Revista Colombiana de Estadística - Theoretical Statistics
January 2026, volume 49, issue 1, pp. 33–63
http://doi.org/10.15446/rce.v49n1.118686

On Parametric Modal Beta Regression Model

Modelo de regresión Beta Modal

Erika Fernandes[a], Francisco Medeiros[b], Marcelo Bourguignon[c]

Department of Statistics, Federal University of Rio Grande do Norte, Natal, Brazil

Abstract

The beta regression model is part of a class of models applied to continuous responses restricted to the standard unit interval, such as rates and proportions. Ferrari & Cribari-Neto (2004) proposed the beta regression model incorporating covariates in the mean of the distribution through a link function. However, for studies in which the response variable presents asymmetry and/or discrepant values, this model may not be appropriate. A more convenient measure of central tendency in this situation is the mode of the distribution because of its robustness to outliers and easy interpretation in the presence of asymmetry. Zhou et al. (2020) proposed a parameterization for the beta distribution in terms of the mode and a precision parameter and presented a modal regression model robust to outliers. In this work, we present a more complete study of the modal beta regression properties and performance and a comparison between this model and the usual beta regression model. We perform Monte Carlo simulation studies to evaluate the maximum likelihood estimators under different scenarios of asymmetry and sensitivity to outliers when some patterns of disturbance are imposed. Furthermore, we propose and evaluate three residuals for this class of models. The numerical results suggest that the modal regression model presents a good performance on symmetrical and asymmetrical data and in most scenarios, it performs better in the presence of outliers than the usual beta regression model. Finally, we present and discuss two empirical applications and a comparative analysis of the mean and modal beta regression models.

***Keywords*:** Beta regression; Maximum likelihood; Mode; Parametric modal regression.

[a]Master's degree. E-mail: erika-rayanne@hotmail.com
[b]Ph.D. E-mail: francisco.medeiros@ufrn.br
[c]Ph.D. E-mail: m.p.bourguignon@gmail.com

Resumen

El modelo de regresión beta es parte de una clase de modelos aplicados a respuestas continuas restringidas al intervalo unitario estándar, como tasas y proporciones. Ferrari & Cribari-Neto (2004) propuso el modelo de regresión beta incorporando covariables en la media de la distribución a través de una función de enlace. Sin embargo, para estudios en los que la variable respuesta presenta asimetría y/o valores discrepantes, este modelo puede no ser apropiado. Una medida de tendencia central más conveniente en esta situación es la moda de la distribución debido a que es robusta para valores atípicos y también a su fácil interpretación en presencia de asimetría. Zhou et al. (2020) propuso una parametrización para la distribución beta en términos de la moda y un parámetro de precisión y presentó un modelo de regresión modal robusto a valores atípicos. En este trabajo, presentamos un estudio más completo de las propiedades y el desempeño de la regresión beta modal y una comparación entre este modelo y el modelo de regresión beta usual. Fueron realizados estudios de simulación de Monte Carlo para evaluar los estimadores de máxima verosimilitud en diferentes escenarios de asimetría y sensibilidad a valores atípicos cuando se imponen algunos patrones de perturbación. Además, proponemos y evaluamos tres residuos para esta clase de modelos. Los resultados numéricos sugieren que el modelo de regresión beta modal presenta un buen desempeño en datos simétricos y asimétricos y, en la mayoría de los escenarios, se desempeña mejor en presencia de valores atípicos que el modelo de regresión beta habitual. Finalmente, fueron presentadas y discutidas dos aplicaciones empíricas y un análisis comparativo de los modelos de regresión beta media y modal.

***Palabras clave*:** Moda; Máxima verosimilitud; Regresión beta; Regresión modal paramétrica.

1. Introduction

There are several studies whose variables of interest have values in the unit interval (0, 1), such as rates and/or proportions. In the literature, there are some probability distributions to model this type of data. For instance, the beta distribution is one of the most used distributions due to its simplicity compared to some other models and the diversity of shapes assumed for different parameter values (Ghitany et al., 2019).

In practice, there are situations where the interest is to evaluate the behavior of a specific variable in relation to other variables. To study the relationship between these variables, we can use regression models. In this context, Kieschnick & McCullough (2003) proposed a regression model for proportion data assuming a beta distribution for the response variable. Nevertheless, this model uses a fixed link function (logit) to introduce covariates in the parameter of interest, which makes the model more restricted. Later, Ferrari & Cribari-Neto (2004) proposed a regression model that assumes a beta distribution for the response variable and allows for the use of a more generic link function. In addition, the authors parameterized the distribution in terms of a mean and a dispersion parameter, including covariates on the mean of the distribution. In this work, we define this model as the mean (or usual) beta regression model. Furthermore, beta regression models with joint modeling of the mean and the dispersion parameter were previously proposed by Cepeda (2001) under a Bayesian approach, using linear predictors of the form logit and log to model the parameters μ and ϕ, respectively. This framework was extended by Cepeda & Gamerman (2005) within the biparametric family of distributions. Additionally, Cepeda-Cuervo (2015) introduced beta regression models that jointly model the mean and the variance, providing greater flexibility for handling proportional data with heteroscedasticity. Several published works have studied this model, for example, Espinheira et al. (2008*a*) proposed diagnostic tools for the model influence, Espinheira et al. (2008*b*) presented two residuals for this class of models, Ospina & Ferrari (2010) developed a zero-and-one inflated beta regression model, and Espinheira et al. (2017) proposed a new residual to be used in linear and non-linear beta regression models. Bourguignon & Gallardo (2025), extend the usual mean beta regression model using a general and unified parameterization of this distribution that is indexed by some central tendency measure, such as median, mode, arithmetic mean, geometric mean or harmonic mean, and a concentration parameter.

However, there are data in the unit interval that present asymmetry and/or outliers. In these situations, the mean is not the most suitable measure of central tendency. Thus, the beta regression model proposed by Ferrari & Cribari-Neto (2004) may not be adequate for data with these characteristics. As an alternative, Bayes et al. (2012) proposed a regression model in which the response variable assumes a rectangular beta distribution. Despite the great performance, the rectangular beta model includes an additional parameter compared to the mean beta model. An alternative approach in such contexts is to model the conditional mode instead of the conditional mean. The mode is a central tendency measure of easy interpretation, robust to outliers, more representative in asymmetric data Oelker et al. (2015), and has also been used in regression models.

The regression models that consider the mode modeling of the response variable conditional on a set of regressors are known in the literature as modal regression models. The modal regression idea was initially proposed by Sager & Thisted (1982), who used the zero-one loss function for the maximum likelihood estimation of isotonic mode regression. They minimized a weighted loss with its weight related to a nonparametric likelihood. Lee (1989) and Lee (1993) developed some pioneering studies on modal regression. The author introduced semiparametric estimators for the conditional mode based on the minimization of the associated loss function. Yao & Li (2014) developed a modal linear regression model to explore high-dimensional data. In this model, the authors consider a linear link function to model the conditional mode of the response variable given a set of covariates. Through studies with real and simulated data, they concluded that the proposed model provides smaller prediction intervals than those obtained through linear regression models for the mean and median. In the parametric context, modal regression has been little explored. Aristodemou (2014) proposed a parametric modal regression model based on the gamma distribution, in which a new parameterization in terms of the mode was considered and a constant precision assumed. Bourguignon et al. (2020) proposed an extension of the model proposed by

Aristodemou (2014) considering a new parameterization for the gamma distribution and enabling the assumption of non-constant precision through the incorporation of a regression structure in the precision parameter.

According to Kemp & Santos Silva (2012), interest in modal regression is much broader, and the mode is indisputably the most intuitive measure of central tendency, being especially useful in modeling right-skewed data commonly found in many econometric applications (e.g., wages, prices, and expenditures). Modal regression models have been applied across various fields of study. For instance, Cao et al. (2023) employed modal regression models with skew-normal distributions to analyze body mass index data. Zhou et al. (2020) used modal regression models in neuroscience data, while Ho et al. (2017) fitted modal regression models to econometric datasets.

The beta distribution does not have a closed expression for the median for arbitrary values of its parameters. In this context, it is more intuitive to model data that present asymmetry and/or outliers assuming the beta distribution for the response variable and modeling the mode, since this measure has a closed-form expression in this distribution. Then, Zhou et al. (2020) established a modal regression model assuming the beta distribution using a general link function. For this, they proposed a new parameterization for the beta distribution in terms of a mode and a precision parameter. However, the authors did not study the inferential aspects of this model, which motivates us to write this paper. In this work, we perform a more in-depth study of the modal beta model and explore the following aspects: (1) obtaining the Fisher information matrix; (2) conducting a more complete study on the performance of the model parameters estimates in symmetrical and asymmetrical cases; (3) analyzing the coverage probability of the approximated confidence intervals; (4) evaluating and comparing, by simulation, the sensitivity of the modal and mean beta regression models to outliers; (5) proposing three different residuals and their numerical evaluation; and (6) applying the modal and mean beta regression models to two sets of real data.

This paper is organized as follows. Section 2 presents some information about the beta distribution and the others parameterizations studied in this work, such as the density probability function, properties, and some inferential aspects. Section 3 brings the definition of the modal beta regression model, some inferential aspects of the model parameters, and the proposition of three residuals for this class of models. Section 4 presents the simulation studies conducted to evaluate the maximum likelihood estimators, coverage probability, residuals, and the sensitivity of the mean and modal beta regression models to outliers. In Section 5, two applications to real data are presented and discussed. In this section, the fitted mean and modal beta regression models are compared. In addition, the adequacy of both models to the datasets is evaluated by analyzing the residuals. Concluding remarks and future works are presented in Section 6.

2. Modal Beta Distribution

A random variable Y follows a beta distribution with shape parameters $\alpha > 0$ and $\beta > 0$, denoted by $Y \sim \text{Beta}(\alpha, \beta)$, if its probability density function (pdf) is given by

$$f(y; \alpha, \beta) = \frac{1}{B(\alpha, \beta)} y^{\alpha-1}(1-y)^{\beta-1}, \qquad 0 < y < 1, \tag{1}$$

where $B(\alpha, \beta) = \Gamma(\alpha)\Gamma(\beta)/\Gamma(\alpha+\beta)$ is the beta function and $\Gamma(\alpha) = \int_0^\infty \omega^{\alpha-1}\mathrm{e}^{-\omega}\mathrm{d}\omega$ is the gamma function.

The mean and variance of Y are given, respectively by

$$\mathrm{E}(Y) = \frac{\alpha}{\alpha+\beta}, \quad \mathrm{Var}(Y) = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}. \tag{2}$$

This distribution has no general closed-form expression for the quantile function and consequently has no closed-form expression for the median for arbitrary values of α and β. However, besides the mean, the beta distribution has a closed-form expression for the mode, which is given by

$$\text{Mode}(Y) = \frac{\alpha - 1}{\alpha + \beta - 2}, \quad \alpha > 1 \quad \text{and} \quad \beta > 1.$$

When $\alpha = \beta = 1$, the Beta distribution corresponds to the $Uniform(0, 1)$ distribution, characterized by a constant density function and consequently no well-defined mode. In the case where $\alpha < 1$ and $\beta < 1$, the distribution exhibits a bimodal shape with modes located near the boundaries 0 and 1. Furthermore, for $\alpha < 1$ and $\beta \geq 1$, the mode is degenerate at $m = 0$, while for $\beta < 1$ and $\alpha \geq 1$, the mode is degenerate at $m = 1$.

