Article

Modelling Interval Data with Random Intercepts: A Beta Regression Approach for Clustered and Longitudinal Structures

Olga Usuga-Manco [1],*, Freddy Hernández-Barajas [2] and Viviana Giampaoli [3]

1 Departamento de Ingeniería Industrial, Universidad de Antioquia, Medellín 050010, Colombia
2 Departamento de Estadística, Universidad Nacional de Colombia sede Medellín, Medellín 050034, Colombia; fhernanb@unal.edu.co
3 Departamento de Estatística, Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo 05508-090, Brazil; vivig@ime.usp.br
* Correspondence: olga.usuga@udea.edu.co

Abstract

Beta regression models are a class of models used frequently to model response variables in the interval $(0, 1)$. Although there are articles in which these models are used to model clustered and longitudinal data, the prediction of random effects is limited, and residual analysis has not been implemented. In this paper, a random intercept beta regression model is proposed for the complete analysis of this type of data structure. We proposed some types of residuals and formulate a methodology to obtain the best prediction of random effects. This model is developed through the parameterisation of beta distribution in terms of the mean and dispersion parameters. A log-likelihood function is approximated by the Gauss–Hermite quadrature to numerically integrate the distribution of random intercepts. A simulation study is used to investigate the performance of the estimation process and the sampling distributions of the residuals.

Keywords: Gauss–Hermite quadrature; maximum likelihood estimation; beta distribution; beta regression

1. Introduction

The class of models known as beta regression models is used to model variables that assume values in the interval $(0, 1)$, e.g., rates, proportions, or percentages, as this type of data appears so frequently that these models have received the attention of many researchers. Paolino [1], Kieschnick and McCullough [2], and Smithson and Verkuilen [3] proposed a beta regression model with applications for political science, market shares, and psychology. Ferrari and Cribari-Neto [4] defined a parameterisation of the beta distribution in terms of the mean and precision parameter and proposed a beta regression model with a regression structure for the mean.

Extensions of the regression model proposed by Ferrari and Cribari-Neto [4] have been developed. Venezuela [5] developed a generalised estimating equation to analyse longitudinal data by considering marginal beta regression models. Simas et al. [6] proposed a regression structure for the precision parameter and nonlinear regression structure for the mean. Ferrari and Pinheiro [7] considered regression structure for the precision parameter and performed small-sample inference for beta regression models.

Cribari-Neto and Souza [8] performed testing inference in beta regression models based on a second parameterisation of the beta distribution that considers the mean and

dispersion parameters. Rigby and Stasinopoulos [9] incorporated this parameterisation into the framework of generalised additive models for location, scale, and shape (GAMLSS).

Some more recent theoretical works include those by Karlsson [10], Abonazel and Dawoud [11], and Algamal [12] in which different estimators are proposed to address the problem of multicollinearity in the regression model. In applied publications, the work of Mullen [13] illustrates the use of these models in solar radiation predictions; the work of Douma [14] highlights the use of this type of models in ecology and evolution; the work of Geissinger [15] emphasises their application in the natural sciences; the work of Abonazel [16] shows a comparative study in medical science; and the work of Cribari-Neto [17] shows a beta regression analysis of COVID-19 mortality in Brazil.

A mixed beta regression model provides a class of models for the analysis of clustered and longitudinal data for beta-distributed responses. Early contributions of this class of models include a logit random intercept beta regression model implemented by Rigby and Stasinopoulos [9] on the R gamlss package, a logit random intercept beta regression model via Template Model Builder (TBM) implemented by Brooks et al. [18] on the R glmmTMB package, and a mixed beta regression model with logit and log link functions, a special case of which is the finite mixture model proposed by Verkuilen and Smithson [19].

The main goal of this article is to propose a random intercept beta regression model using a parameterisation scheme defined in terms of the mean and dispersion parameters by considering different link functions and utilising an estimation method different from the GAMLSS model. In this model, the parameters related to fixed and random effects are estimated jointly using the maximum likelihood method. The advantages of using random intercept beta regression model instead of the usual beta regression model are that by introducing random effects into the mean, we describe the heterogeneity of the means between clusters, and by introducing random effects into the dispersion, we can describe the dispersions between clusters. According to Lee, Nelder, and Pawitan (2006) [20], introducing random effects in the dispersion can describe abrupt changes among repeated measures.

Also, with the proposed model, we can analyse hierarchical datasets, such as clustered data, repeated measurements, and longitudinal data. We also develop predictions of random effects and propose residuals for the proposed model. An application is provided with the longitudinal data from a prospective ophthalmology study described in the work of Meyers et al. [21], in which the percentage of gas left in the eye of each patient was recorded on different follow-up days.

The rest of this paper unfolds as follows. In Section 2, we define the random intercept beta regression model and present the associated log-likelihood function. Further, we obtain closed-form expressions for the score function and observed information matrix. In the third section, we define the mean, variance, and covariance of the marginal distribution of beta random variables and use them to build the marginal residual. In Section 4, we describe randomised quantiles and conditional and marginal residuals to detect outlying observations and standardised random effect estimates to detect outlying clusters. The prediction of random effects used to build the proposed residuals is presented in Section 5. The results of the simulations regarding the performance of the estimation method and the residuals are presented in Section 6. In Section 7, we analyse a dataset with the proposed model. Finally, concluding remarks are presented in Section 8. Some technical details are collected in the Supplementary Material.

2. Random Intercept Beta Regression Model

In this section, we present two different parameterisations of the beta distribution, the random intercept model, the likelihood inference, and the estimation procedure.

2.1. Beta Distribution

Suppose Y is a response variable that follows a beta distribution. Ferrari and Cribari-Neto [4] proposed a parameterisation of its density indexed by the mean μ and precision parameter ϕ. The parameterisation is given by

$$f(y; \mu, \phi) = \frac{\Gamma(\phi)}{\Gamma(\mu\phi)\Gamma((1-\mu)\phi)} y^{\mu\phi-1}(1-y)^{(1-\mu)\phi-1}, \tag{1}$$

where $0 < y < 1$, $0 < \mu < 1$, and $\phi > 0$. The mean and variance of Y are given, respectively, by $\mathrm{E}(Y) = \mu$ and $\mathrm{Var}(Y) = \mu(1-\mu)/(1+\phi)$.

Alternatively, the beta distribution can be parameterised in terms of the mean μ and the dispersion parameter σ.

$$\begin{aligned} f(y; \mu, \sigma) = &\frac{\Gamma((1-\sigma^2)/\sigma^2)}{\Gamma(\mu((1-\sigma^2)/\sigma^2))\Gamma((1-\mu)((1-\sigma^2)/\sigma^2))} y^{\mu((1-\sigma^2)/\sigma^2)-1} \\ &\times (1-y)^{(1-\mu)((1-\sigma^2)/\sigma^2)-1}, \end{aligned} \tag{2}$$

with $0 < y < 1$, $0 < \mu < 1$, and $0 < \sigma < 1$. In this parameterisation, the mean of Y is $\mathrm{E}(Y) = \mu$, and the variance of Y is $\mathrm{Var}(Y) = \sigma^2\mu(1-\mu)$.

The relationship between the precision parameter in (1) and the dispersion parameter in (2) is $\phi = (1-\sigma^2)/\sigma^2$.

2.2. Random Intercept Model

Without loss of generality, we assume that $(0,1) = (a,b)$, where a and b are known scalars, $a < b$. If the response variable is constrained to the interval (a,b), we model $(y-a)/(b-a)$ instead of y.

