, \mathbf{d}), \quad (12)$$

Where  $\mathbf{d} = (d_1, d_2, \dots, d_k)$  are the associated unknown regression parameters. Assuming a joint prior distribution for  $\alpha, \beta, \mathbf{d}$  and  $\delta$ , where  $\mathbf{d} = (d_1, d_2, \dots, d_k)$  given by  $\pi_0(\alpha, \beta, \mathbf{d}, \delta)$ , Buckle (1995) shows that the joint posterior distribution for parameters  $\alpha, \beta, \mathbf{d}$  and  $\delta$ , is given by

$$\begin{aligned} \pi(\alpha, \beta, \mathbf{d}, \delta / \mathbf{x}) \propto & \int \left( \frac{\alpha}{|\alpha - 1| \delta} \right)^n \exp \left\{ - \sum_{i=1}^n \left| \frac{z_i}{t_{\alpha, \beta}(Y_i)} \right|^\theta \right\} \prod_{i=1}^n \left| \frac{z_i}{t_{\alpha, \beta}(Y_i)} \right|^\theta \\ & \cdot \frac{1}{|z_i|} \pi_0(\alpha, \beta, \mathbf{d}, \delta) d\mathbf{y}, \end{aligned} \quad (13)$$

where  $\theta = \frac{\alpha}{\alpha - 1}$ ,  $z_i = \frac{x_i - \gamma_i}{\delta}$ , for  $i = 1, \dots, n$ ,  $\alpha \in (0, 2]$ ,  $\beta \in [-1, 1]$  and  $\delta \in (0, \infty)$ ;  $\mathbf{x} = (x_1, x_2, \dots, x_n)$  and  $\mathbf{y} = (y_1, y_2, \dots, y_n)$  are respectively, the observed and non-observed data vectors. Observe that the joint distribution in expression (13) 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  $\sigma^{-1}$  multiplied by the right-hand-side of expression (10)).

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  $\delta$  and variance equals to  $2\sigma^2$ .

For a Bayesian analysis of the proposed model, we assume uniform  $U(a, b)$  priors for  $\alpha, \beta$  and  $\delta$  where the hyperparameters  $a$  and  $b$  are assumed to be known in each application following the restrictions  $a \in (0, 2]$ ,  $\beta \in [-1, 1]$  and  $\delta \in (0, \infty)$ . We also assume Normal  $N(a, b^2)$  prior distributions for the regression parameters  $\mathbf{d} = (d_1, d_2, \dots, d_k)$  considering known hyper parameter 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  $\alpha, \beta, \mathbf{d}$  and  $\delta$ , having the joint posterior distribution (13), we assume a uniform  $U(-0.5, 0.5)$  prior distribution for the latent random quantities  $Y_i$  for  $i = 1, \dots, n$ . 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 Spiegelhalter, 2003) which gives great

simplification to obtain the simulated Gibbs samples for the joint posterior distribution of interest.

From expression (13), the joint posterior probability distribution for  $\alpha, \beta, \mathbf{d}, \delta$  and  $\mathbf{y} = (y_1, y_2, \dots, y_n)$  is given by

$$\pi(\alpha, \beta, \mathbf{d}, \delta, \mathbf{y} | x) \propto \left(\frac{\alpha}{|\alpha-1|\delta}\right)^n \exp\left\{\sum_{i=1}^n \left|\frac{z_i}{t_{\alpha,\beta}(y_i)}\right|\right\} \prod_{i=1}^n \left|\frac{z_i}{t_{\alpha,\beta}(y_i)}\right|^{\theta} \frac{1}{|z_i|} \cdot \prod_{i=1}^n h(y_i) \pi_0(\alpha, \beta, \mathbf{d}, \delta), \quad (14)$$

where  $t_{\alpha,\beta}(\cdot)$  are respectively defined in (10) and (11) and  $h(y_i)$  is a  $U(-0.5, 0.5)$  density function, for  $i = 1, \dots, n$ .

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.