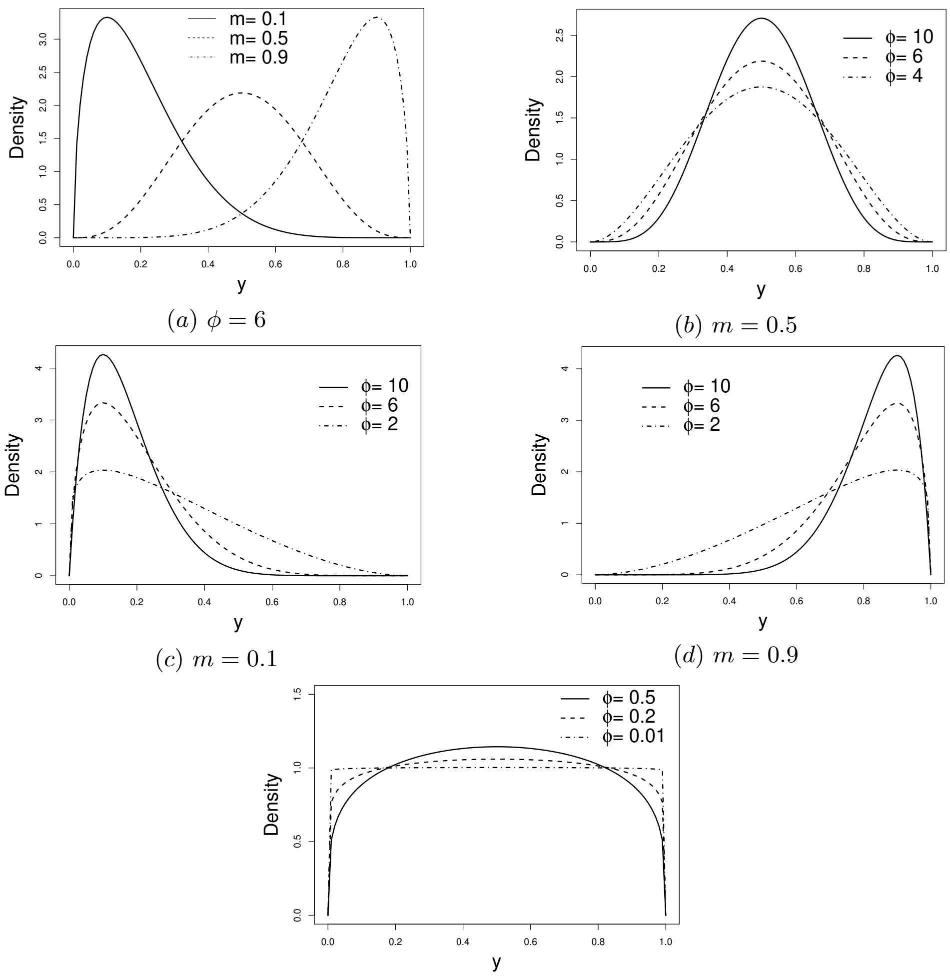
In this work, we consider a different parameterization for the beta distribution proposed by Zhou et al. (2020). For α and $\beta > 1$, we let $m = (\alpha - 1)/(\alpha + \beta - 2)$ and $\phi = \alpha + \beta - 2$, i.e., $\alpha = m\phi + 1$ and $\beta = (1 - m)\phi + 1$. Thus, it follows that

$$\text{Mode}(Y) = m, \quad \mathrm{E}(Y) = \frac{m\phi + 1}{\phi + 2}, \quad \text{and} \quad \mathrm{Var}(Y) = \frac{[m\phi + 1][(1 - m)\phi + 1]}{(\phi + 2)^2(\phi + 3)}.$$

Then, the beta density in (1) can also be rewritten as

$$f(y; m, \phi) = \frac{y^{m\phi}(1-y)^{(1-m)\phi}}{B(m\phi + 1, (1 - m)\phi + 1)}, \quad 0 < y < 1, \tag{3}$$

where $0 < m < 1$ is the mode and $\phi > 0$ is the precision parameter. If a random variable has pdf given in (3), we use the notation $Y \sim \text{MB}(m, \phi)$. From now on, we call this distribution as modal beta distribution. Figure 1 shows the pdf of $Y \sim \text{MB}(m, \phi)$ for different combinations of parameter values. Figure 1(a) illustrates that, for a fixed value of ϕ, the probability density function exhibits different patterns of skewness depending on the value of m, ranging from left-skewed to right-skewed, or adopting an approximately symmetric profile. Figures 1(b), 1(c), and 1(d) indicate that, as ϕ increases, the MB distribution's pdf becomes more concentrated around the mode, whereas smaller ϕ values result in more dispersed pdfs. Figure 1(e) illustrates that, as ϕ approaches zero, the MB distribution's pdf converges to a uniform distribution.



$(a)\ \phi = 6$ $(b)\ m = 0.5$

$(c)\ m = 0.1$ $(d)\ m = 0.9$

$(c)\ m = 0.5$

FIGURE 1: Density functions under different parameter settings

The skewness and kurtosis coefficients of $Y \sim \mathrm{MB}(m, \phi)$ are respectively given by

$$\delta = \frac{2\phi(1-2m)\sqrt{\phi+3}}{(\phi+4)\sqrt{[m\phi+1][(1-m)\phi+1]}} \tag{4}$$

and

$$\kappa = \frac{6\left\{\phi^2(2m-1)^2(\phi+3) - [m\phi+1][(1-m)\phi+1](\phi+4)\right\}}{[m\phi+1][(1-m)\phi+1](\phi+4)(\phi+5)}.$$

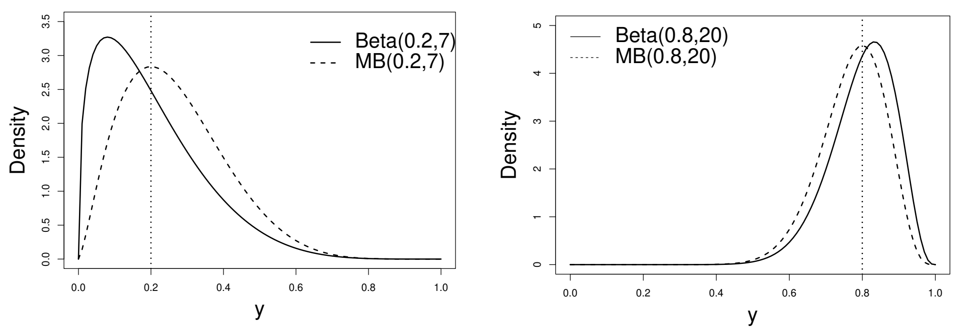
The factor $(1-2m)$ controls the sign in (4), in other words, we confirm what was seen at Figure 1, if $m = 0.5$, $m > 0.5$, and $m < 0.5$, the distribution is symmetric, left-skewed, and right-skewed, respectively.

The modal beta density function, Equation (3), can be rewritten in the following form:

$$f(y; m, \phi) = \exp\{m\phi \log(y/(1-y)) + \phi \log(1-y) - \log B(m\phi + 1, (1-m)\phi + 1)\},$$

that is, the modal beta distribution belongs to the two-parameter canonical exponential family ($k = 2$) with $\tau(m, \phi) = (\tau_1, \tau_2) = (m\phi, \phi)$, $T = (T_1, T_2) = (\log(y/(1-y)), \log(1-y))$, $d(m, \phi) = \log(B(m\phi+1, (1-m)\phi+1))$, $h(y) = 1$, and $\mathcal{A} = (0, 1)$.

Figure 2 shows a comparison between the mean and modal beta distributions. For this, the same values were chosen for the first and second parameters of both distributions under positive and negative asymmetry. The modal beta distribution's first parameter is more informative than that of the mean beta distribution in this scenario because it represents the point of higher mass concentration of the distribution.



(a) Beta(μ, ϕ) (b) MB(m, ϕ)

FIGURE 2: Densities of Beta(μ, ϕ) and MB(m, ϕ).

3. Modal Beta Regression Model: Inference and Diagnostics

Let $Y_1, \ldots, Y_n$ be independent random variables with $Y_i \sim \text{MB}(m_i, \phi)$, where $\text{Mode}(Y_i) = m_i, i = 1, \ldots, n$. We introduce a regression structure in the mode satisfying the following functional relation:

$$g(m_i) = \eta_i = \mathbf{x}_i^\top \boldsymbol{\beta}, \quad i = 1, \ldots, n, \tag{5}$$

where $g : (0, 1) \to \mathbb{R}$ is a link function strictly monotone and twice differentiable, $\boldsymbol{x}_i^\top = (x_{i1}, \ldots, x_{ip})$ is the covariate vector associated with the i-th response, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\top$ is the p-vector ($p < n$) of unknown parameters associated with the covariates, and η_i is the linear predictor. We assume the design matrix $\boldsymbol{X} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)^\top$ with dimension $n \times p$ has full rank ($\text{rank}(\boldsymbol{X}) = p$). Besides this, the precision parameter ϕ is unknown and constant for all observations.

The $g(\cdot)$ function defined in (5) enables the use of several functions, making the model more flexible. If we consider that the link function is logit, probit, or complement log-log, we obtain, respectively, $g(m_i) = \log[m_i/(1-m_i)]$, $g(m_i) = \Phi^{-1}(m_i)$, and $g(m_i) = \log\{-\log(1-m_i)\}$, where $\Phi(\cdot)$ is the cumulative distribution function of a standard normal random variable. Based on a sample of n independent observations, the log-likelihood function of the regression parameter vector $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \phi)^\top$ has the form

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^{n} \ell(m_i, \phi), \tag{6}$$

where

$$\ell(m_i, \phi) = m_i\phi\log(y_i) + (1-m_i)\phi\log(1-y_i) - \log[B(m_i\phi + 1, (1-m_i)\phi + 1)].$$

The score vector, obtained by differentiating the log-likelihood function given in (6) with respect to each element of $\boldsymbol{\theta}$, is given by $\mathbf{U}(\boldsymbol{\theta}) = (\mathbf{U}_{\boldsymbol{\beta}}, \mathrm{U}_\phi)^\top$, with

$$\mathbf{U}_{\boldsymbol{\beta}} = \phi\mathbf{X}^\top\mathbf{D}(\mathbf{y}^* - \mathbf{m}^*)$$

and

$$U_\phi = \sum_{i=1}^{n}\{m_i(y_i^* - m_i^*) + \log(1-y_i) - \Psi^{(0)}((1-m_i)\phi + 1) + \Psi^{(0)}(\phi + 2)\},$$

where $\boldsymbol{D} = \text{diag}\{g'(m_1)^{-1}, \ldots, g'(m_n)^{-1}\}$, $\boldsymbol{y^*} = (y_1^*, \ldots, y_n^*)^\top$, $\boldsymbol{m^*} = \{m_1^*, \ldots, m_n^*\}$, with $y_i^* = \log\left(\frac{y_i}{1-y_i}\right)$, $m_i^* = \Psi^{(0)}(m_i\phi + 1) - \Psi^{(0)}((1-m_i)\phi + 1)$, and $\Psi^{(0)}$ is the digamma function.

The maximum likelihood (ML) estimators $\widehat{\boldsymbol{\beta}} = (\widehat{\beta_1}, \ldots, \widehat{\beta_p})^\top$ and $\widehat{\phi}$ of $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\top$ and ϕ, respectively, can be obtained by solving simultaneously the nonlinear system of equations $\boldsymbol{U_\beta} = \mathbf{0}_{p\times 1}$ and $U_\phi = 0$. However, no closed-form expressions for the ML estimators are obtained. Therefore, we must use an iterative method for nonlinear optimization.

The Fisher information matrix for the modal beta regression model can be expressed in matrix form as

$$\mathbf{K}(\boldsymbol{\theta}) = \begin{bmatrix} \mathbf{K}_{\boldsymbol{\beta\beta}} & \mathbf{K}_{\boldsymbol{\beta}\phi} \\ \mathbf{K}_{\boldsymbol{\beta}\phi}^\top & \mathrm{K}_{\phi\phi} \end{bmatrix} = \begin{bmatrix} \phi\mathbf{X}^\top\boldsymbol{W}\boldsymbol{X} & \mathbf{X}^\top\boldsymbol{D}\boldsymbol{c} \\ \boldsymbol{c}^\top\boldsymbol{D}\boldsymbol{X} & \text{tr}(\mathbf{V}) \end{bmatrix},$$

where $\mathbf{W} = \text{diag}\{w_1, \ldots, w_n\}$, $\boldsymbol{c} = (c_1, \ldots, c_n)^\top$, $\mathbf{V} = \text{diag}\{v_1, \ldots, v_n\}$, with $w_i = \phi m_i^{**} g'(m_i)^{-2}$, $c_i = \phi\{m_i m_i^{**} - \Psi^{(1)}((1-m_i)\phi + 1)\}$, $v_i = m_i^2 m_i^{**} - m_i\Psi^{(1)}((1-m_i)\phi + 1) + (1-m_i)^2\Psi^{(1)}((1-m_i)\phi + 1) - \Psi^{(1)}(\phi + 2)$, $m_i^{**} = \Psi^{(1)}(m_i\phi + 1) + \Psi^{(1)}((1-m_i)\phi + 1)$, and $\Psi^{(1)}(x) = \frac{d\Psi^{(0)}(x)}{dx}$. Note that the parameters $\boldsymbol{\beta}$ and ϕ

are not orthogonal because the Fisher information matrix is not a block diagonal matrix.

