Bayesian inference for zero-and/or-one augmented rectangular beta regression models

Ana R. S. Santos, [1], Caio L N Azevedo[1]*, Jorge L. Bazan[2], Juvêncio S. Nobre[3]

[1] Department of Statistics, State University of Campinas, Brazil
[2] Department of Applied Mathematics and Statistics, University of São Paulo, Brazil
[2] Department of Statistics and Applied Mathematics, Federal University of Ceará, Brazil

Abstract

In this paper, we developed a Bayesian inference for a zero-and/or-one augmented rectangular beta regression model to analyze limited-augmented data, under the presence of outliers. The proposed Bayesian tools were parameter estimation, model fit assessment, model comparison, residual analysis and case influence diagnostics, developed through MCMC algorithms. In addition, we adapted available methods of posterior predictive checking using appropriate discrepancy measures. Also, a comparison with the maximum likelihood estimation, previously proposed in the literature was performed, in terms of parameter recovery. We noticed that the results are quite similar, but the Bayesian approach is more easily implemented, including influence diagnostics tools, besides also allowing incorporating prior information. We conducted several simulation studies, considering some situations of practical interest, in order to evaluate the parameter recovery of the proposed model and estimation method, as well as the impact of transforming the observed zeros and ones along the use of non-augmented models. A psychometric real data set was analyzed to illustrate the performance of the developed tools.

keywords: Augmented rectangular beta distribution; Bayesian inference; diagnostic analysis; MCMC algorithms; generalized linear models

*Corresponding author: Caio L N Azevedo, Department of Statistics, State University of Campinas, Mailbox 6065, SP, Brazil. Email: cnaber@ime.unicamp.br

1 Introduction

The literature related to the limited response regression models is very extensive. Several well-consolidated models are available. Among them, we can cite Paolino (2001), Ferrari and Cribari-Neto (2004), Smithson and Verkuilen (2006), Cribari-Neto and Zeileis (2010), Ma et al. (2011), Galvis et al. (2014) and Silva et al. (2017), for example. For most of the models, only the frequentist approach was developed. Within the related literature, some works provide evidence that the Bayesian inference can overcome, in some aspects, the frequentist approach, see Buckley (2003), Branscum et al. (2007), Figueroa-Zúñiga et al. (2013) and Nogarotto et al. (2015). Furthermore, Bayesian inference in terms of parameter, predictive model checking and influence are to be implemented more straightforwardly. Also, prior information can be easily included.

Wang and Luo (2015) and Wang and Luo (2016) generalize the model proposed by Bayes et al. (2012) and developed, under the Bayesian perspective, an augmented rectangular beta regression model to account for the occurrence of boundary values 0 and 1 for (0,1) data. Moreover, they account for the within-subject correlation in a longitudinal setup by introducing random effects under the generalized linear mixed models framework. Also, Bandyopadhyay et al. (2017) introduced a class of general proportion density, and further augmented the probabilities of zero and one to this general proportion density, controlling for the clustering.

The developed inference tools are parameter estimation, residual analysis, statistics for model comparison, case influence diagnostics and posterior predictive checking. MCMC algorithms are used to develop these tools. We present a comparison, in terms of parameter recovery, between the Bayesian inference here developed and the maximum likelihood inference developed by Silva et al. (2017). Also, the impact of some factors of interest (sample size, regression models and modeled parameters) on the estimates, are measured. In addition, we evaluate the impact of transforming the discrete values, in order to use non-augmented regression models, on the parameter estimation. Also, the performance of some usual statistics of model comparison are studied, concerning the selection between our model and the ZOAB regression model, using simulated data. Finally, Bayesian influence diagnostics tools and posterior predictive checking are proposed and studied.

It is noteworthy that, unlike the works of Wang and Luo (2015), Wang and Luo (2016) and Bandyopadhyay et al. (2017), in this paper, as well as in Silva et al. (2017), we present a new parametrization of the rectangular beta distribution. We also include regression structures for more parameters, and

develop techniques of residual analysis, by using the randomized quantile residuals. Finally, we conduct simulation studies in order to measure the parameter recovery, to study the behavior of the residuals and to evaluate the performance of the statistics of model comparison, considering different scenarios, defined by crossing the levels of some factors of interest. Also in this paper, we conduct simulation studies in order to analyze the behavior of the K-L divergence measure and study the posterior predictive checking techniques.

The remaining of the paper is organized as follows. In Section 2, we present the zero-and/or-one augmented rectangular beta (ZOABR) regression model. In Section 3, we discuss about the prior choice, the obtaining of the full conditional distributions and the MCMC algorithm. In Section 4, we develop techniques of residual analysis, by using the randomized quantile residuals, and we present some statistical tools for model selection, case influence diagnostics and posterior predictive checking. In Section 5, we present some simulation studies. In Section 6, we present the analysis of a psychometric data set using the developed methodology and finally, in Section 7, we present some discussion and suggestions for future research.

2 Zero-and/or-one augmented rectangular beta regression model

The ZOABR model is based on the zero-and/or-one augmented rectangular beta distribution with parameters $(p_0, p_1, \gamma, \phi, \alpha)^\top$, whose density is of the form

$$f(y; p_0, p_1, \gamma, \phi, \alpha) = p_0^{1-y} p_1^{y} \mathbb{1}_{\{0,1\}}(y) + (1 - p_0 - p_1) h(y|\gamma, \phi, \alpha) \mathbb{1}_{(0,1)}(y)$$

where $h(y; \cdot)$ the reparameterized density of the rectangular beta distribution, denoted by $\text{BRr}(\gamma, \phi, \alpha)$, may be expressed as

$$\begin{aligned} h(y; \gamma, \phi, \alpha) &= \left(1 - \sqrt{1 - 4\alpha\gamma(1-\gamma)}\right) \mathbb{1}_{(0,1)}(y) + \sqrt{1 - 4\alpha\gamma(1-\gamma)} \times \\ &\times \; b\left(\frac{\gamma - \frac{1}{2} + \frac{1}{2}\sqrt{1 - 4\alpha\gamma(1-\gamma)}}{\sqrt{1 - 4\alpha\gamma(1-\gamma)}}, \phi\right) \mathbb{1}_{(0,1)}(y) \end{aligned} \tag{2.1}$$

with $b(\cdot, \phi)$ representing the density beta distribution with parameterization used in Ferrari and Cribari-Neto (2004). Note that, this parameterization induces a restriction in the parameter space given by $0 < p_0 + p_1 < 1$. Also, when $\alpha = 0$, we obtain the augmented beta distribution proposed by Ospina

and Ferrari (2010). On the other hand, the cumulative distribution function (cdf) of ZOABR, which is useful to define the so-called quantile residuals, can be defined as follows

$$\mathrm{F}(y;\tau,\eta,\gamma,\phi,\alpha) = \tau\mathrm{Ber}(y;\eta) + (1-\tau)\mathrm{BRr}(y;\gamma,\phi,\alpha),$$

where $\tau = p_0+p_1$, $\mathrm{Ber}(y;\eta)$ is the cdf of a Bernoulli with parameter $\eta = p_1/\tau$ and $\mathrm{BRr}(y;\gamma,\phi,\alpha)$ is the cdf of the rectangular beta given in (2.1). For more details see Silva et al. (2017).

Under the Bayesian paradigm, the ZOABR model is defined by considering a set of random variables, let us say $Y_1,\ldots,Y_n$, such that $Y_t \overset{iid}{\sim} \mathrm{ZOABR}(p_{0t},p_{1t},\gamma_t,\phi_t,\alpha), t = 1,\ldots,$ with

$$\begin{aligned} g_1(\gamma_t) &= \sum_{i=1}^{p} x_{ti}\beta_i = \mathbf{x}_t^\top\boldsymbol{\beta}, g_2(\phi_t) = \sum_{j=1}^{k} -w_{tj}\delta_j = -\mathbf{w}_t^\top\boldsymbol{\delta} \qquad (2.2)\\ H(p_{0t},p_{1t}) &= (h_0(p_{0t},p_{1t}),h_1(p_{0t},p_{1t})) = (\zeta_{0t},\zeta_{1t}) = (\mathbf{v}_t^\top\boldsymbol{\rho},\mathbf{z}_t^\top\boldsymbol{\psi}), \end{aligned}$$

where $\gamma_t = \mathbb{E}(Y_t|Y_t \in (0,1)), p_{0t} = \mathbb{P}(Y_t = 0), p_{1t} = \mathbb{P}(Y_t = 1)$ and $1-p_{0t}-p_{1t} = \mathbb{P}(Y_t \in (0,1))$; $\boldsymbol{\beta} = (\beta_1,\ldots,\beta_p)^\top, \boldsymbol{\delta} = (p_1,\ldots,\delta_k)^\top, \boldsymbol{\rho} = (\rho_1,\ldots,\rho_{k_0})^\top, \boldsymbol{\psi} = (\psi_1,\ldots,\psi_{k_1})^\top$ are vectors of unknown regression parameters such that $\boldsymbol{\beta} \in \mathbb{R}^p, \boldsymbol{\delta} \in \mathbb{R}^k, \boldsymbol{\rho} \in \mathbb{R}^{k_0}$ and $\boldsymbol{\psi} \in \mathbb{R}^{k_1}$. Here, $\mathbf{x}_t = (x_{t1},\ldots,x_{tp})^\top, \mathbf{w}_t = (w_{t1},\ldots,w_{tk})^\top, \mathbf{v}_t = (v_{t1},\ldots,v_{k_0})^\top$ and $\mathbf{z}_t = (z_{t1},\ldots,z_{k_1})^\top$ are vectors with p,k,k_0 e k_1 covariates, respectively.

According to Ospina (2008), we can consider H such that

$$\begin{aligned} H(p_{0t},p_{1t}) &= (h_0(p_{0t},p_{1t}),h_1(p_{0t},p_{1t}))^\top\\ &= \left(h\left(\frac{p_{0t}}{1-p_{0t}-p_{1t}}\right), h\left(\frac{p_{1t}}{1-p_{0t}-p_{1t}}^\top\right)\right), \end{aligned}$$

where $h : \mathbb{R}^+ \to \mathbb{R}$ is strictly monotonic and twice differentiable. Notice that h_0 and h_1 are functions of $\mathbb{R}^2$ in $\mathbb{R}$. We consider the logit link functions for g_1, that is, $g_1(\gamma_t) = \log(\gamma_t/1-\gamma_t)$ and the log link for g_2, this is, $g_2(\phi_t) = \log(\phi_t)$. Following Ospina (2008), we chose h as the log link, that is, $h_0(p_{0t},p_{1t}) = \log(p_{0t}/(1-p_{0t}-p_{1t})) = \zeta_{0t}$ and $h_1(p_{0t},p_{1t}) = \log(p_{1t}/(1-p_{0t}-p_{1t})) = \zeta_{1t}$.