Let y_{ij}, $i = 1, 2, \ldots, N$, be the observed repeated measurements on the i-th clusters, and let t_{ij}, $j = 1, 2, \ldots, n_i$, be the corresponding times on which the measurements are taken on each cluster. In the beta random intercept model, it is assumed that the conditional distribution of y_{ij} given $\gamma_i = (\gamma_{i1}, \gamma_{i2})^{\mathrm{T}}$ follows a beta distribution with a density given by Equation (2). Given γ_i, the repeated measurements, $y_{i1}, y_{i2}, \ldots, y_{in_i}$ are independent. We will assume the following model:

$$\begin{aligned} y_{ij} \mid \gamma_{i1}, \gamma_{i2} &\overset{\mathrm{ind}}{\sim} Be(\mu_{ij}, \sigma_{ij}), \\ g_1(\mu_{ij}) &= \eta_{ij1} = \mathbf{x}_{ij1}^{\mathrm{T}}\boldsymbol{\beta}_1 + \gamma_{i1}, \\ g_2(\sigma_{ij}) &= \eta_{ij2} = \mathbf{x}_{ij2}^{\mathrm{T}}\boldsymbol{\beta}_2 + \gamma_{i2}, \end{aligned} \tag{3}$$

where $\mathbf{x}_{ij1} = (x_{ij11}, x_{ij21}, \ldots, x_{ijp_11})^{\mathrm{T}}$ and $\mathbf{x}_{ij2} = (x_{ij12}, x_{ij22}, \ldots, x_{ijp_22})^{\mathrm{T}}$ contain values of explanatory variables, possibly defined in terms of t_{ij}. The vectors $\boldsymbol{\beta}_1 = (\beta_{11}, \beta_{21}, \ldots, \beta_{p_11})^{\mathrm{T}}$ and $\boldsymbol{\beta}_2 = (\beta_{12}, \beta_{22}, \ldots, \beta_{p_22})^{\mathrm{T}}$ are fixed parameters that do not depend on the time. The random intercepts γ_{i1} and γ_{i2} are shared among measurements within the same cluster. The link functions $g_1 : (0,1) \to \Re$ and $g_2 : (0,1) \to \Re$ are strictly monotonic and twice differentiable. The same or different link function may be used for the mean and the dispersion parameter, e.g., logit, probit, clog-log, log-log, or cauchit. For a discussion of these link functions, see McCullagh and Nelder [22].

The random intercepts γ_{i1} and γ_{i2} are independent and identically distributed normal random variables.

$$\begin{aligned} \gamma_{i1} &\overset{\mathrm{i.i.d}}{\sim} N(0, \lambda_1^2), \\ \gamma_{i2} &\overset{\mathrm{i.i.d}}{\sim} N(0, \lambda_2^2), \end{aligned} \tag{4}$$

where λ_1 and λ_2 are the standard deviations of random intercepts. Larger estimates of these parameters, λ_1 and λ_2, imply greater differences between the groups. In the particular case where $\lambda_1 = 0$, the mean of the response variable can be modelled without a random intercept.

2.3. Likelihood Inference

Let $f(y_{ij} \mid \gamma_{i1}, \gamma_{i2}; \beta_1, \beta_2)$ denote the probability density function of $y_{ij} \mid \gamma_{i1}, \gamma_{i2}$. Let $f(\gamma_{i1}; \lambda_1)$ and $f(\gamma_{i2}; \lambda_2)$ denote the probability density functions of γ_{i1} and γ_{i2}, respectively. The parameter vector is $\boldsymbol{\theta} = (\beta_1^{\mathrm{T}}, \beta_2^{\mathrm{T}}, \lambda_1, \lambda_2)^{\mathrm{T}}$. The marginal distribution of $\boldsymbol{y}_i = (y_{i1}, y_{i2}, \dots, y_{in_i})^{\mathrm{T}}$ is given by

$$f(\boldsymbol{y}_i; \boldsymbol{\theta}) = \int \int_{\Re^2} \prod_{j=1}^{n_i} f(y_{ij} \mid \gamma_{i1}, \gamma_{i2}; \beta_1, \beta_2) \cdot f(\gamma_{i1}; \lambda_1) f(\gamma_{i2}; \lambda_2) \, \mathrm{d}\gamma_{i1} \mathrm{d}\gamma_{i2}, \tag{5}$$

and the likelihood function of $\boldsymbol{\theta}$ given the observed data $\boldsymbol{y} = (\boldsymbol{y}_1, \boldsymbol{y}_2, \dots, \boldsymbol{y}_N)^{\mathrm{T}}$ is

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{N} \int \int_{\Re^2} \prod_{j=1}^{n_i} f(y_{ij} \mid \gamma_{i1}, \gamma_{i2}; \beta_1, \beta_2) \cdot f(\gamma_{i1}; \lambda_1) f(\gamma_{i2}; \lambda_2) \, \mathrm{d}\gamma_{i1} \mathrm{d}\gamma_{i2}. \tag{6}$$

Unlike a Gaussian linear model, the marginal distribution (5) and the likelihood function (6) do not have closed-form expressions. The main difficulty of likelihood inference for a random intercept beta regression model is the evaluation of intractable integrals in likelihood function (6). Common approaches include Monte Carlo integration, EM algorithms, and approximation methods (see Wu [23]). In this work, we use multivariate Gauss–Hermite quadrature to approximate the integrals. Thus, the likelihood function is

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{N} \left(\sum_{k_1=1}^{Q_1} \sum_{k_2=1}^{Q_2} \prod_{j=1}^{n_i} f(y_{ij} \mid \sqrt{2}\lambda_1 z_{k_1}, \sqrt{2}\lambda_2 z_{k_2}; \beta_1, \beta_2) \frac{w_{k_1} w_{k_2}}{\pi} \right),$$

and the log-likelihood function can be written as

$$\ell(\boldsymbol{\theta}) = \sum_{i=1}^{N} \log \left(\sum_{k_1=1}^{Q_1} \sum_{k_2=1}^{Q_2} \prod_{j=1}^{n_i} f_{ij}(y_{ij} \mid \sqrt{2}\lambda_1 z_{k_1}, \sqrt{2}\lambda_2 z_{k_2}; \beta_1, \beta_2) \frac{w_{k_1} w_{k_2}}{\pi} \right), \tag{7}$$

where Q_1 and Q_2 denote the numbers of quadrature points, z_{k_1} and z_{k_2} are the quadrature points, and w_{k_1} and w_{k_2} are the corresponding quadrature weights. For more details on multivariate Gauss–Hermite quadrature, see Fahrmeir and Tutz [24].