When n is large and under conditions that are fulfilled for parameters in the interior of the parameter space but not on the boundary, Cox & Hinkley (1983), we have that

$$\begin{pmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\phi} \end{pmatrix} \overset{a}{\sim} N_{p+1}\left(\begin{pmatrix} \boldsymbol{\beta} \\ \phi \end{pmatrix}, \boldsymbol{K}(\boldsymbol{\beta}, \phi)^{-1}\right),$$

where $\overset{a}{\sim}$ means "approximately distributed" and $\mathbf{K}(\boldsymbol{\beta}, \phi)^{-1}$ is given by

$$\mathbf{K}(\boldsymbol{\theta})^{-1} = \begin{bmatrix} \mathbf{K}^{\boldsymbol{\beta}\boldsymbol{\beta}} & \mathbf{K}^{\boldsymbol{\beta}\phi} \\ \mathbf{K}^{\phi\boldsymbol{\beta}} & \mathrm{K}^{\phi\phi} \end{bmatrix}.$$

By defining a confidence level of $100(1-\alpha)\%$, with $\alpha \in (0,1)$, for $\boldsymbol{\theta}$, the approximate confidence intervals β_r, $r = 1, \ldots, p$ and ϕ are given, respectively, by

$$\left(\widehat{\beta}_r - z_{1-\alpha/2}\sqrt{\mathrm{K}(\widehat{\boldsymbol{\theta}})^{\boldsymbol{\beta}\boldsymbol{\beta}}_{rr}}, \widehat{\beta}_r + z_{1-\alpha/2}\sqrt{\mathrm{K}(\widehat{\boldsymbol{\theta}})^{\boldsymbol{\beta}\boldsymbol{\beta}}_{rr}}\right)$$

and

$$\left(\widehat{\phi} - z_{1-\alpha/2}\sqrt{\mathrm{K}(\widehat{\boldsymbol{\theta}})^{\phi\phi}}, \widehat{\phi} + z_{1-\alpha/2}\sqrt{\mathrm{K}(\widehat{\boldsymbol{\theta}})^{\phi\phi}}\right),$$

where $\mathrm{K}(\widehat{\boldsymbol{\theta}})^{\boldsymbol{\beta}\boldsymbol{\beta}}_{rr}$ and $\mathrm{K}(\widehat{\boldsymbol{\theta}})^{\phi\phi}$ are, respectively, the asymptotic variances of $\widehat{\beta}_r$ and $\widehat{\phi}$, and $\mathrm{K}(\widehat{\boldsymbol{\theta}})^{\boldsymbol{\beta}\boldsymbol{\beta}}_{rr}$ represents the r-th main diagonal element of matrix $\mathrm{K}(\widehat{\boldsymbol{\theta}})^{\boldsymbol{\beta}\boldsymbol{\beta}}$. Besides that, $z_{1-\alpha/2}$ denotes the quantile $1-\alpha/2$ from the standard normal distribution.

After fitting the model, we must perform a diagnostic analysis to verify the assumptions of the model, such as outliers and/or model misspecification. The main diagnostic technique used in this context is the analysis of residuals. An important general formulation to residuals is defined in Cox & Snell (1968). The randomized quantile residuals were initially proposed by Dunn & Smyth (1996) and have been used so far. These residuals do not require specific distribution for the response variable and follow a standard normal distribution since the specified distribution to the model is correct. The randomized quantile residual proposed to this class of models is defined by

$$r_i^q = \Phi^{-1}(F(y_i; \widehat{m}_i, \widehat{\phi})),$$

where $\Phi(\cdot)$ is the cumulative distribution function of the standard normal distribution, $F(y; m, \phi)$ is the cumulative distribution function of the modal beta distribution, and $\widehat{m}$ and $\widehat{\phi}$ are the ML estimators of m and ϕ, respectively.

Ospina (2007) proposed a standardized weighted residual to the class of usual beta regression models. This residual is based on the difference $y_i^* - \widehat{\mu}_i^*$, where y_i^* is the logit of the observation y_i and $\widehat{\mu}_i^*$ is the ML estimation of $\mathrm{E}(Y_i^*)$, and also from the fact that the beta distribution belongs to the two-parameter canonical exponential family. It follows that

$$\mathrm{E}(T_1) = \mathrm{E}(Y^*) = \frac{\partial d(\boldsymbol{\theta})}{\partial \tau_1} = \Psi^{(0)}(m_i\phi + 1) - \Psi^{(0)}((1-m_i)\phi + 1) = m_i^*$$

and

$$\text{Var}(T_1) = \text{Var}(Y^*) = \frac{\partial^2 d(\boldsymbol{\theta})}{\partial \tau_1^2} = \frac{\partial m_i^*}{\partial \tau_1} = \Psi^{(1)}(m_i\phi + 1) + \Psi^{(1)}((1 - m_i)\phi + 1) = m_i^{**}.$$

Thus, we can define the following residual to the modal beta regression model:

$$r_i^p = \frac{y_i^* - \widehat{m}_i^*}{\sqrt{\widehat{m}_i^{**}}},$$

which consists of standardization of y_i^*, subtracting its expected value and dividing by its standard deviation, where $\text{E}(r_i^p) = 0$ and $\text{Var}(r_i^p) = 1$

Another residual widely used is the Cox-Snell residual proposed by Cox & Snell (1968). This residual follows an exponential distribution with a parameter equal one (standard exponential) since the model is correctly fitted. The Cox-Snell residual to this class of models is given by

$$r_i^c = -\log(1 - F(y_i; \widehat{m}_i, \widehat{\phi})),$$

where $F(y_i; m_i, \phi)$ is the cumulative distribution function of the modal beta distribution and $\widehat{m}$ and $\widehat{\phi}$ are the ML estimators of m and ϕ, respectively.

To evaluate the adequacy of the fitted model, simulated envelopes were incorporated into the QQ-plots, generated via Monte Carlo simulations as suggested by Atkinson (1985). Residuals (r^q, r^p, and r^c) were computed based on the fitted model, and three confidence bands were constructed from these values: the outermost band, defined by the minimum and maximum residuals; the intermediate band, derived from the 0.025 and 0.975 percentiles, representing a 95% confidence level; and the innermost band, derived from the 0.005 and 0.995 percentiles, representing a 99% confidence level. Under a well-specified model, residuals are expected to fall predominantly within the outermost band, while the inner bands provide a means to assess the degree of concordance between observed and theoretical quantiles, offering a rigorous visual tool for model diagnostic evaluation.

4. Simulation Study

In this section, we present a Monte Carlo (MC) simulation study to investigate and evaluate the model parameters estimates, the coverage probability of asymptotic confidence intervals, sensitivity and residuals of the model. All simulations were performed using R (R Core Team, 2020), and maximizations of the log-likelihood function for the model parameters were performed considering the BFGS method through the `optim` function.

4.1. Estimates and Coverage Probability

To evaluate the performance of the ML estimators of the proposed model, we generated n observations of a random variable $Y_i \sim \text{MB}(m_i, \phi)$ and applied the modal beta model with structure given by

$$g(m_i) = \eta_i = \log\left(\frac{m_i}{1-m_i}\right) = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_p x_{ip}, \quad i = 1, \ldots, n,$$

where the values of $x_{i1}, \ldots, x_{ip}$ are taken as random draws from a uniform distribution $U(0,1)$ and all regression parameters were chosen such that scenarios with left and right asymmetry are obtained, and $\phi = 50$. We considered different values for the number of regression parameters (p) and sample size. The number of Monte Carlo replicates was 10 000 and all parameters were estimated by the ML method.

For each n, we obtained the following quantities: the mean of the r estimates, bias, standard deviation, the square root of the mean square error, the coefficient of variation, and the asymptotic standard error, which are given, respectively, by

$$\widehat{\lambda}_{mc} = \frac{1}{r}\sum_{i=1}^{r}\widehat{\lambda}_i, \qquad \mathrm{b}(\widehat{\lambda}_{mc}) = \widehat{\lambda}_{mc} - \lambda,$$

$$\mathrm{sd}(\widehat{\lambda}_{mc}) = \sqrt{\frac{1}{r-1}\sum_{i=1}^{r}(\widehat{\lambda}_i - \widehat{\lambda}_{mc})^2}, \qquad \sqrt{\mathrm{mse}(\widehat{\lambda}_{mc})} = \sqrt{\mathrm{sd}^2(\widehat{\lambda}_{mc}) + \mathrm{b}^2(\widehat{\lambda}_{mc})},$$

$$\mathrm{cv}(\widehat{\lambda}_{mc}) = \frac{\mathrm{sd}(\widehat{\lambda}_{mc})}{|\widehat{\lambda}_{mc}|} \times 100, \qquad \mathrm{SE}_{as} = \sqrt{K(\widehat{\lambda}_{mc})^{\lambda\lambda}}.$$

where λ is the true value of the parameter vector, $\widehat{\lambda}_i$ is the ML estimate of the i-th Monte Carlo replicate, and $\widehat{\lambda}_{mc}$ is the Monte Carlo estimate (mean of the r estimates).

Tables 1, 2, and 3 show the estimates for the parameters of the modal beta regression model under left asymmetry, symmetry, and right asymmetry, respectively. We observe that the estimates for the square root of the mean square error, standard deviation, asymptotic standard error, and the coefficient of variation of $\boldsymbol{\beta}$ decrease as the sample size increases for all scenarios. In most scenarios, the bias for these estimators decreases as the sample size increases. We also note that $\hat{\beta}_1$ tends to underestimate the true value of the parameter when the data are left-skewed, and overestimate when the data are right-skewed. Similarly, for $\widehat{\phi}$, all the quantities decrease as the sample size increases. Additionally, $\widehat{\phi}$ tends to overestimate the true value of the parameter regardless of the scenario. We can see that adding more covariates to the model slightly decreases the accuracy of the estimates. The last column of Tables 1, 2, and 3 presents, in percentage terms, the absolute value of the coefficient of variation (cv) estimates, which represents the proportion of the estimate explained by variability.

Figure 3 shows the boxplots of the ML estimates for the model parameters considering $\phi = 50$. Making an analogy of Figure 3 with a matrix, the lines represent the parameters β_0, β_1, β_2, β_3, and ϕ, respectively, and the columns the type of asymmetry considered, being asymmetry on the left, symmetry, and asymmetry on the right, respectively. For all scenarios, we observe that as the sample size increases, the bias of the estimators decreases, as expected. In addition, in all scenarios for $n = 50$ and $n = 100$, $\hat{\phi}$ tends to overestimate, in median, the true value of the parameter. However, as expected, as the sample size increases, the estimates converge in median to the true value of the parameter.

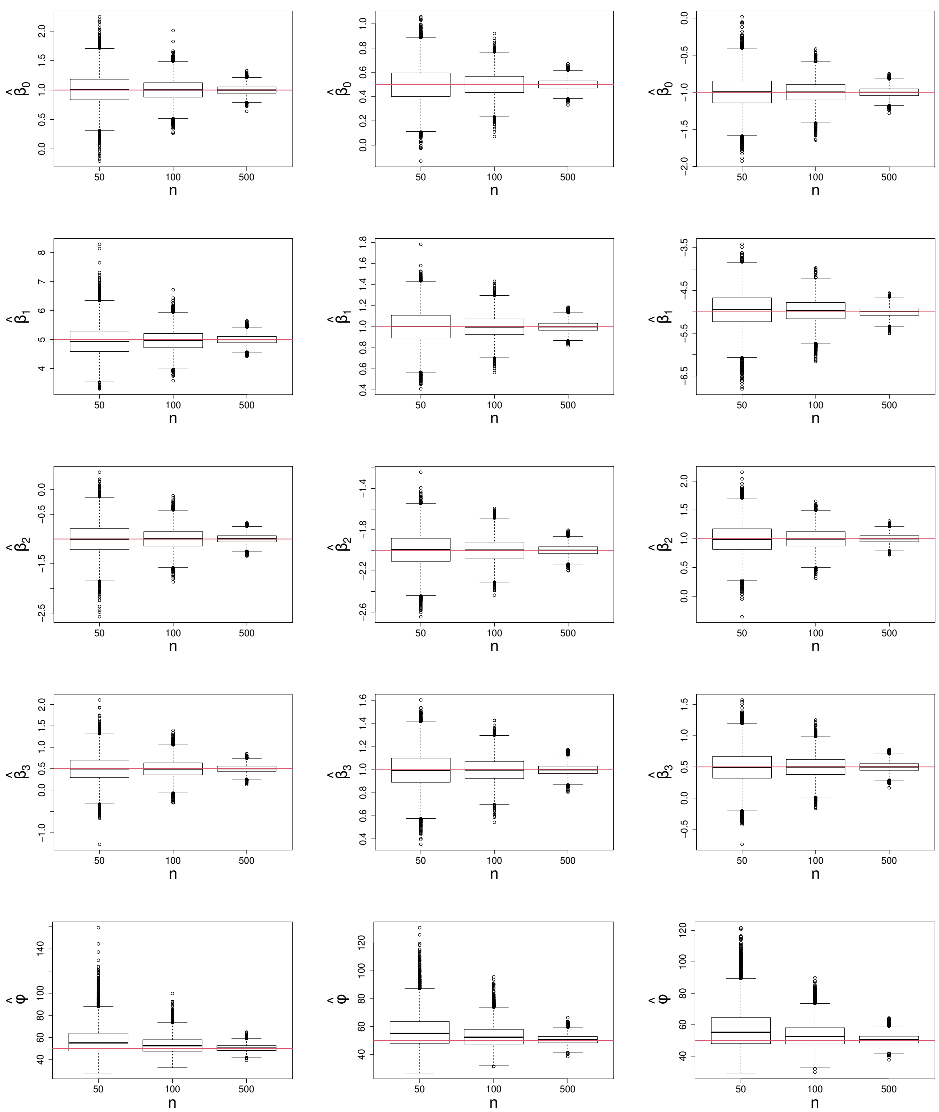


Figure 3: Boxplots of 10 000 ML estimates for parameters β_0, β_1, β_2, β_3, and ϕ by sample size, with $\phi = 50$. Left asymmetry scenario (first column), symmetry scenario (second column), right asymmetry scenario (third column).