In order to structure it in a clearer way, the mixture between the reparameterized rectangular beta distribution and Bernoulli, can be define as the following augmented observable (indicator) variable z_t which assumes the value 0 if $y_t \in (0,1)$ or 1 if $y_t \in \{0,1\}$. Therefore, the joint distribution

of $(y_t, z_t^*)^\top$ is given by:

$$\begin{aligned} f(y_t, z_t^*; p_{0t}, p_{1t}, \gamma_t, \phi_t, \alpha) &= f_1(y_t; \gamma_t, \phi_t, \alpha)^{1-z_t^*} f_2(y_t; \eta_t)^{z_t^*} \times \\ &\times (p_{0t} + p_{1t})^{z_t^*} (1 - p_{0t} - p_{1t})^{1-z_t^*} \mathbb{1}_{\{y_t, z_t^*\}} \end{aligned} \quad (2.3)$$

where $\mathbb{1}_{\{y_t, z_t^*\}} = \mathbb{1}_{(0,1)}(y_t)\mathbb{1}_{\{0\}}(z_t^*) + \mathbb{1}_{\{0,1\}}(y_t)\mathbb{1}_{\{1\}}(z_t^*)$. The $f_1(y_t; \gamma_t, \phi_t, \alpha)$ is the density of the rectangular beta distribution as defined in (2.1) and $f_2(y_t; \eta_t)$ is the probability function of a Bernoulli with parameter $\eta_t = p_{1t}/(p_{0t}+p_{1t})$. Let us denote $f(y_t, z_t^*; p_{0t}, p_{1t}, \gamma_t, \phi_t, \alpha) \equiv f(y_t, z_t^*)$, for short.

The likelihood for the ZOABR regression model considering is given by

$$L(\boldsymbol{\Upsilon}) = \prod_{t=1}^{n} f(y_t, z_t^*; p_{0t}, p_{1t}, \gamma_t, \phi_t, \alpha) = L_1(\boldsymbol{\rho}, \boldsymbol{\psi}) L_2(\boldsymbol{\beta}, \boldsymbol{\delta}, \alpha), \quad (2.4)$$

where

$$\begin{aligned} L_1(\boldsymbol{\rho}, \boldsymbol{\psi}) &= \prod_{t=1}^{n} (p_{0t}^{1-y_t} p_{1t}^{y_t})^{z_t^*} (1 - p_{0t} - p_{1t})^{1-z_t^*} \quad \text{and} \\ L_2(\boldsymbol{\beta}, \boldsymbol{\delta}, \alpha) &= \prod_{t=1}^{n} h(y_t; \gamma_t, \phi_t, \alpha)^{1-z_t^*}, \end{aligned}$$

where p_{0t}, p_{1t}, γ_t and ϕ_t are defined by (2.2) as functions of $\boldsymbol{\rho}, \boldsymbol{\psi}, \boldsymbol{\beta}$ and $\boldsymbol{\delta}$, respectively.

Therefore, the likelihood is partially separable. In the next section, we present a discussion about the prior choice, posterior distribution and the related MCMC algorithms.

3 Prior and posterior distributions and the related MCMC algorithms

The marginal posterior distributions comprise the main tool to perform Bayesian inference. Unfortunately, it is not possible to obtain closed-form expressions of the marginal posterior distributions for our model, regardless the prior structure adopted and/or the likelihood considered. MCMC algorithms will be used to obtain samples from the marginal posteriors, see Gamerman and Lopes (2006), for example. Also, since none of the so-called full conditional distributions are known (as we show ahead), some auxiliary algorithm needs to be used to sample from them, as the Metropolis-Hastings, slice sampling or adaptive rejection sampling. We made all implementations in the R program; see R Core Team (2008).

The prior structure adopted here, for $\boldsymbol{\Upsilon}$, is $\pi(\boldsymbol{\Upsilon}) = \pi(\boldsymbol{\rho})\pi(\boldsymbol{\psi})\pi(\boldsymbol{\beta})\pi(\boldsymbol{\delta})\pi(\alpha)$, where $\boldsymbol{\rho} \sim \mathcal{N}_{k_0}(\boldsymbol{\mu}_\rho, \Sigma_\rho)$, $\boldsymbol{\psi} \sim \mathcal{N}_{k_1}(\boldsymbol{\mu}_\psi, \Sigma_\psi)$, $\boldsymbol{\beta} \sim \mathcal{N}_p(\boldsymbol{\mu}_\beta, \Sigma_\beta)$, $\boldsymbol{\delta} \sim \mathcal{N}_k(\boldsymbol{\mu}_\delta, \Sigma_\delta)$ and $\alpha \sim$ Beta(a,b).

Combining the likelihood (2.4) and the prior distribution $\pi(\boldsymbol{\Upsilon})$, the joint posterior distribution is given by

$$\begin{aligned}
\pi(\boldsymbol{\Upsilon}|\mathbf{y}) \quad \propto \quad & \left\{\prod_{t=1}^{n}(p_{0t}^{1-y_t}p_{1t}^{y_t})^{z_t^*}(1-p_{0t}-p_{1t})^{1-z_t^*}\right\}\left\{\prod_{t=1}^{n}\{\theta_t+(1-\theta_t)b(y_t;\mu_t,\phi_t)\}^{1-z_t^*}\right\} \\
\times \quad & \left\{\exp\left[-\frac{1}{2}(\boldsymbol{\rho}-\mu_\rho)^\top\boldsymbol{\Sigma}_\rho^{-1}(\boldsymbol{\rho}-\mu_\rho)\right]\right\}\left\{\exp\left[-\frac{1}{2}(\boldsymbol{\psi}-\mu_\psi)^\top\boldsymbol{\Sigma}_\psi^{-1}(\boldsymbol{\psi}-\mu_\psi)\right]\right\} \\
\times \quad & \left\{\exp\left[-\frac{1}{2}(\boldsymbol{\beta}-\mu_\beta)^\top\boldsymbol{\Sigma}_\beta^{-1}(\boldsymbol{\beta}-\mu_\beta)\right]\right\}\left\{\exp\left[-\frac{1}{2}(\boldsymbol{\delta}-\mu_\delta)^\top\boldsymbol{\Sigma}_\delta^{-1}(\boldsymbol{\delta}-\mu_\delta)\right]\right\} \\
\times \quad & \left\{\alpha^{a-1}(1-\alpha)^{b-1}\right\}. \qquad (3.1)
\end{aligned}$$

The joint posterior distribution (3.1) have intractable form, but the full conditionals are easy to sample from, even though they are not known (do not correspond to a particular distribution), which are:

$$\pi(\boldsymbol{\rho}|\mathbf{y},\boldsymbol{\psi}) \quad \propto \quad \prod_{t=1}^{n}(p_{0t}^{1-y_t}p_{1t}^{y_t})^{z_t^*}(1-p_{0t}-p_{1t})^{1-z_t^*}\exp\left\{-\frac{1}{2}(\boldsymbol{\rho}-\mu_\rho)^\top\boldsymbol{\Sigma}_\rho^{-1}(\boldsymbol{\rho}-\mu_\rho)\right\},$$

$$\pi(\boldsymbol{\psi}|\mathbf{y},\boldsymbol{\rho}) \quad \propto \prod_{t=1}^{n}(p_{0t}^{1-y_t}p_{1t}^{y_t})^{z_t^*}(1-p_{0t}-p_{1t})^{1-z_t^*}\exp\left\{-\frac{1}{2}(\boldsymbol{\psi}-\mu_\psi)^\top\boldsymbol{\Sigma}_\psi^{-1}(\boldsymbol{\psi}-\mu_\psi)\right\}.$$

$$\pi(\boldsymbol{\beta}|\mathbf{y},\boldsymbol{\delta},\alpha) \quad \propto \quad \prod_{t=1}^{n}\{\theta_t+(1-\theta_t)b(y_t;\mu_t,\phi_t)\}^{1-z_t^*}\exp\left\{-\frac{1}{2}(\boldsymbol{\beta}-\mu_\beta)^\top\boldsymbol{\Sigma}_\beta^{-1}(\boldsymbol{\beta}-\mu_\beta)\right\}, \qquad (3.2)$$

$$\pi(\boldsymbol{\delta}|\mathbf{y},\boldsymbol{\beta},\alpha) \quad \propto \quad \prod_{t=1}^{n}\{\theta_t+(1-\theta_t)b(y_t;\mu_t,\phi_t)\}^{1-z_t^*}\exp\left\{-\frac{1}{2}(\boldsymbol{\delta}-\mu_\delta)^\top\boldsymbol{\Sigma}_\delta^{-1}(\boldsymbol{\delta}-\mu_\delta)\right\}, \qquad (3.3)$$

$$\pi(\alpha|\mathbf{y},\boldsymbol{\beta},\boldsymbol{\delta}) \quad \propto \quad \prod_{t=1}^{n}\{\theta_t+(1-\theta_t)b(y_t;\mu_t,\phi_t)\}^{1-z_t^*}\,\alpha^{a-1}(1-\alpha)^{b-1}. \qquad (3.4)$$

where $\theta_t = 1 - \sqrt{1 - 4\alpha\gamma_t(1-\gamma_t)}$, $\mu_t = \dfrac{\gamma_t - \frac{1}{2} + \frac{1}{2}\sqrt{1 - 4\alpha\gamma_t(1-\gamma_t)}}{\sqrt{1 - 4\alpha\gamma_t(1-\gamma_t)}}$ and $b(y_t; \mu_t, \phi_t)$ the density beta distribution used by Ferrari and Cribari-Neto (2004).

In order to perform the Metropolis-Hastings steps, we need to consider kernel densities for the parameters $\boldsymbol{\rho}, \boldsymbol{\psi}, \boldsymbol{\beta}$, $\boldsymbol{\delta}$ and α, given by

$$\begin{aligned} q(\boldsymbol{\rho}^{(t-1)}, \boldsymbol{\rho}) &= \mathcal{N}_{k_0}(\boldsymbol{\rho}^{(t-1)}, \boldsymbol{\Sigma}_1), \quad q(\boldsymbol{\psi}^{(t-1)}, \boldsymbol{\psi}) = \mathcal{N}_{k_1}(\boldsymbol{\psi}^{(t-1)}, \boldsymbol{\Sigma}_2) \\ q(\boldsymbol{\beta}^{(t-1)}, \boldsymbol{\beta}) &= \mathcal{N}_p(\boldsymbol{\beta}^{(t-1)}, \boldsymbol{\Sigma}_3), \quad q(\boldsymbol{\delta}^{(t-1)}, \boldsymbol{\delta}) = \mathcal{N}_k(\boldsymbol{\delta}^{(t-1)}, \boldsymbol{\Sigma}_4). \end{aligned} \quad (3.5)$$

Concerning the parameter α, we will consider $q(\alpha^{(t-1)}, \alpha) = \mathcal{U}\left(\nu_1(\alpha^{(t-1)}), \nu_2(\alpha^{(t-1)})\right)$, where $\nu_1(\alpha^{(t-1)}) = \max\left\{0, \alpha^{(t-1)} - \Delta_\alpha\right\}$ and $\nu_2(\alpha^{(t-1)}) = \min\{1, \alpha^{(t-1)} + \Delta_\alpha\}$, with $\Delta_\alpha > 0$ is a constant to be previously defined. This kernel density was based on Gonçalves et al. (2013). We can notice that the distributions (3.5) are symmetric, in the sense that, for example, $q(\boldsymbol{\rho}^{(t-1)}, \boldsymbol{\rho}) = q(\boldsymbol{\rho}, \boldsymbol{\rho}^{(t-1)})$, while $q(\alpha^{(t-1)}, \alpha)$ is not. Denoting the set of all other parameters by $(\cdot)$, the MCMC algorithm, for $t = 1, 2, \ldots, B, \ldots, M$, where B is the burn-in and M is the generated sample size, simulates iteratively all unknown quantities in the following order:

1. Start the algorithm by choosing suitable initial values.

   Repeat steps 2–6.