#### 4 An example with intrauterine growth data for preterm infants in USA

In Table 1, we have data set related to intrauterine growth data (IGD) for preterm infants in USA (see Olsen et al, 2010). In this data set we have the weight of the preterms and the gestational age (GA) of the mothers (55445 girls and 73995 boys).

From a preliminary data analysis, we see that a nonlinear regression model (1) is suitable for the data set in the transformed scales (see Figure 1).

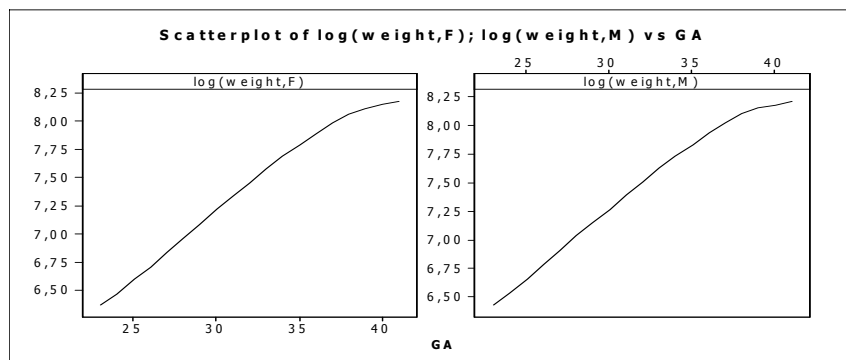


Figure 1 - Plots of observed log(weight) versus gestational age.

Using the information of Table 1, since we do not have all data (intrauterine weights of 55445 girls and 73995 boys) reported in the paper of Olsen et al (2010) but we have the sample means and the standard deviations for each group of infants in a fixed gestational age, we simulated samples of the weights considering normal distributions with means and standard-deviations equal to the sample means and sample standard deviations (simulation of 553 weights for girls and 738 weights for boys) with gives a sample size close to 10% of all data introduced in Table 1.

Table 1 - Intrauterine growth data for preterm infants (55445 girls and 73995 boys)

#	GA	size F	mean, F	size M	mean, M
123	133	587	153	622	
224	438	649	451	689	
325	603	738	722	777	
426	773	822	881	888	
527	966	934	1030	1001	
628	1187	1058	1281	1138	
729	1254	1199	1505	1277	
830	1606	1376	1992	1435	
931	2044	1548	2460	1633	
1032	3007	1730	3677	1823	
1133	4186	1960	5014	2058	
1234	5936	2194	7291	2288	
1335	5082	2420	6952	2529	
1436	4690	2675	7011	2798	
1537	4372	2946	6692	3058	
1638	5755	3184	8786	3319	
1739	5978	3342	8324	3476	
1840	5529	3461	7235	3582	
1941	1906	3546	2538	3691	

(GA=gestational age; F=female; M=male; size=sample size)

#### 4.1 Gompertz model with normal errors (classical approach)

Assuming standard classical approach for nonlinear models considering the Gompertz growth model ( $v$ ), that is,

$$f(v, \theta) = \theta_1 \exp[-\exp(\theta_2 - \theta_3 v)], \quad (15)$$

where  $v = \text{GA}$  (gestational age), introduced in section 2, with the responses in the logarithmic scale  $\log(\text{weight})$ , and the software Minitab version 16, we have in Table 2, the obtained maximum likelihood estimates (MLE) for the parameters of the Gompertz growth model for each sex. We have used initial values equal to zero for all parameters  $\theta_1$ ,  $\theta_2$  and  $\theta_3$ , of the Gompertz model in the Gauss-Newton iterative procedure to get the MLE.

Table 2 - MLE estimates for the Gompertz growth model

Gender	$\theta_1$	$\theta_2$	$\theta_3$
F (girls)	9.06173 (0.15523)	0.76876 (0.14414)	0.07561 (0.00736)
M(boys)	9.14384 (0.14679)	0.77103 (0.13259)	0.07488 (0.00674)

(In parentheses: SE of estimate)

From residuals plots (see Appendix 1), we see that the standard assumptions for the model are verified.