Table 4 presents the coverage probabilities of asymptotic intervals for the model parameters to asymmetric data on the left, right, and symmetric data, considering three levels of confidence (90%, 95%, and 99%). As a general analysis, we note that the estimated coverage probability is close to the confidence level established in all scenarios. The larger the sample size, the closer the estimated coverage probability is to the determined confidence level. Although the number of covariates influences the coverage probability, even when including more covariates we observe that the estimates of the coverage probability are close to the nominal levels.

Table 1: Numerical results for the modal beta regression model with left asymmetry.

n	Parameter	True value	$\widehat{\theta}_{mc}$	b($\widehat{\theta}_{mc}$)	sd($\widehat{\theta}_{mc}$)	SE$_{as}$	$\sqrt{\text{mse}(\widehat{\theta}_{mc})}$	cv($\widehat{\theta}_{mc}$)
	β_0	1.0	0.9996	−0.0004	0.1500	0.1454	0.1501	15.01
	β_1	5.0	4.9855	−0.0146	0.6169	0.5933	0.6170	12.37
	ϕ	50.0	54.4219	4.4219	11.8928	10.9613	12.6883	21.85
	β_0	1.0	1.0052	0.0052	0.2001	0.1887	0.2001	19.90
	β_1	5.0	4.9673	−0.0327	0.4695	0.4424	0.4706	9.45
50	β_2	−1.0	−1.0013	−0.0013	0.2921	0.2751	0.2921	29.17
	ϕ	50.0	55.7621	5.7621	12.1652	11.2834	13.4608	21.82
	β_0	1.0	1.0094	0.0094	0.2738	0.2541	0.2740	27.13
	β_1	5.0	4.9513	−0.0487	0.5387	0.5015	0.5409	10.88
	β_2	−1.0	−1.0043	−0.0043	0.3254	0.3065	0.3254	32.39
	β_3	0.5	0.4951	−0.0049	0.3175	0.3011	0.3176	64.13
	ϕ	50.0	56.9308	6.9308	12.7301	11.4751	14.4945	22.36
	β_0	1.0	0.9999	−0.0001	0.1036	0.1021	0.1036	10.36
	β_1	5.0	4.9930	−0.0070	0.4240	0.4200	0.4240	8.49
	ϕ	50.0	51.9697	1.9697	7.5562	7.3986	7.8087	14.54
	β_0	1.0	0.9978	−0.0022	0.1373	0.1328	0.1373	13.76
	β_1	5.0	4.9895	−0.0105	0.3218	0.3154	0.3220	6.45
100	β_2	−1.0	−0.9949	0.0052	0.1985	0.1931	0.1985	19.95
	ϕ	50.0	52.5599	2.5599	7.8083	7.5179	8.2172	14.86
	β_0	1.0	1.0033	0.0033	0.1842	0.1761	0.1842	18.36
	β_1	5.0	4.9710	−0.0290	0.3682	0.3543	0.3693	7.41
	β_2	−1.0	−0.9961	0.0039	0.2183	0.2113	0.2183	21.91
	β_3	0.5	0.4952	−0.0048	0.2143	0.2079	0.2143	43.26
	ϕ	50.0	53.3349	3.3349	7.9402	7.5957	8.6121	14.89
	β_0	1.0	0.9999	−0.0001	0.0451	0.0454	0.0451	4.52
	β_1	5.0	4.9966	−0.0034	0.1859	0.1875	0.1859	3.72
	ϕ	50.0	50.5264	0.5264	3.3316	3.2153	3.3729	6.59
	β_0	1.0	1.0001	0.0001	0.0596	0.0590	0.0596	5.96
	β_1	5.0	4.9980	−0.0020	0.1433	0.1411	0.1433	2.87
500	β_2	−1.0	−0.9995	0.0005	0.0860	0.0854	0.0860	8.60
	ϕ	50.0	50.5757	0.5757	3.2632	3.2341	3.3136	6.45
	β_0	1.0	0.9997	−0.0003	0.0786	0.0781	0.0786	7.87
	β_1	5.0	4.9951	−0.0049	0.1615	0.1597	0.1615	3.23
	β_2	−1.0	−0.9973	0.0027	0.0928	0.0934	0.0928	9.31
	β_3	0.5	0.4993	−0.0007	0.0913	0.0917	0.0913	18.28
	ϕ	50.0	50.5940	0.5940	3.2637	3.2193	3.3174	6.45

Table 2: Numerical results for the modal beta regression model with symmetry.

n	Parameter	True value	$\widehat{\theta}_{mc}$	$\mathrm{b}(\widehat{\theta}_{mc})$	$\mathrm{sd}(\widehat{\theta}_{mc})$	SE_{as}	$\sqrt{\mathrm{mse}(\widehat{\theta}_{mc})}$	$\mathrm{cv}(\widehat{\theta}_{mc})$
	β_0	0.5	0.5000	0.0000	0.0899	0.0864	0.0899	17.98
	β_1	1.0	0.9987	−0.0013	0.1652	0.1597	0.1652	16.54
	ϕ	50.0	54.5977	4.5977	12.1370	11.2407	12.9786	22.23
	β_0	0.5	0.4995	−0.0006	0.1138	0.1093	0.1138	22.78
	β_1	1.0	0.9981	−0.0019	0.1517	0.1470	0.1517	15.20
50	β_2	−2.0	−1.9980	0.0020	0.1573	0.1507	0.1573	7.87
	ϕ	50.0	55.8075	5.8075	12.5754	11.4702	13.8516	22.53
	β_0	0.5	0.4994	−0.0006	0.1436	0.1353	0.1436	28.76
	β_1	1.0	1.0009	0.0009	0.1629	0.1522	0.1629	16.28
	β_2	−2.0	−1.9965	0.0035	0.1657	0.1565	0.1658	8.30
	β_3	1.0	0.9956	−0.0044	0.1590	0.1524	0.1591	15.97
	ϕ	50.0	56.7654	6.7654	12.4983	11.6674	14.2119	22.02
	β_0	0.5	0.4996	−0.0004	0.0625	0.0616	0.0625	12.51
	β_1	1.0	0.9997	−0.0003	0.1153	0.1137	0.1153	11.53
	ϕ	50.0	52.1531	2.1531	7.7945	7.6026	8.0864	14.95
	β_0	0.5	0.4996	−0.0004	0.0794	0.0776	0.0794	15.90
	β_1	1.0	0.9995	−0.0005	0.1057	0.1044	0.1057	10.58
100	β_2	−2.0	−1.9995	0.0005	0.1111	0.1071	0.1111	5.56
	ϕ	50.0	52.8516	2.8516	7.9666	7.6926	8.4616	15.07
	β_0	0.5	0.4998	−0.0002	0.0988	0.0963	0.0988	19.76
	β_1	1.0	0.9988	−0.0012	0.1112	0.1082	0.1112	11.13
	β_2	−2.0	−1.9987	0.0013	0.1152	0.1114	0.1152	5.76
	β_3	1.0	0.9989	−0.0011	0.1104	0.1082	0.1104	11.05
	ϕ	50.0	53.1074	3.1074	8.0613	7.7327	8.6395	15.18
	β_0	0.5	0.4997	−0.0003	0.0275	0.0276	0.0275	5.51
	β_1	1.0	1.0005	0.0005	0.0511	0.0510	0.0511	5.11
	ϕ	50.0	50.4049	0.4049	3.3238	3.2894	3.3484	6.59
	β_0	0.5	0.5003	0.0003	0.0354	0.0349	0.0354	7.08
	β_1	1.0	0.9995	−0.0005	0.0476	0.0470	0.0476	4.77
500	β_2	−2.0	−2.0003	−0.0003	0.0486	0.0482	0.0486	2.43
	ϕ	50.0	50.5244	0.5244	3.2821	3.2930	3.3237	6.50
	β_0	0.5	0.4999	−0.0001	0.0432	0.0431	0.0432	8.64
	β_1	1.0	0.9999	−0.0000	0.0489	0.0485	0.0489	4.89
	β_2	−2.0	−1.9995	0.0006	0.0501	0.0500	0.0501	2.50
	β_3	1.0	0.9994	−0.0006	0.0484	0.0485	0.0484	4.84
	ϕ	50.0	50.6310	0.6310	3.3398	3.3015	3.3989	6.60

TABLE 3: Numerical results for the modal beta regression model with right asymmetry.

n	Parameter	True value	$\widehat{\theta}_{mc}$	$\mathrm{b}(\widehat{\theta}_{mc})$	$\mathrm{sd}(\widehat{\theta}_{mc})$	SE_{as}	$\sqrt{\mathrm{mse}(\widehat{\theta}_{mc})}$	$\mathrm{cv}(\widehat{\theta}_{mc})$
50	β_0	-1.0	-1.0005	-0.0005	0.1521	0.1452	0.1521	15.20
	β_1	-5.0	-4.9789	0.0211	0.6187	0.5907	0.6191	12.43
	ϕ	50.0	54.4859	4.4859	11.5910	10.9763	12.4288	21.27
	β_0	-1.0	-1.0002	-0.0002	0.1984	0.1881	0.1984	19.84
	β_1	-5.0	-4.9689	0.0311	0.4714	0.4429	0.4724	9.49
	β_2	1.0	0.9952	-0.0048	0.2867	0.2749	0.2868	28.81
	ϕ	50.0	55.6592	5.6592	12.4714	11.2636	13.6953	22.41
	β_0	-1.0	-0.9973	0.0027	0.2285	0.2167	0.2285	22.91
	β_1	-5.0	-4.9616	0.0384	0.4192	0.3905	0.4210	8.45
	β_2	1.0	0.9937	-0.0063	0.2683	0.2552	0.2684	27.00
	β_3	0.5	0.4940	-0.0060	0.2671	0.2516	0.2672	54.06
	ϕ	50.0	57.2101	7.2101	12.8789	11.6120	14.7598	22.51
100	β_0	-1.0	-1.0005	-0.0005	0.1042	0.1020	0.1042	10.42
	β_1	-5.0	-4.9923	0.0077	0.4288	0.4193	0.4289	8.59
	ϕ	50.0	52.1448	2.1448	7.7104	7.4227	8.0031	14.79
	β_0	-1.0	-1.0002	-0.0002	0.1379	0.1327	0.1379	13.78
	β_1	-5.0	-4.9821	0.0179	0.3225	0.3145	0.3230	6.47
	β_2	1.0	0.9951	-0.0049	0.1969	0.1929	0.1970	19.79
	ϕ	50.0	52.6490	2.6490	7.8663	7.5323	8.3003	14.94
	β_0	-1.0	-1.0001	-0.0001	0.1562	0.1520	0.1562	15.62
	β_1	-5.0	-4.9796	0.0205	0.2837	0.2776	0.2845	5.70
	β_2	1.0	0.9970	-0.0030	0.1819	0.1780	0.1819	18.24
	β_3	0.5	0.4990	-0.0010	0.1815	0.1754	0.1815	36.37
	ϕ	50.0	53.3457	3.3457	7.8644	7.6562	8.5465	14.74
500	β_0	-1.0	-1.0006	-0.0006	0.0453	0.0454	0.0453	4.53
	β_1	-5.0	-4.9942	0.0058	0.1861	0.1875	0.1862	3.73
	ϕ	50.0	50.5495	0.5495	3.3290	3.2170	3.3741	6.59
	β_0	-1.0	-0.9987	0.0013	0.0597	0.0590	0.0597	5.98
	β_1	-5.0	-4.9982	0.0018	0.1427	0.1411	0.1427	2.85
	β_2	1.0	0.9983	-0.0017	0.0866	0.0855	0.0866	8.67
	ϕ	50.0	50.5310	0.5310	3.2407	3.2315	3.2839	6.41
	β_0	-1.0	-0.9999	0.0001	0.0679	0.0676	0.0679	6.79
	β_1	-5.0	-4.9958	0.0042	0.1256	0.1250	0.1257	2.51
	β_2	1.0	0.9993	-0.0007	0.0791	0.0790	0.0791	7.92
	β_3	0.5	0.4992	-0.0008	0.0788	0.0779	0.0788	15.78
	ϕ	50.0	50.6425	0.6425	3.2860	3.2502	3.3482	6.49

Table 4: Coverage probability of asymptotic confidence intervals for the model parameters under different asymmetry scenarios and confidence levels.

