THE ZERO, ONE AND ZERO-AND-ONE-INFLATED NEW UNIT-LINDLEY DISTRIBUTIONS

André Bisca FERREIRA[1]
Josmar MAZUCHELI[2]

▪ *ABSTRACT: In this paper, we propose the zero, one and zero-and-one-inflated New unit-Lindley distributions as natural extensions of the New unit-Lindley distribution to model continuous responses measured at the following intervals* $[0,1)$, $(0,1]$ *and* $[0,1]$*. They were constructed based on convex combinations between the New unit-Lindley distribution and the distributions degenerate at zero, one, and Bernoulli distribution. They also have a number of interesting properties, such as being members of the exponential family. Besides, they have closed forms for the cumulative distribution functions, quantiles, and moments. Inferential aspects and regression structures are discussed in this work as well as a Monte Carlo simulation study to evaluate the performance of the regressors. Finally, we bring an application to real data on the suicide rate in the year 2016.*
Keywords*: Inflated Models, Maximum Likelihood Estimation, Monte Carlo Simulation Study, Regression Models.*

1 Introduction

In several areas of knowledge, data in the form of rates, ratios or proportions are continuously measured or observed within the unit range zero to one. In these cases, given the continuous and restricted nature of the data, parametric analysis demands probability distributions that have the same characteristics. The beta distribution is widely used in this type of analysis because of the different forms its

[1]Universidade Estadual de Maringá - UEM, Departamento de Estatística, CEP: 87020-900, Maringá, PR, Brasil. E-mail: *abferreira@yahoo.com*
[2]Universidade Estadual de Maringá - UEM, Departamento de Estatística, CEP: 87020-900, Maringá, PR, Brasil. E-mail: *jmazucheli@uem.br*

density takes, on the other hand, in many everyday situations it may not present a satisfactory fit, which leads to searches for new distributions with unit support.

Among the unit models recently proposed, we have the unit family distributions: unit-Weibull, unit-Gompertz and unit-Inverse Gaussian proposed from the transformation $Y = e^{-X}$ in which X follows, respectively, a variable Weibull, Gompertz and Inverse Gaussian (MAZUCHELI; MENEZES; GHITANY, 2018; MAZUCHELI; MENEZES; DEY, 2019; GHITANY et al., 2019) and the unit-lindley and New unit-lindley distributions proposed respectively from the transformations $Y = X\,(1+X)^{-1}$ and $Y = (1+X)^{-1}$ where X follows a Lindley variable (MAZUCHELI; MENEZES; CHAKRABORTY, 2019; MAZUCHELI; BAPAT; MENEZES, 2020) among others.

Although the unit family distributions are solutions to several practical problems, the presence of zeros and/or ones are generally not captured by these models, which consequently makes their use unfeasible. In this case, a viable alternative from the literature is the adoption of mixing models with the Degenerate at Zero, One and Bernoulli distributions to assign mass points at the extremes as needed. Constructed from convex sums, these mixed models have both continuous and discrete natures, have the support defined in the intervals $[0, 1)$, $(0, 1]$ and $[0, 1]$ accordingly with the distributions adopted in the mix and can also deal with problems of excess zeros and/or ones, see (XIE; HE; GOH, 2001; OSPINA; FERRARI, 2010; CRIBARI-NETO; SANTOS, 2019; TOMARCHIO; PUNZO, 2019; LIU et al., 2020; RIVAS; CAMPOS, 2021; BAPAT; BHARDWAJ, 2021) for more details on mixed models.

Within the regression structure, mixed nature models allow evaluating the influence of covariates on the continuous and discrete components as distinct events and also directly on the marginal mean of the model. Ospina e Ferrari (2012) for example, treats the proportion of fatal accidents in Brazilian municipalities considering a mixed nature model in which the discrete component describes the probability of occurrence (or not) of fatalities in traffic accidents in these municipalities while, the continuous component deals with the proportion of fatal traffic accidents considering that deaths were registered. Chai et al. (2018) present in their work a detailed discussion comparing the regression models reparametrized on the marginal mean and the conditional mean (continuous component). For more details on regression structure involving mixed models see (SANTOS; BOLFARINE, 2015, 2018; CHAI et al., 2018; MENEZES; MAZUCHELI; BOURGUIGNON, 2021; QUEIROZ; LEMONTE, 2021; SILVA et al., 2021).

In view of this, we propose in this work versions inflated by zero, one and zero and one for the New unit-Lindley distribution, as well as a brief study of their respective regression structures. The rest of this article is organized as follows: in Sections 2 and 3, we present the construction of zero or one-inflated New unit-Lindley distributions, as well as important properties and inferential aspects related to the models. Section 4 addresses the regression structure for these distributions, modeling the conditional mean and probability mass parameters of these models through a set of predictor variables. Sections 5 and 6 present the construction, some important properties and inferential aspects concerning the zero-and-one inflated New unit-Lindley distribution. A regression structure for this distribution along the same lines as the previous models is proposed in Section 7. In Section 8, we evaluate all these regression models based on a Monte Carlo simulation study. Finally, Section 9 brings the application of the models to real data in order to illustrate their applicability.

2 The Zero or One-inflated New unit-Lindley distribution

The New unit-Lindley (NUL) distribution was proposed by Mazucheli, Bapat e Menezes (2020), based on the transformation $X = (1+Y)^{-1}$, where $Y \sim \text{Lindley}(\theta)$ (LINDLEY, 1958). Therefore, since $Y \in (0, +\infty)$, the X variable is restricted to the open interval $(0, 1)$, and their cumulative distribution (c.d.f.) and probability density (p.d.f.) functions are given, respectively, by

$$\mathrm{F}(x \mid \theta) = \frac{(\theta + x)}{x\,(1+\theta)} \exp\left(-\frac{\theta\,(1-x)}{x}\right), \tag{1}$$

and

$$\mathrm{f}(x \mid \theta) = \frac{\theta^2}{(1+\theta)\,x^3} \exp\left(-\frac{\theta\,(1-x)}{x}\right) \tag{2}$$

where $\theta > 0$. On the other hand, since $\mathbb{E}(X) = \mu = \theta(1+\theta)^{-1}$, Equations 1 and 2 can be rewritten as follows

$$\mathrm{F}(x \mid \mu) = \frac{x\,(1-\mu) + \mu}{x} \exp\left(-\frac{\mu\,(1-x)}{(1-\mu)\,x}\right), \tag{3}$$

and

$$\mathrm{f}(x \mid \mu) = \frac{\mu^2}{(1-\mu)\,x^3} \exp\left(-\frac{\mu\,(1-x)}{(1-\mu)\,x}\right), \tag{4}$$

where $0 < \mu < 1$. The NUL density describes unimodal asymmetric forms having a density-reflexive behavior of the unit-Lindley (UL) distribution proposed by Mazucheli, Menezes e Chakraborty (2019) through the transformation $X' = Y(1+Y)^{-1}$, where $Y \sim \text{Lindley}(\theta)$. Thus, $X = 1 - X'$, and, for that reason, inflated versions of the UL distribution can be obtained directly through the same transformation over the inflated versions of the NUL distribution, and vice versa.