2.4. Estimating Procedure

The maximum likelihood estimators (MLEs) of $\boldsymbol{\theta} = (\beta_1^{\mathrm{T}}, \beta_2^{\mathrm{T}}, \lambda_1, \lambda_2)^{\mathrm{T}}$ are obtained as the solutions of the nonlinear system $\boldsymbol{U}(\boldsymbol{\theta}) = \boldsymbol{0}$, where $\boldsymbol{U}(\boldsymbol{\theta})$ denotes the p-dimensional gradient of $\ell(\boldsymbol{\theta})$, with $p = p_1 + p_2 + 2$. The score function is $\boldsymbol{U}(\boldsymbol{\theta}) = (\boldsymbol{U}_{\beta_1}^{\mathrm{T}}(\boldsymbol{\theta}), \boldsymbol{U}_{\beta_2}^{\mathrm{T}}(\boldsymbol{\theta}), U_{\lambda_1}(\boldsymbol{\theta}), U_{\lambda_2}(\boldsymbol{\theta}))^{\mathrm{T}}$, where the vectors $\boldsymbol{U}_{\beta_1}^{\mathrm{T}}(\boldsymbol{\theta})$ and $\boldsymbol{U}_{\beta_2}^{\mathrm{T}}(\boldsymbol{\theta})$ and the quantities $U_{\lambda_1}(\boldsymbol{\theta})$ and $U_{\lambda_2}(\boldsymbol{\theta})$ are defined in the Supplementary Material.

The maximum likelihood estimators do not have closed forms and must be computed by numerically maximising the log-likelihood function (7) using a nonlinear optimisation algorithm, e.g., Newton's method or a quasi-Newton algorithm. As initial values for β_1^{T} and β_2^{T}, we suggest $\hat{\beta}_1^{\mathrm{T}}$ and $\hat{\beta}_2^{\mathrm{T}}$, the estimates from the beta regression model without random intercepts, i.e.,

$$y_{ij} \overset{\text{ind}}{\sim} \text{Be}(\mu_{ij}, \sigma_{ij}),$$
$$g_1(\mu_{ij}) = x_{ij1}^{\text{T}} \beta_1,$$
$$g_2(\sigma_{ij}) = x_{ij2}^{\text{T}} \beta_2.$$

These models can be fitted in R, for example, using the packages gamlss or betareg, developed by Rigby and Stasinopoulos [9] and Cribari-Neto and Zeileis [25], respectively. For the parameters λ_1 and λ_2, the standard deviations of random intercepts, we suggest starting with initial values of 0, which corresponds to the usual beta regression model.

To obtain the covariance matrix of $\hat{\theta}$, the $(p \times p)$ observed information matrix is required, which can be written as

$$\mathbf{J}(\theta) = \begin{pmatrix} \mathbf{J}_{\beta_1\beta_1} & \mathbf{J}_{\beta_1\beta_2} & \mathbf{J}_{\beta_1\lambda_1} & \mathbf{J}_{\beta_1\lambda_2} \\ \mathbf{J}_{\beta_2\beta_1} & \mathbf{J}_{\beta_2\beta_2} & \mathbf{J}_{\beta_2\lambda_1} & \mathbf{J}_{\beta_2\lambda_2} \\ \mathbf{J}_{\lambda_1\beta_1} & \mathbf{J}_{\lambda_1\beta_2} & \mathbf{J}_{\lambda_1\lambda_1} & \mathbf{J}_{\lambda_1\lambda_2} \\ \mathbf{J}_{\lambda_2\beta_1} & \mathbf{J}_{\lambda_2\beta_2} & \mathbf{J}_{\lambda_2\lambda_1} & \mathbf{J}_{\lambda_2\lambda_2} \end{pmatrix}$$

where the elements of $\mathbf{J}(\theta)$ are defined in the Supplementary Material.

Under the usual regularity conditions for maximum likelihood estimation, the MLEs $\hat{\theta}$ of θ are such that $\sqrt{n}(\hat{\theta} - \theta) \overset{D}{\to} \mathcal{N}_{p_1+p_2+2}(0, \mathbf{J}(\theta)^{-1})$, where $\overset{D}{\to}$ denotes convergence in distribution.

3. Moments

The moments of the marginal distribution of y in model (3) are given by

$$\begin{aligned} \text{E}(y_{ij}) &= \text{E}(\mu_{ij}), \\ \text{Var}(y_{ij}) &= \text{E}(\mu_{ij}^2) - (\text{E}(\mu_{ij}))^2 + \text{E}(\sigma_{ij}^2) \cdot \text{E}(\mu_{ij}(1 - \mu_{ij})), \\ \text{Cov}(y_{ij}, y_{ij'}) &= \text{E}(\mu_{ij} \cdot \mu_{ij'}) - \text{E}(\mu_{ij}) \cdot \text{E}(\mu_{ij'}), \quad j \neq j' \end{aligned} \tag{8}$$

where the expectations involving μ_{ij} and σ_{ij} are taken with respect to the distributions of the random intercepts γ_{i1} and γ_{i2}. Note that the results depend on the link functions used for μ and σ. To obtain these moments, the required integrals must be solved using numerical integration methods.

4. Prediction of Random Effects

In practice, we are usually primarily interested in estimating the fixed-effect parameters, but it is often useful to obtain predictions of the random effects γ_{i1} and γ_{i2} as well. Because the random effects reflect between-cluster variability, they are particularly useful for detecting unusual response profiles or groups of clusters whose profiles evolve differently over time. Moreover, estimates of the random effects are needed whenever cluster-specific predictions are of interest (Fitzmaurice et al. [26]). We propose using an empirical Bayes method to obtain the best predictor (BP) of the random intercepts, which takes the following form:

$$\tilde{\gamma}_{i1} = \text{E}[\gamma_{i1} \mid y_{ij}; \hat{\theta}] = \frac{\int \int_{\Re^2} \gamma_{i1} \prod_{j=1}^{n_i} f(y_{ij} \mid \gamma_{i1}, \gamma_{i2}; \hat{\beta}_1, \hat{\beta}_2) f(\gamma_{i1}; \hat{\lambda}_1) f(\gamma_{i2}; \hat{\lambda}_2) \, d\gamma_{i1} d\gamma_{i2}}{\int \int_{\Re^2} \prod_{j=1}^{n_i} f(y_{ij} \mid \gamma_{i1}, \gamma_{i2}; \hat{\beta}_1, \hat{\beta}_2) f(\gamma_{i1}; \hat{\lambda}_1) f(\gamma_{i2}; \hat{\lambda}_2) \, d\gamma_{i1} d\gamma_{i2}},$$

$$\tilde{\gamma}_{i2} = \mathrm{E}[\gamma_{i2} \mid y_{ij}; \hat{\boldsymbol{\theta}}] = \frac{\int \int_{\Re^2} \gamma_{i2} \prod_{j=1}^{n_i} f(y_{ij} \mid \gamma_{i1}, \gamma_{i2}; \hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2) f(\gamma_{i1}; \hat{\lambda}_1) f(\gamma_{i2}; \hat{\lambda}_2) \, \mathrm{d}\gamma_{i1} \mathrm{d}\gamma_{i2}}{\int \int_{\Re^2} \prod_{j=1}^{n_i} f(y_{ij} \mid \gamma_{i1}, \gamma_{i2}; \hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2) f(\gamma_{i1}; \hat{\lambda}_1) f(\gamma_{i2}; \hat{\lambda}_2) \, \mathrm{d}\gamma_{i1} \mathrm{d}\gamma_{i2}},$$

where $i = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, n_i$.

5. Residual Analysis

Authors such as Hilden-Minton [27], Verbeke and Lesaffre [28], Pinheiro and Bates [29], and Nobre and Singer [30] have presented different types of residuals that accommodate the additional sources of variability present in linear mixed models. To study departures from these assumptions, outlying observations, and clusters, we adopt three types of residuals that account for the extra variability in the proposed model: randomised quantile residuals (see Dunn and Smyth [31]), standardised conditional residuals and standardised marginal residuals. In addition, we define standardised random effect estimates for μ and σ to assess the distribution of random effects and to study the presence of outlying clusters.