2. Simulate $\boldsymbol{\rho}$ from $\boldsymbol{\rho} \mid (.)$.

3. Simulate $\boldsymbol{\psi}$ from $\boldsymbol{\psi} \mid (.)$.

4. Simulate $\boldsymbol{\beta}$ from $\boldsymbol{\beta} \mid (.)$.

5. Simulate $\boldsymbol{\delta}$ from $\boldsymbol{\delta} \mid (.)$.

6. Simulate α from $\alpha \mid (.)$.

Our approach also applies for the inflated beta regression model proposed by Ospina and Ferrari (2012). In this case, we have to fix $\alpha = 0$ which implies that $\theta_t = 0$ in expressions of full conditionals on (3.2) and (3.3).

4 Model fit assessment and model comparison

The residual analysis is an important tool for model fit assessment. It is possible, through the residual analysis, checking the presence of outliers, as

well as the departing from specific model assumptions. In this work, we use the randomized quantile residual (RQR). Also, we developed Bayesian influence measures by using the measure of Kullback-Leibler. In addition, we adapt methods of posterior predictive checking available in the literature, to our model, using appropriate discrepancy measures and finally, we present Bayesian model comparison criteria.

4.1 Randomized quantile residual

We adapted the randomized quantile residual (Dunn and Smyth (1996)) for our model, which is a randomized version of Cox and Snell (1968) residual, and it is given by

$$r_t^q = \Phi^{-1}(W_t), \quad t = 1, \ldots, n,$$

where $\Phi(\cdot)$ denotes cumulative distribution function of the standard normal distribution, W_t is a uniform random variable on the interval $(a_t, b_t]$, with $a_t = \lim_{y \uparrow y_t} \mathrm{F}(y; \widehat{\tau}_t, \widehat{\eta}_t, \widehat{\gamma}_t, \widehat{\phi}_t, \widehat{\alpha})$ and $b_t = \mathrm{F}(y_t; \widehat{\tau}_t, \widehat{\eta}_t, \widehat{\gamma}_t, \widehat{\phi}_t, \widehat{\alpha})$ where $\widehat{\eta}_t = \frac{\widehat{p}_{1t}}{\widehat{\tau}_t}$ and $\widehat{\tau}_t = \widehat{p}_{0t} + \widehat{p}_{1t}$. Here, $\mathrm{F}(y; \widehat{\tau}_t, \widehat{\eta}_t, \widehat{\gamma}_t, \widehat{\phi}_t, \widehat{\alpha})$ is the cdf of ZOABR distribution. For example, for zero and one augmented rectangular beta regression model, W_t is a uniform random variable on $(0, \widehat{\tau}(1 - \widehat{\eta})]$ if $y_t = 0$, is a uniform random variable on $(1 - \widehat{\tau}\widehat{\eta}, 1]$ if $y_t = 1$ and W_t =$\mathrm{F}(Y_t; \widehat{\tau}, \widehat{\eta}, \widehat{\gamma}, \widehat{\phi}, \widehat{\alpha})$ if $y_t \in (0, 1)$. Since the variable W_t is no longer continuous, we need to simulate several set of values of r_t^q and take, for example, the respective medians, for each observation. However, since the Bayesian estimators are consistent in the frequentist sense, these medians are expected to follow, approximately, a standard normal distribution. In practice, it is important so simulate at least four sets of RQR.

A plot of RQR against the index of the observations (t) or against the predicted values should present a random pattern. A systematic behavior may suggest a misspecification of the model. Also, a quantile-quantile plot based on the standard normal distributions, with simulated envelopes, are a helpful diagnostic tool, see to details Atkinson (1985). The simulated envelopes were constructed simulating standard Normal distribution values.

4.2 Bayesian model comparison criteria

Under the Bayesian framework, especially when MCMC algorithms are used to obtain the posterior distributions, some statistics for model comparison can be easily calculated, see Spiegelhalter et al. (2002). To facilitate the definition of these statistics, we will first define $D(\boldsymbol{\Upsilon}) = -2\log[L(\boldsymbol{\Upsilon}; y)]$, where

$L(\boldsymbol{\Upsilon}; y)$ is the likelihood given in (2.4). Also, let $\boldsymbol{\Upsilon}^{(m)}$, $m = 1, ..., M$, be the $m-$th value of the valid simulated MCMC sample, that is, the MCMC sample obtained after discarding the burn-in and a proper spacing (lag) between the values. In addition, let $\overline{\boldsymbol{\Upsilon}}$ be the vector with the posterior expectation, of each parameter, based on the valid MCMC sample, and $\overline{D(\boldsymbol{\Upsilon})} = \frac{1}{M}\sum_{m=1}^{M} D(\boldsymbol{\Upsilon}^{(m)})$. Denote also the deviance by $D(\overline{\boldsymbol{\Upsilon}}) = -2\log[L(\overline{\boldsymbol{\Upsilon}}; y)]$, and the deviance information criterion (DIC) by $DIC = D(\overline{\boldsymbol{\Upsilon}}) + 2p_D$, where $p_D = \overline{D(\boldsymbol{\Upsilon})} - D(\overline{\boldsymbol{\Upsilon}})$.

The EAIC (posterior expectation of AIC) and EBIC (posterior expectation of BIC) are given, respectively, by $EAIC = D(\bar{\boldsymbol{\Upsilon}}) + 2p_{\boldsymbol{\Upsilon}}$ and $EBIC = D(\bar{\boldsymbol{\Upsilon}}) + p_{\boldsymbol{\Upsilon}}\log(n)$, where $p_{\boldsymbol{\Upsilon}}$ is the total number of parameters of the model. Finally, let $L(\boldsymbol{\Upsilon}|y_i)$, $i = 1, ..., n$, be the likelihood (see Equation (2.4)) related to the ith observation. Then, the LPML (logarithm of the pseudo-marginal likelihood) is calculated as LPML $= \sum_{i=1}^{n} \ln(\widehat{CPO_i})$,

$$\widehat{CPO_i} = \left\{ \frac{1}{M} \sum_{m=1}^{M} \frac{1}{L(\boldsymbol{\Upsilon}^{(m)}|y_i)} \right\}^{-1}, \tag{4.1}$$

represents the conditional predictive ordinate, see Ibrahim et al. (2001) and Gelfand et al. (1992). The smaller the values of DIC, EAIC, EBIC and deviance, the better the model fit, but the opposite occurs with the LPML statistic. The EAIC and EBIC statistics tend to select the model with the smallest number of parameters ($p_{\boldsymbol{\Upsilon}}$) since it gives more penalty to models with more parameters. On the other hand, the DIC tends to select the most complex (or the most general) model, that is, it tends to select the overfitted model, see Ando (2007). Finally, the LPLM and the Deviance statistics tend to select the model that presents the largest likelihood. This corresponds to the most general model, when the competing models are nested.

4.3 Bayesian case influence diagnostics

Since regression models are sensitive to the underlying model assumptions, it is important to perform sensitivity analysis. Here, we consider the measure of divergence within the Bayesian context as in Cho et al. (2009), who developed case deletion influence diagnostics for both joint and marginal posterior distributions based on the Kullback-Leibler (K-L) divergence, and presented a simple way of calculating such influence measure by using MCMC outputs. Let $K(P, P_{(-i)})$ be the K-L divergence between P and $P_{(-i)}$, where P stands for the posterior distribution of $\boldsymbol{\Upsilon}$ for the full data and $P_{(-i)}$ stands for the

posterior distribution of $\boldsymbol{\Upsilon}$ without the ith observation. Then, we have

$$K(P; P_{(-i)}) = \int \pi(\boldsymbol{\Upsilon}|\mathbf{y}) \log \left[\frac{\pi(\boldsymbol{\Upsilon}|\mathbf{y})}{\pi(\boldsymbol{\Upsilon}|\mathbf{y}^{(-i)})} \right] d\boldsymbol{\Upsilon}.$$

where $\mathbf{y}_{(-i)}$ corresponds to $\mathbf{y}$ without the ith observation. Also, using the notation introduced earlier in Section 4.2, the MCMC estimate of $K(P, P_{(-i)})$ is $\widehat{K}(P, P_{(-i)}) = \frac{1}{M} \sum_{m=1}^{n} \log[L(\boldsymbol{\Upsilon}^{(m)}|y_i)] - \log(\widehat{CPO_i})$, where $\widehat{CPO_i}$ is as in (4.1).

Cho et al. (2009) also show a way to calibrate the divergence $K(P; P_{(-i)})$. The calibration is given by

$$p_i = \frac{1}{2} \left[1 + \sqrt{1 - \exp\left\{-2K(P; P_{(-i)})\right\}} \right].$$

This expression implies that $0.5 \leq p_i \leq 1$. The authors suggest considering the $i-$th observation as influent if p_i is much greater than 0.5, $i = 1, \ldots, n$.

4.4 Posterior Predictive Checking

Under the Bayesian perspective, a way to check the goodness of the model fit, is to compare the predictive distribution with the distribution of the observed data. Let $\mathbf{y}^{obs}$ be the observed response, and $\mathbf{y}^{rep}$ the replicated response generated from its posterior predictive distribution, which is given by

$$p(\mathbf{y}^{rep}|\mathbf{y}^{obs}) = \int p(\mathbf{y}^{rep}|\boldsymbol{\Upsilon}) p(\boldsymbol{\Upsilon}|\mathbf{y}^{obs}) d\boldsymbol{\Upsilon}. \tag{4.2}$$

Discrepancy measures $D(\mathbf{y}, \boldsymbol{\Upsilon})$ are defined (Gelman et al. (1996)), and the posterior distribution of $D(\mathbf{y}^{obs}, \boldsymbol{\Upsilon})$ is compared to the posterior predictive distribution of $D(\mathbf{y}^{rep}, \boldsymbol{\Upsilon})$, an substantial differences between them indicating model misfit. Gelman et al. (2003) suggest several graphs to compare the replicated and the observed data, under the given measure of divergence.