To inflate the NUL distribution to zero or one, in which the term inflation comes from the high probability at these points, we adopted the methodology proposed by OSPINA, which consists of a convex combination with a degenerate distribution at c, where c is equal to zero or one, depending on the case. Thus, the c.d.f. and p.d.f. of these models are given, respectively, by

$$\mathrm{FCINUL}(y \mid \mu, \sigma) = \sigma\, \Delta_c(y) + (1-\sigma)\, \mathrm{F}(y \mid \mu),$$

and

$$\mathrm{fcinul}(y \mid \mu, \sigma) = \sigma\, \delta_c(y) + (1-\sigma)\, \mathrm{f}(y \mid \mu), \tag{5}$$

where $y \in (0,1) \cup \{c\}$, $0 < \sigma < 1$ corresponds to the probability mass in $y = c$, $\delta_c(y)$ is an indicator function given by $\delta_c(y) = 1$ if $y = c$, and $\delta_c(y) = 0$, otherwise, $\Delta_c(y)$ is the cumulative function $\delta_c(y)$ given by $\Delta_c(y) = 1$, if $y \geq c$ and $\Delta_c(y) = 0$ if $y < c$, and terms $\mathrm{F}(\,\cdot \mid \mu)$ and $\mathrm{f}(\,\cdot \mid \mu)$ are, respectively, the cumulative distribution and NUL probability density functions, given by Equations 3 and 4.

Definition 2.1. Let Y be a random variable whose density is given by Equation 5, then

- If $c = 0$, Y is said to be a Zero-inflated New unit-Lindley distribution (ZINUL), and we denote by $Y \sim ZINUL(\mu, \sigma)$, where $\sigma = P(Y = 0)$.
- If $c = 1$, Y is said to be a One-inflated New unit-Lindley distribution (OINUL), and we denote by $Y \sim OINUL(\mu, \sigma)$, where $\sigma = P(Y = 1)$.

Note that one of the disadvantages of this mixing model, and consequently of the ZINUL and OINUL distributions, is a discontinuity in the mass point at $x = c$. Although the support of the NUL distribution has been extended at c, at this point, the variable has a discrete nature, while, for the rest of the observations, its nature is continuous. Thus, our model has both a continuous and discrete nature and, therefore, when the maximum densities of the continuous and discrete components are equal, the model takes on a bimodal form. That said, the ZINUL distribution is

- bimodal with mode at $y = 0$ and $y = \mu\,(3(1-\mu))^{-1}$ if $\mu < 3/4$ and $\sigma = \mathbf{b}(\mu)/(\mathbf{b}(\mu) + \mu)$
- bimodal with mode at $y = 0$ and $y = 1$ if $\mu \geq 3/4$ and $\sigma = \mu^2\,(\mu^2 - \mu + 1)^{-1}$
- unimodal with mode at $y = 0$ if $\mu \geq 3/4$ and $\sigma > \mu^2\,(\mu^2 - \mu + 1)^{-1}$ or if $\mu < 3/4$ and $\sigma > \mathbf{b}(\mu)/(\mathbf{b}(\mu) + \mu)$
- unimodal with mode at $y = \mu\,(3(1-\mu))^{-1}$ if $\mu < 3/4$ and $\sigma < \mathbf{b}(\mu)/(\mathbf{b}(\mu)+\mu)$
- unimodal with mode at $y = 1$ if $\mu \geq 3/4$ and $\sigma < \mu^2\,(\mu^2 - \mu + 1)^{-1}$

whereas the OINUL distribution is

- bimodal with mode at $y = (3(1-\mu))^{-1}$ and $y = 1$ if $\mu < 3/4$ and $\sigma = \mathbf{b}(\mu)/(\mathbf{b}(\mu) + \mu)$

- unimodal with mode at $y = \mu\,(3(1-\mu))^{-1}$ if $\mu < 3/4$ and $\sigma < \mathbf{b}(\mu)/(\mathbf{b}(\mu)+\mu)$

- unimodal with mode at $y = 1$ if $\mu \geq 3/4$ or if $\mu < 3/4$ and $\sigma > \mathbf{b}(\mu)/(\mathbf{b}(\mu)+\mu)$

where $\mathbf{b}(\mu) = 27 \exp\left[(4\mu - 3)(1-\mu)^{-1}\right](\mu - 1)^2$.

Note that, if Y is a random variable whose p.d.f. is given by Equation 5, then $\mathbb{E}[Y^r] = \mu_r^{'} = \mathbb{E}[(1-\sigma)\,X^r + \sigma\,Z^r] = (1-\sigma)\,\mathbb{E}[X^r] + \sigma\,\mathbb{E}[Z^r] = (1-\sigma)\,\mu_r + \sigma c$, where r is an integer greater than zero; X is a NUL random variable with moment of order r equal to μ_r, and Z is a degenerate random variable at c. In particular, for $r = 1, 2$ the mean and variance are respectively given by

$$\begin{aligned} \mu_1^{'} &= (1-\sigma)\,\mu + \sigma c \\ \mu_2^{'} &= (1-\sigma)\left[\mu^2\,\mathbf{d}(\mu)(1-\mu)^{-1}\right] + \sigma c \end{aligned}$$

where $\mathbf{d}(\mu) = \mathrm{Ei}[1,\, \mu(1-\mu)^{-1}] \exp\left(\mu(1-\mu)^{-1}\right)$ and $\mathrm{Ei}[a, z] = \int_1^\infty z^{-a} e^{-yz}\, dz$ is the exponential integral function (ABRAMOWITZ; STEGUN, 1974). That said, depending on the type of inflation, the marginal mean of the model, $\mathbb{E}[Y]$, moves left or right. In Figures 1 and 2, we present the graphic behavior of ZINUL and OINUL densities for different values of μ and σ. The blue and red dotted lines mark the respective means of the NUL distribution and the mixing model.