To assess the overall adequacy of the random intercept beta regression model for the data, we proposed the randomised quantile residual, which is given by

$$r_{\mathrm{q}ij} = \Phi^{-1}\big(F(y_{ij}; \hat{\mu}_{ij}, \hat{\sigma}_{ij})\big), \tag{9}$$

where $\Phi(\cdot)$ denotes the cumulative distribution function of the standard normal distribution, and $\mathrm{F}(y_{ij}; \hat{\mu}_{ij}, \hat{\sigma}_{ij})$ denotes the cumulative distribution function of $\mathrm{Be}(\hat{\mu}_{ij}, \hat{\sigma}_{ij})$.

To study outlying observations, we consider the standardised conditional and marginal residuals. The standardised conditional residual is defined as

$$r_{\mathrm{c}ij} = \frac{y_{ij} - \widehat{\mathrm{E}}(y_{ij} \mid \gamma_{i1}, \gamma_{i2})}{\sqrt{\widehat{\mathrm{Var}}(y_{ij} \mid \gamma_{i1}, \gamma_{i2})}}, \tag{10}$$

where $\widehat{\mathrm{E}}(y_{ij} \mid \gamma_{i1}, \gamma_{i2}) = \hat{\mu}_{ij}$ and $\widehat{\mathrm{Var}}(y_{ij} \mid \gamma_{i1}, \gamma_{i2}) = \hat{\sigma}_{ij}^2 \hat{\mu}_{ij}(1 - \hat{\mu}_{ij})$, with $\hat{\mu}_{ij} = \mathrm{g}_1^{-1}(x_{ij1}^{\mathrm{T}} \hat{\boldsymbol{\beta}}_1 + \tilde{\gamma}_{i1})$ and $\hat{\sigma}_{ij} = \mathrm{g}_2^{-1}(x_{ij2}^{\mathrm{T}} \hat{\boldsymbol{\beta}}_2 + \tilde{\gamma}_{i2})$; $\hat{\boldsymbol{\beta}}_1$ and $\hat{\boldsymbol{\beta}}_2$ are the maximum likelihood estimators of $\boldsymbol{\beta}_1$ and $\boldsymbol{\beta}_2$; and $\tilde{\gamma}_{i1}$ and $\tilde{\gamma}_{i2}$ denote the best predictors (BPs) of γ_{i1} and γ_{i2}, respectively.

The standardised marginal residual is given by

$$r_{\mathrm{m}ij} = \frac{y_{ij} - \widehat{\mathrm{E}}(y_{ij})}{\sqrt{\widehat{\mathrm{Var}}(y_{ij})}}, \tag{11}$$

where $\widehat{\mathrm{E}}(y_{ij}) = \mathrm{E}(\hat{\mu}_{ij})$ and $\widehat{\mathrm{Var}}(y_{ij}) = \mathrm{E}(\hat{\mu}_{ij}^2) - \big(\mathrm{E}(\hat{\mu}_{ij})\big)^2 + \mathrm{E}(\hat{\sigma}_{ij}^2) \cdot \mathrm{E}\big(\hat{\mu}_{ij}(1 - \hat{\mu}_{ij})\big)$.

Thus, when the logit link function is used for μ and σ, the mean and the variance are given by

$$\mathrm{E}(y_{ij}) = 1 - \mathrm{E}\left(\frac{1}{1 + \mathrm{ae}^{\gamma_{i1}}}\right),$$

$$\mathrm{Var}(y_{ij}) = \mathrm{E}\left(\frac{1}{(1 + \mathrm{ae}^{\gamma_{i1}})^2}\right) - \left[\mathrm{E}\left(\frac{1}{1 + \mathrm{ae}^{\gamma_{i1}}}\right)\right]^2$$
$$+ \mathrm{ab}^2 \mathrm{E}\left(\frac{\mathrm{e}^{2\gamma_{i2}}}{(1 + \mathrm{be}^{\gamma_{i2}})^2}\right) \mathrm{E}\left(\frac{\mathrm{e}^{\gamma_{i1}}}{(1 + \mathrm{ae}^{\gamma_{i1}})^2}\right),$$

where $\mathrm{a} = \mathrm{e}^{x_{ij1}^{\mathrm{T}} \boldsymbol{\beta}_1}$ and $\mathrm{b} = \mathrm{e}^{x_{ij2}^{\mathrm{T}} \boldsymbol{\beta}_2}$. Replacing $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \gamma_{i1}$, and γ_{i2}, we obtain $\hat{\boldsymbol{\beta}}_1, \hat{\boldsymbol{\beta}}_2, \hat{\gamma}_{i1}$, and $\hat{\gamma}_{i2}$, as well as $\hat{\mathrm{E}}(y_{ij})$ and $\hat{\mathrm{Var}}(y_{ij})$. Note that to calculate these estimates, multivariate

Gauss–Hermite quadrature is used to approximate the integrals in parameter estimation and random intercept prediction.

To study outlying clusters and assess the random effect distribution, we consider the standardised random effect estimates for μ and σ, defined as

$$r_{\mathrm{r}\mu i} = \frac{\tilde{\gamma}_{i1}}{\hat{\lambda}_1}, \tag{12}$$

$$r_{\mathrm{r}\sigma i} = \frac{\tilde{\gamma}_{i2}}{\hat{\lambda}_2}, \tag{13}$$

with $i = 1, 2, \ldots, N$, where $\tilde{\gamma}_{i1}$ and $\tilde{\gamma}_{i2}$ are the BPs of γ_{i1} and γ_{i2}, and $\hat{\lambda}_1$ and $\hat{\lambda}_2$ are the maximum likelihood estimates of λ_1 and λ_2.

Atkinson [32] and Kutner et al. [33] suggested the use of probability plots with simulated half-normal envelopes as diagnostic tools. Such plots are useful for identifying outliers and examining the adequacy of the fitted model, even when the residual distribution is unknown. Half-normal plots with a simulated envelope can be produced as follows:

1. Fit the beta random intercept model and generate a sample of n independent observations, treating the fitted model as the true model.
2. Fit the beta random intercept model to the generated sample and compute the ordered absolute residuals.
3. Repeat steps (1) and (2) k times.
4. For each of the n positions, collect the k order statistics and compute their average, minimum, and maximum values.
5. Plot these values together with the ordered residuals of the original sample against the half-normal scores $\phi^{-1}((i + n - 0.125)/(2n + 0.5))$, where i is the ith order statistic, $1 \le i \le n$, and n is the sample size.

The minimum and maximum values of the k-order statistics define the envelope. Observations with absolute residuals outside the simulated limits warrant further investigation. Moreover, if a considerable proportion of points falls outside the envelope, there is evidence against the adequacy of the fitted model.

6. Simulation Study

The purpose of this simulation study is twofold. First, we study the behaviour of the estimates of $\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \lambda_1$, and λ_2 as the number of clusters, cluster size, and the standard deviations of the random intercepts in μ and σ for clustered and longitudinal data. Second, we investigate the sampling distribution of the residuals for identifying outlying observations and clusters.

