ISSN 0104-0499

# PROGRAMA INTERINSTITUCIONAL DE PÓS-GRADUAÇÃO EM ESTATÍSTICA UFSCar-USP

## DEs-UFSCar e SME-ICMC-USP

# L-LOGISTIC DISTRIBUTION: PROPERTIES, INFERENCE AND AN APPLICATION TO STUDY POVERTY AND INEQUALITY IN BRAZIL

**Rosineide F. da Paz**
**Narayanaswamy Balakrishnan**
**Jorge Luis Bazán**

## RELATÓRIO TÉCNICO

## TEORIA E MÉTODO – SÉRIE A

**Abril/2016**
**nº 262**

# L-logistic distribution: Properties, Inference and an Application to Study Poverty and Inequality in Brazil

Rosineide F. da Paz* & Narayanaswamy Balakrishnan[1] & Jorge Luis Bazán[2]

April 7, 2016

**Abstract**

A two-parameter distribution on a bounded domain is studied in this work. This distribution, called l-logistic distribution, provides great flexibility and has the uniform distribution as a particular case. In addition, it has an explicit distribution function that facilitates easy random number generation. Several properties of the distribution are studied including skewness and kurtosis. Bayesian inference is discussed with non informative and informative prior distributions. Simulation studies considering prior sensitivity analysis and parameter recovery studies show the robustness of the proposed estimation method and the efficiency of the algorithm adopted. In addition, a regression model considering with the response following the l-logistic distribution is introduced. Applications to Study Poverty and Inequality in Brazil are performed showing a comparison of results obtained from the beta and l-logistic distributions. The obtained results show that when the contain potential outliers, the l-logistic model provides a better fit for these data than the beta model.

**Key words:** Bayesian analysis, l-logistic distribution, Regression analysis, Gini index, Beta distribution.

## 1 Introduction

Despite many alternatives and generalizations present in the literature, the beta distribution is still a popular family of continuous distributions with bounded support. Recently, Gómez-Déniz et al. (2014) and Mitnik and Baek (2013) proposed new alternatives to beta distribution. However, there are still continuous distributions with continuous bounded support that need further study. For example, Kotz and van Dorp (2004) discuss such distributions with attractive statistical properties that can be useful in various applied fields. In this work, we discuss a continuous distribution on

* *Universidade Federal and Universidade de São Paulo, São Carlos-SP, Brazil. E-mail: rfpaz@icmc.usp.br*

[1] *McMaster University, Department of Mathematics and Statistics, Hamilton, Ontario, Canada.*

[2] *Universidade de São Paulo, Instituto de Ciências Matemáticas e de Computação, São Carlos-SP, Brazil.*

unit interval, which we refer to as l-logistic distribution. We discuss some of its properties concerning the shape of the probability density function (pdf), the cumulative distribution function (cdf), the percentiles, and some re-parameterizations. The generation of random numbers from the l-logistic distribution is also discussed. The moments and some descriptive measures such as mode, skewness, and kurtosis are also presented. The l-logistic distribution was originally proposed by Tadikamalla and Johnson (1982), but a new parameterization is presented here, with an emphasis on regression analysis, based on this model.

We adopt a Bayesian approach using Markov chain Monte Carlo (MCMC) algorithm for our modeling framework. The issues of model fitting is addressed via Gibbs sampling suitable, choice of prior distributions, and model comparison criteria. In this context, we report two studies with simulated data sets to investigate the performance of the proposed estimation method concerning parameter recovery and influence of the prior distributions. Results obtained from the simulation studies display that the proposed estimation works very well.

Finally, we apply the model to social data, in which the proportion of children vulnerable to poverty and the Gini index of the municipalities of the state of Alagoas in Brazil for the 2010 season are modeled. The Gini index is modeled as a function of the percentage of people employed in the manufacturing industry. For the case of the l-logistic distribution, we use the median regression model in the context of quantile regression. Quantile regression (QR) models were introduced by Koenker and Bassett (1978) and can model conditional quantiles as functions of predictors. The median regression model accomplishes the same goal in order to represent the relationship between the median (central location) of the response and a set of covariates. If the data are highly skewed, since the median remains robust is a natural measure of the center, the conditional median modeling is more useful than conditional mean modeling in this context. This fact is seen in our first application.

The rest of the paper is organized as follows. In Section 2, we present the pdf, the cdf, the quantile function, and also describe the generation of the l-logistic distribution. In Section 3, we study some characteristics of the distribution, other parameterizations, some related distributions, the moments, and the skewness and kurtosis of the l-logistic distribution. Section 4 is dedicated to the Bayesian estimation of the model parameters. Some methods for model comparison and diagnosis are also discussed in this section. Section 5 presents the results of a simulation study that examines a prior sensitivity analysis and the estimation of the model parameters. Section 6 discusses an application of the model to a on social data from the state of Alagoas, Brazil (proportion of children vulnerable to poverty and Gini index as function of the percentage of people employed in manufacturing). Finally, some concluding comments are made in Section 7.

## 2 The L-logistic Distribution

We say that the r.v. $Y$ follows a l-logistic distribution if its probability density function (pdf) is given by

$$f(y|m,b) = \frac{b(1-m)^b m^b y^{b-1}(1-y)^{b-1}}{\left[(1-m)^b y^b + m^b(1-y)^b\right]^2}, \quad 0 < y < 1, 0 < m < 1, \quad b > 0. \tag{1}$$

The parameters $m$ and $b$ allow l-logistic distribution, denoted by $Y \sim LL(m,b)$, to take on a variety of density shapes (see Figure 1 and 2). Note that when we set $m = 0.5$ and $b = 1$ in (1), then the pdf of the l-logistic distribution simply becomes the pdf of the standard uniform distribution. $m$ is a location parameter, which simply shifts the graph to the left or right on the horizontal axis. On the other hand, $b$ is a shape parameter that governs the shape of the distribution. The l-logistic density is uni-modal (or "uni-antimodal"), increasing, decreasing, or constant, depending on the values of its parameters. More details on this issue are presented with another parameterization of the l-logistic model in Section 3.

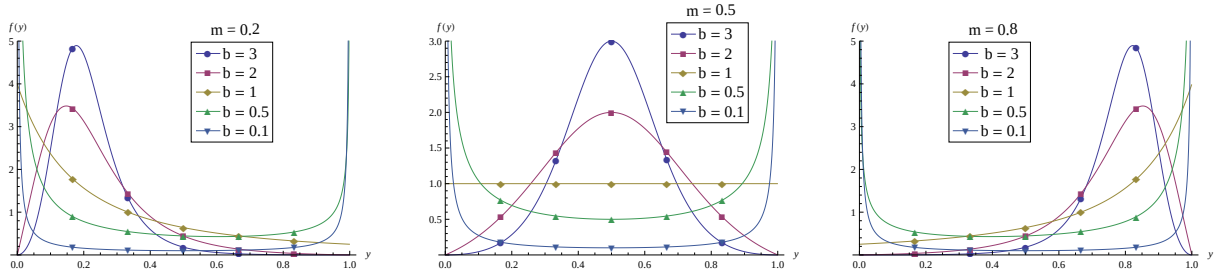


Figure 1: L-logistic probability density function for scale parameter $m = 0.2, 0.5$ and $0.8$ and some values of parameter $b$.

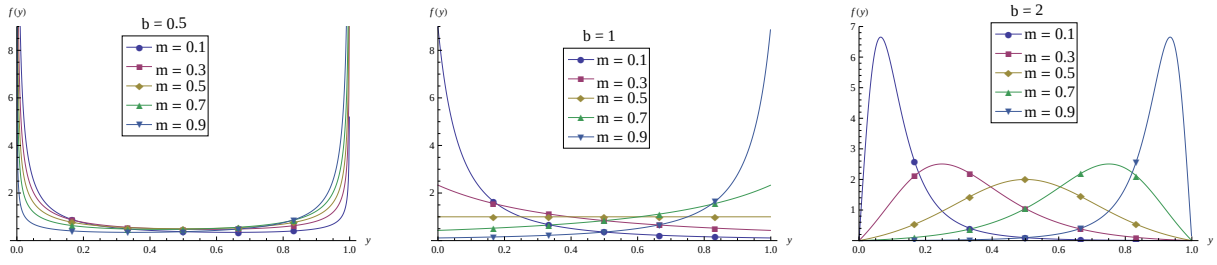


Figure 2: L-logistic probability density function for shape parameter $b = 0.5, 1$ and $2$ and some values of scale parameter $m$.

The cumulative distribution function of the l-logistic distribution is given by

$$F_Y(y|m,b) = \left(1 + \left(\frac{m(1-y)}{y(1-m)}\right)^b\right)^{-1}, \quad 0 < y < 1. \tag{2}$$

which can be readily inverted to yield the quantile function.

$$Q_Y(p) = F_Y^{-1}(p) = \frac{mp^{1/b}}{(1-p)^{1/b}(1-m) + p^{1/b}m}, \quad 0 \le p \le 1. \tag{3}$$

This would readily enable a quantile-based analysis of this model; for example, see Nair et al. (2013) for pertinent details. Note that if $p = (1-p) = 0.5$, then $Q(p) = m$, which means that the location parameter $m$ is indeed the $50^{th}$ percentile or the median of the l-logistic distribution.
Equation (3) facilitates simple random variate generation. If $U \sim Uniform(0,1)$, then

$$X = Q(U) = \frac{mU^{1/b}}{(1-U)^{1/b}(1-m) + U^{1/b}m} \sim LL(m,b) \tag{4}$$

We can also express the inter-quartile range (IQR) as

$$IQR = Q(0.75) - Q(0.25) = \frac{m3^{1/b}}{(1-m) + 3^{1/b}m} - \frac{m}{3^{1/b}(1-m) + m}. \tag{5}$$

The IQR has a breakdown point of 50%, and is often preferred over range. When the distribution is symmetric, half IQR equals the median absolute deviation (MAD), and it is often to detect outliers in data.

# 3 Properties of the L-logistic distribution

This section discuss some properties of the L-logistic distribution such as some related distributions, measures of skewness and kurtosis, mode and the moments. Some pertinent derivations are presented in Appendix A.

## 3.1 Related Distributions

The following property shows the relation between l-logistic distribution and logistic distribution.

**Property 3.1.** *If* $Y \sim LL(m,b)$ *then* $Z = b\log\left(\frac{Y(1-m)}{m(1-Y)}\right)$ *has the standard logistic distribution.*

Next, we present two alternative parameterizations for the l-logistic distribution.

**Property 3.2.** *If* $Y \sim LL(m,b)$*, then, with* $m = \frac{e^{-\frac{\delta}{b}}}{1+e^{-\frac{\delta}{b}}}$ *(*$\delta > 0$*), the pdf and cdf of the l-logistic distribution become*

$$f(y|\delta,b) = \frac{be^{\delta}y^{b-1}(1-y)^{b-1}}{\left[y^b e^{\delta} + (1-y)^b\right]^2}, \;\; 0 < y < 1, \tag{6}$$

*and*

$$F_Y(y|\delta,b) = \left(1 + e^{-\delta}\left(\frac{1-y}{y}\right)^b\right)^{-1}, \;\; 0 < y < 1, \tag{7}$$

*respectively, where* $b > 0$ *and* $\delta \in \mathbb{R}$ *are both shape parameters.*

The parameterization given in (6) and (7) is denoted by $Y \sim LL(\delta, b)$. The pdf of $Y \sim LL(\delta, b)$ was introduced by Tadikamalla and Johnson (1982) and Wang and Rennolls (2005), which extended

this pdf for any bounded interval by introducing two extra parameters. These authors referred to this distribution as logit-logistic distribution.