		Left asymmetry			Symmetry			Right asymmetry		
n	Parameter	90%	95%	99%	90%	95%	99%	90%	95%	99%
50	β_0	88.8	94.2	98.6	88.1	93.8	98.4	88.2	93.7	98.5
	β_1	87.4	92.3	97.2	88.1	93.9	98.6	86.9	92.3	97.0
	ϕ	89.6	95.1	99.2	89.5	95.5	99.1	90.8	95.8	99.2
	β_0	88.3	93.8	98.3	88.0	93.8	98.4	87.8	93.6	98.2
	β_1	86.4	92.2	97.1	88.2	93.9	98.5	86.4	91.8	96.9
	β_2	87.8	93.6	98.3	88.3	93.3	98.2	88.0	93.8	98.4
	ϕ	89.9	95.4	99.2	89.0	94.7	99.2	89.3	94.9	98.9
	β_0	87.8	93.4	98.3	87.3	93.1	98.1	87.7	93.4	98.3
	β_1	85.9	91.4	96.3	87.0	92.7	98.0	85.9	91.7	97.0
	β_2	88.1	93.5	98.3	87.6	93.3	97.9	87.7	93.4	98.3
	β_3	87.8	93.2	98.3	87.6	93.2	98.3	87.6	93.2	98.4
	ϕ	88.6	94.9	99.2	89.1	95.5	99.4	88.2	94.7	99.2
100	β_0	89.1	94.3	98.9	89.4	94.3	98.8	88.6	94.3	98.7
	β_1	88.8	93.6	98.2	89.3	94.3	98.8	88.4	93.2	97.7
	ϕ	90.5	95.4	99.0	89.9	95.3	99.1	90.0	94.9	99.1
	β_0	88.8	94.4	98.6	89.0	94.3	98.7	88.8	94.1	98.6
	β_1	88.5	93.9	98.3	89.4	94.6	98.8	88.3	93.8	98.1
	β_2	89.3	94.1	98.6	88.4	93.8	98.8	89.1	94.2	98.7
	ϕ	89.5	95.1	99.1	89.8	95.3	99.1	89.7	95.1	99.1
	β_0	88.7	93.8	98.6	89.0	94.4	98.6	88.7	94.4	98.7
	β_1	87.6	92.8	97.9	88.8	94.2	98.5	88.4	94.0	98.3
	β_2	88.6	93.9	98.7	88.4	94.0	98.7	89.2	94.3	98.8
	β_3	88.7	94.0	98.7	88.9	94.4	98.8	88.8	94.1	98.7
	ϕ	89.1	94.8	99.3	89.4	95.0	99.1	89.7	95.3	99.3
500	β_0	90.0	95.0	99.2	90.3	95.0	98.9	90.0	94.8	98.9
	β_1	90.1	95.0	98.9	89.7	94.9	98.9	89.9	94.8	98.8
	ϕ	88.8	94.6	99.1	89.9	94.9	98.9	88.9	94.3	98.9
	β_0	89.4	94.8	99.2	89.3	94.5	98.8	89.4	94.7	98.9
	β_1	88.9	94.5	98.9	89.4	94.7	98.9	89.4	94.3	98.9
	β_2	89.6	94.9	99.0	89.5	94.7	98.9	89.7	94.8	98.8
	ϕ	89.8	95.1	99.2	90.4	95.4	99.2	90.1	95.3	99.2
	β_0	89.7	94.7	99.0	90.0	95.1	99.1	89.9	94.8	98.8
	β_1	89.7	94.6	98.8	89.8	94.8	98.9	89.7	94.7	99.0
	β_2	90.0	95.4	99.1	89.9	94.9	99.0	89.9	94.8	98.9
	β_3	90.2	95.2	99.2	89.8	95.1	99.1	89.3	94.6	99.0
	ϕ	89.6	94.8	99.0	89.9	95.3	99.0	89.7	94.8	99.0

4.2. Sensitivity Study

In this section, we conduct a simulation study to compare the sensitivity of the mean and modal beta regressions in the presence of outliers. The procedure used was based on the one described by Bayes et al. (2012). In order to evaluate the influence of outliers on the regression parameter estimates, we considered a contamination of 2% in the response variable for the simulated data, that is, 2% of the observations ($y_i^\star$) were replaced by their contaminated values $y_i^\star(\Delta) = y_i^\star \pm \Delta$. Three cases were considered for the pattern of disturbance:

(i) A decrease of Δ units in the response values for higher values of x;

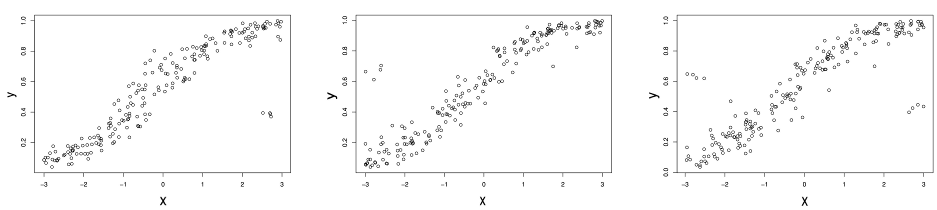
(ii) An increase of Δ units in the response values for lower values of x;

(iii) A decrease and increase of Δ units in the response values for higher and lower values of x, respectively.

Figure 4 illustrates each of these three cases described above and shows how is the shift of the atypical observations. For this illustration, we used the same sample size and combination of parameters that we used in the simulation we shall describe later. Figure 4(a) illustrates the first disturbance pattern, where 2% of observations for higher values of x were randomly selected and shifted down Δ units in the response variable. Figure 4(b) illustrates the second disturbance pattern, where 2% of observations for lower values of x were randomly selected and shifted up Δ units in the response variable. Finally, Figure 4(c) illustrates the third disturbance pattern, where cases i and ii occur simultaneously.

The Δ values vary from 0 to 0.6 with increments of 0.05 (13 cases) for cases (i) and (ii). For case (iii) , Δ varies from 0 to 0.55 with increments of 0.05 (12 cases).

We considered $n = 200$, $\beta_0 = 0.5$, $\beta_1 = 1$, $\phi = 30$, and the logit link function. The covariate was taken as random draws from $U(-3, 3)$.



(a) Case (i) (b) Case (ii) (c) Case (iii)

FIGURE 4: Contaminated datasets under three different outlier patterns in the response variable.

Initially, we generated data from a mean beta distribution with parameters μ and ϕ and fitted the mean beta regression model to the observed data. An MC simulation study with 5000 replications was carried out for each case of perturbation previously described. Next, we generated realizations of a modal beta distribution with parameters m and ϕ and fitted the modal beta regression model to the data. An MC simulation study with 5000 replications was also performed for each case of perturbation.

In order to compare the performance of the mean and modal beta regression models in an outlier context, we obtained the MC estimates of the regression parameters for both models with different values of Δ. In addition, we obtained the relative bias, $\text{RB} = |\frac{\widehat{\lambda}-\lambda}{\lambda}| \times 100$, for each case and model, where $\widehat{\lambda}$ is the MC estimate and λ is the true value of the parameter vector. The closer the RB is to zero, the closer the estimate is to the true value.

Figures 5, 6, and 7 show the curves produced by the average of the 10,000 estimates under the mean and modal beta regression models for parameters β_0 and β_1, in relation to each value of Δ, and cases (i), (ii), and (iii), respectively, as well as the curves of relative bias for each parameter. For case (i), we observe that the parameter estimates generated from the modal beta regression model show better performance, considering that they are closer to the true value of the parameter than the estimates from the mean beta regression model and, consequently, they present smaller relative bias, as for β_0 and β_1 (see Figure 5). In case (ii), see Figure 6, it is observed that the estimates for β_1 from the modal beta regression model are closer to the true parameter value and exhibit a smaller range of variation than the parameters estimates of the mean beta regression model. As a consequence, the estimates of β_1 for the modal beta regression model show smaller relative bias. However, for β_0, the mean beta regression model shows a better performance than the modal beta model in this scenario. In case (iii), the modal beta model estimates show, in general, a smaller relative bias up to $\Delta = 0.5$. For parameter β_1, the modal beta regression model presents a similar performance to that seen in case (i), that is, a better performance than the mean beta regression model for any value of Δ, see Figure 7. Therefore, we conclude that the modal beta regression model presents, in most scenarios, better performance than the mean beta regression model, mainly for the parameter associated with the covariate x (β_1).

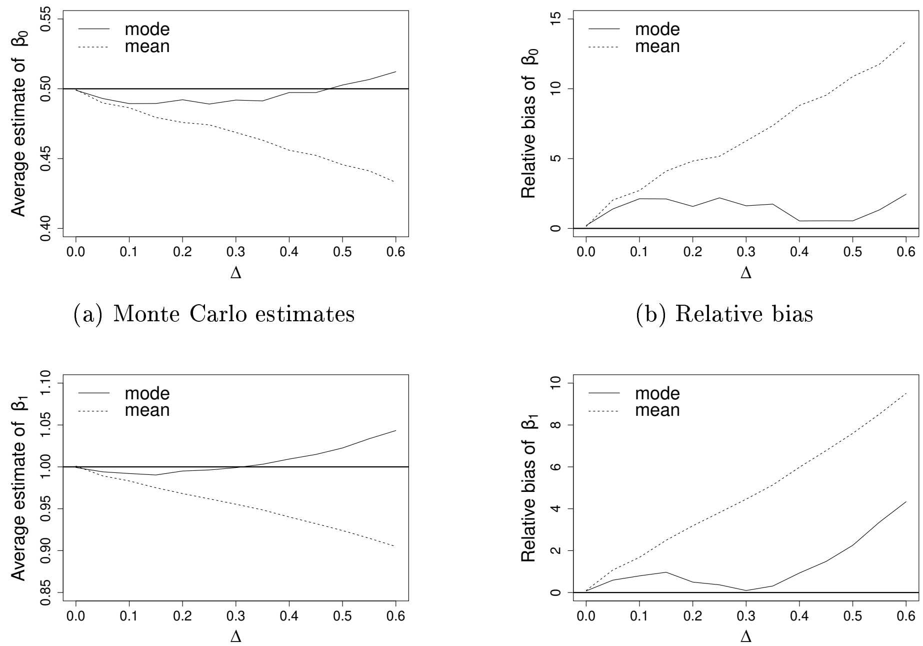
4.3. Residuals Analysis

The residual simulation study was developed to examine the residuals r_i^q, r_i^p, and r_i^c. We generated 20 observations of a random variable $Y_i \sim \text{MB}(m_i, \phi = 50)$ with

$$g(m_i) = \log\left(\frac{m_i}{1-m_i}\right) = \beta_0 + \beta_1 x_{i1} = \eta_i, \quad i = 1, \ldots, 20,$$

where $x_{i1} = (x_{11}, \ldots, x_{n1})^\top$ were generated from a standard uniform distribution.

In the first stage, the $\boldsymbol{\beta}$ vector was $\boldsymbol{\beta}^\top = (0.5, 1)$ and in the second stage, we chose three different combinations for $\boldsymbol{\beta}$, $\boldsymbol{\beta}^\top = (-2.2, 1.8)$, $\boldsymbol{\beta}^\top = (-0.4, 0.8)$, and $\boldsymbol{\beta}^\top = (0.4, 2)$, so that $m \in [0.1, 0.4)$, $m \in [0.4, 0.6)$, and $m \in [0.6, 0.9]$, respectively.