Proposition 2.2. *The ZINUL and OINUL distributions belong to a bi-parametric exponential family.*

Proof. In fact, since $\eta = (\eta_1, \eta_2)$, with $\eta_1 = \log(\sigma) - \log(1-\sigma) - \log\left(\dfrac{\mu^2}{1-\mu}\right)$ and $\eta_2 = -\dfrac{\mu}{1-\mu}$; and $T(y) = (t_1(y), t_2(y))$, with $t_1(y) = \delta_c(y)$ and $t_2(y) = \dfrac{1-y}{y}$ if $y \in (0,1)$ and 0 if $y = c$, we obtain Equation 5 as follows

$$\mathrm{fcinul}(y \mid \mu, \sigma) = \exp\{\eta^\top\, T(y) - \zeta(\eta)\,\}\, h(y)\,,$$

where $\zeta(\eta) = \eta_1 - \log(\sigma)$ and $h(y) = y^{-3}$ if $y \in (0,1)$ and 1 if $y = c$. □

Note that, considering Propositions 2.2 and 5.2, the zero, one, and zero-and-one-inflated versions of the NUL distribution can be used in the regression structure as an object of the Generalized Linear Models family (NELDER; WEDDERBURN, 1972).

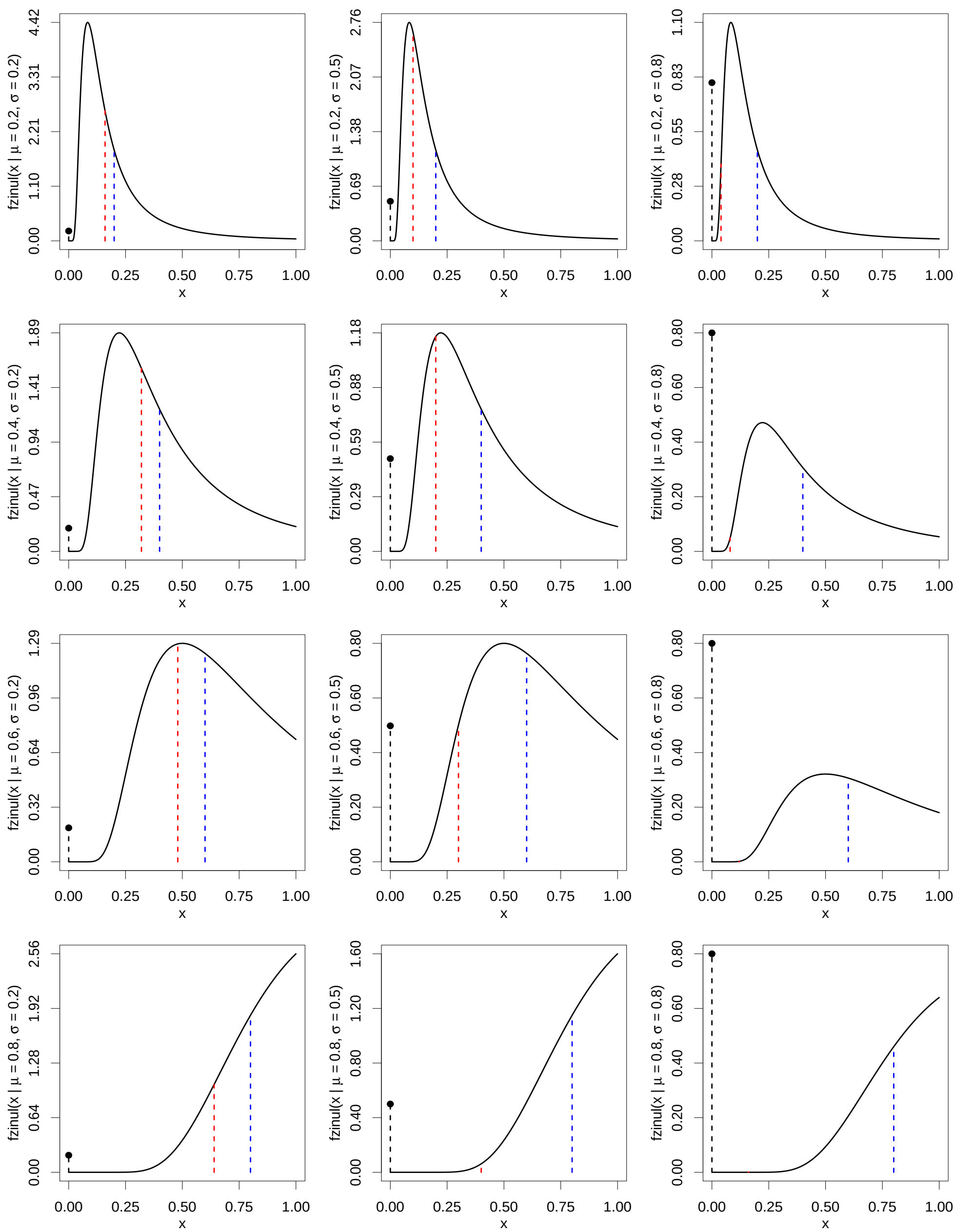

Figure 1 - Behavior of the ZINUL density assuming different values for the parameters vector (black dotted line: proportion of zeros; blue dotted line: mean of the New unit-Lindley distribution; red dotted line: marginal mean of the mixing model).