**Property 3.3.** *If* $Y \sim LL(m, b)$ *and* $\mu = \frac{1}{1+\left(\frac{m}{1-m}\right)^b}$, *then the pdf and cdf the alternative parametrization of l-logistic distribution, denoted by* $Y \sim LL(\mu, b)$, *is given by*

$$f(y|\mu, b) = \frac{b\mu(1-\mu)y^{b-1}(1-y)^{b-1}}{\left[y^b\mu + (1-y)^b(1-\mu)\right]^2}, \;\; 0 < y < 1, \tag{8}$$

*and*

$$F_Y(y|\mu, b) = \left(1 + \left(\frac{1-\mu}{\mu}\right)\left(\frac{1-y}{y}\right)^b\right)^{-1}, \;\; 0 < y < 1, \tag{9}$$

*with* $\mu \in (0, 1)$ *and* $b > 0$.

Note that, although the expressions of the alternative parameterization of the l-logistic distribution have simple expresions for the pdf and cdf, they do not have the median as location parameter. An advantage of the parametrization in (1) and (2) is that the parameter $m$ is the median of the distribution and, consequently, the interpretation of this parametrization becomes more easier.

**Property 3.4.** *If* $Y \sim LL(\delta, b)$, *then* $Z' = \delta + b\log\left(\frac{Y}{1-Y}\right)$ *has the standard logistic distribution.* The reciprocal property was first introduced by Tadikamalla and Johnson (1982), based on an equivalent transformation described by Johnson et al. (1994, pg. 34-49) and first investigated by Johnson (1949), for the case of the standard normal distribution. In this case, $Y$ follows the SB-Johnson distribution.

**Property 3.5.** *A more general form of the l-logistic distribution can be given as*

$$f(y|m, b, c, d) = \frac{(d-c)b(1-m)^b m^b (y-c)^{b-1}(d-y)^{b-1}}{\left[(1-m)^b(y-c)^b + m^b(d-y)^b\right]^2}, \tag{10}$$

$c < y < d$, *with* $c, d \in \mathbb{R}$. To see justifications of these propiertes, see Appendix A.

## 3.2 Mode

**Property 3.6.** *For* $b > 1$, *the mode of the l-logistic distribution is the solution of the equation*

$$\left(\frac{1-m}{m}\right)^b = \left(\frac{1-y_0}{y_0}\right)^b \frac{b + 2y_0 - 1}{b - 2y_0 + 1}. \tag{11}$$

Note that, upon taking $\delta = -b\log\left(\frac{m}{1-m}\right)$, the mode $y_0$ can be obtained by solving the equation

$$\delta = \log\left(\left(\frac{1-y_0}{y_0}\right)^b \frac{b + 2y_0 - 1}{b - 2y_0 + 1}\right). \tag{12}$$

In addition, from (11) and (12), if $y_0 = m = 0.5$, then $\delta = 0$ for all values of $b$. Thus, we can study the behavior of the mode by studying the function given in (12). For this purpose, we take the derivative

of the right-hand side of (12) with respect to $y_0$ to obtain the equation

$$\frac{\partial \delta}{\partial y_0} = \frac{b\left(b^2-1\right)}{(y_0-1)y_0\left\{(b^2-1)+4y_0-4{y_0}^2\right\}}. \tag{13}$$

(11) is negative for $b > 1$, the situation where $\delta$ decreases as $y_0$ (mode) increases (first derivative test), then the mode lies in $(0, 1/2)$ if $\delta > 0$ (or $m < 1/2$) and for $\delta < 0$ (or $m > 1/2$) the mode is in $(1/2, 0)$. If $b < 1$, (13) is positive whenever $\left\{(b^2-1)+4y_0-4y_0^2\right\} > 0$, that is, whenever $\frac{1-b}{2} < y < \frac{1+b}{2}$, the situation where $\delta$ increases as $y$ increases. Thus, from (12) and (13) the minimum of the pdf lies in $(\frac{1-b}{2}, 1/2)$ for $\delta < 0$ or $m > 1/2$, and in $(1/2, \frac{1+b}{2})$ for $\delta > 0$ or $m < 1/2$.

### 3.3 Skewness and Kurtosis

First, we have the following symmetric property.

**Property 3.7.** *The l-logistic density is symmetric when $m = 0.5$ whatever the value of $b$ is.*

For the case when the l-logistic density is asymmetric, the degree of skewness can be quantified by some measures of skewness. Since the l-logistic distribution is related to the logistic distribution, the skewness measure introduced by Arnold and Groeneveld (1995), and denoted by $\gamma_M$, seems to be an appropriate skewness measure. The measure $\gamma_M$ is based on the mode of distribution and is given by

$$\gamma_M = 1 - 2F(M), \tag{14}$$

where $M$ is the mode of the distribution and $F(.)$ is the distribution function. The value of $\gamma_M$ lies in $(-1, 1)$, and if $\gamma_M$ is near 1, it indicates extreme right skewness. On the other hand, if $\gamma_M$ is near -1, it indicates extreme left skewness.

We also consider another measure of skewness called octile skewness (denoted here by $\gamma_p$), first proposed by Hinkley (1975) and discussed further Brys et al. (2003). This skewness measure is given by

$$\gamma_p = \frac{Q(1-p)+Q(p)-2m}{Q(1-p)-Q(p)}, \tag{15}$$

which is a function of high and low percentiles defined by $p \in (0, 0.5)$ with $Q(.)$ as in (3). The maximum value of $\gamma_p$ is 1, representing extreme right skewness, and the minimum is -1, representing extreme left skewness. This measure is also zero for any symmetric distribution. However, the function in (15) depends on the value of $p$. We can remove this dependence by integrating over $p$, (see Groeneveld and Meeden, 1984), or to decide which value of $p$ is appropriate for use. In Brys et al. (2003), there is a comparison between several robust skewness measures in which accuracy, robustness, and computational complexity are considered. The most interesting skewness measure of the measures

investigated is octile skewness. Octile skewness takes $p = 0.125$ in (15), that is, it is given by

$$\gamma_{125} = \frac{Q(0.875) + Q(0.125) - 2m}{Q(0.875) - Q(0.125)}. \tag{16}$$

For the l-logistic distribution, we mase use of this particular skewness measure instead of removing the dependence over $p$ thorough integration.

Moreover, the kurtosis of the l-logistic distribution can also be derived easily by using the quantiles. The kurtosis measure introduced by Moors (1988) is given by

$$k_Q = \frac{Q(0.875) - Q(0.625) - Q(0.375) + Q(0.125)}{Q(0.75) - Q(0.25)}, \tag{17}$$

with $k_Q \in (0, \infty)$.

Figure 3 presents the results of the measures of skewness and kurtosis described here for some values of the location parameter $m$ as a function of the shape parameter $b$, $b > 1$. In this figure, we can see that the two measures of skewness become close as $b$ increases, as intuition suggests. Moors (1988) justified the use of the kurtosis measure in 17 by the interpretation that the two terms in the numerator of (17) are large (small) if relatively little (much) probability mass is concentrated in the neighborhood of $Q(0.75)$ and $Q(0.25)$. This corresponds to large (small) dispersion around (roughly) $E_Y[Y] \pm Var_Y[Y]$ where $E_Y[Y]$ and $Var_Y[Y]$ are the mean and variance of r.v. $Y$, respectively.

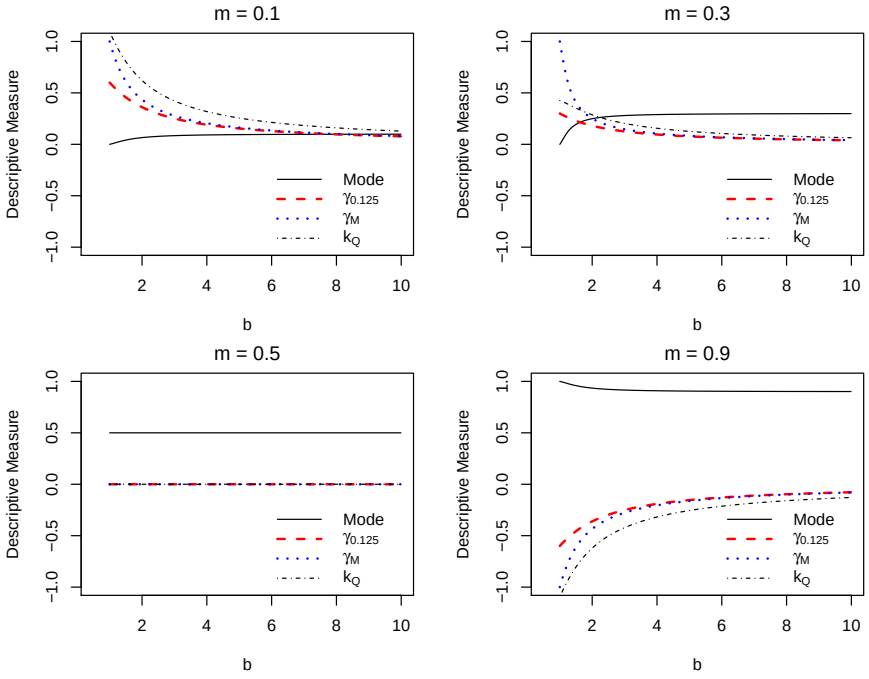


Figure 3: The mode, skewness ($\gamma_M$ and $\gamma_{0.125}$) and kurtosis ($k_Q$) of the l-logistic distribution for some values of the parameters.

### 3.4 Moments

The following propositions gives an expression for the moments of the l-logistic distribution.

**Property 3.8.** *If $Y \sim LL(m, b)$, then the t-th moment of $Y$ about zero is given by*

$$E[Y^t] = \int_0^1 \left[1 + \left(\frac{1-v}{v}\right)^{1/b} \left(\frac{1-m}{m}\right)\right]^{-t} dv. \tag{18}$$

The integral in (18) cannot be expressed in an analytical form. However, we can use numerical integration to evaluate some moments as $E_Y(Y)$, $E_Y(Y^2)$ and $Var_Y(Y) = E_Y(Y^2) - E_Y(Y)^2$. Table 1 shows some values of the first and second moments and the variance of the l-logistic distribution. In addition, Figure 4 shows the graphs of the mean and the variance as functions of the shape parameter $b$, for some values of the location parameter $m$. For this purpose, the integral in (18) was evaluated by the Gauss quadrature through the 'statmod' R package (Smyth et al., 2015).