Another measure used to quantify the goodness of fit, is the Bayesian p-value, which for an adopted discrepancy measure, and is defined as

$$\begin{aligned} \mathbb{P}(D(\mathbf{y}^{rep}, \boldsymbol{\Upsilon})) &\geq D(\mathbf{y}^{obs}, \boldsymbol{\Upsilon})|\mathbf{y}^{obs}) \\ &= \int_{D(\mathbf{y}^{rep}, \boldsymbol{\Upsilon})) \geq D(\mathbf{y}^{obs}, \boldsymbol{\Upsilon})} p(\mathbf{y}^{rep}|\boldsymbol{\Upsilon}) p(\boldsymbol{\Upsilon}|\mathbf{y}^{obs}) d\mathbf{y}^{rep} d\boldsymbol{\Upsilon}. \end{aligned} \tag{4.3}$$

Due to the difficulty in dealing with equation (4.2) or (4.3) analytically, Rubin (1984) suggest simulating replicated data sets from the PPD. One draws M simulations $\boldsymbol{\Upsilon}_1, \boldsymbol{\Upsilon}_2, \ldots, \boldsymbol{\Upsilon}_M$ from the posterior distribution $p(\boldsymbol{\Upsilon}|\mathbf{y})$ of $\boldsymbol{\Upsilon}$ and then draws $\mathbf{y}^{rep,n}$ from the distribution $p(\mathbf{y}|\boldsymbol{\Upsilon}^n)$ for $n = 1, \ldots, M$. The proportion of the M replications for which $D(\mathbf{y}^{rep,n}, \boldsymbol{\Upsilon}^n)$ exceeds $D(\mathbf{y}, \boldsymbol{\Upsilon}^n)$ provides an estimate of the Bayesian p-value. Extreme values of the p-value Bayesian (less than 0.05 or greater than 0.95, depending on the nature of the discrepancy measure) indicate model misfit; see Sinharay et al. (2006). The Bayesian p-value is not necessarily uniformly distributed under the well fitting of the model, and there is some evidence that Bayesian p-value under the correct model tend to be closer to 0.5 more often than would be expected under a uniform distribution (Bayarri and Berger (2000); Sinharay and Stern (2003)).

Based on Gelman et al. (1996), the measure of discrepancy used was

$$D(\mathbf{y}; \boldsymbol{\Upsilon}) = \sum_{i=1}^{n} \frac{(y_i - \mathbb{E}(Y_i|\boldsymbol{\Upsilon}))^2}{\mathbb{V}\text{ar}(Y_i|\boldsymbol{\Upsilon})}.$$

The conditional mean and variance can be found in Silva et al. (2017).

5 Simulation studies

In this section, we present six simulation studies. The first is related to the parameter recovery of the MCMC algorithm with a comparison with the ML estimates (Study 1). Also, we compare the impact of transforming the values zero and one, using the non-augumented beta regression models, instead using our ZOABR model, and also to compare the Bayesian and ML estimates (Study 2). The four other studies are the analysis of the behavior of the RQR (Study 3), the analysis of the behavior of the K-L divergence measure (Study 4), the study of the performance of the statistics of model comparison (Study 5) and the study of the posterior predictive checking techniques (Study 6). From the results related to a convergence study (not presented for the sake of simplicity), we observed that to set a burn-in of 1,000, a spacement of 100 and generating a total of 101,000 values was enough to have valid MCMC samples of 1,000 values for each parameter.

In Study 1, several relevant scenarios were considered, which correspond to the combination of the levels of some factors of interest. The factors (with the respective levels within parenthesis) are sample size (n) (50, 100, 500), regression models (rectangular beta, zero augmented rectangular beta, one

augmented rectangular beta, zero and one augmented rectangular beta), parameters to be modeled (mean, mean and dispersion parameter, mean, dispersion and the probabilities of occurrences of zeros and/or ones). Therefore, 33 scenarios were considered, since for the BRr model, the probabilities of occurrence of zeros and ones are null. For each scenario, we simulated the response and estimated the parameters under the same model. More details are provided in Subsection 5.1.

In Study 2, two factors were fixed: the sample size (n) (50, 100, 500) and the probabilities of zeros and ones ($p_0 = p_1 = 1\%$, $p_0 = 5\%$ and $p_1 = 3\%$, $p_0 = 10\%$ and $p_1 = 8\%$, $p_0 = p_1 = 20\%$). The data sets were simulated from the ZOABR model, modeling only the mean. In addition, we fitted the ZOABR and BRr models, in both modeling only the mean.

The Study 3 was subdivided in two studies. In the first study, we considered four scenarios, in which we used the ZOABR regression model to simulate the data according to the regression structure defined for each situation and we fitted the ZOABR regression model. The regression structure for this study is $g_1(\gamma_t) = \beta_0 + \beta_1 x_t$ and $g_2(\phi_t) = \delta_0 + \delta_1 w_t$, where $t = 1, \ldots, n$. In Table 1, we present the four scenarios considered, where in scenario C3, $F(\cdot)$ is the cumulative distribution function of the t-Student distribution with $\nu = 4$ degrees of freedom. In the second scenario, we simulated the data from the ZOABR model, modeling the mean and dispersion parameter, and we fitted the ZOABR and ZOAB models, in both modeling the mean and the dispersion parameter.

Table 1: Scenarios considered in the Study 3.

Scenarios	Simulation of data	Model fit
C1	$g_1(\gamma_t) = \log(\gamma_t/(1-\gamma_t))$ $g_2(\phi_t) = -\log(\phi_t)$	$g_1(\gamma_t) = \log(\gamma_t/(1-\gamma_t))$ $g_2(\phi_t) = -\log(\phi_t)$
C2	$g_1(\gamma_t) = \log(\gamma_t/(1-\gamma_t))$ $g_2(\phi_t) = -\log(\phi_t)$	$g_1(\gamma_t) = \log(\gamma_t/(1-\gamma_t))$
C3	$g_1(\gamma_t) = F^{-1}(\gamma_t)$ $g_2(\phi_t) = -\log(\phi_t)$	$g_1(\gamma_t) = \log(\gamma_t/(1-\gamma_t))$ $g_2(\phi_t) = -\log(\phi_t)$
C4	$g_1(\gamma_t) = \log[-\log(1-\gamma_t)]$ $g_2(\phi_t) = -\log(\phi_t)$	$g_1(\gamma_t) = \log(\gamma_t/(1-\gamma_t))$ $g_2(\phi_t) = -\log(\phi_t)$

The Studies 4, 5 and 6 were also subdivided in two studies. In the fourth study, we simulated the data from the ZOABR model, modeling the mean, the dispersion parameter and the probabilities of occurrences of zeros and ones and, then, we fitted the ZOABR and ZOAB models modeling the

same parameters. In the fifty study, we simulated the data from the ZOAB model, modeling all parameters, and then we fitted the ZOABR and ZOAB models modelling the mean, dispersion parameter and the probabilities of occurrences of zeros and ones. The sample sizes considered were: 200 (Study 4); 100 and 500 (Study 5); 50, 100 and 500 (Study 6).

The results of Study 1 will be presented only for the ZOABR model, modeling all parameters, except α. In addition, the results of Study 2 will be presented only for the probabilities of zeros and ones in the sample ($p_0 = 5\%$ and $p_1 = 3\%$, $p_0 = 10\%$ and $p_1 = 8\%$). The results regarding the residual analysis, model selection, influence analysis and posterior predictive checking will not be presented, for the sake of simplicity. However, they indicated that the RQR perform well in detecting the departing from some model assumptions (to detect a heteroscedasticity not accommodated by the model, to detect a misspecification of the link function and validity of the distributional assumption of the response variable). Regarding the influence analysis, as expected, we verified that none of the observations appeared as influentials for the ZOABR model. Regarding the model selection study, the results indicate that the true underlying model was chosen in at least 99% of the 100 replicas. Finally considering study posterior predictive checking, we notice that in the first study the ZOABR model presents a better fit compared to the ZOAB model. While in the second study, the ZOAB model was more indicated. In this second study, the estimate of α obtained in the fit of the ZOABR model were very close to zero.

5.1 Study 1

The true parameters were fixed as: $\rho_0 = -1.8, \rho_1 = 1.5, \psi_0 = -1.8, \psi_1 = 1.5, \beta_0 = -1.5, \beta_1 = 1.5, \delta_0 = -3.0, \delta_1 = -1.8$ and $\alpha = 0.5$. We generated 100 replicas from $Y_t \sim$ZOABR$(p_{0t}, p_{1t}, \gamma_t, \phi_t, \alpha)$ considering $\log(p_{0t}/(1-p_{0t}-p_{1t})) = \rho_0 + \rho_1 v_t, \log(p_{1t}/(1-p_{0t}-p_{1t})) = \psi_0 + \psi_1 z_t, \log(\gamma_t/(1-\gamma_t)) = \beta_0 + \beta_1 x_t, \log(\phi_t) = -\delta_0 - \delta_1 w_t$ and $t = 1, \ldots, n$, where x_t, w_t, v_t and z_t were generated independently from a uniform distribution on $(0, 1)$. We fixed three sample sizes namely $n = 50, 100, 500$.

Using the estimates obtained in the replicas, we calculated the usual statistics to measure the accuracy of the estimates: mean, variance (Var), bias, root mean squared error (RMSE) and absolute value of relative bias (AVRB). Let υ be the parameter of interest and let $\hat{\upsilon}_r$ be some estimate related to the replica r. The formulas of the adopted statistics are: Mean$=\sum_{r=1}^{R} \frac{\hat{\upsilon}_r}{R} =$

$\hat{\bar{\upsilon}}_R$, Var$=\sum_{r=1}^{R}\frac{(\widehat{\upsilon}_r-\widehat{\bar{\upsilon}}_R)^2}{R-1}$, Bias$=\widehat{\bar{\upsilon}}_R-\upsilon$, RMSE$=\sqrt{\frac{1}{R}\sum_{r=1}^{R}(\upsilon-\hat{\upsilon}_r)^2}$, AVRB$=\frac{|\widehat{\bar{\upsilon}}_R-\upsilon|}{|\upsilon|}$.
The smaller their value, the more accurate the estimate is, except for the mean.

The results of Study 1 are shown in Table 2. We will refer to the maximum likelihood estimates by ML and the Bayesian estimates, as Bayesian. Note that as the sample size increases, the better is the accuracy of the estimates, as expected. We can observe that the Bias, Variance and AVRB for β_0, β_1 and α tend to approach to zero as the sample increases (n) for both estimation methods. The same occurs with the other parameters, even though the respective Bias and AVRB are higher compared with the other parameters. In addition, the Variance and RMSE are smaller for the Bayesian estimates compared to ML, for all parameters, and sample sizes equal to 50 and 100. In a general way, we can say that the parameters were properly recovered by the two estimation methods. In addition, we can notice certain equivalence between the Bayesian and ML approach, since the Bayesian and ML estimates are equivalent.

Table 2: Mean, Variance, Bias, RMSE and AVRB for the parameters of ZOABR model, under different sample size - Study 1.