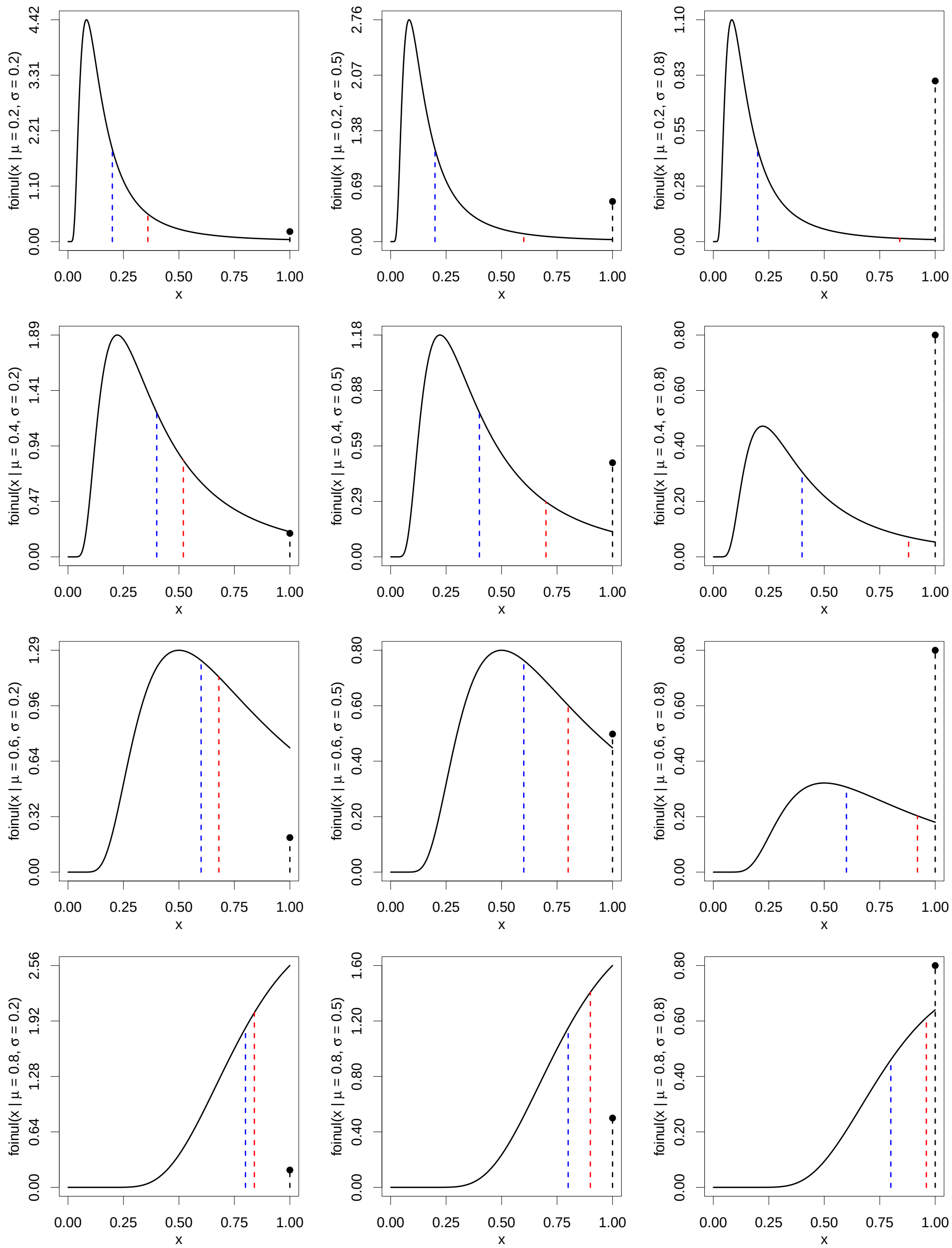

Figure 2 - Behavior of the OINUL density assuming different values for the parameters vector (black dotted line: proportion of ones; blue dotted line: mean of the New unit-Lindley distribution; red dotted line: marginal mean of the mixing model).

The ZINUL and OINUL distributions have closed forms for the quantiles respectively given by

$$\begin{cases} \dfrac{\mu}{\left(W_{-1}\left[-\dfrac{(p-\sigma)\,\mathbf{e}(\mu)}{(1-\sigma)}\right]+1\right)(-1+\mu)} & \text{if } \sigma < p < 1 \\ 0 & \text{if } 0 < p \leq \sigma \end{cases}$$

and

$$\begin{cases} \dfrac{\mu}{\left(W_{-1}\left[-\dfrac{p\,\mathbf{e}(\mu)}{(1-\sigma)}\right]+1\right)(-1+\mu)} & \text{if } 0 < p \leq 1-\sigma \\ 1 & \text{if } 1-\sigma < p < 1 \end{cases}$$

where $\mathbf{e}(\mu) = (1-\mu)^{-1}\exp(-(1-\mu)^{-1})$ and W_{-1} denote the negative branch of the Lambert `W` function (CORLESS et al., 1996; JODRÁ, 2010).

3 Inferential Aspects of the Zero or One-inflated New unit-Lindley distributions

This section brings general expressions of the estimator and the Fisher's expected information matrix for the parameter vector $\Psi = (\mu, \sigma)$ of the NUL distribution inflated at c. The expressions relating to the ZINUL and OINUL distributions are obtained by considering, respectively, c equal to zero and one.

Let $Y = Y_1, Y_2 \ldots, Y_n$ be a random sample taken from a population with p.d.f. given by Equation 5. Without loss of generality, the density of the Y variable can be expressed as follows

$$\text{fcinul}(y \mid \mu, \sigma) = \{(1-\sigma)^{1-\delta_c(y)}\sigma^{\delta_c(y)}\}\{\text{f}(y \mid \mu)\}^{1-\delta_c(y)},$$

where $\text{fcinul}(\,\cdot \mid \mu, \sigma)$, as a function of the parameters, can be factored into two independent terms. In such conditions, for $\mathbf{y} = (y_1, y_2, \ldots, y_n)$, the likelihood function of $\Psi = (\mu, \sigma)$ is given by

$$L(\Psi \mid \mathbf{y}) = L_1(\sigma \mid \mathbf{y}) \times L_2(\mu \mid \mathbf{y}), \tag{6}$$

where

$$L_1(\sigma \mid \mathbf{y}) = \prod_{i=1}^{n} \sigma^{\delta_c(y_i)}(1-\sigma)^{1-\delta_c(y_i)} \tag{7}$$

and

$$L_2(\mu \mid \mathbf{y}) = \prod_{i=1}^{n} \left\{ \frac{\mu^2}{(1-\mu)\, y_i^3} \exp\left(- \frac{\mu\,(1-y_i)}{(1-\mu)\, y_i} \right) \right\}^{1-\delta_c(y_i)}. \tag{8}$$

Regarding Equations 6, 7 and 8, the logarithm of the likelihood function is given by

$$\ell(\Psi \mid \mathbf{y}) = \ell_1(\sigma \mid \mathbf{y}) + \ell_2(\mu \mid \mathbf{y}),$$

in which

$$\ell_1(\sigma \mid \mathbf{y}) = (n - \mathbf{s}(\mathbf{y})) \log(1-\sigma) + \mathbf{s}(\mathbf{y}) \log(\sigma) \tag{9}$$

and

$$\ell_2(\mu \mid \mathbf{y}) = (n - \mathbf{s}(\mathbf{y})) \log \left(\frac{\mu^2}{(1-\mu)} \right) + \mathbf{r}(\mathbf{y}) - \frac{\mu}{(1-\mu)} \mathbf{t}(\mathbf{y}), \tag{10}$$

where $\mathbf{s}(\mathbf{y}) = \sum_{i=1}^{n} \delta_c(y_i)$, $\mathbf{r}(\mathbf{y}) = -3 \sum_{\substack{i=1 \\ y_i \in (0,1)}} \log(y_i)$ e $\mathbf{t}(\mathbf{y}) = \sum_{\substack{i=1 \\ y_i \in (0,1)}} \left(\frac{(1-y_i)}{y_i} \right)$.