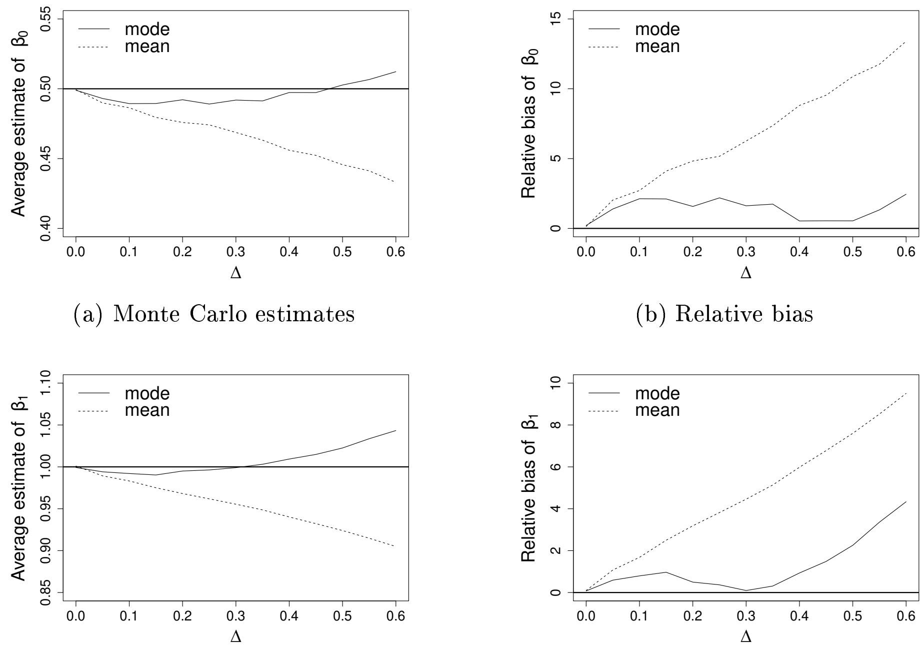


(a) Monte Carlo estimates (b) Relative bias

(c) Monte Carlo estimates (d) Relative bias

FIGURE 5: Monte Carlo estimates and relative bias of the regression parameters for case (i).

Table 5 reports the mean, standard deviation, kurtosis, and skewness of the residuals computed for each j-th residual corresponding to the i-th observation ($i = 1, \ldots, n$; $j = 1, \ldots, M$), where $M = 10\,000$ denotes the number of Monte Carlo replications. The residuals considered include the randomized quantile residuals ($r^q_{j,i}$), the standardized weighted residuals ($r^p_{j,i}$), and the Cox-Snell residuals ($r^c_{j,i}$). The standard normal distribution presents mean, kurtosis, and skewness equal to zero and standard deviation equal to one. The standard exponential distribution has mean and standard deviation equal to one, kurtosis equal to six, and skewness equal to two. In general, there is agreement between these samples and the theoretical values.



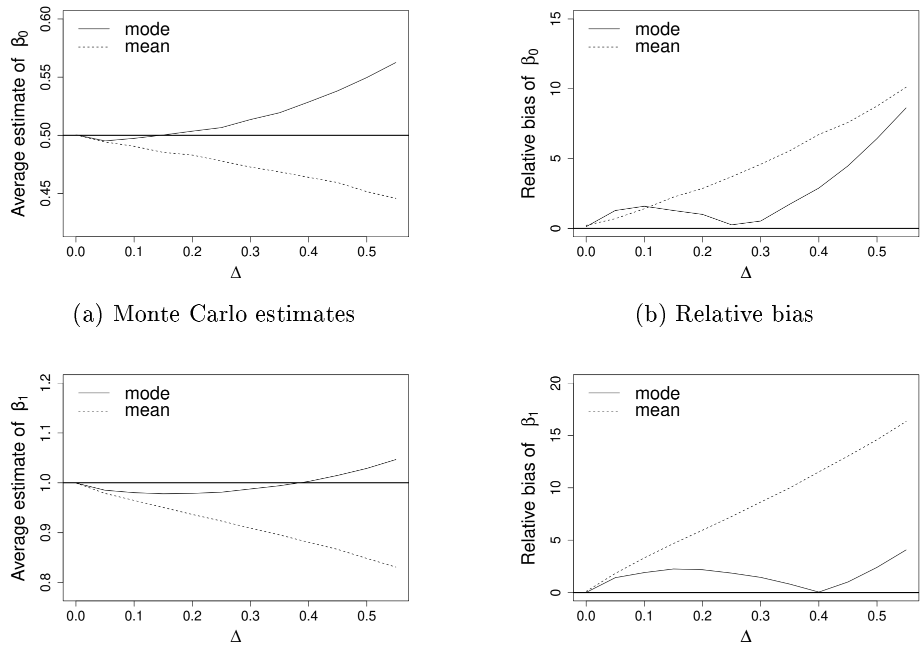
(a) Monte Carlo estimates (b) Relative bias

(c) Monte Carlo estimates (d) Relative bias

Figure 6: Monte Carlo estimates and relative bias of the regression parameters for case (i).

Table 5: Mean, standard deviation, kurtosis and skewness coefficients of the residuals considered in this study.

n	Mean			Standard deviation			Kurtosis			Skewness		
	r_i^q	r_i^p	r_i^c	r_i^q	r_i^p	r_i^c	r_i^q	r_i^p	r_i^c	r_i^q	r_i^p	r_i^c
1	−0.0030	−0.0033	0.9957	0.9998	0.9978	0.9774	−0.0865	−0.0481	4.7940	−0.0462	0.0839	1.8504
2	−0.0116	−0.0118	0.9899	0.9980	0.9981	0.9933	−0.1158	−0.0274	5.7333	0.0379	0.1689	1.9391
3	0.0101	0.0100	1.0095	1.0016	1.0010	0.9861	−0.1745	−0.1299	3.7452	−0.0281	0.0970	1.7260
4	0.0081	0.0083	1.0094	1.0038	1.0036	0.9984	−0.0987	−0.0385	5.0795	−0.0138	0.1193	1.8776
5	−0.0065	−0.0068	0.9917	0.9964	0.9955	0.9747	−0.1628	−0.1037	4.3814	−0.0167	0.1111	1.8072
6	−0.0027	−0.0027	0.9970	0.9983	0.9981	0.9917	−0.0891	−0.0239	4.6876	0.0078	0.1409	1.8713
7	0.0000	−0.0005	0.9938	0.9893	0.9885	0.9751	−0.1351	−0.0669	4.9785	−0.0056	0.1234	1.8655
8	−0.0008	−0.0008	1.0004	1.0015	1.0010	0.9971	−0.0726	−0.0096	5.1669	0.0029	0.1362	1.9170
9	−0.0016	−0.0016	0.9988	1.0014	1.0005	0.9891	−0.0890	−0.0218	5.1682	−0.0205	0.1125	1.9096
10	−0.0076	−0.0074	0.9929	0.9996	0.9984	0.9784	−0.1661	−0.1114	4.7243	−0.0145	0.1135	1.8175
11	−0.0041	−0.0045	0.9920	0.9934	0.9923	0.9759	−0.0941	−0.0459	4.8521	−0.0180	0.1136	1.8471
12	−0.0129	−0.0128	0.9903	1.0033	1.0021	0.9835	−0.1380	−0.0805	4.6413	−0.0104	0.1172	1.8445
13	0.0109	0.0116	1.0200	1.0169	1.0170	1.0147	−0.1417	−0.0915	4.2858	−0.0105	0.1206	1.8033
14	0.0080	0.0081	1.0084	1.0021	1.0018	0.9985	−0.0976	−0.0219	5.6490	−0.0087	0.1267	1.9326
15	−0.0032	−0.0034	0.9962	0.9981	0.9975	0.9825	−0.1753	−0.1023	4.6145	−0.0028	0.1245	1.8311
16	−0.0080	−0.0075	0.9969	1.0058	1.0054	1.0029	−0.0856	0.0032	6.3725	0.0161	0.1509	2.0190
17	0.0062	0.0062	1.0044	0.9967	0.9968	0.9943	−0.1239	−0.0580	4.8698	0.0137	0.1448	1.8668
18	0.0128	0.0123	1.0067	0.9916	0.9914	0.9843	−0.1437	−0.0702	5.6036	−0.0082	0.1208	1.8933
19	0.0058	0.0059	1.0076	1.0018	1.0027	1.0071	−0.1194	−0.0293	5.0098	0.0321	0.1656	1.9054
20	−0.0000	0.0001	1.0016	0.9995	1.0005	1.0108	−0.0672	0.0274	5.4922	0.0534	0.1890	1.9766



(a) Monte Carlo estimates (b) Relative bias

(c) Monte Carlo estimates (d) Relative bias

FIGURE 7: Monte Carlo estimates and relative bias of the regression parameters for case (iii).

Figure 8 shows the QQ-plots of the average of the residual order statistics, calculated from 10 000 replications, by the theoretical distribution for each studied scenario. We can see that regardless of the $\boldsymbol{\beta}$ combination, the randomized quantile and standardized weighted residuals are well approximated by the standard normal distribution, and the Cox-Snell residuals are well approximated by the standard exponential distribution.

Consider now the case of model misspecification where the data are generated from a mean beta regression distribution, but the modal beta regression model is fitted. For this, we generated $n = 100$ observations from $Y_i \sim \text{Beta}(\mu_i, \phi)$, where $\phi = 50$ with $\boldsymbol{\beta}^\top = (1, 5)$, $(-1, -5)$, and $(0.5, 1)$ for left-skewed, right-skewed, and symmetric data, respectively. Figure 9 shows the simulated envelopes of the residuals for this example of misspecification. By analogy with the structure of a matrix, the columns represent the randomized quantile, standardized weighted, and Cox-Snell residuals, respectively, and the lines represent the cases of asymmetry on the left, symmetry, and asymmetry on the right, respectively. As expected, in cases of positive or negative asymmetry, the residuals detected the model misspecification, and in the symmetry case, due to the mean and the mode being the same, the residuals did not detect any misspecification.

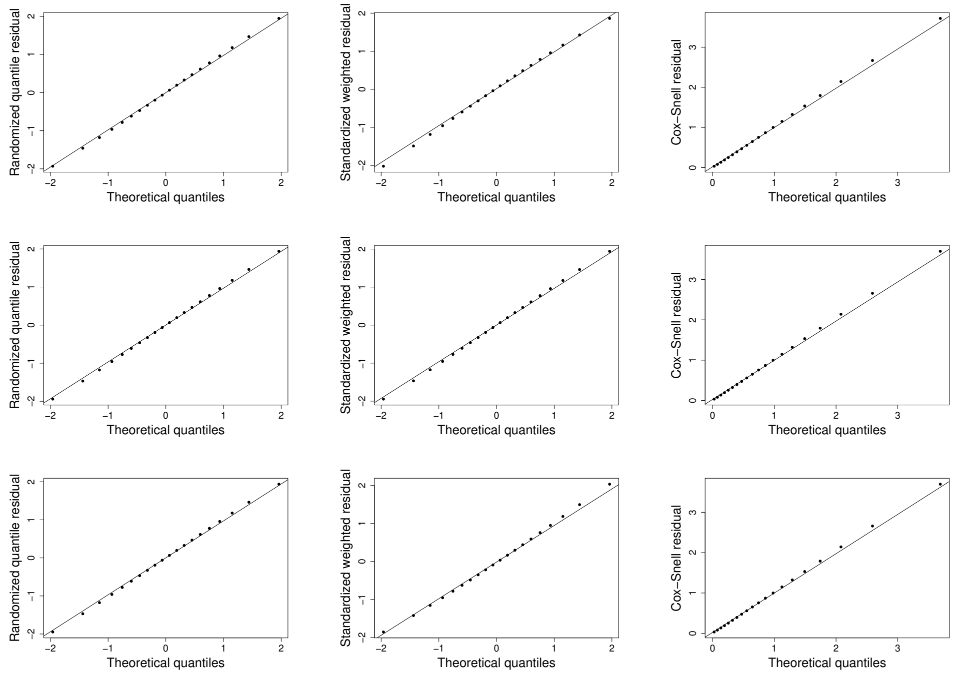


Figure 8: QQ-plots for the randomized quantile, standardized weighted, and Cox-Snell residuals.

5. Applications to Real Data

In this section, we apply the mean and modal beta regression models in two real datasets to compare their performance.