#### 4.2 Gompertz model with normal errors (Bayesian approach)



Under a Bayesian approach, we have in Table 3, the posterior summaries of interest considering the Gompertz growth model with a normal distribution for the error (see (1)) and the OpenBugs software assuming the following prior distributions:  $\theta_j \sim N(0, 1)$ ,  $j = 1, 2, 3, 4$  and  $\zeta = 1/\sigma^2 \sim G(1, 1)$ , where  $G(a, b)$  denotes a gamma distribution with mean  $a/b$  and variance  $a/b^2$ . We simulated 311,000 Gibbs samples, with a “burn-in-sample” of 11,000 samples discarded to eliminate the effects of the initial values in the iterative simulation process and taking a final sample of size 1000 (every 300<sup>th</sup> sample chosen from the 300,000 final simulated samples). Convergence of the Gibbs sampling algorithm was monitored from standard trace plots of the simulated samples. For the girls and boys we obtained respectively, Monte Carlo estimates for the posterior mean of  $\zeta$  equal to 31.46 (1.907) and 31.30 (1.596).

Table 3 - Bayesian estimates for the Gompertz growth model

Gender	$\theta_1$	$\theta_2$	$\theta_3$
F (girls)	8.891 (0.1166)	0.9545 (0.1421)	0.08503 (0.00684)
M(boys)	9.061 (0.1094)	0.8658 (0.1322)	0.07948 (0.00626)

(In parentheses: SE of estimate)

It is interesting to observe that the classical and Bayesian inferences assuming non-informative priors are very similar, but the standard-deviations of the estimates in general are smaller for the Bayesian approach, especially for the parameter  $\theta_1$ .

#### 4.3 Gompertz model with stable distribution (Bayesian approach)

Now, we assume a stable distribution for the intrauterine weights of the preterm infants (girls group or female infants) considering the Gompertz growth curve model (model 5) introduced in section 2. Under a Bayesian approach, we have in Tables 4, the posterior summaries of interest assuming the nonlinear Gompertz regression model, that is, considering the regression model for the location parameter of the stable distribution given by,

$$\gamma_i = d_1 \exp[-\exp(d_2 - d_3 G A_i)]. \quad (16)$$

In the Bayesian analysis of this model, we use the OpenBugs software assuming the following prior distributions:  $\alpha \sim U(0, 2)$ ,  $\beta \sim U(-1, 0)$ ,  $\delta \sim U(0, 3)$ ,  $d_1 \sim N(9, 1)$ ,  $d_2 \sim N(0.8, 1)$  and  $d_3 \sim N(0.07, 1)$ . Observe that we are using some prior information from the results of the classical analysis to choose the prior distributions for the regression parameters (use of empirical Bayesian methods). We also assume a uniform  $U(-0.5, 0.5)$  distribution for the latent variable  $Y_i$ ,  $i=1, 2, \dots, 15$ . We simulated 4,000,000 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 200<sup>th</sup> sample chosen from the 200,000 final samples). Convergence of the Gibbs sampling algorithm was monitored from standard trace plots of the simulated samples.

Table 4 - Bayesian estimates for the Gompertz growth model assuming a stable distribution for the responses (results for girls)

parameter	mean	S.D.	95% credible interval
$\alpha$	1.543	0.00509	(1.536 ; 1.555)
$\beta$	-0.8361	0.00883	(-0.8449 ; -0.8204)
$\delta$	0.1418	0.0051	(0.1356 ; 0.1579)
$d_1$	9.322	0.035	(9.270 ; 9.355)
$d_2$	0.6948	0.00756	(0.680 ; 0.711)
$d_3$	0.0714	0.00073	(0.0707 ; 0.0726)

From the results of Table 4, we observe that in general the Bayesian estimates of the regression parameters ( $d_1, d_2$  and  $d_3$ ) considering the Gompertz model with a stable distribution are similar to the Bayesian estimates of the regression parameters ( $\theta_1, \theta_2$  and  $\theta_3$ ) considering the Gompertz model with normal errors, but in general the standard deviations are smaller than assuming normal errors (see results in Table 3).