From a frequentist point of view, the maximum likelihood estimator (MLE) $\widehat{\Psi} = (\widehat{\mu}, \widehat{\sigma})$ of Ψ can be obtained by maximizing $\ell(\Psi \mid \mathbf{y})$ in relation to σ and μ. The first order partial derivatives of $\ell(\Psi \mid \mathbf{y})$ are given by

$$\frac{\partial}{\partial \sigma} \ell(\Psi \mid \mathbf{y}) = \frac{\mathbf{s}(\mathbf{y})}{\sigma} - \frac{n - \mathbf{s}(\mathbf{y})}{1-\sigma}$$

and

$$\frac{\partial}{\partial \mu} \ell(\Psi \mid \mathbf{y}) = \frac{(n - \mathbf{s}(\mathbf{y}))\mu^2 + (-3n + 3\mathbf{s}(\mathbf{y}) - \mathbf{t}(\mathbf{y}))\mu + 2(n - \mathbf{s}(\mathbf{y}))}{\mu(1-\mu)^2},$$

which leads MLE $\widehat{\Psi}$ to be given by the following components

$$\widehat{\sigma} = \frac{\mathbf{s}(\mathbf{y})}{n} \tag{11}$$

and

$$\widehat{\mu} = \frac{3(n - \mathbf{s}(\mathbf{y})) + \mathbf{t}(\mathbf{y}) - \sqrt{\mathbf{u}(\mathbf{y})}}{n - \mathbf{s}(\mathbf{y})}, \tag{12}$$

where $\mathbf{u}(\mathbf{y}) = n^2 - 2\, n\, \mathbf{s}(\mathbf{y}) + 6\, n\, \mathbf{t}(\mathbf{y}) + \mathbf{s}^2(\mathbf{y}) - 6\, \mathbf{s}(\mathbf{y})\, \mathbf{t}(\mathbf{y}) + \mathbf{t}^2(\mathbf{y})$.

The MLE of the Ψ parameters vector of the ZINUL and OINUL distributions are obtained by replacing c by zero and one, respectively, in Equations 11 and 12.

It is worth mentioning that the $\widehat{\mu}$ estimator of the NUL distribution inflated at c can be obtained directly by expressing the μ MLE of the NUL distribution for a sub-sample of Y with the exclusion of the extremes (zero and one).

The second order partial derivatives of $\ell(\Psi \mid \mathbf{y})$ regarding σ and μ are given, respectively, by

$$\frac{\partial^2}{\partial \sigma^2}\ell(\Psi \mid \mathbf{y}) = -\frac{\mathbf{s}(\mathbf{y})}{\sigma^2} - \frac{n-\mathbf{s}(\mathbf{y})}{(1-\sigma)^2}$$

and

$$\begin{aligned}\frac{\partial^2}{\partial \mu^2}\ell(\Psi \mid \mathbf{y}) = &\ \frac{(-n+\mathbf{s}(\mathbf{y}))\mu^3 + (5[n-\mathbf{s}(\mathbf{y})] + 2\mathbf{t}(\mathbf{y}))\mu^2 + 6(n-\mathbf{s}(\mathbf{y}))\mu}{\mu^2(-1+\mu)^3} \\ &+ \frac{2(n-\mathbf{s}(\mathbf{y}))}{\mu^2(-1+\mu)^3},\end{aligned}$$

which causes the Fisher's expected information matrix to be given by

$$\mathbf{K}(\Psi) = \left[\begin{array}{cc} k_{\mu\,\mu} & 0 \\ 0 & k_{\sigma\,\sigma} \end{array}\right]$$

where $k_{\mu\,\mu} = n\,(2-\mu^2)\,(1-\sigma)\,\mu^{-2}\,(1-\mu)^{-2}$ and $k_{\sigma\,\sigma} = n\,\sigma^{-1}\,(1-\sigma)^{-1}$. Thus, confidence intervals and hypothesis testing of $\Psi = (\mu, \sigma)$ can be constructed based on the assumption of asymptotic normality of the maximum likelihood estimators (CASELLA; BERGER, 2002).

4 Zero or One-inflated New unit-Lindley Regression Model

In this section, we establish a regression structure for the ZINUL and OINUL distributions based on a more general term given by Equation 5. Although these models can be incorporated into generalized linear models, since they belong to the exponential family, we chose to define the zero-inflated New unit-Lindley Regression Model (ZINUL-RM) and the one-inflated New unit-Lindley Regression Model (OINUL-RM) as members of a more general family known as Generalized Additive Models for Location Scale and Shape (GAMLSS) (RIGBY; STASINOPOULOS, 2005). That being considered, and once being aware of the continuous-discrete nature of these distributions, we chose the parameterization $\Psi = (\mu, \sigma)$, which allows to model the continuous and discrete components of these models separately.

Let $Y_1, Y_2, \ldots, Y_n$ be independent random variables such that, for every $i = 1, \ldots, n$, Y_i follows a NUL distribution inflated at c with conditional mean μ_i and probability mass σ_i. In those conditions, the NUL regression model inflated at c is defined by the systematic components

$$h_1(\mu_i) = \sum_{j=1}^{n_1} u_{ij}^{\intercal}\beta_j \text{ e } h_2(\sigma_i) = \sum_{j=1}^{n_2} v_{ij}^{\intercal}\gamma_j,$$

where $u_i = (u_{i1}, u_{i2}, \ldots, u_{in_1})$ and $v - i = (v_{i1}, v_{i2}, \ldots, v_{in_2})$ are the set of fixed covariates $(n_1 + n_2 < n)$, vectors $\beta = (\beta_1, ..., \beta_{n_1})$ and $\gamma = (\gamma_1, ..., \gamma_{n_2})$ are unknown regression coefficients, and $h_1 : (0,1) \mapsto (-\infty, \infty)$ and $h_2 : (0,1) \mapsto (-\infty, \infty)$ are continuous link functions, strictly monotonic and twice differentiable at μ and σ. Some of the potential specifications of $h_1(\cdot)$ and $h_2(\cdot)$ are logit, probit, log-log and complementary log-log functions, which are the most usual. From the classical perspective, β and γ coefficient estimators are, respectively, the solutions for the following equations

$$0 = \frac{\partial}{\partial\sigma_i}\dot{\ell}_1(\gamma \mid \mathbf{y})\frac{\partial}{\partial\gamma_j}\sigma_i$$

and

$$0 = \frac{\partial}{\partial\mu_i}\dot{\ell}_2(\beta \mid \mathbf{y})\frac{\partial}{\partial\beta_j}\mu_i$$

where

$$\dot{\ell}_1(\gamma \mid \mathbf{y}) = \sum_{i=1}^{n} (1 - \delta_c(y_i)) \log(1 - \sigma_i) + \delta_c(y_i) \log(\sigma_i)$$

and

$$\dot{\ell}_2(\beta \mid \mathbf{y}) = \sum_{\substack{i=1 \\ y_i \in (0,1)}} 2\log(\mu_i) - \log(1 - \mu_i) - \frac{(1 - y_i)}{y_i}\mu_i(1 - \mu_i)^{-1}.$$