Parameter	n	Estimation method	Mean	Variance	Bias	RMSE	AVRB
ρ_0	50	Bayesian	-1.997	.584	-0.197	.789	.109
		ML	-1.988	.731	-0.188	.875	.104
	100	Bayesian	-2.038	.483	-0.238	.734	.132
		ML	-2.041	.522	-0.241	.761	.134
	500	Bayesian	-1.841	.058	-0.041	.245	.023
		ML	-1.839	.058	-0.039	.245	.021
ρ_1	50	Bayesian	1.614	1.775	.114	1.337	.076
		ML	1.699	2.092	.199	1.460	.133
	100	Bayesian	1.759	.955	.259	1.011	.173
		ML	1.805	1.031	.305	1.060	.203
	500	Bayesian	1.554	.136	.054	.372	.036
		ML	1.559	.136	.059	.373	.039
ψ_0	50	Bayesian	-1.785	.754	.015	.868	.008
		ML	-1.795	.875	.005	.936	.003
	100	Bayesian	-1.873	.279	-0.073	.533	.040
		ML	-1.869	.293	-0.069	.545	.038
	500	Bayesian	-1.797	.076	.003	.276	.002
		ML	-1.794	.077	.006	.277	.003
ψ_1	50	Bayesian	1.198	2.428	-0.302	1.587	.202
		ML	1.332	2.530	-0.168	1.599	.112
	100	Bayesian	1.584	.712	.084	.848	.056
		ML	1.621	.740	.121	.869	.081
	500	Bayesian	1.499	.182	-0.001	.426	.001
		ML	1.503	.183	.003	.428	.002
β_0	50	Bayesian	-1.521	.053	-0.021	.231	.014
		ML	-1.537	.072	-0.037	.271	.025
	100	Bayesian	-1.539	.023	-0.039	.158	.026
		ML	-1.536	.029	-0.036	.174	.024
	500	Bayesian	-1.498	.007	.002	.081	.001
		ML	-1.497	.007	.003	.083	.002

Continues on the next page

Continuation of Table 2

Parameter	n	Estimation method	Mean	Variance	Bias	RMSE	AVRB
	50	Bayesian	1.518	.081	.018	.285	.012
		ML	1.541	.125	.041	.356	.027
β_1	100	Bayesian	1.534	.037	.034	.194	.022
		ML	1.532	.042	.032	.208	.021
	500	Bayesian	1.493	.009	-0.007	.096	.004
		ML	1.492	.009	-0.008	.097	.005
	50	Bayesian	-3.060	.616	-0.060	.787	.020
		ML	-3.079	1.351	-0.079	1.165	.026
δ_0	100	Bayesian	-2.935	.369	.065	.611	.022
		ML	-2.970	.487	.030	.699	.010
	500	Bayesian	-3.017	.080	-0.017	.284	.006
		ML	-3.024	.082	-0.024	.287	.008
	50	Bayesian	-1.820	2.147	-0.020	1.466	.011
		ML	1.966	4.920	-0.166	2.224	.092
δ_1	100	Bayesian	-1.750	.921	.050	.961	.028
		ML	-1.888	1.366	-0.088	1.172	.049
	500	Bayesian	-1.769	.208	.031	.457	.017
		ML	-1.794	.214	.006	.463	.003
	50	Bayesian	.447	.020	-0.053	.151	.107
		ML	.432	.039	-0.068	.209	.137
α	100	Bayesian	.464	.008	-0.036	.098	.072
		ML	.475	.011	-0.025	.110	.050
	500	Bayesian	.491	.003	-0.009	.054	.017
		ML	.495	.003	-0.005	.055	.010

5.2 Study 2

We fixed the values of the parameters as: $\beta_0 = -1.5, \beta_1 = 1.5, \phi = 50$ and $\alpha = 0.5$. Then, we generated 100 replicas from $Y_t \sim$ZOABR$(p_0, p_1, \gamma_t, \phi, \alpha)$ considering $\log(\gamma_t/(1-\gamma_t)) = \beta_0 + \beta_1 x_t$ and $t = 1, \ldots, n$, where x_t was generated from a uniform distribution on $(0, 1)$. We fixed three sample sizes, $n = 50, 100, 500$ and the following probabilities of zeros and ones: $(a) p_0 = 5\%, p_1 = 3\%$ and $(b) p_0 = 10\%, p_1 = 8\%$. Then, we fitted the ZOABR and the BRr models, modeling only the mean, in both models. In the case of the BR model (non augmented), the observations equal to zero were replaced by 0.001, whereas those equal to one were replaced by 0.999.

For the ZOABR model, we present only the results related to continuous

part $(\boldsymbol{\beta}^\top, \boldsymbol{\delta}^\top, \boldsymbol{\alpha}^\top)^\top$ (once we are comparing them with those obtained by the BRr regression model and these are the unique parameters presented in the two models).

The results are shown in Tables 3 and 4. We will refer to the maximum likelihood and Bayesian estimates for the BRr regression model as ML_1 and Bayesian_1, respectively, and for the ZOABR regression model as ML_2 and Bayesian_2, respectively

We can notice that the higher is the percentage of zeros and ones, the less accurate are the estimates associated to the BRr model. This behavior is expected, since the higher those quantities are, the greater the impact of the transformation applied in the data, on the parameter estimation, see, Silva et al. (2017), Galvis et al. (2014) and Nogarotto et al. (2015). Table 3 shows the results for $p_0 = 5\%$ and $p_1 = 3\%$. Considering the sample sizes equal to 50, 100 and 500, the number of observations equal to zeros and ones in the sample are approximately 4, 8 and 40, respectively. The variance associated with ϕ are much higher for the BRr regression model, under sample sizes equal to 50 and 100, mainly under Bayesian estimation. In addition, the estimates of ϕ differ between the Bayesian and ML methods. For example, under $n = 50$, the mean of the estimates obtained through the the BRr and ZOABR models using the Bayesian approach, are 94.252 and 72.282, respectively, and using the ML method, the mean of the estimates are 33.817 and 58.762, respectively. Moreover, as the sample size increases, the estimates tend to be more accurate for the ZOABR model.

Table 4 presents the results under $p_0 = 10\%$ and $p_1 = 8\%$. Differently from the previous situation, as the sample size increases, the Bias and AVRB increase, for the BRr model. In addition, the estimation of ϕ is substantially affected by the data transformation, since the AVRB were very high compared to those related to the ZOABR model, although the variance associated with ϕ is higher for the ZOABR model when considering the ML method. However, considering the Bayesian method, for sample sizes equal to 50 and 100, the variance associated with ϕ is very high for the BRr model compared to those related to the ZOABR model. Moreover, the ZOABR model performs better according to all other statistics.

Table 3: Mean, Variance, Bias, RMSE and AVRB for the parameters of ZOABR and BRr models, under different sample size - (a) $p_0 = 5\%, p_1 = 3\%$ - Study 2.

Parameter	n	Estimation method	Mean	Variance	Bias	RMSE	AVRB
		Bayesian_1	-1.274	.063	.226	.337	.151
		Bayesian_2	-1.502	.030	-0.002	.174	.001
	50	ML_1	-1.181	.251	.319	.594	.212
		ML_2	-1.520	.034	-0.020	.186	.013
β_0		Bayesian_1	-1.274	.041	.226	.304	.151
		Bayesian_2	-1.492	.014	.008	.117	.005
	100	ML_1	-1.184	.109	.316	.456	.210
		ML_2	-1.499	.015	.001	.122	.001
		Bayesian_1	-1.308	.003	.192	.200	.128
		Bayesian_2	-1.506	.003	-0.006	.054	.004
	500	ML_1	-1.307	.003	.193	.201	.129
		ML_2	-1.507	.003	-0.007	.054	.005
		Bayesian_1	1.288	.064	-0.212	.330	.141
		Bayesian_2	1.506	.042	.006	.205	.004
	50	ML_1	1.169	.385	-0.331	.703	.220
		ML_2	1.525	.044	.025	.212	.017
β_1		Bayesian_1	1.280	.040	-0.220	.297	.147
		Bayesian_2	1.490	.021	-0.010	.146	.006
	100	ML_1	1.146	.198	-0.354	.568	.236
		ML_2	1.497	.022	-0.003	.150	.002
		Bayesian_1	1.314	.004	-0.186	.197	.124
		Bayesian_2	1.508	.004	.008	.066	.005
	500	ML_1	1.313	.004	-0.187	.198	.125
		ML_2	1.508	.004	.008	.066	.006
		Bayesian_1	94.252	5.688.562	44.252	87.446	.885
		Bayesian_2	72.282	754.337	22.282	35.367	.446
	50	ML_1	33.817	1.166.488	-16.183	37.794	.324
		ML_2	58.762	594.277	8.762	25.905	.175
ϕ		Bayesian_1	79.079	9.124.155	29.079	99.849	.582
		Bayesian_2	61.323	243.076	11.323	19.269	.226
	100	ML_1	46.093	795.931	-3.907	28.482	.078
		ML_2	55.880	218.808	5.880	15.918	.118
		Bayesian_1	56.904	29.862	6.904	8.805	.138
		Bayesian_2	51.328	22.739	1.328	4.950	.027
	500	ML_1	55.865	28.470	5.865	7.929	.117
		ML_2	50.434	22.165	0.434	4.728	.009
		Bayesian_1	.643	.010	.143	.175	.286
		Bayesian_2	.493	.014	-0.007	.118	.015
	50	ML_1	.448	.073	-0.052	.274	.104
		ML_2	.479	.020	-0.021	.143	.042
α		Bayesian_1	.639	.006	.139	.158	.278
		Bayesian_2	.495	.005	-0.005	.073	.009
	100	ML_1	.528	.052	.028	.230	.055
		ML_2	.491	.007	-0.009	.082	.017
		Bayesian_1	.634	.001	.134	.137	.268
		Bayesian_2	.495	.001	-0.005	.035	.011
	500	ML_1	.635	.001	.135	.139	.271
		ML_2	.495	.001	-0.005	.036	.011

Table 4: Mean, Variance, Bias, RMSE and AVRB for the parameters of ZOABR and BRr models using different sample size - (b) $p_0 = 10\%, p_1 = 8\%$ - Study 2.