The data were generated according to the following random intercept beta regression model:

$$
\begin{aligned}
y_{ij} \mid \gamma_{i1}, \gamma_{i2} &\overset{\mathrm{ind}}{\sim} \mathrm{Be}(\mu_{ij}, \sigma_{ij}), \\
\mathrm{g}(\mu_{ij}) &= \beta_{11} + \beta_{21}x_{1ij} + \beta_{31}x_{2ij} + \gamma_{i1}, \\
\mathrm{g}(\sigma_{ij}) &= \beta_{12} + \beta_{22}x_{1ij} + \beta_{32}x_{2ij} + \gamma_{i2},
\end{aligned}
$$

where $i = 1, \ldots, N$, $j = 1, \ldots, n_i$, $\mathrm{g}(\cdot)$ is the logit function, and $\boldsymbol{\beta}_1 = (\beta_{11}, \beta_{21}, \beta_{31})^T$ and $\boldsymbol{\beta}_2 = (\beta_{12}, \beta_{22}, \beta_{32})^T$ are parameter vectors. The covariates are generated following simulation studies with longitudinal data, such as those by Park and Wu [34], Guoyou and Zhongyi [35], and Fu and Wang [36]. The covariate x_{1ij} is generated from the uniform distribution, $U(0, 1)$, with a different value for each pair ij. The design time points are taken between 0 and 1 as $x_{2ij} = (j - 1)/n_i$, taking the same values for each cluster. Both

parameter vectors are identical $\boldsymbol{\beta}_1 = \boldsymbol{\beta}_2 = (-0.15, 0.15, -0.15)^{\mathrm{T}}$, and are independent of the random intercepts, which are generated as $\gamma_{i1} \sim N(0, \lambda_1^2)$ and $\gamma_{i2} \sim N(0, \lambda_2^2)$.

We examined all combinations of five clusters, N, $(20, 40, 60, 100, 150)$; four cluster sizes, n_i, $(3, 5, 8, 12)$; and three standard deviations of the random effects, λ_1, λ_2, $(0.5, 1.0, 1.5)$. In total, there were 60 combinations, and we simulated 10,000 datasets for each of them. For the Gauss–Hermite quadrature method, $Q_1 = Q_2 = 8$ quadrature points were used to obtain the estimates and predictions of the random intercepts. All analyses were conducted in R [37]; the R glmmML package proposed by Broström and Holmberg [38] was used to calculate the points and weights needed for Gauss–Hermite quadrature, and the R Stats function nlminb [37] was used to maximise the log-likelihood function (7).

To study the maximum likelihood estimates of $\boldsymbol{\theta} = (\boldsymbol{\beta}_1, \boldsymbol{\beta}_2, \lambda_1, \lambda_2)^{\mathrm{T}}$, we used the root multivariate mean squared error (RMSE) for $\hat{\boldsymbol{\theta}}$, following Wissel [39], which is defined as

$$\mathrm{RMSE} = (\mathrm{trace}(\Sigma(\hat{\boldsymbol{\theta}})) + (\hat{\boldsymbol{\theta}} - \boldsymbol{\theta})^{\mathrm{T}}(\hat{\boldsymbol{\theta}} - \boldsymbol{\theta}))^{1/2}.$$

To analyse the convergence of the optimiser used in the estimation process, we computed the convergence rate (CR), defined as

$$\mathrm{CR} = \frac{10,000 \quad iterations}{Final \quad iterations},$$

where "final iterations" corresponds to the number of iterations required to complete 10,000 simulations. If the convergence rate equals one, this indicates that 10,000 iterations were sufficient. If it is less than one, more than 10,000 iterations were needed to complete the simulation study. Additional iterations occur when the algorithm does not converge because the simulated values of y are very close to zero or one due to large variances in the random effects.

Table 1 presents the RMSE across 10,000 simulated datasets for $\lambda_1 = 0.5$ and 1.5. The RMSE decreases as the cluster size (n_i) increases for fixed (N) and (λ_1, λ_2), indicating improved estimation with larger within-cluster information. For example, with $\lambda_1 = 0.5$, $\lambda_2 = 0.5$, and $N = 150$, the RMSE values are $(0.476, 0.339, 0.252, 0.210)$ for $n_i = 3, 5, 8$, and 12, respectively. RMSE also decreases with a larger number of clusters (N), holding (n_i) and (λ_1, λ_2) fixed. For instance, with $\lambda_1 = 1.5$, $\lambda_2 = 0.5$, and $n_i = 12$, the RMSE values are $(1.075, 0.771, 0.646, 0.553, 0.454)$ for $N = 20, 40, 60, 100$, and 150. Conversely, RMSE increases with higher λ_2 for given fixed values of λ_1, (N), and (n_i), reflecting reduced estimation accuracy under greater random-intercept variability. For example, with $\lambda_1 = 1.5$, $N = 60$, and $n_i = 3$, the RMSE values are $(1.134, 1.244, 1.364)$ for $\lambda_2 = 0.5, 1.0, 1.5$. Exceptions arise for $N = 20$ and $n_i = 3$ with $\lambda_1 = 0.5$ and 1.5, and for $N = 20$ and $n_i = 5$ with $\lambda_1 = 0.5$. A similar pattern is observed for λ_1: RMSE increases as λ_1 rises for fixed n_i, N, and λ_2. For instance, with $\lambda_2 = 1.5$, $N = 60$, and $n_i = 3$, the RMSE values are $(1.062, 1.364)$ for $\lambda_1 = 0.5$ and 1.5.

Table 2 presents the convergence rate (CR) over 10,000 simulated datasets for $\lambda_1 = 0.5, 1.5$ and $\lambda_2 = 0.5, 1.0, 1.5$.

To study the sampling distributions of the proposed residuals, we used the maximum likelihood estimates and the best predictors (BPs) of the random intercepts for μ and σ, computed in each replication of the study, to obtain the residuals r_{q} and r_{c}, as well as the standardised random effect estimates $r_{\mathrm{r}\mu}$ and $r_{\mathrm{r}\sigma}$. In addition, estimates of the mean and variance of the marginal distribution of y were used to compute the residual r_{m} defined in expression (11).

Table 1. RMSE of $\hat{\boldsymbol{\theta}}$ when $\lambda_1 = 0.5, 1.5$ and $\lambda_2 = 0.5, 1.0, 1.5$.

| | | $\lambda_1 = 0.5$ | | | | | $\lambda_1 = 1.5$ | | |
| | | λ_2 | | | | | λ_2 | | |
N	n_i	0.5	1.0	1.5	N	n_i	0.5	1.0	1.5
20	3	2.897	3.188	2.296	20	3	3.577	3.619	3.506
	5	1.388	1.499	1.497		5	1.811	1.822	1.868
	8	0.918	0.995	1.116		8	1.279	1.369	1.543
	12	0.706	0.807	0.950		12	1.075	1.191	1.458
40	3	1.250	1.462	1.472	40	3	1.645	1.737	1.789
	5	0.781	0.860	0.954		5	1.031	1.157	1.301
	8	0.572	0.638	0.782		8	0.857	1.000	1.189
	12	0.435	0.558	0.745		12	0.771	0.928	1.181
60	3	0.937	1.014	1.062	60	3	1.134	1.244	1.364
	5	0.590	0.663	0.756		5	0.816	0.903	1.092
	8	0.430	0.518	0.658		8	0.693	0.816	1.061
	12	0.347	0.443	0.656		12	0.646	0.805	1.050
100	3	0.638	0.688	0.727	100	3	0.777	0.859	1.005
	5	0.432	0.459	0.570		5	0.593	0.692	0.909
	8	0.315	0.387	0.540		8	0.556	0.677	0.901
	12	0.257	0.356	0.537		12	0.553	0.675	0.895
150	3	0.476	0.528	0.566	150	3	0.599	0.682	0.963
	5	0.339	0.368	0.466		5	0.491	0.641	0.861
	8	0.252	0.314	0.478		8	0.490	0.590	0.822
	12	0.210	0.295	0.407		12	0.454	0.573	0.792