#### 4.4 Gompertz model with stable distribution in presence of outliers (Bayesian approach)

Now, let us consider the presence of outliers (female infants) replacing the 5<sup>th</sup> intrauterine weight (732.58) by 73200.58 and the first intrauterine weight (587.00) by 58700.00. In Table 5, we have the obtained posterior summaries assuming the same priors and simulation procedure assumed for the results of Table 4. We also observe in Table 5, that the estimated regression parameters  $\theta_1, \theta_2$  and  $\theta_3$  of the Gompertz non-linear model with normal errors are strongly affected by the presence of outliers (see results of Table 2). Considering a stable distribution, the estimates of the regression parameters ( $d_1, d_2$  and  $d_3$ ) of the Gompertz non-linear model are very close to the estimates of the regression model with normal errors not considering the presence of outliers. This is a great advantage of the non-linear model assuming a stable distribution, that is, the Gompertz regression model with a stable distribution is robust to the presence of outliers.

Table 5 - Bayesian estimates for the Gompertz growth model assuming a stable distribution for the responses (presence of an outlier)

Parameter (stable)	mean	S.D.	95% credible interval
$\alpha$	1.543	0.0051	(1.536 ; 1.555)
$\beta$	-0.8361	0.0088	(-0.8449; -0.8204)
$\delta$	0.1418	0.0051	(0.1356 ; 0.1579)
$d_1$	9.322	0.035	(2.098 ; 3.104)
$d_2$	0.6948	0.0076	(0.680 ; 0.7115)
$d_3$	0.0714	0.00073	(0.0707 ; 0.0726)
Parameter (normal)	mean	S.D.	95% credible interval
$\theta_1$	9.224	0.2116	(8.825 ; 9.675)
$\theta_2$	0.3107	0.1279	(0.0772 ; 0.6047)
$\theta_3$	0.0597	0.0072	(0.0475 ; 0.0760)

#### Concluding Remarks

The presence of outliers or discordant observations, many times due to measure errors is very common in applications of nonlinear regression analysis. In the presence of these discordant observations, the usual obtained classical inferences on the regression parameters or in the predictions 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 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 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 OpenBugs software.

In the illustrative example related to infant intrauterine growth introduced in Sections 4, the use of data augmentation techniques (see, for instance, Damien et al, 1999) 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 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.

ACHCAR, J. A.; OLIVEIRA, B. R. G. Curvas de crescimento de Gompertz assumindo distribuições estáveis: uma aplicação ao crescimento intra-uterino para prematuros. *Rev. Bras. Biom.*, São Paulo, v.33, n.2, p.170-183, 2015.

- RESUMO: Neste artigo, consideramos o uso de métodos bayesianos para analisar alguns modelos de crescimento que pertencem à classe de modelos de regressão não-linear. Para a modelagem não linear assumimos erros normais uma suposição usual e também distribuições estáveis para a variável resposta. Estudamos também alguns aspectos de robustez dos modelos de regressão não linear para a presença de outliers ou observações discordantes considerando o uso de distribuições estáveis para a resposta no lugar da suposição de normalidade habitual. É bem sabido que, em geral, não há nenhuma forma fechada para a função densidade de probabilidade de distribuições estáveis. No entanto, sob uma abordagem Bayesiana, a utilização de uma variável aleatória latente ou auxiliar proporciona uma simplificação para obter qualquer distribuição a posteriori quando relacionado com distribuições estáveis. Para demonstrar a utilidade dos aspectos computacionais, a metodologia é aplicada a um exemplo relacionado com as curvas de crescimento intra-uterino para prematuros. Resumos a posteriori de interesse são obtidos utilizando métodos MCMC (Markov Chain Monte Carlo) e o software OpenBugs.
- PALAVRAS-CHAVE: Distribuição estável; análise Bayesiana; modelos de regressão não-lineares; métodos MCMC; software OpenBugs.