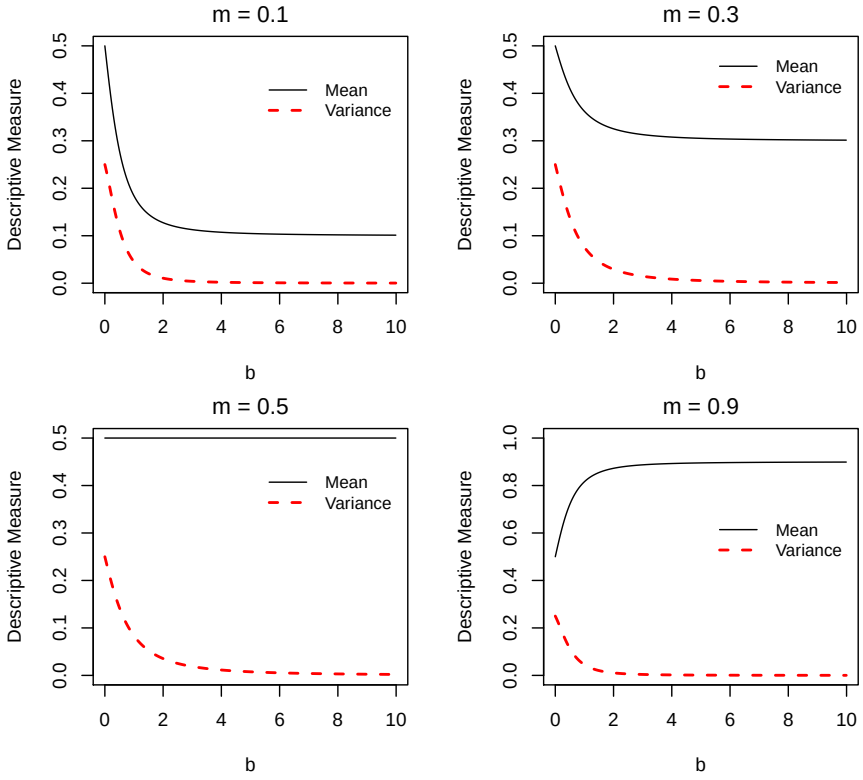


Figure 4: Descriptive measures of the l-logistic distributions for some values of the parameters

Table 1: $E_Y[Y]$, $E_Y[Y^2]$, and $Var_Y(X)$ of the l-logistic distribution for some values of $b$ and $m$.

| $m$ | 0.2 | 0.5 | 0.8 | 0.2 | 0.5 | 0.8 |
|---|---|---|---|---|---|---|
| $b$ | 1 | 1 | 1 | 3 | 3 | 3 |
| $E_Y[Y]$ | 0.283 | 0.5 | 0.717 | 0.216 | 0.5 | 0.784 |
| $E_Y[Y^2]$ | 0.145 | 0.333 | 0.579 | 0.056 | 0.269 | 0.625 |
| $Var_Y[Y]$ | 0.065 | 0.083 | 0.065 | 0.010 | 0.019 | 0.01 |

# 4 Bayesian inference

In this section, we describe the Bayesian approach for the estimation of the parameters of the l-logistic model. If we consider a random sample $\boldsymbol{Y} = (Y_1, ..., Y_n)$ from the distribution in (2), then the likelihood function is given by

$$L(\theta|\boldsymbol{y}) = \prod_{i=1}^{n} \frac{b(1-m)^b m^b y_i^{b-1}(1-y_i)^{b-1}}{\left[(1-m)^b y_i^b + m^b(1-y_i)^b\right]^2}. \tag{19}$$

## 4.1 Prior specification

To complete the Bayesian specification of the model, since parameters $m$ and $b$ have different behavior, we assume independence between them, and the following structure is then considered:

$$\pi(m, b) = \pi(m)\pi(b), \tag{20}$$

where $\pi(m)$ and $\pi(b)$ are the prior densities for $m$ and $b$, respectively.

Assuming that $m \sim uniform(0, 1)$, where $uniform(0, 1)$ represents the uniform distribution on the unit interval, and prior $\pi(b)$ for the parameter $b$, the joint posterior distribution for $(m, b)$ is given by

$$\pi(m, b|\boldsymbol{y}) \propto \prod_{i=1} \frac{b^n(1-m)^b m^b \pi(b)}{\left[(1-m)^b y_i^b + m^b(1-y_i)^b\right]^2}. \tag{21}$$

The prior $\pi(b)$ can be, for example, the pdf of the gamma distribution with parameters vector $(\epsilon, \epsilon)$, $\epsilon$ a small value. Some priors for parameter $b$ are compared and discussed in Section 7.

Since the posterior distribution is not available in a form closed, the Markov Chain Monte Carlo (MCMC) approach (Gelman et al., 2013, pp. 259 – 349) is used to estimate the model parameters. Initially, we consider the full conditional posterior distributions for the parameters $(m, b)$ given by

$$\pi(m|b, \boldsymbol{y}) = K_1^{-1} \frac{(1-m)^{nb} m^{nb}}{\prod_{i=1}^{n}\left[(1-m)^b y_i^b + m^b(1-y_i)^b\right]^2} \tag{22}$$

$$\pi(b|m, \boldsymbol{y}) = K_2^{-1} \prod_{i=1}^{n} \left( \frac{b(1-m)^b m^b y_i^{b-1}(1-y_i)^{b-1}}{\left[(1-m)^b y_i^b + m^b(1-y_i)^b\right]^2} \right) \pi(b) \tag{23}$$

$$\tag{24}$$

where $K_1$ and $K_2$ are normalizing constants.

Thus, a hybrid algorithm that combines Metropolis-Hastings and Gibbs sampling was implemented in R language (R Development Core Team, 2015) to obtain a sample from the posterior distribution of model parameters $(m, b)$. These codes are available upon request. The model can also be implemented easily in Bugs or other software such as Jags or R Stan (Lunn et al., 2000; Plummer, 2003; Guo, 2015).

### 4.2 Model comparison criteria

In order to compare different models, we made use of some model comparison criteria. Specifically, we consider the Expected Akaike information criteria (EAIC), the expected Bayesian information criteria (EBIC), the deviance information criteria (DIC), and the Watanabe-Akaike information criteria (Watanabe, 2010, WAIC). For a review of these criteria, one may refer to Gelman et al. (2013). The EAIC, EBIC and DIC can be estimated as

$$EAIC = \bar{D} + 2 \times p \tag{25}$$

$$EBIC = \bar{D} + 10 \times \log n \tag{26}$$

$$DIC = \bar{D} + p_D, \tag{27}$$

where

$$\bar{D} = G^{-1} \sum_{g=1}^{G} \left( -2 \sum_{i=1}^{n} \log \left( f(y_i | b^g, m^g) \right) \right),$$

$$\hat{D} = -2 \sum_{i=1}^{n} \log \left( f(y_i | \bar{m}, \bar{b} \right),$$

$p_D = \bar{D} - \hat{D}$, and $p$ represents the number of model parameters. Here, the notation $\bar{\theta}$ means the posterior mean of $\theta$ and $\theta^{(g)}$ represents the $g$th parameters of vector $\theta$ of a sequence of size $G$ generated from the posterior distribution by $MCMC$ method.

WAIC is a Bayesian approach for estimating the expected pointwise log predictive density (elppd) for a new dataset, and is given by

$$elppd = \sum_{i=1}^{n} E_y \left[ \log \left( P(\tilde{y}) \right) \right], \tag{28}$$

where the expectancy $E_y[.]$ and the pdf $P(.)$ are related to the predictive distribution (pd) induced by the posterior distribution $P(m, b|\boldsymbol{y})$ and $\tilde{y}$ is a future data. To estimate the elppd, we start by computing the pointwise log predictive density (lppd) as

$$\begin{aligned} lppd &= \sum_{i=1}^{n} \log P(y_i) \\ &= \sum_{i=1}^{n} \log \int f(y_i | m, b) dP(m, b|\boldsymbol{y}) \\ &\approx \sum_{i=1}^{n} \log \left( G^{-1} \sum_{g=1}^{G} f(y_i | m^g, b^g) \right) = \hat{lppd}. \end{aligned} \tag{29}$$

Thus, we can use a factor of correction for the effective number of parameters to adjust for overfitting. In the literature, there are two factors of correction that can be viewed as approximations to cross-validation (Gelman et al., 2013, pp. 169-174). The factor of correction used here makes use of the

variance in each term of the log predictive density and it is used here to obtain the WAIC. This factor is given by

$$p_{WAIC} = \sum_{i=1}^{n} Var_{(m,b)|\boldsymbol{y}} \left[\log\left(f(y_i|(m,b))\right]\right.$$
$$\approx \sum_{i=1}^{n} \left( \frac{1}{G-1} \sum_{g=1}^{G} \left[ \log\left(f(y_i|m^g, b^g)\right) - G^{-1} \sum_{g=1}^{G} \log\left(f(y_i|m^g, b^g)\right) \right]^2 \right) = \hat{p}_{WAIC_2}. \quad (30)$$

Finally, we can compute the WAIC as an approximation to the elppd as

$$WAIC = \hat{lppd} - \hat{p}_{WAIC}. \quad (31)$$

### 4.3 Posterior predictive checking

One point that will be of interest is the predictive distribution for a future observation. The density estimation of the posterior predictive distribution is obtained by integrating the unconditional predictive distribution $P(\tilde{y}|\mathbf{y}) = \int P(\tilde{y}|\boldsymbol{\theta}) dP(\boldsymbol{\theta}|\boldsymbol{y})$, where $P(\boldsymbol{\theta}|\mathbf{y})$ is the density of the posterior distribution of the parameters of the assumed model. In practice, we will be interested in simulating draws from the posterior predictive distribution of unknown observables $\tilde{y}$. Thus, as discussed by Gelman et al. (2013), the predictive distribution can be used to compare the predicted value under the assumed model $y^{rep}$ with the actual data $y$, where $y^{rep}$ can be thought as an estimate of $\tilde{y}$, or as an attempt to replicate the observed data based on the parameters. If the model fits well, then this predicted value should be similar to the observed data.

Using MCMC techniques, we could simulate values of the posterior predictive distribution by generating $y^{rep}$ from the distribution assumed by its model structure with the parameters generated from the posterior distribution. Let us consider $y_1, ..., y_n$ as observations generated independently from the l-logistic distribution. Then, we can generate $y^{rep}$ from the l-logistic distribution with appropriate parameters, that is,

$$y_i^{rep,(s)} \sim LL(m^s, b^s),$$

$s = 1, ..., S$, where $(m^1, b^1), ..., (m^S, b^S)$ is a sample generated from the posterior distribution. After generating $\left(y_1^{rep,(s)}, ..., y_n^{rep,(s)}\right)$, we order the sample to achieve $\left(y_{(1)}^{rep,(s)}, ..., y_{(n)}^{rep,(s)}\right)$, the ordered generated value. We can then compare the distribution of the ordered generated values $y_{(i)}^{rep,(s)}$ with the ordered observed values $y_{(i)}$. Finally, error bar can be constructed or posterior predictive values can be obtained by making use of the discrepancy measure, allowing for an evaluation of the model fit. Details about the predictive model checking are discussed by Gelman et al. (2013), Ntzoufras (2011), and Berkhof et al. (2000).

# 5 Simulation studies

This section is devoted two simulation studies, one that examines a prior sensitivity analysis and another investigates the recovery of the parameters of the model by the proposed estimation method. For this purpose, the Bayesian method is applied on simulated data sets from the l-logistic distribution, considering different scenarios. For the estimation of parameters, we generated $20,000$ samples from the posterior distribution given in (21), then the first $10,000$ samples were discarded and sequences of 10 observations were eliminated, and finally the resulting samples of size $1,000$ were used for inference.

## 5.1 Prior sensitivity analysis

Prior sensitivity analysis plays an important role in applied Bayesian analyses. This is especially true for Bayesian models used for new distribution, where the interpretability of the corresponding parameters becomes important. In this section, we consider a simulation study to evaluate the sensitivity of different choices of prior distributions for parameter $b$ since this is different from parameter $m$, which is clearer in its interpretation. Specifically, we assume prior independence between parameter $b$ and $m$, considering a unit uniform distribution for parameter $m$.