5 The Zero-and-One inflated New unit-Lindley distribution

To add zeros and ones to the NUL distribution, we used the Bernoulli distribution mixing model, thus, assigning positive probabilities to the extremes of the interval. Thus, the cumulative distribution and probability density functions of this mixing are given, respectively, by

$$\mathrm{FZOINUL}(y \mid \mu, \sigma, \rho) = \sigma\,\mathrm{Ber}(y \mid \rho) + (1 - \sigma)\mathrm{F}(y \mid \mu), \tag{13}$$

and

$$\text{fzoinul}(y \mid \mu, \sigma, \rho) = \sigma\,\text{ber}(y \mid \rho) + (1-\sigma)\text{f}(y \mid \mu) \tag{14}$$

where $\text{Ber}(\cdot \mid \rho)$ and $\text{ber}(\cdot \mid \rho)$ denote, respectively, the c.d.f. and p.d.f. of a Bernoulli distribution with probability of success ρ, whereas $\text{F}(\cdot \mid \mu)$ and $\text{f}(\cdot \mid \mu)$ are the NUL c.d.f. and p.d.f., which are respectively given by Equations 3 and 4.

Definition 5.1. A Y random variable is said to be a Zero-and-One-inflated New unit-Lindley distribution, denoted by $Y \sim \text{ZOINUL}(\mu, \sigma, \rho)$, if their c.d.f and p.d.f. are expressed by Equations 13 and 14.

Alternatively to Equation 14, we can express the ZOINUL p.d.f as a function defined by branches given by

$$\text{fzoinul}(y \mid \mu, \sigma, \rho) = \begin{cases} \sigma(1-\rho) & \text{if } y = 0 \\ (1-\sigma)\text{f}(y \mid \mu) & \text{if } 0 < y < 1 \\ \sigma\rho & \text{if } y = 1 \end{cases}, \tag{15}$$

Where the $\sigma(1-\rho)$ and $\sigma\rho$ values respectively correspond to the probabilities of observing zeros and ones. Conversely, if $Y \sim \text{ZOINUL}(\mu, \sigma, \rho)$, then, $\mathbb{E}[Y^r] = \mu_r^{'} = \mathbb{E}[(1-\sigma)\,X^r + \sigma\,W^r] = (1-\sigma)\,\mathbb{E}[X^r] + \sigma\,\mathbb{E}[W^r] = (1-\sigma)\,\mu_r + \sigma\rho$, where $r = 1, 2, 3, \ldots$ corresponds to the order of the moment assessed, X corresponds to a NUL random variable and W to a Bernoulli random variable. Particularly, for $r = 1, 2$, the mean and variance of the ZOINUL model are given, respectively, by

$$\begin{aligned} \mu_1^{'} &= (1-\sigma)\,\mu + \sigma\rho \\ \mu_2^{'} &= (1-\sigma)\left[\mu^2\,\mathbf{d}(\mu)(1-\mu)^{-1}\right] + \sigma\rho \end{aligned}$$

where $\mathbf{d}(\mu) = \text{Ei}[1,\, \mu(1-\mu)^{-1}]\,\exp\left(\mu(1-\mu)^{-1}\right)$ and $\text{Ei}[a, z] = \int_1^{\infty} z^{-a} e^{-yz}$ dz is the exponential integral function (ABRAMOWITZ; STEGUN, 1974).

As with the ZINUL and OINUL distributions, the ZOINUL distribution is a mixing model of a continuous and discrete nature, and it has points of discontinuity in the probability masses at zero and one. Its density, with expanded support at zero and one points, can take many forms, and it is multimodal when $\mu < 3/4$, $\sigma = \mathbf{b}(\mu)/(\mathbf{b}(\mu) + \mu/2)$ and $\rho = 1/2$, where $\mathbf{b}(\mu) = 27\exp\left[(4\mu-3)(1-\mu)^{-1}\right](\mu-1)^2$.

Proposition 5.2. *The ZOINUL distribution belongs to the multiparametric exponential family.*

Proof. In fact, since $\eta = (\eta_1, \eta_2, \eta_3)$, with $\eta_1 = \log(\sigma(1-\rho)) - \log(1-\sigma) - \log\left(\dfrac{\mu^2}{1-\mu}\right)$, $\eta_2 = \log(\sigma\rho) - \log(1-\sigma) - \log\left(\dfrac{\mu^2}{1-\mu}\right)$ and $\eta_3 = -\dfrac{\mu}{1-\mu}$; and $T(y) = (t_1(y), t_2(y), t_3(y))$, with $t_1(y) = \delta_0(y)$, $t_2(y) = \delta_1(y)$ and $t_3(y) = \dfrac{1-y}{y}$ if

$y \in (0,1)$ and equal to 0 otherwise, we obtain Equation 15 , as follows

$$\text{fzoinul}(y \mid \mu, \sigma, \rho) = \exp\{\eta^\top T(y) - \zeta(\eta)\}\, h(y)\,,$$

where $\zeta(\eta) = \eta_2 - \log(\sigma\rho)$ and $h(y) = y^{-3}$ if $y \in (0,1)$ and equal to 1 otherwise. $\square$