Table 2. CR when $\lambda_1 = 0.5, 1.5$ and $\lambda_2 = 0.5, 1.0, 1.5$.

| | | $\lambda_1 = 0.5$ | | | | | $\lambda_1 = 1.5$ | | |
| | | λ_2 | | | | | λ_2 | | |
N	n_i	0.5	1.0	1.5	N	n_i	0.5	1.0	1.5
20	3	0.997	0.849	0.903	20	3	0.700	0.957	0.970
	5	1.000	0.958	0.975		5	0.875	0.985	0.943
	8	1.000	0.990	0.978		8	0.996	0.979	0.893
	12	0.998	0.999	0.969		12	0.992	0.942	0.817
40	3	0.870	0.960	0.977	40	3	0.869	0.983	0.950
	5	0.992	0.999	0.988		5	0.981	0.968	0.867
	8	1.000	1.000	0.970		8	0.993	0.940	0.779
	12	1.000	1.000	0.951		12	0.967	0.883	0.648
60	3	0.953	0.986	0.988	60	3	0.940	0.979	0.913
	5	1.000	0.999	0.987		5	0.992	0.960	0.846
	8	1.000	1.000	0.962		8	0.985	0.917	0.686
	12	1.000	1.000	0.923		12	0.964	0.842	0.524
100	3	0.994	0.999	0.985	100	3	0.987	0.973	0.855
	5	1.000	1.000	0.974		5	0.989	0.929	0.724
	8	1.000	0.998	0.937		8	0.962	0.866	0.511
	12	1.000	0.997	0.860		12	0.935	0.742	0.348
150	3	1.000	1.000	0.991	150	3	0.996	0.959	0.785
	5	1.000	1.000	0.962		5	0.986	0.894	0.610
	8	1.000	0.998	0.909		8	0.964	0.814	0.398
	12	1.000	0.994	0.814		12	0.917	0.647	0.229

Figure 1 presents normal probability plots for the randomised quantile (r_q), standardised conditional (r_c), and standardised marginal residual (r_m) for $N = 20$, $\lambda_1 = 0.5$, $n_i = 3, 5$, and $\lambda_2 = 0.5, 1.0, 1.5$. Each panel includes the three residuals. The plots show that the randomised quantile residuals follow an approximately normal distribution, while the conditional and marginal residuals exhibit symmetry with heavier tails, consistent with Student's t distribution. Similar results were obtained for $n_i = 8$ and 12.
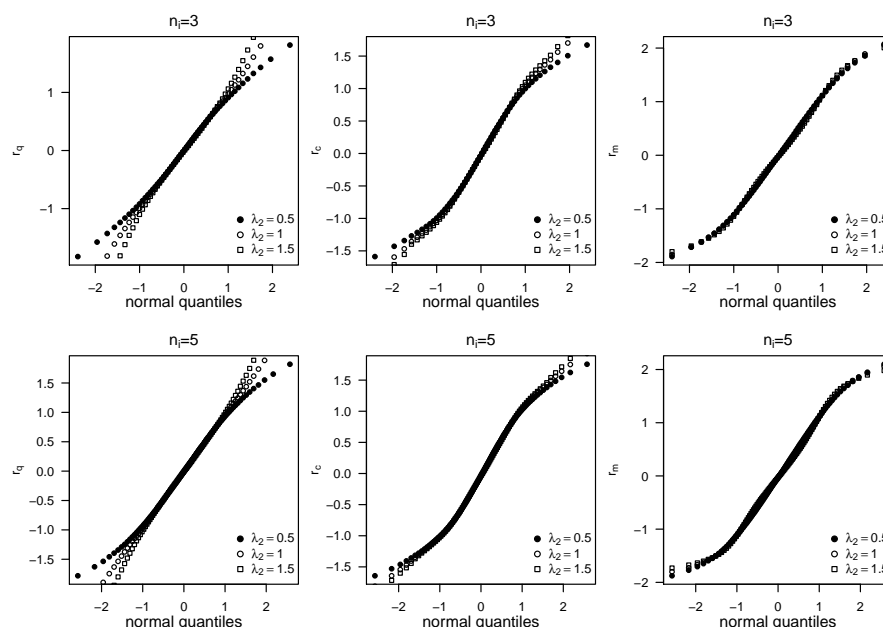
Figure 1. Normal probability plots for randomised quantile (column 1), conditional (column 2), and marginal (column 3) residuals with $N = 20$, $\lambda_1 = 0.5$, $n_i = 3$, and $n_i = 5$.

Figure 2 displays normal probability plots for the standardised random effect estimates r_{r_μ} and r_{r_σ}, with $N = 20$ and $n_i = 3, 5$. Both distributions are well approximated by the normal law. Similar plots were generated for $N = 20$, $\lambda_1 = 0.5$, and $n_i = 8, 12$, and comparable results (not shown) were also obtained for $\lambda_1 = 1.0, 1.5$ and $N = 40, 60, 100, 150$.
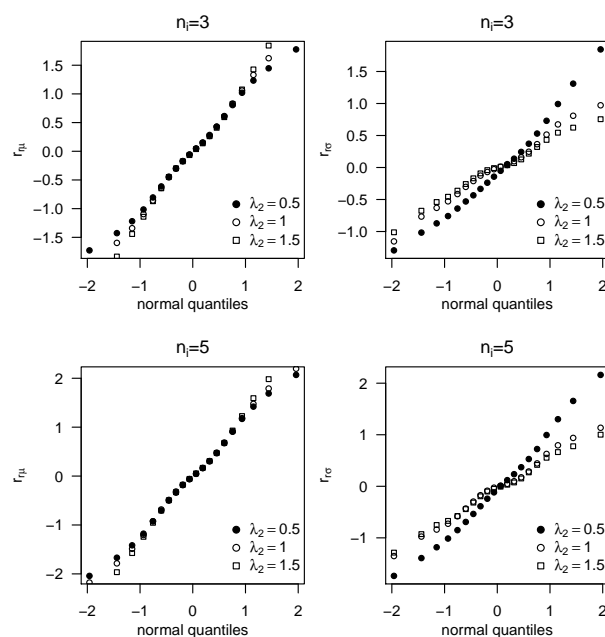
Figure 2. Normal probability plots for standardised random effect estimates for μ and σ with $N = 20$, $\lambda_1 = 0.5$, $n_i = 3$, and $n_i = 5$.

7. Application

In this section, we analyse data from a prospective ophthalmology study reported by Meyers et al. [21], with emphasis on heterogeneous mean and dispersion. Previous analyses of these data have examined homogeneous and heterogeneous dispersion using simplex regression models [40,41]. In the study, intraocular gas (C_3F_8) was used in complex retinal surgeries to provide an internal tamponade for retinal breaks. The concentration levels of C_3F_8 administered were 25%, 20%, and 15%. The primary objective was to assess whether the concentration of injected gas affected its decay rate. The sample comprised 29 patients, each followed-up with between 3 and 15 times over a 3-month period. Let y_{ij} denote the proportion of remaining gas volume relative to the initial injected volume for patient i at time j on follow-up day t_{ij}. Since the response variable lies in $(0, 1)$, the beta regression model is an appropriate choice.