The models estimated with different prior distributions were compared by using WAIC, EAIC, EBIC and DIC. We considered five different prior distributions for parameter $b$, considering simulated data sets from the l-logistic distribution for some pairs of parameters $m$ and $b$. The values for the parameters $m$ and $b$ used are as follows: $b \in \{0.5, 1, 5\}$ and $m \in \{0.2, 0.5, 0.9\}$, leading to nine pairs of parameters. We simulated samples of size $n = 100$, $y_1, ...y_n$, from the l-logistic distribution considering these pairs of parameters, then nine distinct simulated datasets were considered for the analysis.

Based on the works of Gelman (2006) and Figueroa-Zúñiga et al. (2013), we consider for the parameter $b$ three non-informative and two informative prior distributions. The non-informative prior distributions are the gamma distribution with parameters vector $(0.001, 0.001)$ ($b \sim Gamma(0.001, 0.001)$), denoted by prior A, the uniform distribution with parameters vector $(1, 100)$ for $U$ ($U \sim Uniform(0, 100)$) with $b = U^2$, denoted by prior B, and the central Student t distribution with parameters vector $(10, 0, 2)$ ($L \sim St(0, 100)$) for $L$ with $\log(b) = L$, denoted by prior C. The prior B is chosen as it is less informative than the usual gamma with parameter vector $(\epsilon, \epsilon)$. For the informative prior distributions, we consider $b \sim Gamma(2.5, 1)$, denoted by prior D, and $b \sim Gamma(50, 1)$, denoted by prior E. Note that prior E provides incorrect information about parameter $b$, while prior D provides almost correct information. For all the cases, the prior distribution for parameter $m$ is the uniform distribution with parameters 0 and 1, ($m \sim Uniform(0, 1)$).

Table 2 shows the values of WAIC, EAIC, EBIC, and DIC for the fitted models. For all

the simulated datasets, we found that with prior E the estimated model achieves the worst fit among all models fitted for the same dataset. However, for the models using all other prior distributions, the values of WAIC, EAIC, EBIC, and DIC are quite close, showing no significant difference, giving evidence that the estimated models provide almost the same quality of fit for the analyzed samples. Thus, for these cases, the posterior distribution does not seen to be sensitive with respect to the specification of these prior distributions.

Additionally, we chose two non-informative priors A and C and the worst informative prior E. The prior A was chosen for this analysis was it is the simplest among the non-informative priors considered, while prior C is chosen as it is less informative than A. Table 3 reports the posterior mean and the 95% HPD interval (obtained by the package of Martin et al. (2011)). We observe that when the prior is E, the posterior mean is not far from the true value of $b$, showing that the data still dominates the prior information for these data. However, for some cases, the HPD interval does not contain the true value of $b$.

Table 2: Statistics for model comparison, prior distributions for parameter $b$ and the true value of the parameters of the l-logistic distribution used to simulate the data sets.

| Parameter | Prior | Criteria | | | |
|---|---|---|---|---|---|
| $(m, b)$ | | WAIC | EAIC | EBIC | DIC |
| | A, $b \sim Gamma(0.001, 0.001)$ | 41.761 | -77.482 | 10.622 | -82.685 |
| | B, $b = U^2$, $U \sim Uniform(0, 100)$ | 41.763 | -77.483 | 10.620 | -82.688 |
| $(0.2, 0.5)$ | C, $\log(b) = L$, $L \sim St(10, 0, 2)$ | 41.782 | -77.516 | 10.587 | -82.754 |
| | D, $b \sim Gamma(2.5, 1)$ | 41.793 | -77.464 | 10.639 | -82.650 |
| | E, $b \sim Gamma(50, 1)$ | 35.833 | -64.595 | 23.508 | -56.912 |
| | A, $b \sim Gamma(0.001, 0.001)$ | 31.945 | -58.058 | 30.045 | -63.480 |
| | B, $b = U^2$, $U \sim Uniform(0, 100)$ | 31.959 | -57.995 | 30.109 | -63.353 |
| $(0.2, 1)$ | C, $\log(b) = L$, $L \sim St(10, 2)$ | 31.971 | -58.066 | 30.038 | -63.496 |
| | D, $b \sim Gamma(2.5, 1)$ | 31.971 | -58.026 | 30.077 | -63.416 |
| | E, $b \sim Gamma(50, 1)$ | 26.129 | -45.275 | 42.829 | -37.914 |
| | A, $b \sim Gamma(0.001, 0.001)$ | 160.483 | -314.935 | -226.831 | -320.145 |
| | B, $b = U^2$, $U \sim Uniform(0, 100)$ | 160.482 | -315.031 | -226.927 | -320.336 |
| $(0.2, 5)$ | C, $\log(b) = L$, $L \sim St(10, 2)$ | 160.493 | -315.006 | -226.903 | -320.288 |
| | D, $b \sim Gamma(2.5, 1)$ | 160.369 | -314.847 | -226.743 | -319.969 |
| | E, $b \sim Gamma(50, 1)$ | 156.062 | -305.213 | -217.110 | -300.702 |
| | A, $b \sim G(0.001, 0.001)$ | 25.885 | -45.841 | 42.263 | -51.163 |
| | B, $b = U^2$, $U \sim Uniform(0, 100)$ | 25.904 | -45.847 | 42.256 | -51.176 |
| $(0.5, 0.5)$ | C, $\log(b) = L$, $L \sim St(10, 2)$ | 25.921 | -45.844 | 42.259 | -51.170 |
| | D, $b \sim Gamma(2.5, 1)$ | 25.915 | -45.862 | 42.242 | -51.205 |
| | E, $b \sim Gamma(50, 1)$ | 19.934 | -32.760 | 55.344 | -25.001 |

...

Table 2 – Continued

| Parameter | Prior | Criteria | | | |
|---|---|---|---|---|---|
| $(m, b)$ | | WAIC | EAIC | EBIC | DIC |
| | A, $b \sim Gamma(0.001, 0.001)$ | 1.634 | 2.633 | 90.736 | -2.734 |
| | B, $b = U^2$, $U \sim Uniform(0, 100)$ | 1.647 | 2.665 | 90.768 | -2.670 |
| $(0.5, 1)$ | C, $\log(b) = L$, $L \sim St(10, 2)$ | 1.664 | 2.623 | 90.726 | -2.754 |
| | D, $b \sim Gamma(2.5, 1)$ | 1.658 | 2.628 | 90.732 | -2.744 |
| | E, $b \sim Gamma(50, 1)$ | -4.103 | 15.133 | 103.236 | 22.266 |
| | A, $b \sim Gamma(0.001, 0.001)$ | 117.115 | -228.297 | -140.194 | -233.624 |
| | B, $b = U^2$, $U \sim Uniform(0, 100)$ | 117.121 | -228.277 | -140.174 | -233.583 |
| $(0.5, 5)$ | C, $\log(b) = L$, $L \sim St(10, 2)$ | 117.126 | -228.282 | -140.179 | -233.594 |
| | D, $b \sim Gamma(2.5, 1)$ | 116.994 | -228.085 | -139.982 | -233.199 |
| | E, $b \sim Gamma(50, 1)$ | 112.630 | -218.350 | -130.247 | -213.730 |
| | A, $b \sim Gamma(0.001, 0.001)$ | 91.031 | -176.006 | -87.903 | -181.160 |
| | B, $b = U^2$, $U \sim Uniform(0, 100)$ | 91.051 | -176.081 | -87.977 | -181.309 |
| $(0.9, 0.5)$ | C, $\log(b) = L$, $L \sim St(10, 2)$ | 91.069 | -176.135 | -88.031 | -181.418 |
| | D, $b \sim Gamma(2.5, 1)$ | 91.059 | -176.039 | -87.936 | -181.227 |
| | E, $b \sim Gamma(50, 1)$ | 84.975 | -162.958 | -74.854 | -155.064 |
| | A, $b \sim Gamma(0.001, 0.001)$ | 84.183 | -162.356 | -74.253 | -167.553 |
| | B, $b = U^2$, $U \sim Uniform(0, 100)$ | 84.195 | -162.383 | -74.280 | -167.608 |
| $(0.9, 1)$ | C, $\log(b) = L$, $L \sim St(10, 2)$ | 84.240 | -162.419 | -74.316 | -167.680 |
| | D, $b \sim Gamma(2.5, 1)$ | 84.202 | -162.330 | -74.227 | -167.501 |
| | E, $b \sim Gamma(50, 1)$ | 78.539 | -150.041 | -61.938 | -142.923 |
| | A, $b \sim Gamma(0.001, 0.001)$ 9 | 218.348 | -430.817 | -342.714 | -436.198 |
| | B, $b = U^2$, $U \sim Uniform(0, 100)$ | 218.364 | -430.837 | -342.734 | -436.239 |
| $(0.9, 5)$ | C, $\log(b) = L$, $L \sim St(10, 2)$ | 218.362 | -430.666 | -342.563 | -435.897 |
| | D, $b \sim Gamma(2.5, 1)$ | 218.246 | -430.697 | -342.594 | -435.958 |
| | E, $b \sim Gamma(50, 1)$ | 213.936 | -421.048 | -332.945 | -416.660 |

Table 3: Posterior mean with 95% HPD interval, prior distributions for parameter $b$ and true values of the parameters of l-logistic distribution used to simulate the data sets.

| Real | Prior A | | Prior C | | Prior E | |
|---|---|---|---|---|---|---|
| Value | Mean | HPD (95%) | Mean | HPD (95%) | Mean | HPD (95%) |
| m = 0.20 | 0.24 | (0.14, 0.35) | 0.24 | (0.14, 0.35) | 0.24 | (0.16, 0.33) |
| b = 0.50 | 0.58 | (0.49, 0.67) | 0.59 | (0.50, 0.68) | 0.78 | (0.68, 0.90) |
| m = 0.20 | 0.21 | (0.17, 0.26) | 0.21 | (0.17, 0.26) | 0.22 | (0.18, 0.26) |
| b = 1.00 | 1.15 | (0.98, 1.34) | 1.17 | (0.98, 1.35) | 1.55 | (1.34, 1.77) |
| m = 0.20 | 0.20 | (0.19, 0.21) | 0.20 | (0.19, 0.21) | 0.20 | (0.19, 0.21) |
| b = 5.00 | 5.77 | (4.89, 6.76) | 5.77 | (4.83, 6.77) | 7.51 | (6.49, 8.55) |
| m = 0.50 | 0.54 | (0.40, 0.68) | 0.54 | (0.39, 0.68) | 0.54 | (0.43, 0.66) |
| b = 0.50 | 0.58 | (0.48, 0.67) | 0.59 | (0.49, 0.68) | 0.78 | (0.67, 0.88) |
| m = 0.50 | 0.52 | (0.44, 0.58) | 0.52 | (0.45, 0.59) | 0.52 | (0.47, 0.58) |
| b = 1.00 | 1.16 | (0.97, 1.33) | 1.17 | (0.98, 1.36) | 1.55 | (1.33, 1.76) |
| m = 0.50 | 0.50 | (0.49, 0.52) | 0.50 | (0.49, 0.52) | 0.50 | (0.49, 0.52) |
| b = 5.00 | 5.78 | (4.79, 6.77) | 5.79 | (4.83, 6.68) | 7.52 | (6.51, 8.68) |
| m = 0.90 | 0.90 | (0.85, 0.95) | 0.90 | (0.86, 0.95) | 0.91 | (0.87, 0.94) |
| b = 0.50 | 0.58 | (0.49, 0.68) | 0.59 | (0.49, 0.69) | 0.78 | (0.68, 0.89) |
| m = 0.90 | 0.90 | (0.88, 0.93) | 0.90 | (0.88, 0.93) | 0.91 | (0.89, 0.93) |
| b = 1.00 | 1.15 | (0.98, 1.34) | 1.17 | (0.98, 1.36) | 1.55 | (1.34, 1.77) |
| m = 0.90 | 0.90 | (0.90, 0.91) | 0.90 | (0.90, 0.91) | 0.90 | (0.90, 0.91) |
| b = 5.00 | 5.78 | (4.96, 6.78) | 5.79 | (4.80, 6.72) | 7.51 | (6.60, 8.76) |