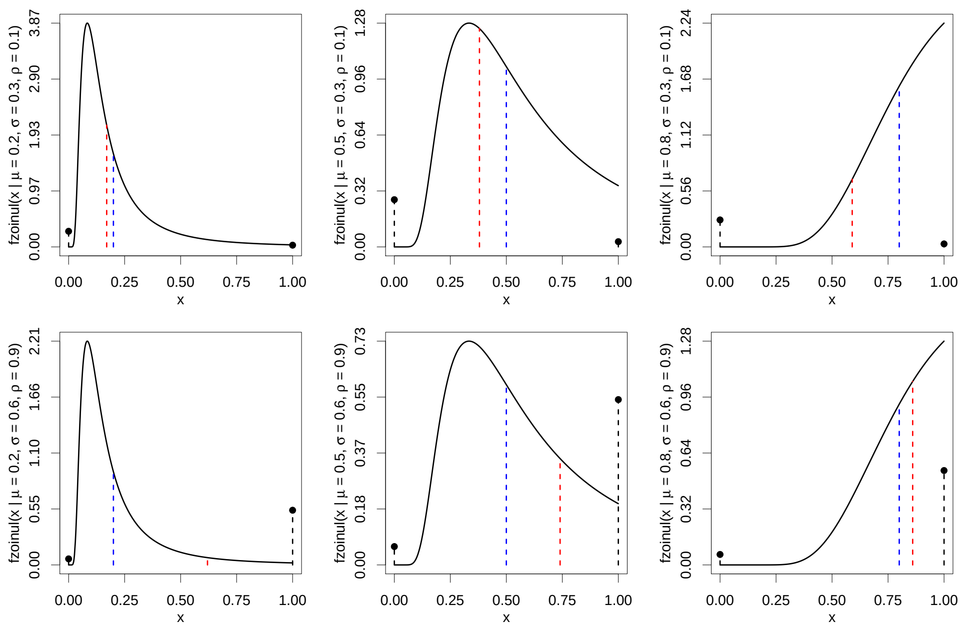

Figure 3 - Behavior of the ZOINUL density assuming different values for the parameters vector (black dotted line: proportion of zeros or ones; blue dotted line: mean of the New unit-Lindley distribution; red dotted line: marginal mean of the mixing model).

The ZOINUL distribution also has a closed expression for the quantile function. Having said that, for $0 < p < 1$, the respective quantiles are given by

$$\begin{cases} 0 & \text{if } \sigma(1-\rho) \geq p \\ 1 & \text{if } (1-\sigma\rho) \leq p \\ \dfrac{\mu}{\left(W_{-1}\left[-\dfrac{(p-\sigma(1-\rho))\,\mathbf{e}(\mu)}{(1-\sigma)}\right]+1\right)(-1+\mu)} & \text{otherwise} \end{cases}$$

where $\mathbf{e}(\mu) = (1-\mu)^{-1}\exp(-(1-\mu)^{-1})$ and W_{-1} denote a negative branch of the Lambert $\mathtt{W}$ function (CORLESS et al., 1996; JODRÁ, 2010).

6 Inferential Aspects of the Zero-and-one-inflated New unit-Lindley distribution

Let $Y = Y_1, Y_2 \ldots, Y_n$ be the random sample taken from a population with p.d.f. given by Equation 15. Without loss of generality, the density of the Y variable can be expressed as follows

$$\text{fzoinul}(y \mid \mu, \sigma, \rho) = (\sigma(1-\rho))^{\delta_0(y)} (\sigma\rho)^{\delta_1(y)} (1-\sigma)^{1-\delta_{\{0,1\}}(y)} \text{f}(y \mid \mu)^{1-\delta_{\{0,1\}}(y)},$$

where $\text{fzoinul}(y \mid \mu, \sigma, \rho)$, as a function of the parameters, can be factored into two independent terms. The first one depends only on σ and ρ, and the second depends only on μ. In such conditions, for $\mathbf{y} = (y_1, y_2, \ldots, y_n)$, the likelihood function of $\Theta = (\mu, \sigma, \rho)$ is given by

$$L(\Theta \mid \mathbf{y}) = L_1(\sigma, \rho \mid \mathbf{y}) \times L_2(\mu \mid \mathbf{y}), \tag{16}$$

where

$$L_1(\sigma, \rho \mid \mathbf{y}) = \prod_{i=1}^{n} (\sigma(1-\rho))^{\delta_0(y_i)} (\sigma\rho)^{\delta_1(y_i)} (1-\sigma)^{1-\delta_{\{0,1\}}(y_i)}$$

and

$$L_2(\mu \mid \mathbf{y}) = \prod_{i=1}^{n} \left\{ \frac{\mu^2}{(1-\mu)\, y_i^3} \exp\left(-\frac{\mu\,(1-y_i)}{(1-\mu)\, y_i} \right) \right\}^{1-\delta_{\{0,1\}}(y_i)}.$$

The logarithm of Equation 16 is given by

$$\ell(\Theta \mid \mathbf{y}) = \ell_1(\sigma, \rho \mid \mathbf{y}) + \ell_2(\mu \mid \mathbf{y}), \tag{17}$$

where

$$\begin{aligned} \ell_1(\sigma, \rho \mid \mathbf{y}) &= \mathbf{s_0}(\mathbf{y}) \log(\sigma(1-\rho)) + (n - \mathbf{s_0}(\mathbf{y}) - \mathbf{s_1}(\mathbf{y})) \log(1-\sigma) + \\ &\quad \mathbf{s_1}(\mathbf{y}) \log(\sigma\rho) \\ \ell_2(\mu \mid \mathbf{y}) &= (n - \mathbf{s_0}(\mathbf{y}) - \mathbf{s_1}(\mathbf{y})) \log\left(\frac{\mu^2}{(1-\mu)} \right) + \mathbf{r}(\mathbf{y}) - \frac{\mu}{(1-\mu)} \mathbf{t}(\mathbf{y}) \end{aligned}$$

with $\mathbf{s_0}(\mathbf{y}) = \sum_{i=1}^{n} \delta_0(y_i)$, $\mathbf{s_1}(\mathbf{y}) = \sum_{i=1}^{n} \delta_1(y_i)$, $\mathbf{r}(\mathbf{y}) = -3 \sum_{\substack{i=1 \\ y_i \in (0,1)}} \log(y_i)$ and $\mathbf{t}(\mathbf{y}) = \sum_{\substack{i=1 \\ y_i \in (0,1)}} \left(\frac{1-y_i}{y_i} \right)$.

The MLE $\widehat{\Theta} = (\widehat{\mu}, \widehat{\sigma}, \widehat{\rho})$ of Θ is given by

$$\widehat{\mu} = \frac{1}{2\,\dot{\mathbf{s}}(\mathbf{y})} \left(3\,\dot{\mathbf{s}}(\mathbf{y}) + \mathbf{t}(\mathbf{y}) + \sqrt{\mathbf{v}(\mathbf{y})}\right),$$
$$\widehat{\sigma} = \frac{\mathbf{s_0}(\mathbf{y}) + \mathbf{s_1}(\mathbf{y})}{n},$$
$$\widehat{\rho} = \frac{\mathbf{s_1}(\mathbf{y})}{\mathbf{s_0}(\mathbf{y}) + \mathbf{s_1}(\mathbf{y})},$$

where $\mathbf{v}(\mathbf{y}) = n^2 + 2n\,[3\mathbf{t}(\mathbf{y}) - \mathbf{s_0}(\mathbf{y}) - \mathbf{s_1}(\mathbf{y})] + \mathbf{t}(\mathbf{y})^2 - 6\mathbf{t}(\mathbf{y})\,[\mathbf{s_0}(\mathbf{y}) + \mathbf{s_1}(\mathbf{y})] + [\mathbf{s_0}(\mathbf{y}) + \mathbf{s_1}(\mathbf{y})]^2$ and $\dot{\mathbf{s}}(\mathbf{y}) = n - \mathbf{s_0}(\mathbf{y}) - \mathbf{s_1}(\mathbf{y})$.