## References

- ACHCAR, J.A.; LOPES, S.R.C. ; MAZUCHELI, J.; LINHARES, R. A Bayesian Approach for Stable Distributions: Some Computational Aspects. *Open Journal of Statistics*, v.3, n.4, p.268-277, 2013.
- BATES, D. M. ; WATTS, D .G. *Nonlinear Regression Analysis and Its Applications*, Wiley, New York, 1988, 392 p.
- BATES, D.M. ; WATTS, D.G. Relative curvature measures of nonlinearity (with discussion), *Journal of the Royal Statistical Society, Series B*, v. 42, n.1, p.1-25, 1980.
- BUCKLE, D.J. Bayesian inference for stable distributions, *Journal of the American Statistical Association*, v.90, n.430, p. 605-613, 1995.
- CASELLA, G.; GEORGE, E. I. Explaining the Gibbs Sampler. *The American Statistician*, v.46, n.3, p.167-174, 1992.
- CHIB, S.; GREENBERG, E. Understanding the Metropolis- Hastings algorithm, *The American Statistician*, v.49, n.4, p.327-333, 1995.
- DAMIEN, P.; WAKEFIELD, J.; WALKER, S. Gibbs sampling for Bayesian non-conjugate and hierarchical models by using auxiliary variables, *Journal of the Royal Statistical Society, Series B*, v.61, n.2, p. 331-344, 1999.
- FELLER, W. *An introduction to probability theory and its applications*, Vol. II. John Wiley, New York, 1971, 626 p.
- GELFAND, A. E.; SMITH, A.F.M. Sampling based approaches to calculating marginal densities, *Journal of the American Statistical Association*, v.85, n.410, p. 398-409, 1990.
- GNEDENKO, B. V. ; KOLMOGOROV, A.N. Limit distributions for sums of independent random variables, Addison-Wesley, Massachusetts, 1968, 282 p.
- IBRAGIMOV, I.A.; ČERNIN, K.E. On the unimodality of stable laws, *Teor. Veroyatnost. i Primenen.*, v.4, p.453-456, 1959.
- KANTER, M. On the unimodality of stable densities, *Annals of Probability*, v.4, n.6, p. 1006-1008, 1976.
- LÉVY, P (1924) Théorie des erreurs la loi de Gauss et les lois exceptionnelles, *Bulletin Society Mathematical*, v. 52, p. 49-85, 1924.
- LUKACS, E. *Characteristic Functions*, Hafner Publishing, New York, 1970, 216 p.
- NOLAN, J. P. *Stable Distributions - Models for Heavy Tailed Data*, Birkhäuser, Boston. In progress, Chapter 1 online at [academic2.american.edu/~jpnolan](http://academic2.american.edu/~jpnolan), 2009.
- OLSEN, I.E.; GROVEMAN, S.A.; LAWSON, L.; CLARK, R.H.; ZEMEL, B.S. New Intrauterine Growth Curves Based on United States Data, *Pediatrics*, doi:10.1542/peds.2009-0913, 2010.
- RATKOWSKY, D.A. *Nonlinear Regression Modelling: A Unified Practical Approach*, Marcel Dekker, New York, 1983, 288 p.
- ROUSSAS, G.G. *Measure-Theoretic probability*, San Diego: Elsevier, 2005, 390 p.
- SEBER, G.A.F. ; WILD, C.J. *Nonlinear Regression*, Wiley, New York, 1989, 792 p.

SKOROHOD, A.V. On a theorem concerning stable distributions. In: *Select. Transl. Math. Statist. and Probability*, v. 1, Inst. Math. Statist. and Amer. Math. Soc., Providence, R.I., p.169-170, 1961.

SPIEGELHALTER, D.J.; THOMAS, A.; BEST, N.G.; LUNN, D. WinBugs User's Manual, *MRC Biostatistics Unit*, Cambridge, 2003.

SPIEGELHALTER, D.J.; BEST, N. G., CARLIN, B. P., BAN DER LINDE, A. Bayesian measures of model complexity and fit, *Journal of the Royal Statistical Society, Series B*, v.64, n.4, p.583-639, 2002.

TANNER. M.A. ; WONG, W.H. The calculation of posterior distributions by data augmentation”, *Journal of American Statistical Association*, v. 82, n.398, p. 528-550, 1987.

Recebido em 16.02.2015

Aprovado após revisão em 26.04.2015

Appendix 1 - residual plots.