### 5.2 Parameter recovery

We carried out an evaluation of the point estimation, based on the $\sqrt{MSE}$ and bias, for simulated data sets from the l-logistic distribution. The mean and variance of an estimator $\hat{\theta}$ can be computed by Monte Carlo simulation, and it can be made done using the approximations

$$E_{\hat{\theta}}\left[\hat{\theta}\right] \approx G^{-1}\sum_{g=1}^{G}\hat{\theta}^{g}, \tag{32}$$

$$Var_{\hat{\theta}}[\hat{\theta}] \approx G^{-1}\sum_{g=1}^{G}\left(\hat{\theta}^{g} - E_{\hat{\theta}}\left[\hat{\theta}\right]\right)^{2}, \tag{33}$$

where $\hat{\theta}^{1}, ..., \hat{\theta}^{G}$ are obtained from $G$ different simulated samples. Thus, the MSE of $\hat{\theta}$ is the function of $\theta$ defined by

$$E_{\hat{\theta}}\left[(\hat{\theta}-\theta)^{2}\right] = Var_{\hat{\theta}}[\hat{\theta}] + \left(E_{\hat{\theta}}[\hat{\theta}] - \theta\right)^{2} \approx G^{-1}\sum_{g=1}^{G}\left(\hat{\theta}^{g} - \theta\right)^{2}, \tag{34}$$

where $E[\hat{\theta}] - \theta$ is the bias of $\hat{\theta}$. Of course, a good estimator should produce mean, standard deviation (square root of variance), and bias close to zero.

For the analysis presented here, we generated samples of size $n = 50$, $n = 100$, and $n = 500$, the values of the parameters were set as $m \in \{0.2, 0.5, 0.9\}$ and $b \in \{0.3, 0.5, 1, 2, 4\}$. For these data sets, we estimated the parameters of the l-logistic model by using the Bayesian method under the assumption of independent priors for $m$ and $b$ with non-informative prior distributions uniform and gamma for $m$ and $b$, respectively, that is,

$$m \sim Uniform(0, 1) \quad \text{and} \quad b \sim Gamma(0.001, 0.001). \tag{35}$$

Gamma distribution is commonly used in the literature for shape parameters (specifically for precision parameters) and, based on the analysis presented in Subsection 5.1, the posterior distribution is not sensitive with respect to the specification of this prior distribution.

Bayes estimator used here is that the mean of the posterior distribution (estimator with respect to squared error loss function).

Table 4 shows the values of the $\sqrt{MSE}$ and bias from the simulated data sets. The estimates for these quantities were obtained by $G = 1,000$ Monte Carlo replications. We can see that the $\sqrt{MSE}$ and bias are close to zero even when the the sample size is $n = 50$. For these samples, the estimator performs very well as both $\sqrt{MSE}$ and bias are very small, for all the analyzed samples.

Table 4: Bias and square root of mean square error ($\sqrt{MSE}$) of the Bayesian estimator of the parameters $m$ and $b$.

| Real value | | | m=0.2 | | m=0.5 | | m=0.9 | |
|---|---|---|---|---|---|---|---|---|
| | n | parameter | Bias | $\sqrt{MSE}$ | Bias | $\sqrt{MSE}$ | Bias | $\sqrt{MSE}$ |
| b=0.3 | 50 | m | -1,5e-02 | 1,5e-02 | -1,7e-01 | 1,7e-01 | -2,2e-01 | 2,2e-01 |
| | | b | 1,9e-02 | 1,9e-02 | 1,9e-02 | 1,9e-02 | 1,8e-02 | 1,8e-02 |
| | 100 | m | -5,3e-02 | 5,3e-02 | -2,4e-01 | 2,4e-01 | -2,8e-01 | 2,8e-01 |
| | | b | -7,0e-03 | 7,0e-03 | -7,1e-03 | 7,1e-03 | -5,6e-03 | 5,6e-03 |
| | 500 | m | -1,8e-02 | 1,8e-02 | -9,8e-02 | 9,8e-02 | -1,0e-01 | 1,0e-01 |
| | | b | -1,2e-02 | 1,2e-02 | -1,1e-02 | 1,1e-02 | -1,1e-02 | 1,1e-02 |
| b=0.5 | 50 | m | -2,4e-02 | 2,4e-02 | -1,3e-01 | 1,3e-01 | -1,3e-01 | 1,3e-01 |
| | | b | 2,8e-02 | 2,8e-02 | 3,2e-02 | 3,2e-02 | 3,2e-02 | 3,2e-02 |
| | 100 | m | -4,3e-02 | 4,3e-02 | -1,7e-01 | 1,7e-01 | -1,6e-01 | 1,6e-01 |
| | | b | -8,8e-03 | 8,8e-03 | -9,5e-03 | 9,5e-03 | -8,1e-03 | 8,1e-03 |
| | 500 | m | -1,6e-02 | 1,6e-02 | -6,1e-02 | 6,1e-02 | -5,5e-02 | 5,5e-02 |
| | | b | -1,9e-02 | 1,9e-02 | -1,8e-02 | 1,8e-02 | -1,9e-02 | 1,9e-02 |
| b=1 | 50 | m | -1,9e-02 | 1,9e-02 | -7,1e-02 | 7,1e-02 | -6,0e-02 | 6,0e-02 |
| | | b | 5,7e-02 | 5,7e-02 | 5,6e-02 | 5,6e-02 | 5,2e-02 | 5,2e-02 |
| | 100 | m | - -2,7e-02 | 2,7e-02 | -9,3e-02 | 9,3e-02 | -7,3e-02 | 7,3e-02 |
| | | b | -1,7e-02 | 1,7e-02 | -1,6e-02 | 1,6e-02 | -2,2e-02 | 2,2e-02 |
| | 500 | m | -9,9e-03 | 9,9e-03 | -3,3e-02 | 3,3e-02 | -2,4e-02 | 2,4e-02 |
| | | b | -3,4e-02 | 3,4e-02 | -3,5e-02 | 3,5e-02 | -3,5e-02 | 3,5e-02 |
| b=2 | 50 | m | -1,2e-02 | 1,2e-02 | -3,7e-02 | 3,7e-02 | -2,7e-02 | 2,7e-02 |
| | | b | 1,2e-01 | 1,2e-01 | 1,2e-01 | 1,2e-01 | 1,3e-01 | 1,3e-01 |
| | 100 | m | -1,5e-02 | 1,5e-02 | -4,9e-02 | 4,9e-02 | -3,3e-02 | 3,3e-02 |
| | | b | -3,4e-02 | 3,4e-02 | -4,3e-02 | 4,3e-02 | -3,4e-02 | 3,4e-02 |
| | 500 | m | -6,5e-03 | 6,5e-03 | -1,7e-02 | 1,7e-02 | -1,1e-02 | 1,1e-02 |
| | | b | -7,2e-02 | 7,2e-02 | -6,9e-02 | 6,9e-02 | -7,2e-02 | 7,2e-02 |
| b=4 | 50 | m | -7,2e-03 | 7,2e-03 | -1,8e-02 | 1,8e-02 | -1,2e-02 | 1,2e-02 |
| | | b | 2,6e-01 | 2,6e-01 | 2,4e-01 | 2,4e-01 | 2,6e-01 | 2,6e-01 |
| | 100 | m | -8,7e-03 | 8,7e-03 | -2,4e-02 | 2,4e-02 | -1,6e-02 | 1,6e-02 |
| | | b | -7,0e-02 | 7,0e-02 | -6,9e-02 | 6,9e-02 | -6,2e-02 | 6,2e-02 |
| | 500 | m | -2,7e-03 | 2,7e-03 | -7,6e-03 | 7,6e-03 | -4,9e-03 | 4,9e-03 |
| | | b | -1,6e-01 | 1,6e-01 | -1,4e-01 | 1,4e-01 | -1,4e-01 | 1,4e-01 |

# 6 Applications to real data

In order to illustrate the proposed methodology, we consider data from municipalities of the state of Alagoas in Brazil, collected in 2010. The state of Alagoas is located in the eastern part of the Northeast Region of Brazil and is made up of 102 municipalities. This state is one of the poorest states of Brazil and its HDI (Human Development Index) is the country's worst, based on

information available in Fundação (2010). Specifically, we are interested in modeling the proportion of children vulnerable to poverty (PPOBC) and the Gini index in the municipalities of Alagoas; see also Fundação (2010). In the case of the Gini index, we model this data as a function of the percentage people employed in the manufacturing industry (EMP) in the municipalities of the Alagoas.

### 6.1 Aplication to PPOBC data

In this subsection, we consider the PPOBC data set, which contains the proportion of children (0-14 years olds) vulnerable to poverty in each municipality of Alagoas. Here, a child is considered vulnerable to poverty if the per capita household income is equal to or less than BRL 255, in 2010. The PPOBC data set comprises 102 observations and is modeled here using the l-logistic model and the beta model that is frequently used to model data when a distribution over some finite interval is needed; see Gupta and Nadarajah (2004). Here, we use the re-parameterized beta distribution discussed by Ferrari and Cribari-Neto (2004) in the context of regression analysis.

The Bayesian methodology was used to estimate the parameters of both models. For the l-logistic distribution with parameters $m$ and $b$, we considered the gamma prior distribution for parameter $b$ ($b \sim Gamma(0.001, 0.001)$) and the uniform prior distribution in the unit interval for parameter $m$ ($m \sim Uniform(0, 1)$). For the beta distribution with parameters $0 < \mu < 1$ and $\phi > 0$, we considered the same prior distributions considered in the l-logistic case, that is, $\phi \sim Gamma(0.001, 0.001)$ and $\mu \sim Uniform(0, 1)$. These priors distributions were chosen based on the discussion in Section 2. In the case of beta distribution, we used the non-informative proper prior distribution commonly used in the literature for the parameters of precision.

Table 5: Estimates and 95% HPD intervals for the parameters of the l-logistic and beta models and the Bayesian information criteria values for these models

| Model | Parameter | | Criteria | | | |
|---|---|---|---|---|---|---|
| | | | WAIC | EAIC | EBIC | DIC |
| L-logistic | $m$ | 0.86(0.85, 0.87) | 155.1322 | -304.2996 | -299.0496 | -306.3422 |
| | $b$ | 4.04(3.42, 4.72) | | | | |
| Beta | $\mu$ | 0.85(0.84, 0.86) | 150.8993 | -295.3312 | -290.0813 | -297.3437 |
| | $\phi$ | 37.81(27.55, 47.83) | | | | |

The final result of the estimation is presented in Table 5. This table also shows the Bayesian information criteria of model comparison in order to evaluate the ability of l-logistic and beta models to fit the data. According to this table, it is clear that the l-logistic model is better for modeling the PPOBC data than the beta model. In addition, Figure 5 shows two graphs with the mean values and error bars with 95% credibility intervals plotted against the corresponding observed value of the data.