As the initial specification, we adopted the quadratic model proposed in [40]:

$$\text{logit}(\mu_{ij}) = \beta_{11} + \beta_{21}\log(t_{ij}) + \beta_{31}\log^2(t_{ij}) + \beta_{41}x_{ij} + \gamma_{i1},$$
$$\text{logit}(\sigma_{ij}) = \beta_{12} + \beta_{22}\log(t_{ij}) + \beta_{32}\log^2(t_{ij}) + \beta_{42}x_{ij} + \gamma_{i2},$$

with $i = 1, \ldots, 29$, where $\gamma_{i1} \sim N(0, \lambda_1^2)$ and $\gamma_{i2} \sim N(0, \lambda_2^2)$. Here, t_{ij} denotes the time covariate (days after surgery) and x_{ij} represents the standardised gas concentration, coded as $1(25\%), 0(20\%)$, or $-1(15\%)$.

The standard beta regression model was first fitted to the data, after which nonsignificant variables were removed. The final specification for μ_{ij} and σ_{ij} was

$$\text{logit}(\mu_{ij}) = \beta_{11} + \beta_{21}\log(t_{ij}) + \beta_{41}x_{ij},$$
$$\text{logit}(\sigma_{ij}) = \beta_{12}. \tag{14}$$

Subsequently, a random-intercept beta regression model for μ was fitted, using the same linear predictors as in (14). The resulting specification for μ_{ij} and σ_{ij} was

$$\text{logit}(\mu_{ij}) = \beta_{11} + \beta_{21}\log(t_{ij}) + \beta_{41}x_{ij} + \gamma_{i1},$$
$$\text{logit}(\sigma_{ij}) = \beta_{12}. \tag{15}$$

Finally, a random-intercept beta regression model for both μ and σ was fitted, using the same linear predictors as in (14). The specification for μ_{ij} and σ_{ij} was

$$\text{logit}(\mu_{ij}) = \beta_{11} + \beta_{21}\log(t_{ij}) + \beta_{41}x_{ij} + \gamma_{i1},$$
$$\text{logit}(\sigma_{ij}) = \beta_{12} + \gamma_{i2}. \tag{16}$$

Table 3 summarises the parameter estimates and standard errors for the beta regression model (BRM), the beta regression model via TBM (TBMBRM), and the mixed beta regression model (MBRM). The Akaike information criterion (AIC) values were -97.2325, -138.2, and -169.034 for BRM, TBMBRM, and MBRM, respectively, leading to the selection of the MBRM as the best-fitting model. The results indicate that gas concentration is statistically significant, showing that higher concentrations are associated with slower decay of gas volume. Moreover, λ_1 was greater than λ_2, suggesting greater variability in the random intercept of the mean.

Table 3. Parameter estimates and standard errors for models (15) and (16).

Model	Est-s.e	β_{11}	β_{21}	β_{41}	β_{12}	λ_1	λ_2
BRM	Est.	1.911	−0.669	0.245	−0.190	–	–
	s.e.	0.187	0.069	0.100	0.078	–	–
TMBBRM	Est.	2.650	−0.958	0.239	−0.187	0.807	–
	s.e.	0.254	0.070	0.243	0.060	0.898	–
MBRM	Est.	2.590	−0.928	0.184	−0.817	0.909	0.411
	s.e.	0.222	0.080	0.191	0.123	0.132	0.111

Figure 3 presents the half-normal probability plot with a simulated envelope for the randomised quantile, standardised conditional, and standardised marginal residuals of the random-intercept beta regression model. As no observations fall outside the simulated envelope, we conclude that the beta regression model with random intercepts provides an adequate fit to the data.
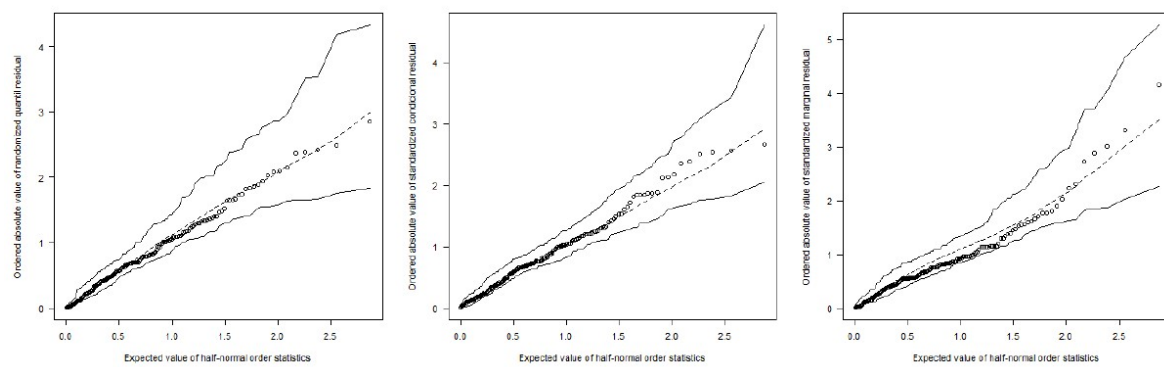
Figure 3. Half-normal probability plot with simulated envelope for randomised quantile, standardised conditional, and standardised marginal residuals.

8. Conclusions

In this paper, we proposed a random-intercept beta regression model for the analysis of clustered and longitudinal data expressed as proportions, rates, or percentages. The model accommodates heterogeneity and heteroscedasticity across clusters by incorporating random effects in both the mean and dispersion structures. Maximum likelihood estimation was carried out using Gauss–Hermite quadrature to approximate the integrals of the log-likelihood, allowing simultaneous estimation of parameters and hyperparameters and straightforward prediction of random effects. We also introduced new residuals to evaluate model adequacy and detect potential outliers at both the observation and cluster levels. Simulation studies confirmed the effectiveness of the methodology, and its application to a real dataset demonstrated its practical usefulness.

Author Contributions: Conceptualisation, O.U.-M. and V.G.; methodology, O.U.-M. and V.G.; software, O.U.-M. and F.H.-B.; validation, O.U.-M., F.H.-B. and V.G.; formal analysis, O.U.-M. and V.G.; investigation, O.U.-M.; resources, V.G.; data curation, O.U.-M. and F.H.-B.; writing—original draft preparation, O.U.-M.; writing—review and editing, O.U.-M., F.H.-B. and V.G.; supervision, V.G.; project administration, V.G.; funding acquisition, V.G. All authors have read and agreed to the published version of the manuscript.