5.1. First Application

In the first application, we deal with data taken from `atlasbrasil.org.br`, considering the municipalities in the state of Mato Grosso in the 2010 Brazilian census. In this application, the response variable is the illiteracy rate of people aged between 25 and 29 years old by municipality, and the covariate is the Human Development Index (HDI) by municipality. In the Brazilian census, the illiteracy rate is calculated as the percentage of people who do not know how to read and write at least one simple note in their native language of the total resident population of the same age group, in a given geographical space and year. The HDI was proposed by the United Nations economists in 1990, and created in order to represent, in a single quantity, the degree of development of each country. This measure is based on the education, longevity, and income index of each country. Pearson's coefficient of skewness obtained for the illiteracy rate is 3.95, indicating that the response variable is right-skewed. In addition, we note the presence of upper outliers in the illiteracy rate (Figure 10), indicating that the modal beta

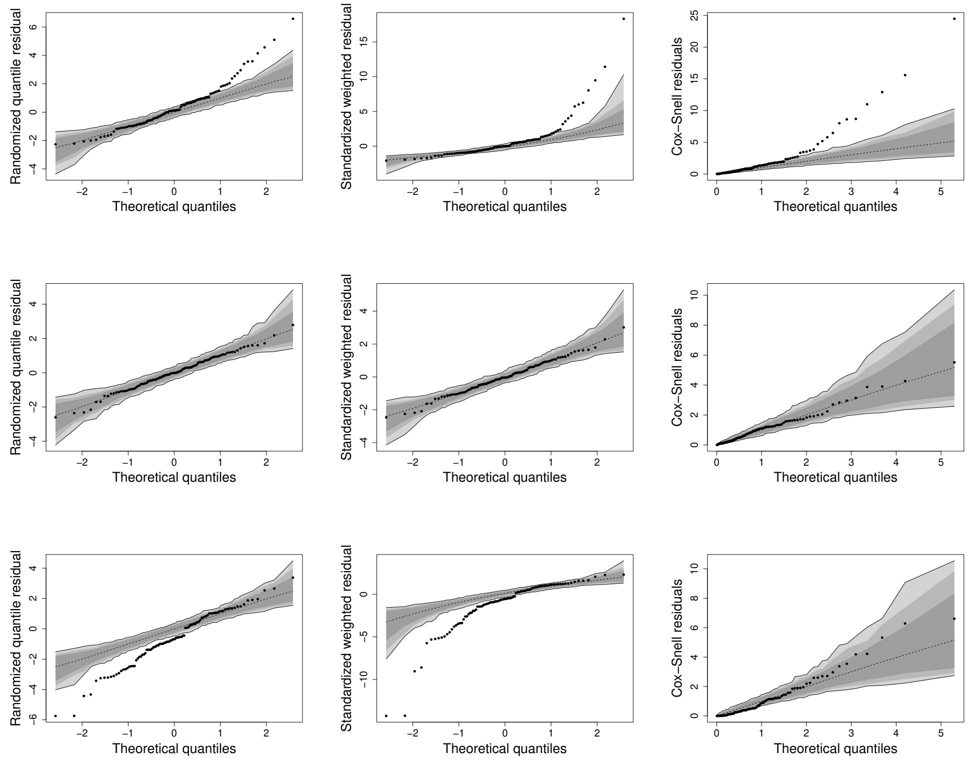
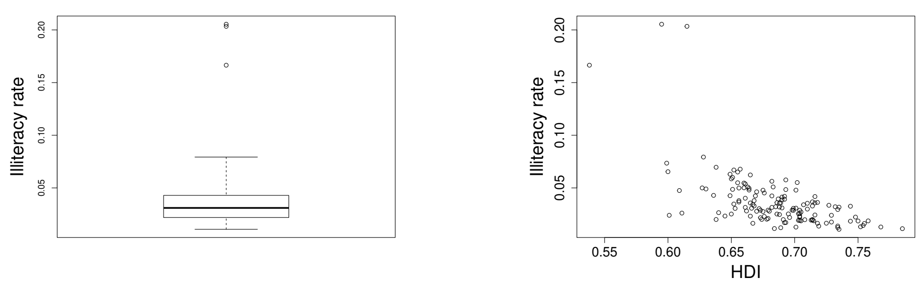


FIGURE 9: QQ-plots for the residuals of the modal beta regression model in the misspecification example.

regression could be a good choice for modeling these data. Moreover, there is a slightly negative relationship between HDI and illiteracy rate, that is, the higher the HDI the lower the illiteracy rate.



(a) Adjusted boxplot (b) Scatterplot

FIGURE 10: Adjusted boxplot and scatterplot for the illiteracy data.

Table 6 shows the summary measures of the study variables. We see that the most frequent illiteracy rate among municipalities in Mato Grosso is 2.4% and the most frequent HDI is 0.704 (high HDI). For these data, we fitted the usual beta ($Y_i \sim \text{Beta}(\mu_i, \phi)$) and the modal beta ($Y_i \sim \text{MB}(m_i, \phi)$) regression models, following the structure defined, respectively, by

$$\log\left(\frac{\mu_i}{1-\mu_i}\right) = \beta_0 + \beta_1 x_{i1}, \quad i = 1, \dots, 141$$

and

$$\log\left(\frac{m_i}{1-m_i}\right) = \beta_0 + \beta_1 x_{i1}, \quad i = 1, \dots, 141,$$

where x_{i1} is the value of HDI in the i-th municipality. Table 7 shows the ML estimates, standard errors (SE), Wald statistics, and p-values for the fitted models. Both models presented similar estimates for the precision parameter; however, the estimates for β_0 and β_1 differ. The AIC and BIC values for the usual beta regression are -790.43 and -781.59, respectively, and for the modal beta regression they are -792.14 and -783.29, respectively. That is, the modal beta regression model has lower values of AIC and BIC than the mean beta regression model. However, it is important to note that these criteria are based on different likelihood functions and the models are not nested, which limits the validity of a direct comparison.

Table 6: Summary statistics for the variables; illiteracy data.

Variable	Min	1st quartile	Mode	Median	Mean	3rd quartile	Max	sd
Illiteracy rate	0.0107	0.0219	0.0240	0.0310	0.0369	0.0429	0.2054	0.0275
HDI	0.5380	0.6610	0.7040	0.6860	0.6843	0.7070	0.7850	0.0383

Table 7: Summary results for the fitted models; illiteracy data.

Model	Parameter	Estimate	SE	Wald statistic	p-value
Modal beta regression	β_0	4.3270	0.6938	6.2368	< 0.0001
	β_1	-11.5189	1.0431	-11.0431	< 0.0001
	ϕ	142.1481	17.3960	8.1713	< 0.0001
Mean beta regression	β_0	3.1337	0.5891	5.3192	< 0.0001
	β_1	-9.4320	0.8762	-10.7645	< 0.0001
	ϕ	142.2817	17.1956	8.2743	< 0.0001

Importantly, in the modal Beta regression model, the estimated coefficients can be interpreted analogously to those in the mean Beta regression; however, they pertain to the mode rather than the expected value. The mode corresponds to the point of maximum concentration of the conditional distribution given the predictors, thereby reflecting the most frequently observed outcome conditional on the covariates.

Specifically, based on the results presented in Table 7, one aspect that can be noted is that the parameter associated with HDI (β_1) has a negative effect on the illiteracy rate, that is, the higher the HDI, the lower the rate of illiteracy. This behavior is consistent with that observed in the scatterplot (Figure 10(b)). Under the modal beta model, the estimated value of the most frequent illiteracy rate is 11.8% for municipalities with an HDI equal to 0.55 (moderate) and 2.3% for municipalities with an HDI equal to 0.7 (high).

Figure 11 shows the QQ-plots with generated bands for the randomized quantile, standardized weighted, and Cox-Snell residuals for the fitted models. Figure 11 suggests that the modal beta regression model provides the best fit to the observed data.

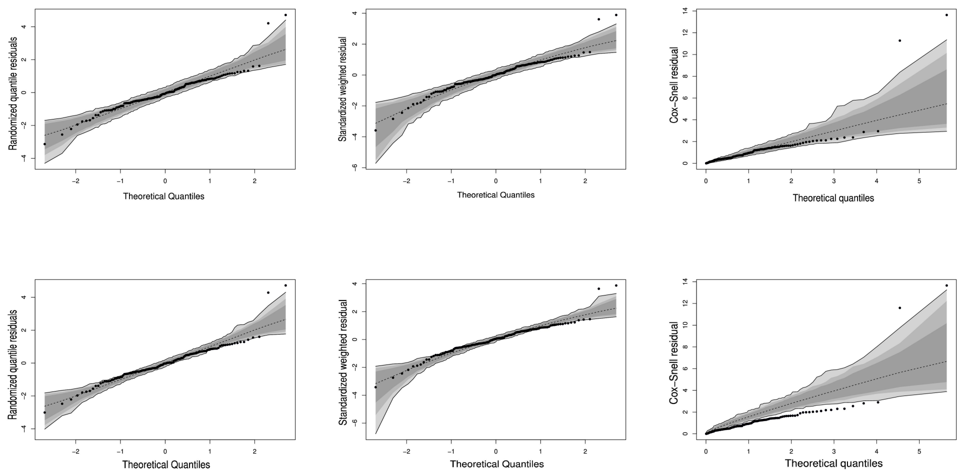


Figure 11: QQ-plots of residuals for the modal (top) and mean (bottom) beta regression models; illiteracy data.

5.2. Second Application

The second application considers the dataset from the Brazilian national high school exam (ENEM) in 2017. This exam measures the knowledge acquired in high school and is also used for admission into universities. The participants are students who completed high school in 2017, took the exam, and were admitted into the Federal University of Rio Grande do Norte (UFRN) in 2018.

The dataset has 2070 observations and 6 variables, described in Table 8. We considered the percentage of correct answers in language, codes and its technologies as the response variable, and the other variables are included as covariates in the regression model. Figure 12 shows the histogram and boxplot of the dependent variable, with the black dot representing the mode of the response. The skewness coefficient of the response variable is -0.33, indicating light asymmetry on the left.

Table 9 shows summary statistics, including the standard deviation (SD), of the response. The most frequent percentage of hits in language, codes and their technologies among the participants that were admitted into the UFRN is 60%. The rate ranges from 24.4% to 93.3%, with a standard deviation of 12.1%. Note that only 25% of approved students obtained more than 70% of hits in the test.

Table 8: Description of the variables; ENEM data.

Variable (code)	Description	Categories
Gender	Participant's gender	1 - Male 2 - Female
Ethnic	Ethnic group of the participant	1 - White 2 - Asian 3 - Indigenous 4 - Brown 5 - Black
FE	Education level of the participant's father	1 - Did not finish elementary school 2 - Finished elementary school but did not finish middle school 3 - Finished middle school but did not finish high school 4 - Finished high school but did not finish college 5 - Finished college but did not finish graduate course 6 - Finished graduate course
ME	Education level of the participant's mother	Same as for education level of the participant's father
TS	Type of school that the participant studied in high school	1 - Only in a public school 2 - Only in a private school 3 - Partly in a public school and partly in a private school

Table 9: Summary statistics of percentage of hits in language, codes and its technologies; ENEM data.

Min	1st quartile	Mode	Median	Mean	3rd quartile	Max	sd
0.2444	0.5333	0.6000	0.6222	0.6105	0.7056	0.9333	0.1211

For these data, we fitted the mean beta regression model ($Y_i \sim \text{Beta}(\mu_i, \phi)$ with $\log\{\mu_i/(1-\mu_i)\}$) and the modal beta regression model ($Y_i \sim \text{MB}(m_i, \phi)$ with $\log\{m_i/(1-m_i)\}$), for $i = 1, \ldots, 2070$. The covariates gender, ethnic group, education level of the participant's father and mother, and type of school are represented as dummy variables in which the reference level is always the first category of the variable.