Parameter	n	Estimation method	Mean	Variance	Bias	RMSE	AVRB
		Bayesian_1	-0.865	.327	.635	.854	.423
		Bayesian_2	-1.486	.032	.014	.180	.009
	50	ML_1	-0.629	.370	.871	1.063	.581
		ML_2	-1.509	.037	-0.009	.194	.006
β_0		Bayesian_1	-0.786	.182	.714	.832	.476
		Bayesian_2	-1.498	.017	.002	.131	.001
	100	ML_1	-0.484	.190	1.016	1.106	.678
		ML_2	-1.506	.019	-0.006	.137	.004
		Bayesian_1	-0.446	.057	1.054	1.081	.703
		Bayesian_2	-1.507	.004	-0.007	.065	.005
	500	ML_1	-0.442	.071	1.058	1.091	.705
		ML_2	-1.508	.004	-0.008	.065	.005
		Bayesian_1	.826	.491	-0.674	.972	.450
		Bayesian_2	1.483	.051	-0.017	.226	.012
	50	ML_1	.596	.582	-0.904	1.183	.603
		ML_2	1.507	.059	.007	.243	.005
β_1		Bayesian_1	.755	.222	-0.745	.882	.497
		Bayesian_2	1.494	.028	-0.006	.167	.004
	100	ML_1	.415	.232	-1.085	1.187	.724
		ML_2	1.501	.029	.001	.172	.001
		Bayesian_1	.396	.071	-1.104	1.136	.736
		Bayesian_2	1.507	.005	.007	.073	.005
	500	ML_1	.394	.084	-1.106	1.143	.737
		ML_2	1.508	.005	.008	.074	.005
		Bayesian_1	218.279	89.532.854	168.279	343.294	3.366
		Bayesian_2	77.043	953.693	27.043	41.049	.541
	50	ML_1	4.394	149.581	-45.606	47.217	.912
		ML_2	59.027	727.490	9.027	28.443	.181
ϕ		Bayesian_1	73.242	28.088.461	23.242	169.200	.465
		Bayesian_2	60.323	173.257	10.323	16.728	.206
	100	ML_1	2.726	90.759	-47.274	48.225	.945
		ML_2	53.977	141.181	3.977	12.530	.080
		Bayesian_1	1.353	16.290	-48.647	48.815	.973
		Bayesian_2	51.106	30.085	1.106	5.595	.022
	500	ML_1	.932	.061	-49.068	49.068	.981
		ML_2	50.116	28.682	.116	5.357	.002
		Bayesian_1	.713	.027	.213	.270	.426
		Bayesian_2	.509	.012	.009	.111	.018
	50	ML_1	.386	.120	-0.114	.365	.227
		ML_2	.487	.019	-0.013	.139	.026
α		Bayesian_1	.632	.025	.132	.207	.264
		Bayesian_2	.498	.008	-0.002	.087	.004
	100	ML_1	.464	.114	-0.036	.340	.072
		ML_2	.493	.009	-0.007	.094	.013
		Bayesian_1	.443	.051	-0.057	.234	.115
		Bayesian_2	.491	.002	-0.009	.048	.018
	500	ML_1	.426	.097	-0.074	.319	.149
		ML_2	.491	.002	-0.009	.048	.018

6 Application

The analyzed data set was obtained from Carlstrom et al. (2000) and it is available from `http://www.stat.ucla.edu/projects/datasets/risk_perception.html`. It corresponds to a psychometric study of risk perception. The part that we are interested in corresponds to the so-called subjective part, where subjects were asked about the risk perceived by them, related to several financial and health activities. Each subject were asked to provide a number in the interval [0,100], such that the higher the value is, the higher the risk perceived is, being 0 non-risk and 100 the maximum risk. In order to use the ZOABR model, the observations were transformed to the interval [0,1]. Also, several covariables were measured and we aim to study their impact on the risk perception. The covariables are age (measured in years), gender (male and female), world view (wvcat), classified as hierarchicalist, individualist, egalitarian or other (unclassifiable) and ethnicity (Caucasian, African-American, Mexican-American or Taiwanese-American).

We analyzed the perception of the subjects about the risk related to a screening for genes that may predispose subjects to heart disease. We have a total of 588 observations, being 86 equal to zero and 21 to one. That is, approximately 14.63% of the participants provide a null risk perception whereas 3.57% provide a maximum risk.

We notice in the simulation studies an equivalence between the Bayesian and ML approaches. In addition to comparing methods, the Bayesian approach is easier to implement (if we use OpenBUGS) and for developing tools for case influence diagnostics and it allows to incorporating prior information.

The MCMC algorithm was implemented in *software* `R`. The prior distributions considered for the ZOAB regression model were: $\boldsymbol{\beta} = (\mu_1, (\beta_2)_2)^\top \sim \mathcal{N}_p(\mathbf{0},\, 25\mathbf{I}_p)$, $\boldsymbol{\delta} = (\mu_2, (\delta_2)_2)^\top \sim \mathcal{N}_k(\mathbf{0},\, 25\mathbf{I}_k)$ and for the ZOABR regression model were: $\boldsymbol{\beta} = (\mu_1, (\beta_2)_2, (\beta_2)_4)^\top \sim \mathcal{N}_p(\mathbf{0},\, 25\mathbf{I}_p)$, $\boldsymbol{\delta} = (\mu_2, (\delta_2)_2, (\delta_2)_4)^\top \sim \mathcal{N}_k(\mathbf{0},\, 25\mathbf{I}_k)$, $\alpha \sim \text{Beta}(1,1)$. Moreover, for the discrete part of the ZOAB and ZOABR regression models, we consider the following prior distributions: $\boldsymbol{\rho} = (\rho_0, \rho_1)^\top \sim \mathcal{N}_{k_0}(\mathbf{0},\, 25\mathbf{I}_{k_0})$ and $\boldsymbol{\psi} = (\psi_0, \psi_1)^\top \sim \mathcal{N}_{k_1}(\mathbf{0},\, 25\mathbf{I}_{k_1})$, where I_j represent the indetity matrix of order j.

We started fitting the ZOABR and ZOAB models, including all covariables that were apparently significant through a descriptive analysis, (the related results will not be presented for the sake of simplicity) without interactions. Assuming $Y_{tijk} \overset{ind.}{\sim} ZOABR(p_{0tijk}, p_{1tijk}, \gamma_{tijk}, \phi_{tijk}, \alpha)$ for the ZOABR model and $Y_{tijk} \overset{ind.}{\sim} ZOAB(p_{0tijk}, p_{1tijk}, \gamma_{tijk}, \phi_{tijk})$ for the ZOAB

model, the initial structure for the linear predictors, for both models, is:

$$\begin{aligned}
\log(\gamma_{tijk}/(1-\gamma_{tijk})) &= \mu_1 + (\beta_1)_i + (\beta_2)_j + (\beta_3)_k, \\
\log(\phi_{tijk}) &= -\mu_2 - (\delta_1)_i - (\delta_2)_j - (\delta_3)_k, \\
\log(p_{0tijk}/(1-p_{0tijk}-p_{1tijk})) &= \mu_3 + \rho_1 x_{tijk} + (\rho_2)_k, \\
\log(p_{1tijk}/(1-p_{0tijk}-p_{1tijk})) &= \mu_4 + \psi_1 x_{tijk} + (\psi_2)_k, \qquad (6.1)
\end{aligned}$$

where $t = 1, \ldots, n$; i=1 (female), 2 (male); j=1 (Caucasian), 2 (African-American), 3 (Mexican-American), 4 (Taiwanese-American); k=1 (unclassifiable), 2 (individualist), 3 (hierarchicalist), 4 (egalitarian) and $(\beta_1)_0 = (\beta_2)_1 = (\beta_3)_0 = 0$, $(\delta_1)_0 = (\delta_2)_1 = (\delta_3)_0 = 0$, $(\rho_2)_0 = 0$, $(\psi_2)_0 = 0$. The parameters $(\beta_1, \delta_1), (\beta_2, \delta_2)$ and $(\beta_3, \delta_3, \rho_2, \psi_2)$ are related to gender, ethnicity and wvcat, respectively and x_{tijk} is the age subject t, from gender i, ethnicity j and world view k.

Considering the regression structure defined in (6.1), we fitted the ZOAB model. Then we excluded all non-significant covariables, combining the equivalent levels within each significant covariable (according to the non-significance of the respective parameters). The final selected regression structure is:

$$\begin{aligned}
\log(\gamma_{tijk}/(1-\gamma_{tijk})) &= \mu_1 + (\beta_2)_j, \\
\log(\phi_{tijk}) &= -\mu_2 - (\delta_2)_j, \\
\log(p_{0tijk}/(1-p_{0tijk}-p_{1tijk})) &= \rho_0 + \rho_1 x_{tijk}, \\
\log(p_{1tijk}/(1-p_{0tijk}-p_{1tijk})) &= \psi_0 + \psi_1 x_{tijk},
\end{aligned}$$

where $(\beta_2)_1 = (\beta_2)_3 = (\beta_2)_4 = 0$, $(\delta_2)_1 = (\delta_2)_3 = (\delta_2)_4 = 0$. Therefore, for the mean and the dispersion parameters, we have only the effect of ethnicity, being the African-American group was different from the others and these mutually equivalent. For p_{0tijk} and p_{1tijk}, only "age" was significant.

Also considering the regression structure defined in (6.1), we fitted the ZOABR model, following the same steps as those for the ZOAB model. The final regression structure was:

$$\begin{aligned}
\log(\gamma_{tijk}/(1-\gamma_{tijk})) &= \mu_1 + (\beta_2)_j, \\
\log(\phi_{tijk}) &= -\mu_2 - (\delta_2)_j, \\
\log(p_{0tijk}/(1-p_{0tijk}-p_{1tijk})) &= \rho_0 + \rho_1 x_{tijk}, \\
\log(p_{1tijk}/(1-p_{0tijk}-p_{1tijk})) &= \psi_0 + \psi_1 x_{tijk},
\end{aligned}$$

where $(\beta_2)_1 = (\beta_2)_3 = 0$, $(\delta_2)_1 = (\delta_2)_3 = 0$, $j = 1, 2, 3, 4$. Therefore, for the mean and the dispersion parameters, only the covariable ethnicity

was significant, being the African-American and Taiwanese-American levels different from the others and these mutually equivalent. For p_{0tijk} and p_{1tijk}, only the covariable age was significant. We can see that the final regression structures for the two models were different, pointing out that different inferences can be drawn from these two models.

In Table 5, we present the criteria for model selection and the Bayesian p-value. Notice that all criteria, except BIC, in decimal places, indicate that the ZOABR model presented the best fit. When comparing the Bayesian p-values, we can say that the models present a similar and a good fitting. A possible explanation for this would be the fact that the posterior predictive checking methods are conservative (indicating that the model is well fitted when it is not).

Table 5: Criteria of model comparison and Bayesian p-value.

Criteria	ZOAB	ZOABR
AIC	500.24	487.41
BIC	535.25	535.55
EAIC	500.23	486.52
EBIC	517.73	508.40
DIC	1,468.75	1,418.72
LPML	-250.08	-244.03
Bayesian p-value	0.497	0.509

In Figure 1, we present the observed and predicted responses under the ZOAB and ZOABR models. In the histograms, the bar with the dot above represents the zeros and ones. From Figure 1(a) we can notice that the ZOABR model predicts better on the right tail when compared to the ZOAB model and as good as in the remaining of the distribution.