The errors bars were constructed from 1000 samples (ordered, and of size 102) generated from the l-logistic and beta distributions, respectively for each graph, with the estimated parameters. In the case of the l-logistic model, the bars crossed by the diagonal line $y = x$ indicate that the model is quite suitable for the data. On the other hand, in the case of the beta model, we observe high deviations between the predicted and observed data, mainly in the tail of the distribution. In this case, an observation is flagged as an outlier, since the corresponding posterior interval does not contain these values that are situated between 0 and 0.6. Thus, Figure 5 provides evidence that the beta model is unsuitable for these data. Finally, the estimated and the observed histogram of the PPOBC data are presented in Figure 6.1, which confirms that the l-logistic model provides a better fit for these data than the beta model.

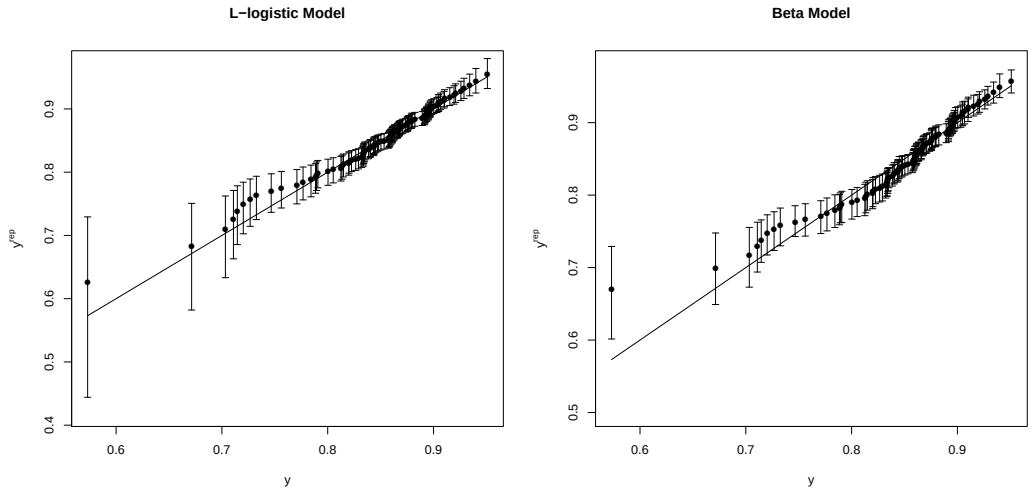


Figure 5: Posterior predictive error bars with 95% confidence intervals of the generated values $y_{(i)}^{rep}$ versus ordered observed data $y_{(i)}$ for the PPOBC data, using l-logistic and beta models.

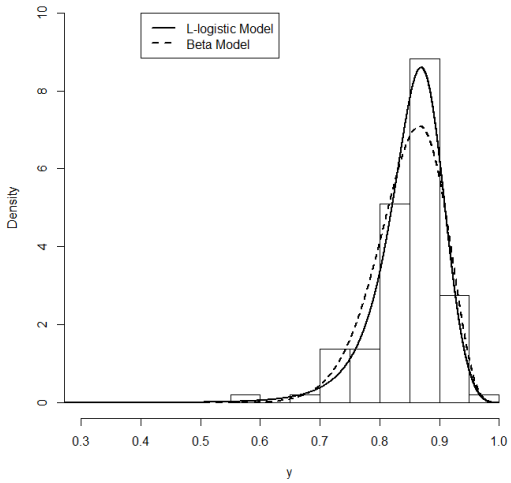


Figure 6: Estimated density of PPOBC data.

## 6.2 Regression Analysis with l-logistic model

Regression analysis estimates the potential differential effect of a covariate on mean or quantiles in the conditional distribution (Hao and Naiman, 2007). Here, we are interested in studying the conditional (or regression) median as a function of the covariates, when the response variable takes values in a bounded interval. In the analysis with the l-logistic distribution, we assume that conditional on the explanatory variables (covariates), the random variable $Y_i$, $i = 1, ..., n$, are mutually independent with l-logistic distribution, $Y_i \sim LL(m_i, b_i)$. Thus, given $\boldsymbol{x}_{1i}^T$ and $\boldsymbol{x}_{2i}^T$ ($q$ and $d$-dimensional vectors, respectively, containing the explanatory variables both with 1 as in the first component), the likelihood of the observed sample $\boldsymbol{y} = (y_1, .., y_n)$ can be written as

$$L(\boldsymbol{\beta}, \boldsymbol{\delta}|\boldsymbol{y}, \boldsymbol{X} = \prod_{i=1}^{n} \frac{b_i(1-m_i)^{b_i} m_i^{b_i} y_i^{b_i-1}(1-y_i)^{b_i-1}}{\left[(1-m_i)^{b_i} y_i^b + m^{b_i}(1-y_i)^{b_i}\right]^2}, \tag{36}$$

where $\boldsymbol{X}$ is the matrix containing all the explanatory variables, and

$$logit(m_i) = \boldsymbol{x}_{1i}^T \boldsymbol{\beta} \quad \text{and} \quad log(b_i^2) = \boldsymbol{x}_{2i}^T \boldsymbol{\delta}. \tag{37}$$

In (37), $\boldsymbol{\beta} = (\beta_0, ..., \beta_{q-1})$ and $\boldsymbol{\delta} = (\delta_0, ..., \delta_{d-1})$ represent, respectively, the $q$ and d-dimensional vectors of unknown regression parameters, $logit(.)$ is the logit function, and $log(.)$ is the natural logarithm function.

In addition, we adopt the following proper prior distributions with large range as is frequently considered in the literature:

$$\begin{aligned} \beta_j &\sim Normal(0, 100), \text{ for } j = 0, .., q-1, \\ \delta_l &\sim Normal(0, 100), \text{ for } l = 0, .., d-1. \end{aligned} \tag{38}$$

Thus, samples of the joint posterior distribution of $\boldsymbol{\beta}$ and $\boldsymbol{\delta}$ can be obtained by considering the MCMC method to simulate from the posterior distribution, with pdf given by

$$\pi(\beta, \delta|\boldsymbol{y}) \propto L(\boldsymbol{\beta}, \boldsymbol{\delta}|\boldsymbol{y}, \boldsymbol{x})\pi(\boldsymbol{\beta})\pi(\boldsymbol{\delta}). \tag{39}$$

### 6.2.1 Modeling Gini index in Brazil

In this subsection, we consider the Gini index data of the municipalities of the state of Alagoas in Brazil. The Gini index is used to measure how evenly income is distributed throughout a country. For more details, one may refer to Lambert and Aronson (1993). In Brazil, the Gini index of the municipalities is elaborated by IBGE ( Portuguese, Instituto Brasileiro de Geografia e Estatística); see IBGE (2010). Since Alagoas comprises 102 municipalities, our sample is composed of 102 observations ($n = 102$). The scatterplot of the Gini index data versus EMP data and the histograms of these data are presented in Figure 8.

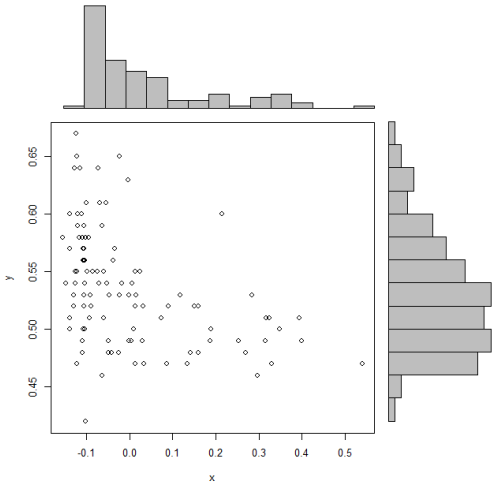


Figure 7: Scatterplot and histograms of the real data.

Considering the l-logistic model, we still consider four models with the following specification, for $i = 1, ..., n$:

$$Y_i \sim LL(m_i, b_i) \Longrightarrow \begin{cases} M_0: & logit(m_i) = \beta_0 \text{ and } \log(b_i^2) = \boldsymbol{\delta} \\ M_1: & logit(m_i) = \beta_0 \text{ and } \log(b_i^2) = \boldsymbol{x}_{1i}^T\boldsymbol{\delta} \\ M_2: & logit(m_i) = \boldsymbol{x}_{1i}^T\boldsymbol{\beta} \text{ and } log(b_i^2) = \delta_0 \\ M_3: & logit(m_i) = \boldsymbol{x}_{1i}^T\boldsymbol{\beta} \text{ and } \log(b_i^2) = \boldsymbol{x}_{2i}^T\boldsymbol{\delta}. \end{cases}$$

In addition, we consider the models $M_0, M_1, M_2$ and $M_3$ for the case where the response follows the re-parameterized beta distribution, that is, for $i = 1, .., n,$:

$$Y_i \sim Beta(\mu_i, \phi_i) \Longrightarrow \begin{cases} M_0: & logit(\mu_i) = \beta_0 \text{ and } \log(\phi_i^2) = \boldsymbol{\delta} \\ M_1: & logit(\mu_i) = \beta_0 \text{ and } \log(\phi_i^2) = \boldsymbol{x}_{1i}^T\boldsymbol{\delta} \\ M_2: & logit(\mu_i) = \boldsymbol{x}_{1i}^T\boldsymbol{\beta} \text{ and } log(b_i^2) = \delta_0 \\ M_3: & logit(\mu_i) = \boldsymbol{x}_{1i}^T\boldsymbol{\beta} \text{ and } \log(\phi_i^2) = \boldsymbol{x}_{2i}^T\boldsymbol{\delta}, \end{cases}$$

where $\phi_i$ is a parameter of precision.

For both, beta and l-logistic models, the Bayesian approach is considered for the inference process with prior distribution for the unknown regression parameters as given in (6.2). All the algorithms were prepared in R language and we report the results corresponding to $10,000$ iterations following a burn-in period also of $10,000$ iterations. In order to eliminate dependence, we eliminated a sequence of 10 observations every 11 simulations in the sample of size $10,000$, resulting in a final sample of $1,000$ elements. Finally, the convergence of MCMC chain was assessed by using the separated partial means test of Geweke (1992), which provided evidence for the chains to have converged.

Table 6 shows the estimated parameters for the l-logistic and beta cases for all the models. This table also shows the 95% HPD intervals for all the coefficients, where we can see that the amplitude of the intervals are considered small. Additionally, since the HPD intervals for $\beta_1$ and $\delta_1$ do not contain zero, we can conclude that the EPM covariate is significant to explain the Gini index in both the l-logistic and beta models.