References

1. Paolino, P. Maximum Likelihood Estimation of Models with Beta-Distributed Dependent Variables. *Political Anal.* **2001**, *9*, 325–346. [CrossRef]
2. Kieschnick, R.; McCullough, B. Regression analysis of variates observed on (0, 1): Percentages, proportions and fractions. *Stat. Model.* **2003**, *3*, 193–213. [CrossRef]
3. Smithson, M.; Verkuilen, J. A Better Lemon Squeezer? Maximum-Likelihood Regression with Beta-Distributed Dependent Variables. *Psychol. Methods* **2006**, *11*, 54–71. [CrossRef] [PubMed]
4. Ferrari, S.; Cribari-Neto, F. Beta regression for modeling rates and proportions. *J. Appl. Stat.* **2004**, *31*, 799–815. [CrossRef]
5. Venezuela, M. Equação de Estimação Generalizada e Influência Local para Modelos de Regressão Beta com Medidas Repetidas. Ph.D. Thesis, Instituto de Matemática e Estatística, Universidade de São Paulo, São Paulo, Brasil, 2008.
6. Simas, A.; Barreto-Souza, W.; Rocha, A. Improved estimators for a general class of beta regression models. *Comput. Stat. Data Anal.* **2010**, *54*, 348–366. [CrossRef]
7. Ferrari, S.; Pinheiro, E. Improved likelihood inference in beta regression. *J. Stat. Comput. Simul.* **2011**, *81*, 431–443. [CrossRef]
8. Cribari-Neto, F.; Souza, T. Testing inference in variable dispersion beta regressions. *J. Stat. Comput. Simul.* **2012**, *82*, 1827–1843. [CrossRef]
9. Rigby, R.; Stasinopoulos, D. Generalized additive models for location, scale and shape. *Appl. Stat.* **2005**, *54*, 507–554. [CrossRef]
10. Karlsson, P.; Månsson, K.; Kibria, B.G. A Liu estimator for the beta regression model and its application to chemical data. *J. Chemom.* **2020**, *34*, e3300. [CrossRef]
11. Abonazel, M.R.; Dawoud, I.; Awwad, F.A.; Lukman, A.F. Dawoud–Kibria estimator for beta regression model: Simulation and application. *Front. Appl. Math. Stat.* **2022**, *8*, 775068. [CrossRef]
12. Algamal, Z.Y.; Abonazel, M.R. Developing a Liu-type estimator in beta regression model. *Concurr. Comput. Pract. Exp.* **2022**, *34*, e6685. [CrossRef]
13. Mullen, R.; Marshall, L.; McGlynn, B. A beta regression model for improved solar radiation predictions. *J. Appl. Meteorol. Climatol.* **2013**, *52*, 1923–1938. [CrossRef]
14. Douma, J.C.; Weedon, J.T. Analysing continuous proportions in ecology and evolution: A practical introduction to beta and Dirichlet regression. *Methods Ecol. Evol.* **2019**, *10*, 1412–1430. [CrossRef]
15. Geissinger, E.; Khoo, C.; Richmond, I.; Faulkner, S.; Schneider, D. A case for beta regression in the natural sciences. *Ecosphere* **2022**, *13*. [CrossRef]
16. Abonazel, M.R.; Said, H.A.; Tag-Eldin, E.; Abdel-Rahman, S.; Khattab, I.G. Using beta regression modeling in medical sciences: A comparative study. *Commun. Math. Biol. Neurosci.* **2023**, *2023*, 18. [CrossRef]
17. Cribari-Neto, F. A beta regression analysis of COVID-19 mortality in Brazil. *Infect. Dis. Model.* **2023**, *8*, 309–317. [CrossRef] [PubMed]
18. Brooks, M.E.; Kristensen, K.; van Benthem, K.J.; Magnusson, A.; Berg, C.W.; Nielsen, A.; Skaug, H.J.; Maechler, M.; Bolker, B.M. glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. *R J.* **2017**, *9*, 378–400. [CrossRef]
19. Verkuilen, J.; Smithson, M. Mixed and Mixture Regression Models for Continuous Bounded Responses Using the Beta Distribution. *J. Educ. Behav. Stat.* **2012**, *37*, 82–113. [CrossRef]
20. Lee, Y.; Nelder, J.; Pawitan, Y. *Generalized Linear Models with Random Effects: Unified Analysis via H-Likelihood*; Chapman and Hall/CRC: Orange, CA, USA, 2017.
21. Meyers, S.; Ambler, J.; Tan, M.; Werner, J.; Huang, S. Variation of perfluorproane disappearance after vitrectomy. *Retina* **1992**, *12*, 359–363. [CrossRef]
22. McCullagh, P.; Nelder, J. *Generalized Linear Models*, 2nd ed.; Chapman and Hall: London, UK, 1989.
23. Wu, L. *Mixed Effects Models for Complex Data*; Chapman & Hall/CRC: Boca Raton, FL, USA, 2010.
24. Fahrmeir, L.; Tutz, G. *Multivariate Statistical Modelling Based on Generalized Linear Models*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2001.
25. Cribari-Neto, F.; Zeileis, A. Beta Regression in R. *J. Stat. Softw.* **2010**, *34*, 1–24. [CrossRef]
26. Fitzmaurice, G.; Davidian, M.; Verbeke, G.; Molenberghs, G. *Longitudinal Data Analysis*, 2nd ed.; Chapman and Hall/CRC: Boca Raton, FL, USA, 2009.

27. Hilden-Minton, J. Multilevel Diagnostics for Mixed and Hierarchical Linear Models. Ph.D. Thesis, University of California, Los Angeles, CA, USA, 1995.

28. Verbeke, G.; Lesaffre, E. A linear mixed-effects model with heterogeneity in the random-effects population. *J. Am. Stat. Assoc.* **1996**, *91*, 217–221. [CrossRef]

29. Pinheiro, J.; Bates, D. *Mixed-Effects Models in S ans S-PLUS*; Springer: Berlin/Heidelberg, Germany, 2000.

30. Nobre, J.; Singer, J. Residual Analysis for Linear Mixed Models. *Biom. J.* **2007**, *49*, 863–875. [CrossRef]

31. Dunn, P.; Smyth, G. Randomized Quantile Residuals. *J. Comput. Graph. Stat.* **1996**, *5*, 236–244. [CrossRef]

32. Atkinson, A. *Plots, Transformations and Regression: An Introduction to Graphical Methods of Diagnostic Regression Analysis*; Oxford University Press: New York, NY, USA, 1985.

33. Kutner, M.; Nachtsheim, C.; Neter, J.; Li, W. *Applied Linear Statistical Models*, 5th ed.; McGraw-Hill: New York, NY, USA, 2005.

34. Park, J.G.; Wu, H. Backfitting and local likelihood methods for nonparametric mixed-effects models with longitudinal data. *J. Stat. Plan. Inference* **2006**, *136*, 3760–3782. [CrossRef]

35. Guoyou, Q.; Zhongyi, Z. Robust estimation in partial linear mixed model for longitudinal data. *Acta Math. Sci.* **2008**, *28*, 333–347. [CrossRef]

36. Fu, L.; Wang, Y.G. Quantile regression for longitudinal data with a working correlation model. *Comput. Stat. Data Anal.* **2012**, *56*, 2526–2538. [CrossRef]

37. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2025; ISBN 3-900051-07-0.

38. Broström, G.; Holmberg, H. *glmmML: Generalized Linear Models with Clustering*; R Package Version 0.82-1; R Core Team: Vienna, Austria, 2011. Available online: https://CRAN.R-project.org/package=glmmML (accessed on 20 September 2024).

39. Wissel, J. A New Biased Estimator for Multivariate Regression Models with Highly Collinear Variables. Ph.D. Thesis, Institut für Mathematik, Universität Würzburg, Würzburg, Germany, 2009.

40. Song, P.X.K.; Tan, M. Marginal models for longitudinal continuous proportional data. *Biometrics* **2000**, *56*, 496–502. [CrossRef]

41. Song, P.; Qiu, Z.; Tan, M. Modelling Heterogeneous Dispersion in Marginal Models for Longitudinal Proportional Data. *Biom. J.* **2004**, *46*, 540–553. [CrossRef]