In Figures 2 and 3, we present some residual analysis for the ZOAB and ZOABR models, respectively. Concerning the ZOAB model, we observed in Figure 2(a) a different variability along the fitted values, indicating that only the model captured a part of the variability of the data. From Figure 2(d) we can notice that, although most of the residuals are within the confidence bands, some are outside or close to the limits, especially in the tails of the distribution. This behavior is, probably, due to the observations in the tails that were not well accommodated by the ZOAB model. Concerning the ZOABR model, from Figure 3(a), we can notice a behavior similar to that for the residuals related to the ZOAB model. However, from Figure 3(d) we can see that all residuals are well within the confidence bands and those

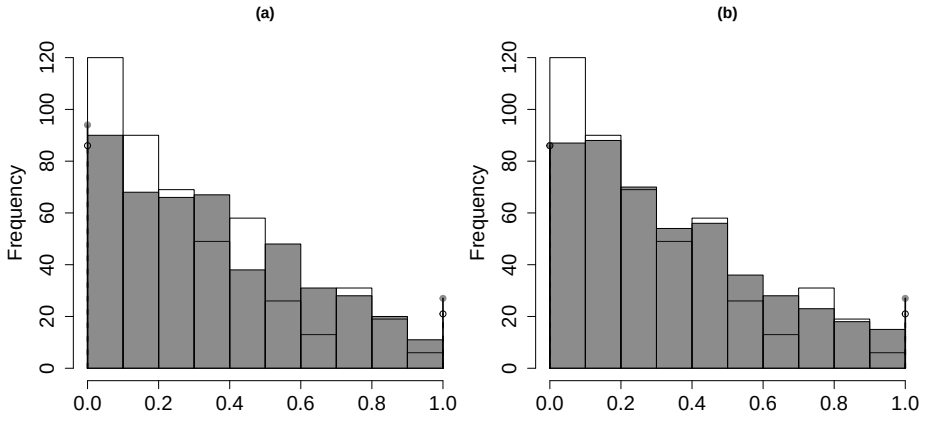

Figure 1: Histogram of the predicted distribution for the regression model: (a) ZOAB (b) ZOABR.

observations that highlighted in the ZOAB model, were well accommodated by the ZOABR model.

Furthermore, we analyzed the impact of transforming the extreme risks (zero and one) on the estimates, that is, replacing these values by 0.001 and 0.999, respectively, and fitting the beta and rectangular beta models. Figure 4 presents the simulated envelopes for the two models. We can notice that many of the residuals are out of the confidence bands and they present a systematic behavior. Then, we can conclude that the non augmented models are not suitable to analyze this data, even transforming the zero/one observations.

Also, depending on the model, some of the parameters are not significant (for the sake of simplicity, we did not present the results for the non augmented models). This is also an important aspect that illustrates how the use of non augmented models, in the transformed data, can lead to misleading inference.

Figure 5 indicates that potentially influential observations under the ZOAB model are not influential under the ZOABR model. Therefore, we can conclude that the ZOABR model is again more appropriate, under this criterion.

Table 6 presents the expectation a posteriori (EAP), the posterior standard deviation (PSD) and the respective 95% equi-tailed credibility intervals CI(95%) of the final models. We can see that the estimate of α indicates

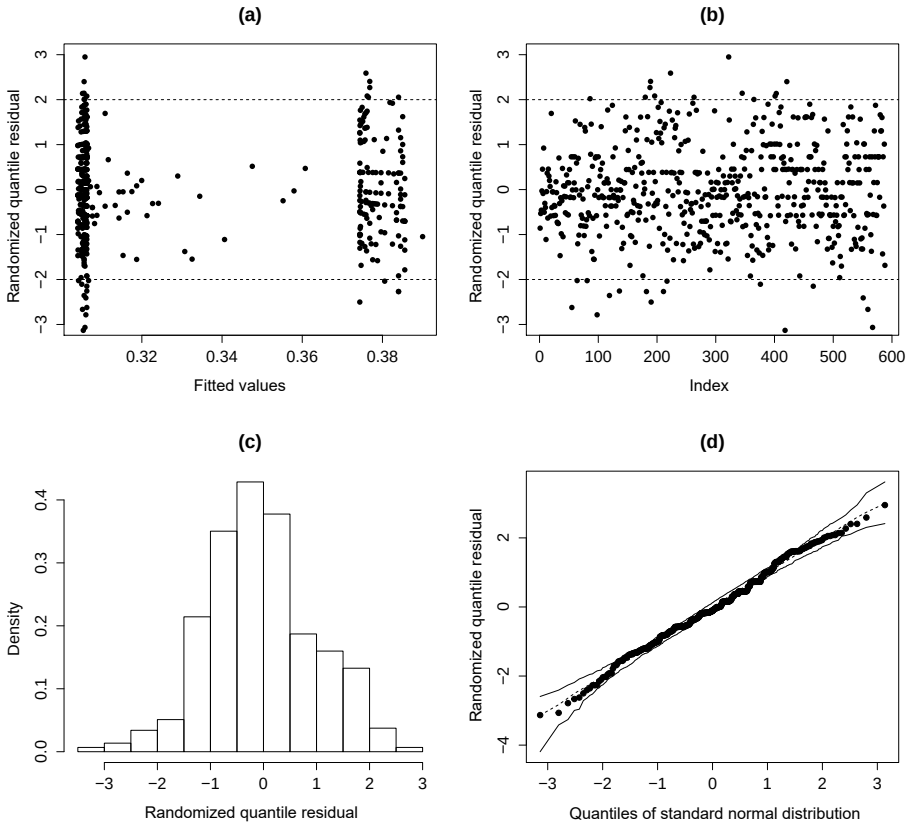

Figure 2: Residual plots for the ZOAB model.

that its true value is around 0.5 (see Figure 6), which, in its turn, suggest heavy tails for the conditional distribution of the response. According to the ZOABR model and the estimates of $((\beta_2)_2, (\beta_2)_4)^\top$ it is possible to conclude that the risk perceived is higher for the African-American and Taiwanese-American groups compared with the Caucasian and Mexican-American ones. Also, from the estimates of $((\delta_2)_2, (\delta_2)_4)^\top$ it is possible to conclude that the dispersion of the risk perceived is smaller for the African-American and Taiwanese-American groups compared with the Caucasian and Mexican-American groups. Finally, from the estimates of $(\rho_0, \rho_1)^\top$ and $(\psi_0, \psi_1)^\top$ we obtain that the proportion of the subjects that provide a null risk perceived is equals to 0.1454 (14.54%) whereas 0.0349 (3.49%) provide a maximum risk.

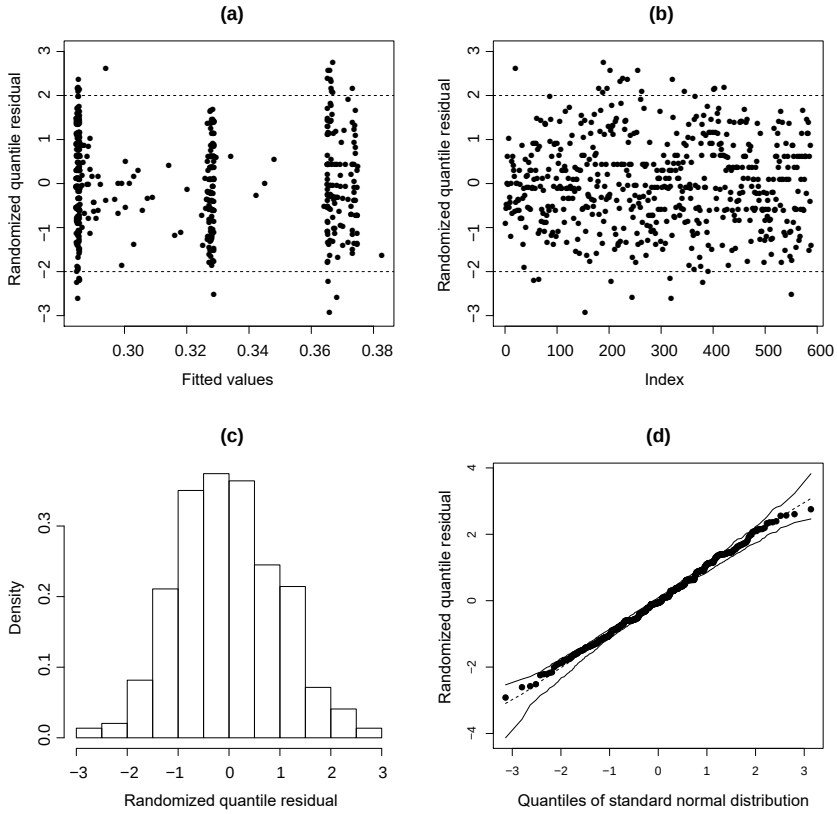

Figure 3: Residual plots for the ZOABR model.

7 Concluding remarks

In this paper we developed Bayesian inference for the ZOABR regression model. Parameter estimation, model fit assessment, model comparison, influence diagnostics and posterior predictive checking were developed through MCMC algorithms. The results from the simulation studies indicated that the estimation methods (including the maximum likelihood) recovery all parameters properly and that the Bayesian paradigm is equivalent, in terms of the accuracy of the estimates, to the ML method. However, the computational implementation of the MCMC algorithm and the influence diagnostics are easier than the EM algorithm based approach developed by Silva et al. (2017) and the related influence diagnostic analysis. Therefore, we can conclude that our approach is as good as that developed by Silva et al. (2017).

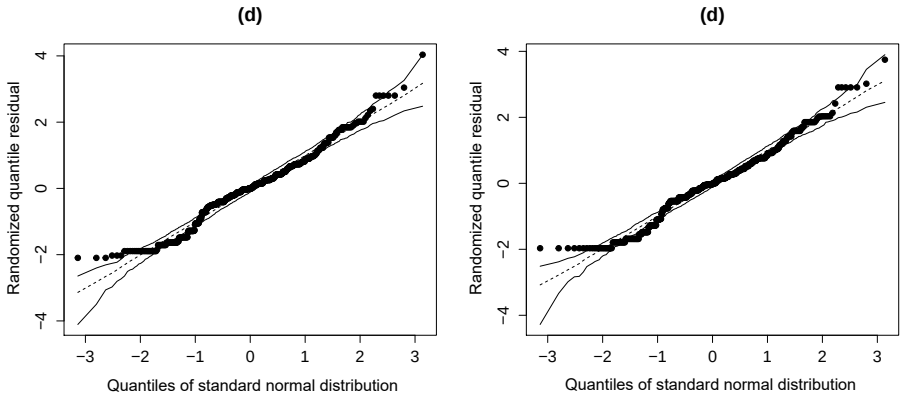

Figure 4: Residual plots for the (a) beta and (b) rectangular beta models.