Table 6: Parameter estimates and 95% HPD intervals for the l-logistic and beta models.

| Model | | Coefficient | | | |
|---|---|---|---|---|---|
| | | $\beta_0$ | $\beta_1$ | $\delta_0$ | $\delta_1$ |
| L-logistic | $M_3$ | 0.13(0.090, 0.164) | -0.48(-0.620,-0.305) | 4.59(4.243, 4.920) | 3.71(1.170, 6.038) |
| | $M_2$ | 0.13(0.098, 0.167) | -0.56(-0.781, -0.347) | 4.50(4.144, 4.801) | - |
| | $M_1$ | 0.11(0.050, 0.161) | - | 4.28(3.985, 4.630) | 1.42(-1.225, 4.316) |
| | $M_0$ | 0.13(0.078, 0.162) | - | 4.27(3.956, 4.604) | - |
| Beta | $M_3$ | 0.13(0.101, 0.175) | -0.49(-0.652, -0.339) | 9.66(9.109, 10.216) | 6.88(3.068, 10.681) |
| | $M_2$ | 0.14(0.102, 0.176) | -0.56(-0.792, -0.330) | 9.47(8.902, 9.979) | - |
| | $M_1$ | 0.12(0.065, 0.167) | - | 9.14(8.576, 9.665) | 2.59(-1.897, 6.550) |
| | $M_0$ | 0.14(0.04, 0.176) | - | 9.10(8.571, 9.670) | - |

In addition, in both cases, the models $M_0, M_1, M_2$ and $M_3$ were compared by the use of EAIC, EBIC, DIC and WAIC criteria described earlier. Results are presented in Table 7, in which we can see that the model $M_3$ is the best for both distributions. Note that values of coefficients $\beta_0$ and $\beta_1$ are similar for both models and that values of $\delta_0$ and $\delta_1$ are different. Thus, there is no significant difference between the models for both distributions. Therefore, based on the obtained results, both models explain equally the phenomenon considered here. So, we can choose either of the two distributions in this case.

Table 7: Model comparison criteria for model comparison.

| | L-logistic model | | | | Beta model | | | |
|---|---|---|---|---|---|---|---|---|
| Sub-model | WAIC | EAIC | EBIC | DIC | WAIC | EAIC | EBIC | DIC |
| $M_3$ | 175.22 | -339.09 | -328.59 | -343.29 | 175.07 | -338.24 | -327.74 | -342.26 |
| $M_2$ | 170.22 | -331.99 | -324.12 | -335.15 | 169.73 | -330.74 | -322.87 | -333.67 |
| $M_1$ | 159.62 | -310.10 | -302.22 | -313.25 | 160.61 | -312.29 | -304.41 | -315.33 |
| $M_0$ | 158.82 | -311.87 | -306.62 | -313.96 | 159.89 | -313.96 | -308.71 | -316.06 |

Finally, for the l-logistic and beta models $M_3$ , we performed a predictive check based on the ordered data $y_{(i)}$ and generated values $y_{(i)}^{rep}$ of the posterior predictive distribution, as mentioned in Subsection 4.3. Error bars of the $y_{(i)}^{rep}$ against the correspondent observed data are presented in Figure 8, in which we can see that the diagonal line $y = x$ crosses the error bars for all observations for both

distributions. Thus, both models seem to fit the data well in this case. Figure 8 also presents the realized residual versus adjusted values ( $\hat{y}_i$, $i = 1, .., n$ ). The realized residual ($r_i = y_i - \hat{y}_i$) is based on a posterior draw of the parameters rather than point estimations, for more details, see for example, (Gelman et al., 2013). Based on Figure 8, we can see that the spread of the residuals is quite very similar for the beta and the l-logistic models.

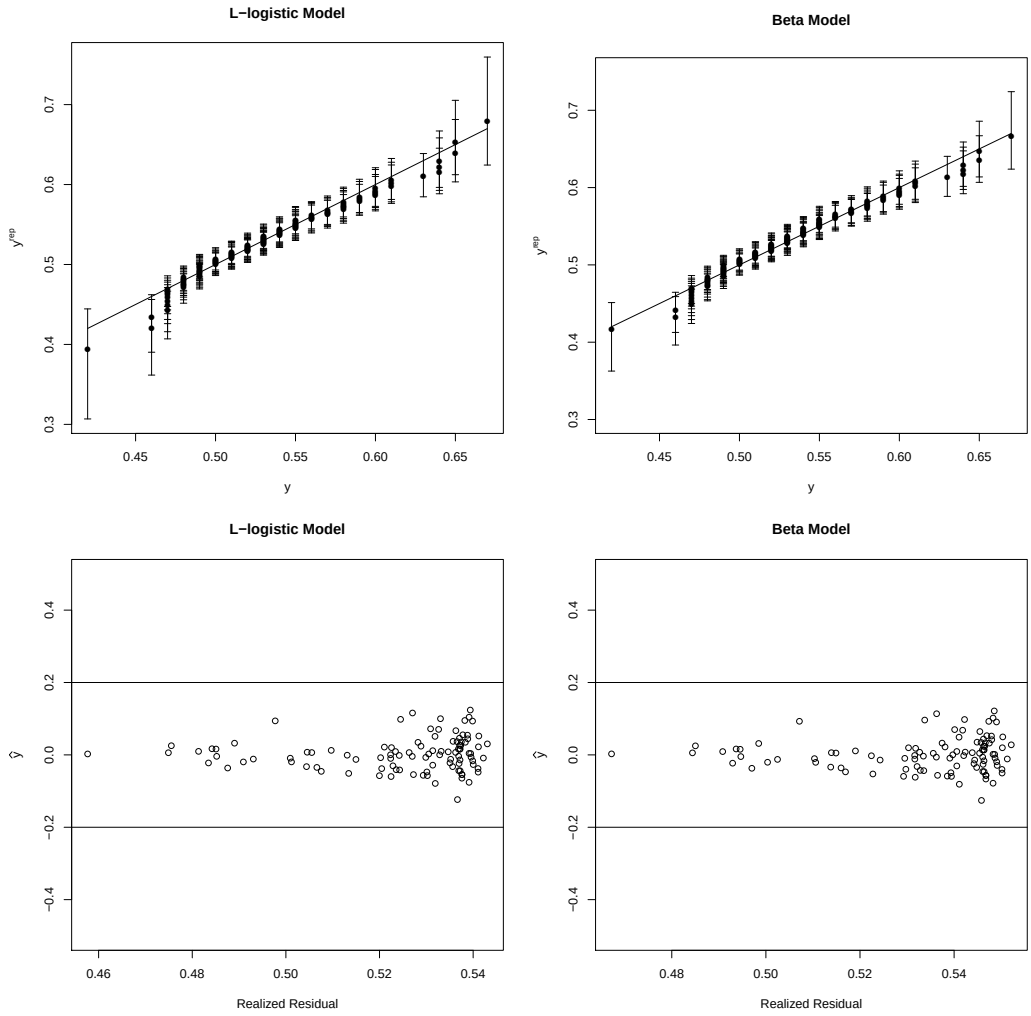


Figure 8: Posterior predictive error bars of generated values $y_{(i)}^{rep}$ versus ordered observed data $y_{(i)}$ and realized residual versus adjusted values for the model l-logistic $M_3$ and beta model $M_3$ for the Gini index data.

# 7 Final remarks

The l-logistic distribution is a bounded continuous distribution that possesses some properties, as discussed in Section 3. In the Bayesian context, a non-informative prior distribution can be adopted for the location parameter $m$ since it lies in the unit interval, enabling the use of unit uniform distribution as a non-informative prior distribution. Two simulation studies are presented in Section 6 for evaluating the posterior distribution with respect to the specification of the prior distribution

for the shape parameter $b$ and to evaluate the performance of the Bayesian estimator chosen. In the first study, for the studied cases, we observe that the non-informative prior distributions provide correct information about parameter $b$, based on the results of WAIC, EAIC, EBIC, and DIC. Thus, the posterior distribution is not sensitive with respect to the specification of these prior distributions. Some characteristics of the posterior distribution are also calculated where two non-informative and one informative prior distributions for the parameter $b$ are considered. In this study, we observe that the prior information is dominated by the sample information. In the second study, we evaluate the estimates of the parameters of the l-logistic distribution obtained by using Bayesian method upon considering the prior gamma distribution with parameters vector $(0.001, 0.001)$. We observe that the $\sqrt{MSE}$ and bias lie always close to zero even when the sample size is small. Hence, for the samples analyzed, the estimator seems to provide reasonable estimates.

In addition, we introduce the l-logistic distribution in the context of conditional median regression models. Conditional median regression is a special case of quantile regression in which the conditional 0.5th quantile is modeled as a function of covariates. An advantage of this approach is the possibility of modeling other quantiles in order to describe a non-central position of a distribution. So, one may choose a position a specific for his needs

inquiries. For example, poverty studies concern the low-income population and tax-policy studies concern the rich people. Conditional quantile models offer the flexibility to focus on these population segments, whereas conditional mean models do not. Thus, multiple quantiles can be modeled. However, since quantile regression curves are estimated individually, the quantile curves can cross, leading to an invalid distribution for the response. Thus, this problem, referred to as crossing in the literature, needs to be studied carefully. Some authors have proposed methods to deal with this problem, see, for example, Cai and Jiang (2015).

Finally, for the future, we aim to develop techniques for mixed quantile regression for the l-logistic distribution. Moreover, we intend to explore mixtures of L-logistic distributions in a Bayesian framework as well as a multivariate version of this distribution.

# A Appendix

If $Z = b\log\left(\frac{Y(1-m)}{m(1-Y)}\right)$ and $Y \sim LL(m,b)$, then the pdf of $Z$ can be obtained by using the transformation technique as

$$\begin{aligned} f_Z(z) &= f_Y\left(\left(1+e^{-z/b}\left(\tfrac{1-m}{m}\right)\right)^{-1}\right)\left|\frac{\partial\left(1+e^{-z/b}\left(\frac{1-m}{m}\right)\right)^{-1}}{\partial z}\right| \\ &= \frac{b(1-m)^b m^b\left(\frac{1}{1+(\frac{1-m}{m})e^{-z/b}}\right)^{b-1}\left(\frac{(\frac{1-m}{m})e^{-z/b}}{1+(\frac{1-m}{m})e^{-z/b}}\right)^{b-1}}{\left[(1-m)^b\left(\frac{1}{1+(\frac{1-m}{m})e^{-z/b}}\right)^b+m^b\left(\frac{(\frac{1-m}{m})e^{-z/b}}{1+(\frac{1-m}{m})e^{-z/b}}\right)^b\right]^2}\left(\frac{(\frac{1-m}{m})e^{-z/b}}{b\left[1+(\frac{1-m}{m})e^{-z/b}\right]^2}\right) \\ &= \frac{(1-m)^b m^b\left((\frac{1-m}{m})e^{-z/b}\right)^{b-1}}{\left[(1-m)^b+m^b(\frac{1-m}{m})^b e^{-z}\right]^2}\left((\tfrac{1-m}{m})e^{-z/b}\right) \\ &= \frac{(1-m)^b m^b(1-m)^{b-1}e^{-z(b-1)/b}}{m^{b-1}\left[(1-m)^b+(1-m)^b e^{-z}\right]^2}(\tfrac{1-m}{m})e^{-z/b} = \frac{e^{-z}}{[1+e^{-z}]^2} \end{aligned} \tag{40}$$

$$\Rightarrow f_Z(z) = \frac{e^z}{[1+e^z]^2} I_{\mathbb{R}}(z), \tag{41}$$

that is, $Z$ has the standard logistic distribution.

Using the transformation technique, as in (40), we can achieve the pdf and cdf of the distribution of $Z' = \delta + b\log\left(\frac{Y}{1-Y}\right)$ with $Y \sim LL(\delta,b)$ as the pdf and cdf of the standard logistic distribution.