The significant variables at the 5% significance level by the Wald test in both models were: FatherEduc, MotherEduc, and TypeSchool. Excluding the covariates that did not show any statistical significance, we obtained the following final models:

$$\log\left(\frac{\mu_i}{1-\mu_i}\right) = \beta_0 + \beta_6\text{FE}_{2i} + \beta_7\text{FE}_{3i} + \beta_8\text{FE}_{4i} + \beta_9\text{FE}_{5i} + \beta_{10}\text{FE}_{6i} + \beta_{11}\text{ME}_{2i} + \beta_{12}\text{ME}_{3i} + \beta_{13}\text{ME}_{4i} + \beta_{14}\text{ME}_{5i} + \beta_{15}\text{ME}_{6i} + \beta_{16}\text{TS}_{2i} + \beta_{17}\text{TS}_{3i},$$

and

$$\log\left(\frac{m_i}{1-m_i}\right) = \beta_0 + \beta_6\text{FE}_{2i} + \beta_7\text{FE}_{3i} + \beta_8\text{FE}_{4i} + \beta_9\text{FE}_{5i} + \beta_{10}\text{FE}_{6i} + \beta_{11}\text{ME}_{2i} + \beta_{12}\text{ME}_{3i} + \beta_{13}\text{ME}_{4i} + \beta_{14}\text{ME}_{5i} + \beta_{15}\text{ME}_{6i} + \beta_{16}\text{TS}_{2i} + \beta_{17}\text{TS}_{3i}.$$

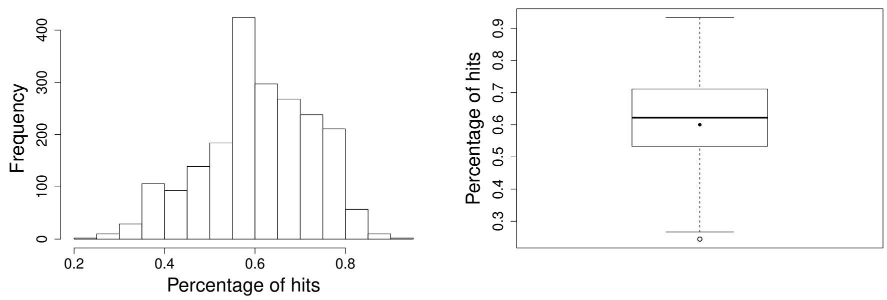


FIGURE 12: Histogram and boxplot of the percentage of hits in language, codes and its technologies; ENEM data.

Tables 10 and 11 show the estimates, standard errors, Wald test statistics, and corresponding p-values for each regression coefficient for both models. The coefficient estimates for both models are similar, as expected, because the response variable for the modal beta model is nearly symmetrical. The AIC and BIC values for the mean beta regression are -3115.55 and -3036.66, respectively, and for the modal beta regression they are -3115.66 and -3036.76, respectively. However, it is important to note that these criteria are based on different likelihood functions and the models are not nested, which limits the validity of a direct comparison.

TABLE 10: Coefficients estimates for the mean beta regression model; ENEM data.

Covariate: level	Parameter	Estimate	SE	Wald statistic	p-value
Intercept	β_0	0.1683	0.0535	3.1483	0.0016
FE_{2i}	β_6	0.0438	0.0496	0.8834	0.3770
FE_{3i}	β_7	0.0946	0.0494	1.9166	0.0553
FE_{4i}	β_8	0.1688	0.0412	4.0931	0.0000
FE_{5i}	β_9	0.2893	0.0468	6.1788	0.0000
FE_{6i}	β_{10}	0.3025	0.0503	6.0134	0.0000
ME_{2i}	β_{11}	-0.0089	0.0679	-0.1304	0.8963
ME_{3i}	β_{12}	-0.0294	0.0651	-0.4526	0.6509
ME_{4i}	β_{13}	0.0461	0.0592	0.7783	0.4364
ME_{5i}	β_{14}	0.1449	0.0628	2.3076	0.0210
ME_{6i}	β_{15}	0.1675	0.0643	2.6067	0.0091
TS_{2i}	β_{16}	0.0608	0.0236	2.5755	0.0100
TS_{3i}	β_{17}	0.0253	0.0564	0.4492	0.6533
Precision	ϕ	17.1509	0.5190	-	-

Figure 13 shows the QQ-plots with generated bands for each residual and fitted model. All residuals are within the generated bands for both models, indicating that the models are well adjusted. Considering the modal beta regression model, it is estimated that the most frequent percentage of correct answers of an individual who studied in a private school and their parents finished college is 68.1%, considering a significance level of 5%.

TABLE 11: Coefficients estimates for the modal beta regression model; ENEM data.

Covariate: level	Parameter	Estimate	SE	Wald statistic	p-value
Intercept	β_0	0.1905	0.0607	3.1365	0.0017
FE_{2i}	β_6	0.0498	0.0564	0.8821	0.3777
FE_{3i}	β_7	0.1075	0.0562	1.9123	0.0558
FE_{4i}	β_8	0.1925	0.0470	4.0998	0.0000
FE_{5i}	β_9	0.3321	0.0536	6.1955	0.0000
FE_{6i}	β_{10}	0.3475	0.0578	6.0154	0.0000
ME_{2i}	β_{11}	−0.0104	0.0773	−0.1347	0.8929
ME_{3i}	β_{12}	−0.0343	0.0740	−0.4639	0.6427
ME_{4i}	β_{13}	0.0520	0.0674	0.7710	0.4407
ME_{5i}	β_{14}	0.1658	0.0716	2.3166	0.0205
ME_{6i}	β_{15}	0.1922	0.0734	2.6191	0.0088
TS_{2i}	β_{16}	0.0701	0.0272	2.5804	0.0099
TS_{3i}	β_{17}	0.0303	0.0647	0.4686	0.6394
Precision	ϕ	15.1508	0.5190	-	-

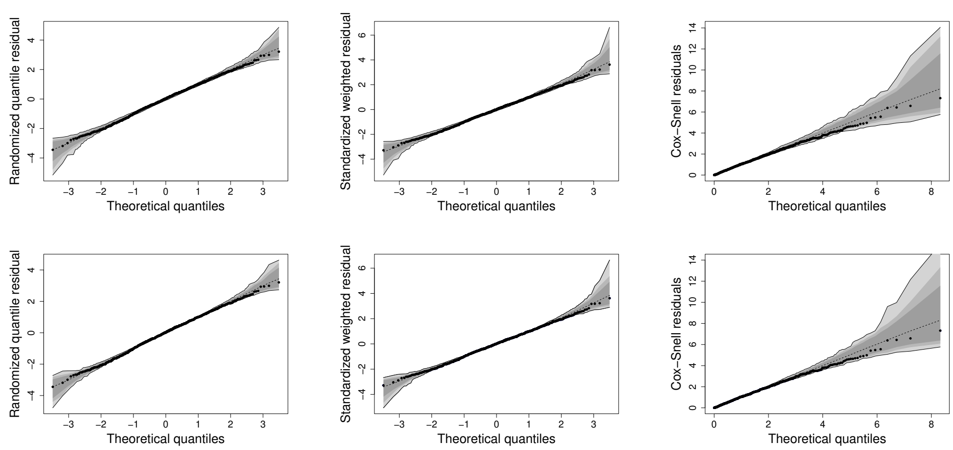


FIGURE 13: QQ-plots of residuals for the modal (top) and mean (bottom) beta regression models; ENEM data.

6. Concluding Remarks

In this work, we studied the parameterization of the beta distribution proposed by Zhou et al. (2020), which considers the mode in the regression structure and a precision parameter. Some properties of this distribution were presented and parameter estimation was conducted through the maximum likelihood approach. Assuming the response variable follows a modal beta distribution, we studied the modal beta regression model proposed by Zhou et al. (2020). We performed more comprehensive simulation studies to evaluate the maximum likelihood estimators and concluded that the estimators of the modal beta regression model exhibited good properties in symmetric and asymmetric data. In addition, we evaluated the performance of the mean and modal beta regressions in data with outliers using a procedure adapted from Bayes et al. (2012), in which three patterns of disturbance in the response variable were considered. The Monte Carlo simulation results showed that the modal beta regression model, in most scenarios, performed better than the mean beta regression model, especially when we evaluated the parameter associated with the covariate. Furthermore, three different residuals were proposed for this class of models, as well as the construction of confidence bands for them. The numerical evidence showed that the three residuals presented a good performance. Additionally, in the case where the data come from the mean beta distribution and the modal beta regression is fitted, the residuals detected misspecification for asymmetry on the right or left. However, when the data are symmetrical, the models are equivalent, so there is no case of misspecification.

We also presented and discussed two applications in real data. Our analysis illustrated that in the case of symmetrical data, the modal approach is as suitable as the mean approach, and with asymmetrical data (and in the presence of outliers) the modal beta regression model presented a more adequate fit.

In future work, we want to extend the model to consider cases of varying precision by including a regression structure in the precision parameter.

Acknowledgements

We gratefully acknowledge the grants from the Brazilian agency CNPq. We also thank the reviewers and the Associate Editor for their valuable comments and suggestions. Additionally, we thank COMPERVE for providing the data from the second application.

[Received: February 2025 — Accepted: September 2025]

References

Aristodemou, K. (2014), New regression methods for measures of central tendency, PhD thesis.

Atkinson, A. C. (1985), *Plots, Transformations, and Regression: An Introduction to Graphical Methods of Diagnostic Regression Analysis*, Oxford University Press, Oxford, UK.

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

Bourguignon, M. & Gallardo, D. I. (2025), 'A general and unified parameterization of the beta distribution: A flexible and robust beta regression model', *Statistica Neerlandica* **79**(2), e70007.

Bourguignon, M., Leão, J. & Gallardo, D. (2020), 'Parametric modal regression with varying precision', *Biometrical Journal* **62**(1), 202–220.

Cao, X., Wang, D. & Wu, L. (2023), 'Performance of ridge estimator in skew-normal mode regression model', *Communications in Statistics - Simulation and Computation* **52**(3), 1164–1177.

Cepeda-Cuervo, E. (2015), 'Beta regression models: Joint mean and variance modeling', *Journal of Statistical Theory and Practice* **9**, 134–145.

Cepeda, E. C. (2001), Variability Modeling in Generalized Linear Models, Ph.d. thesis, Mathematics Institute, Universidade Federal do Rio de Janeiro. Unpublished. https://repositorio.unal.edu.co/handle/unal/11839

Cepeda, E. & Gamerman, D. (2005), 'Bayesian methodology for modeling parameters in the two parameter exponential family', *Revista Estadística* **57**(168-169), 93–105.

Cox, D. & Hinkley, D. (1983), *Theoretical Statistic*, Chapman and Hall, London.

Cox, D. R. & Snell, E. J. (1968), 'A general definition of residuals', *Journal of the Royal Statistical Society. Series B (Methodological)* **30**, 248–275. http://www.jstor.org/stable/2984505

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

Espinheira, P. L., Ferrari, S. L. P. & Cribari-Neto, F. (2008*a*), 'Influence diagnostics in beta regression', *Computational Statistics & Data Analysis* **52**(9), 4417–4431.

Espinheira, P. L., Ferrari, S. L. P. & Cribari-Neto, F. (2008*b*), 'On beta regression residuals', *Journal of Applied Statistics* **35**(4), 407–419.

Espinheira, P. L., Santos, E. G. & Cribari-Neto, F. (2017), 'On nonlinear beta regression residuals', *Biometrical Journal* **59**(3), 445–461.

Ferrari, S. L. P. & Cribari-Neto, F. (2004), 'Beta regression for modelling rates and proportions', *Journal of applied statistics* **31**(7), 799–815.

Ghitany, M., Mazucheli, J., Menezes, A. & Alqallaf, F. (2019), 'The unit-inverse Gaussian distribution: A new alternative to two-parameter distributions on the unit interval', *Communications in Statistics-Theory and Methods* **48**(14), 3423–3438.

Ho, C.-s., Damien, P. & Walker, S. (2017), 'Bayesian mode regression using mixtures of triangular densities', *Journal of Econometrics* **197**(2), 273–283. https://www.sciencedirect.com/science/article/pii/S0304407616302184

Kemp, G. C. & Santos Silva, J. (2012), 'Regression towards the mode', *Journal of Econometrics* **170**(1), 92–101.

Kieschnick, R. & McCullough, B. D. (2003), 'Regression analysis of variates observed on (0, 1): percentages, proportions and fractions', *Statistical modelling* **3**(3), 193–213.

Lee, M. (1989), 'Mode regression', *Journal of Econometrics* **42**(3), 337–349.

Lee, M. (1993), 'Quadratic mode regression', *Journal of Econometrics* **57**(1-3), 1–19.

Oelker, M., Sobotka, F., Klein, N. & Kneib, T. (2015), 'Semiparametric Bayesian mode regression', *Computational Statistics & Data Analysis* **90**, 18–30.

Ospina, P. L. E. (2007), Regressão beta, PhD thesis, Universidade de São Paulo.

Ospina, R. & Ferrari, S. L. P. (2010), 'Inflated beta distributions', *Statistical papers* **51**(1), 111.

R Core Team (2020), *R: A Language and Environment for Statistical Computing*, R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Sager, T. W. & Thisted, R. A. (1982), 'Maximum likelihood estimation of isotonic modal regression', *The Annals of Statistics* pp. 690–707.

Yao, W. & Li, L. (2014), 'A new regression model: modal linear regression', *Scandinavian Journal of Statistics* **41**(3), 656–671.

Zhou, H., Huang, X. & Initiative, A. D. N. (2020), 'Parametric mode regression for bounded responses', *Biometrical Journal* **62**(7), 1791–1809.