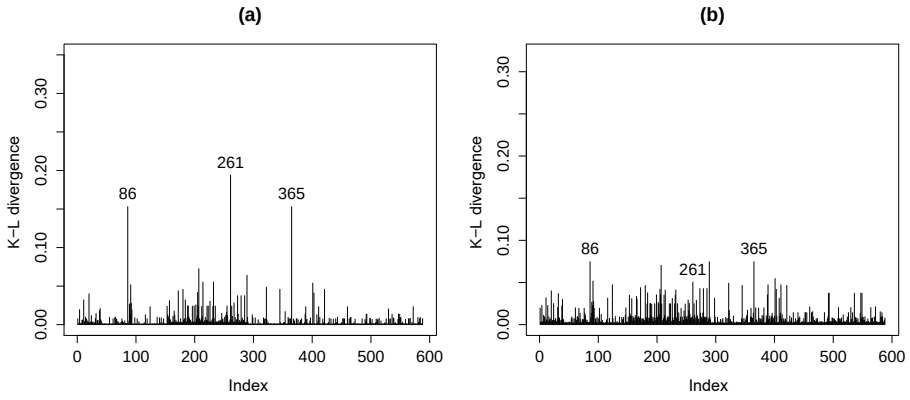

Figure 5: K-L divergence measure for the models: (a) ZOAB and (b) ZOABR.

Also, the Bayesian tools for model comparison and model fit assessment indicated that the ZOABR regression model fitted to the data well and better than the ZOAB model. In addition, it shows that misleading inference can be obtained, when the non-augmented models are used in the transformed data.

As future developments, we suggest the use of Jeffreys-rule prior and independence Jeffreys prior. Other auxiliary algorithms as the Hamiltonian Monte Carlo (see Homan and Gelman (2014) and Carpenter et al. (2016)), adaptive reject sampling and slice sampling (see Gamerman and

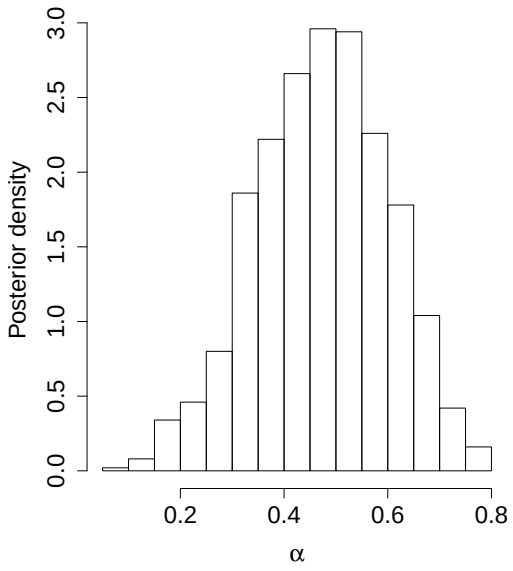

Figure 6: Histogram of the posterior density of parameter α.

Table 6: Bayesian estimates for the ZOAB and ZOABR final models.

	ZOAB			ZOABR		
Parameter	EAP	PSD	CI(95%)	EAP	PSD	CI(95%)
μ_1	-0.700	.053	[-0.801; -0.594]	-0.810	.069	[-0.933; -0.673]
$(\beta_2)_2$	.392	.123	[.154; .635]	.451	.155	[.122; .736]
$(\beta_2)_4$	-	-	-	.225	.113	[.016; .450]
μ_2	-1.112	.067	[-1.240; -0.983]	-1.746	.181	[-2.104; -1.401]
$(\delta_2)_2$	.315	.143	[.040; .608]	.706	.287	[.051; 1.218]
$(\delta_2)_4$	-	-	-	.562	.218	[.147; 1.003]
α	-	-	-	.475	.125	[.219; .705]
ρ_0	-2.413	.269	[-2.960; -1.890]	-2.409	.261	[-2.938; -1.916]
ρ_1	.023	.008	[.008; .038]	.023	.008	[.007; .038]
ψ_0	-4.547	.508	[-5.550; -3.600]	-4.541	.522	[-5.633; -3.591]
ψ_1	.043	.012	[.019; .065]	.043	.012	[.018; .067]

Lopes (2006)) could be used and compared. Also, other numerical methods to obtain approximation for the marginal posterior distributions, as the INLA algorithm, can be useful, see Rue and Martino (2009). Finally, Item Response Theory (IRT) models for continuous-limited responses, augmented

in zeros and ones, can be developed by using the results presented in this work.

Acknowledgments

We gratefully acknowledge São Paulo Research Foundation (FAPESP), for the financial support of this project, through a Master's scholarship, grant number 2013/07850-0, granted to the first author under the guidance of the second.

References

Ando, T. Bayesian predictive information criterion for the evaluation of hierarchical Bayesian and empirical Bayes models, *Biometrika*, 94, 2, 443–458, (2007).

Atkinson, A. C. *Plots, transformations, and regression: an introduction to graphical methods of diagnostic regression analysis*. Clarendon Press Oxford,(1985).

Bandyopadhyay, D., Galvis, D.M. and Lachos, V.H. Augmented mixed models for clustered proportion data, *JStatistical methods in medical research*, 26, 2, 880–897, (2017).

Bayarri, M. J. and Berger, J. O. P-values for composite null models, *Journal of the American Statistical Association*, 95, 452, 1127–1142, (2000).

Bayes, C. L., Bazán, J. L., and García, C. A new robust regression model for proportions, *Bayesian Analysis*, **7**, 4, 841–866, (2012).

Buckley, J. Estimation of models with beta-distributed dependent variables: A replication and extension of Paolino's Study. *Political Analysis*, 11, 2, 204–205, (2003).

Branscum, A. J., Johnson, W. O., Thurmond, M. C. Bayesian beta regression: Application to household data and genetic distance between foot-and-mouth disease viruses. *Australian & New Zealand Journal of Statistics*, 49, 3, 287–301, (2007).

Carlstrom, L., Woodward, J. and Palmer, C. Evaluating the Simplified Conjoint Expected Risk Model: Comparing the Use of Objective and Subjective Information, *Risk Analysis*, 20, 3, 385-392, (2000).

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M. A., Guo, J., Li, P. and Riddell, A. Stan: A probabilistic programming language. *Journal of Statistical Software (in press)*, (2016).

Cho, H., Ibrahim, J. G., Sinha, D. and Zhu, H. Bayesian case influence diagnostics for survival models, *Biometrics*, 65, 1, 116–124, (2009).

Cox, D. R. and Snell, E. J. A general definition of residuals, *Journal of the Royal Statistical Society. Series B (Methodological)*, 248–275, (1968).

Dunn, P. K. and Smyth, G. K. Randomized quantile residuals, *Journal of Computational and Graphical Statistics*, 5, 3, 236–244, (1996).

Cribari-Neto, F. and Zeileis, A. Beta regression in R, *J Stat Software*, 34, 1–24, (2010).

Ferrari, S. L. P. and Cribari-Neto, F. Beta Regression for Modelling Rates and Proportions, *Journal odf applied Statistics*, 31, 7, 799–815, (2004).

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

Galvis, D. M., Bandyopadhyay, D. and Lachos, V. H. Augmented mixed beta regression models for periodontal proportion data, *Statistics in Medicine*, 33, 21, 3759–3771, (2014).

Gamerman, D. and Lopes, H. *Stochastic simulation for bayesian inference, second edition*, Chapman & Hall/CRC, New York-NY, (2006).

Gelfand, A. E., Dey, D. K. and Chang, H. *Model determination using predictive distributions with implementation via sampling-based methods*, DTIC Document, (1992).

Gelman, A., Meng, Xiao-Li and Stern, H. Posterior predictive assessment of model fitness via realized discrepancies, *Statistica sinica*, 6, 4, 733–760, (1996).

Gelman, A., Carlin, J. B., Stern, H. S. and Rubin, D. B. *Bayesian data analysis*. 2ed. Taylor & Francis, (2003).

Gonçalves, F.B., Gamerman, D. and Soares, T.M. Simultaneous multifactor DIF analysis and detection in Item Response Theory, *Computational Statistics & Data Analysis*, 59, 144–160, (2013).

Homan, M. D. and Gelman, A. The no-U-turn sampler: Adaptively setting path lengths in Hamiltonian Monte Carlo, *The Journal of Machine Learning Research*, 15, 1, 1593–1623, (2014).

Ibrahim, J. G., Chen, M. H. and Sinha, D. *Bayesian Survival Analysis*, Springer, New York, (2001).

Ma, Z., Leijon, A. Bayesian Estimation of Beta Mixture Models with Variational Inference, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33, 11, 2160–2173, (2011).

Nogarotto, D. C., Azevedo, C. L. N., Bazán. *Bayesian estimation, residual analysis and prior sensitivity study for zero-one augmented beta regression model with an application to psychometric data*. Technical Report, `http://www.ime.unicamp.br/sites/default/files/rp14_2016.pdf`, University of Campinas, (2015).

Ospina, R. *Inflated Beta Regression Model. Doctoral's Thesis (In Portuguese)* IME- USP, (2008).

Ospina, R. and Ferrari, S.L.P. Inflated beta distributions. *Statistical Papers*, 51, 1, 111–126, (2010).

Ospina, R. and Ferrari, S.L.P. A general class of zero-or-one inflated beta regression models. *Computational Statistics & Data Analysis*, 56, 6, 1609–1623, (2012).

Paolino, P. Maximum likelihood estimation of models with beta-distributed dependent variables, *Political Analysis*, 9, 4, 325–346, (2001).

R Development Core Team. *R: A Language and Environment for Statis tical Computing. R Foundation for Statistical Computing, Vienna, Austria.* `http://www.R-project.org`, (2008).

Rubin, D. B. Bayesianly justifiable and relevant frequency calculations for the applies statistician, *The Annals of Statistics*, 12, 4, 1151–1172, (1984).

Rue, H. and Martino, S. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations, *Journal of the Royal Statistical Society B*, 71, 2, 319–392, (2009).

Silva, A. R. S., Azevedo, C. L. N., Bazán, J. L., Nobre, J. S. N. *Likelihood-based inference for zero-and/or-one augmented rectangular beta regression models*. Technical Report, `https://www.ime.unicamp.br/sites/`

default/files/pesquisa/relatorios/rp-2017-07.pdf, University of Campinas, (2017).

Sinharay, S. and Stern, H. S. Posterior predictive model checking in hierarchical models, *Journal of Statistical Planning and Inference*, 111, 1, 209–221, (2003).

Sinharay, S., Johnson, M. S. and Stern, H. S. Posterior predictive assessment of item response theory models, *Applied Psychological Measurement*, 30, 4, 298–321, (2006).

Smithson, M. and Verkuilen, J. A better lemon squeezer? Maximum-likelihood regression with beta-distributed dependent variables, *Psychological methods*, 11, 1, 54, (2006).

Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and Van Der Linde, A. Bayesian measures of model complexity and fit, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64, 4, 583–639, (2002).

Wang, J. and Luo, S. Augmented Beta rectangular regression models: A Bayesian perspective, *Biometrical Journal*, **58**, 1, 206–221, (2016).

Wang, J. and Luo, S. Bayesian multivariate augmented Beta rectangular regression models for patient-reported outcomes and survival data, *Statistical methods in medical research*, 1–20, (2015).