Since $m = \frac{e^{-\frac{\delta}{b}}}{1+e^{-\frac{\delta}{b}}}$ with $\delta > 0$, and $Y \sim LL(m,b)$, we can replace the parameter $m = \frac{e^{-\frac{\delta}{b}}}{1+e^{-\frac{\delta}{b}}}$ in (1) and (2) to obtain

$$\begin{aligned} f(y|\delta,b) &= \frac{b\left(1-\frac{e^{-\frac{\delta}{b}}}{1+e^{-\frac{\delta}{b}}}\right)^b\left(\frac{e^{-\frac{\delta}{b}}}{1+e^{-\frac{\delta}{b}}}\right)^b y^{b-1}(1-y)^{b-1}}{\left[\left(1-\frac{e^{-\frac{\delta}{b}}}{1+e^{-\frac{\delta}{b}}}\right)^b y^b+\left(\frac{e^{-\frac{\delta}{b}}}{1+e^{-\frac{\delta}{b}}}\right)^b(1-y)^b\right]^2} \\ &= \frac{b\left(\frac{1}{1+e^{-\frac{\delta}{b}}}\right)^b\left(\frac{e^{-\frac{\delta}{b}}}{1+e^{-\frac{\delta}{b}}}\right)^b y^{b-1}(1-y)^{b-1}}{\left[\left(\frac{1}{1+e^{-\frac{\delta}{b}}}\right)^b y^b+\left(\frac{e^{-\frac{\delta}{b}}}{1+e^{-\frac{\delta}{b}}}\right)^b(1-y)^b\right]^2} \\ &= \frac{be^{-\delta}y^{b-1}(1-y)^{b-1}}{\left[y^b+e^{-\delta}(1-y)^b\right]^2} = \frac{be^{\delta}y^{b-1}(1-y)^{b-1}}{\left[y^b e^{\delta}+(1-y)^b\right]^2} \end{aligned} \tag{42}$$

and

$$\begin{aligned} F_Y(y|\delta,b) &= \left(1+\left(\frac{\frac{e^{-\frac{\delta}{b}}}{1+e^{-\frac{\delta}{b}}}(1-y)}{y(\frac{1}{1+e^{-\frac{\delta}{b}}})}\right)^b\right)^{-1} \\ &= \left(1+\left(\frac{e^{-\frac{\delta}{b}}(1-y)}{y}\right)^b\right)^{-1} \\ &= \left(1+e^{-\delta}\left(\frac{1-y}{y}\right)^b\right)^{-1} \end{aligned}$$

which are the pdf and cdf of the l-logistic distribution with parameters $\delta$ and $b$.

Assuming $\mu = \frac{1}{1+\left(\frac{m}{1-m}\right)^b}$ with $Y \sim LL(m,b)$, as in (42), by replacing parameter $m = \frac{1}{1+\left(\frac{\mu}{1-\mu}\right)^{1/b}}$ in (1) and (2), we obtain the corresponding pdf and cdf as

$$f(y|\mu,b) = \frac{b\mu(1-\mu)y^{b-1}(1-y)^{b-1}}{\left[y^b\mu + (1-y)^b(1-\mu)\right]^2}, \ 0<y<1, \tag{43}$$

and

$$F_Y(y|\mu,b) = \left(1+\left(\frac{1-\mu}{\mu}\right)\left(\frac{1-y}{y}\right)^b\right)^{-1}, \ 0<y<1, \tag{44}$$

with $\mu \in (0,1)$ and $b>0$.

If $Y = X(d-c)+c$ with $c,d \in \mathbb{R}$ and $X \sim LL(m,b)$, then $Y \in (c,d)$. In this case, we can obtain the pdf of $Y$ by transformation technique as

$$f(y|m,b,c,d) = \frac{(d-c)b(1-m)^b m^b (y-c)^{b-1}(d-y)^{b-1}}{\left[(1-m)^b(y-c)^b + m^b(d-y)^b\right]^2}. \tag{45}$$

Thus, $Y$ has the l-logistic distribution with support on $(c,d)$, $c,d \in \mathbb{R}$.

Assuming that $Y \sim LL(m,b)$, then the t-th moment of the random variable $Y$ about zero can be obtained from (1) and (41) as

$$\begin{aligned} E[Y^t] &= \int_0^1 y^t \frac{b(1-m)^b m^b y^{b-1}(1-y)^{b-1}}{\left[(1-m)^b y^b + m^b(1-y)^b\right]^2} dy \\ &= \int_{-\infty}^{\infty} \left(\frac{1}{1+\left(\frac{1-m}{m}\right)e^{-\frac{z}{b}}}\right)^t \frac{e^z}{[1+e^z]^2} dz. \end{aligned}$$

Let $v = \frac{e^z}{1+e^z} \Rightarrow z = \log\left(\frac{v}{1-v}\right)$ and $dv = \frac{e^z}{(1+e^z)^2}dz$. Then, $E[Y^t] = \int_0^1 \left[1+\left(\frac{1-v}{v}\right)^{1/b}\left(\frac{1-m}{m}\right)\right]^{-t} dv$.

In order to find the mode of the l-logistic distribution, we can obtain the derivative of the $f(y|m,b)$, given in (1), relative to $y$ as

$$\frac{\partial f}{\partial y} = \frac{b((1-m)m)^b((1-y)y)^{-2+b}\left\{m^b(1-y)^b(-1+b+2y) - (1-m)^b(1+b-2y)y^b\right\}}{\left(m^b(1-y)^b + (1-m)^b y^b\right)^3}. \tag{46}$$

Thus, $\frac{\partial f}{\partial y}(y_0) = 0 \Leftrightarrow \left\{m^b(1-y)^b(-1+b+2y) - (1-m)^b(1+b-2y)y^b\right\} = 0$. Therefore, the mode $y_0$ is a solution of the equation

$$\left(\frac{1-m}{m}\right)^b = \left(\frac{1-y_0}{y_0}\right)^b \frac{b+2y_0-1}{b-2y_0+1}, \tag{47}$$

when $b>1$. For $b \leqslant 1$, the curve is convex and does not have mode.

To show that the density of l-logistic distribution is symmetric when $m=0.5$ whatever the value of $b$ is, let $f(y)$ be the pdf of the l-logistic distribution with parameters $b$ and $m=0.5$:

$$\begin{aligned} f(m-y) &= \frac{b(1-m)^b m^b (m-y)^{b-1}(1-m+y)^{b-1}}{\left[(1-m)^b(m-y)^b + m^b(1-m+y)^b\right]^2} \\ &= \frac{m^{2b}}{m^{2b}} \frac{b(m-y)^{b-1}(m+y)^{b-1}}{\left[(m-y)^b + (m+y)^b\right]^2} \\ &= \frac{b(m^2-y^2)^{b-1}}{\left[(m-y)^b + (m+y)^b\right]^2} \\ &= f(m+y). \end{aligned}$$

## References


Arnold, B. C. and R. A. Groeneveld (1995). Measuring skewness with respect to the mode. *The American Statistician 49*(1), 34–38.

Berkhof, J., I. van Mechelen, and H. Hoijtink (2000). Posterior predictive checks: Principles and discussion. *Computational Statistics 15*(3), 337–354.

Brys, G., M. Hubert, and A. Struyf (2003). A comparison of some new measures of skewness. In R. Dutter, P. Filzmoser, U. Gather, and P. J. Rousseeuw (Eds.), *Developments in Robust Statistics*, pp. 98–113. Springer-Verlag.

Cai, Y. and T. Jiang (2015). Estimation of non-crossing quantile regression curves. *Australian and New Zealand Journal of Statistics 57*(1), 139–162.

Ferrari, S. and F. Cribari-Neto (2004). Beta regression for modelling rates and proportions. *Journal of Applied Statistics 31*(7), 799–815.

Figueroa-Zúñiga, J. I., R. B. Arellano-Valle, and S. L. P. Ferrari (2013). Mixed beta regression: A Bayesian perspective. *Computational Statistics Data Analysis 61*(1), 137–147.

Fundação, J. o. P. (2010). Atlas do desenvolvimento humano no brasil. `ttp://www.atlasbrasil.org.br/2013/pt/ranking/`. [Online: accessed 30-setember-2015].

Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models. *Bayesian Analysis 1*(3), 515–533.

Gelman, A., J. Carlin, H. Stern, D. Dunson, A. Vehtari, and D. Rubin (2013). *Bayesian Data Analysis, Third Edition* (3 ed.). Philadelphia: Taylor & Francis.

Geweke, J. (1992). Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments. In *Bayesian Statistics*, pp. 169–193. University Press.

Gómez-Déniz, E., M. A. Sordo, and E. Calderín-Ojeda (2014). The log-lindley distribution as an alternative to the beta regression model with applications in insurance. *Insurance: Mathematics and Economics 54*(1), 49–57.

Groeneveld, R. A. and G. Meeden (1984). Measuring skewness and kurtosis. *Journal of the Royal Statistical Society, Series D (The Statistician) 33*(4), 391–399.

Guo, J. (2015). *R Interface to Stan, Version 2.8.0*.

Gupta, A. and S. Nadarajah (2004). *Handbook of Beta Distribution and Its Applications*. Philadelphia: Taylor & Francis.

Hao, L. and D. Naiman (2007). *Quantile Regression*. New Jersey: SAGE Publications.

Hinkley, D. V. (1975). On power transformations to symmetry. *Biometrika 62*(1), 101–111.

IBGE, D. d. P. (2010). *Pesquisa nacional por amostra de domicílios, PNAD.: Síntese de indicadores*. Fundação Instituto Brasileiro de Geografia e Estatística - IBGE.

Johnson, N., S. Kotz, and N. Balakrishnan (1994). *Continuous Univariate Distributions*, Volume 1. New York: John Wiley & Sons.

Johnson, N. L. (1949). Systems of frequency curves generated by methods of translation. *Biometrika 36*(1/2), 149–176.

Koenker, R. and J. G. Bassett (1978). Regression quantiles. *Econometrica, 46*(1), 33–50.

Kotz, S. and J. R. van Dorp (2004). *Beyond Beta: Other Continuous Families of Distributions with Bounded Support and Applications*. Singapore: World Scientific Publishing.

Lambert, P. J. and J. R. Aronson (1993). Inequality decomposition analysis and the gini coefficient revisited. *Economic Journal 103*(420), 1221–27.

Lunn, D. J., A. Thomas, N. Best, and D. Spiegelhalter (2000, October). Winbugs – a Bayesian modelling framework: Concepts, structure, and extensibility. *Statistics and Computing 10*(4), 325–337.

Martin, A. D., K. M. Quinn, and J. H. Park (2011). MCMCpack: Markov chain monte carlo in R. *Journal of Statistical Software 42*(9), 0–22.

Mitnik, P. A. and S. Baek (2013). The kumaraswamy distribution: median-dispersion re-parameterizations for regression modeling and simulation-based estimation. *Statistical Papers 54*(1), 177–192.

Moors, J. J. A. (1988). A quantile alternative for kurtosis. *Journal of the Royal Statistical Society 37*(1), 25–32. Series B.

Nair, N. U., P. Sankaran, and N. Balakrishnan (2013). *Quantile-Based Reliability Analysis*. Boston: Birkhäuser Basel.

Ntzoufras, I. (2011). *Bayesian Modeling Using WinBUGS*. New Jersey: John Wiley & Sons.

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In *Proceedings of the 3rd International Workshop on Distributed Statistical Computing*.

R Development Core Team (2015). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing.

Smyth, G. K., Y. Hu, P. Dunn, B. Phipson, and Y. Chen (2015). statmod: statistical modeling.

Tadikamalla, P. R. and N. L. Johnson (1982). Systems of frequency curves generated by transformations of logistic variables. *Biometrika 69*(2), 461–465.

Wang, M. and K. Rennolls (2005). Three diameter distribution modelling: Introducing the logit-logistic. *Canadian Journal of Forest Research 35*(6), 1305–1313.

Watanabe, S. (2010). Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research 11*, 3571–3594.