The second order derivatives of Equation 17 with respect to the Θ parameters vector are given by

$$\frac{\partial^2}{\partial\mu^2}\,\ell(\Theta \mid \mathbf{y}) = \frac{\mu^3\,\dot{\mathbf{s}}(\mathbf{y}) - \mu^2\,[5\,\dot{\mathbf{s}}(\mathbf{y}) + 2\,\mathbf{t}(\mathbf{y})] + 6\,\mu\,\dot{\mathbf{s}}(\mathbf{y}) - 2\,\dot{\mathbf{s}}(\mathbf{y})}{\mu^2\,(1-\mu)^3},$$
$$\frac{\partial^2}{\partial\sigma^2}\,\ell(\Theta \mid \mathbf{y}) = \frac{-n\,\sigma^2 + 2\,\sigma\,[\mathbf{s_0}(\mathbf{y}) + \mathbf{s_1}(\mathbf{y})] - \mathbf{s_0}(\mathbf{y}) - \mathbf{s_1}(\mathbf{y})}{\sigma^2\,(1-\sigma)^2},$$
$$\frac{\partial^2}{\partial\rho^2}\,\ell(\Theta \mid \mathbf{y}) = -\frac{\mathbf{s_0}(\mathbf{y})}{(1-\rho)^2} - \frac{\mathbf{s_1}(\mathbf{y})}{\rho^2},$$
$$\frac{\partial^2}{\partial\mu\,\partial\sigma}\,\ell(\Theta \mid \mathbf{y}) = 0,$$
$$\frac{\partial^2}{\partial\mu\,\partial\rho}\,\ell(\Theta \mid \mathbf{y}) = 0,$$
$$\frac{\partial^2}{\partial\sigma\,\partial\rho}\,\ell(\Theta \mid \mathbf{y}) = 0.$$

Thus, the expected Fisher information matrix is given by

$$\mathbf{K}(\Theta) = \begin{bmatrix} k_{\mu\,\mu} & 0 & 0 \\ 0 & k_{\sigma\,\sigma} & 0 \\ 0 & 0 & k_{\rho\,\rho} \end{bmatrix}$$

where $k_{\mu\,\mu} = n\,(2-\mu^2)\,(1-\sigma)\,\mu^{-2}\,(1-\mu)^{-2}$, $k_{\sigma\,\sigma} = n\,\sigma^{-1}\,(1-\sigma)^{-1}$ and $k_{\rho\,\rho} = n\,\sigma\,\rho^{-1}\,(1-\rho)^{-1}$. Confidence intervals and hypothesis testing of $\Theta = (\mu, \sigma, \rho)$ are constructed based on the assumption of asymptotic normality of the maximum likelihood estimators (CASELLA; BERGER, 2002).

7 Zero-and-One-inflated New unit-Lindley Regression Model

In this section, we establish a regression structure for the ZOINUL distribution. Like ZINUL and OINUL, the ZOINUL distribution can be incorporated into generalized linear models. Yet, in this section, we define the Zero-and-One-inflated

New unit-Lindley Regression Model (ZOINUL-RM) as a member of the GAMLSS family.

Let $Y_1, Y_2, \ldots, Y_n$ be independent random variables such that, for every $i = 1, \ldots, n$, Y_i follows a ZOINUL distribution, with parameters μ_i, σ_i and ρ_i. In such conditions, the ZOINUL-RM is defined by the systematic components

$$h_1(\mu_i) = \sum_{j=1}^{n_1} u_{ij}\beta_j,\; h_2(\sigma_i) = \sum_{j=1}^{n_2} v_{ij}\gamma_j \text{ and } h_3(\rho_i) = \sum_{j=1}^{n_3} w_{ij}\lambda_j,$$

where $u_i = (u_{i1}, u_{i2}, \ldots, u_{in_1})$, $v_i = (v_{i1}, v_{i2}, \ldots, v_{in_2})$ and $w_i = (w_{i1}, w_{i2}, \ldots, w_{in_3})$ are the set of fixed covariates $(n_1 + n_2 + n_3 < n)$, the vectors $\beta = (\beta_1, ..., \beta_{n_1})$, $\gamma = (\gamma_1, ..., \gamma_{n_2})$ and $\lambda = (\lambda_1, ..., \lambda_{n_3})$ are unknown regression coefficients, and $h_1 : (0,1) \mapsto (-\infty, \infty)$, $h_2 : (0,1) \mapsto (-\infty, \infty)$ and $h_3 : (0,1) \mapsto (-\infty, \infty)$ are continuous link functions, strictly monotonic, twice differentiable at μ, σ and ρ.

Again through a classic approach, the estimator of β is the solution for the equation

$$0 = \frac{\partial}{\partial \mu_i} \ddot{\ell}_2(\beta \mid \mathbf{y}) \frac{\partial}{\partial \beta_j} \mu_i \,,$$

While the estimators of γ and λ are solutions for the equations system

$$\begin{aligned} 0 &= \frac{\partial}{\partial \sigma_i} \ddot{\ell}_1(\gamma, \lambda \mid \mathbf{y}) \frac{\partial}{\partial \gamma_j} \sigma_i \\ 0 &= \frac{\partial}{\partial \rho_i} \ddot{\ell}_1(\gamma, \lambda \mid \mathbf{y}) \frac{\partial}{\partial \lambda_j} \rho_i \end{aligned}$$

where

$$\begin{aligned} \ddot{\ell}_1(\gamma, \lambda \mid \mathbf{y}) = \sum_{i=1}^{n} & \delta_0(y_i) \log(\sigma_i(1-\rho_i)) + \delta_1(y_i) \log(\sigma_i \rho_i) + (1 - \delta_{\{0,1\}}(y_i)) \\ & \times \log(1 - \sigma_i) \end{aligned}$$

and

$$\ddot{\ell}_2(\beta \mid \mathbf{y}) = \sum_{\substack{i=1 \\ y_i \in (0,1)}} 2\log(\mu_i) - \log(1-\mu_i) - \frac{(1-y_i)}{y_i} \mu_i (1-\mu_i)^{-1}.$$

8 Simulation Study

In this section, we conduct a Monte Carlo simulation study to evaluate the performance of the regression coefficients of the zero, one and zero-and-one-inflated NUL regression models. The experiment was carried out based on 10000 replicates for the sample sizes $n = 50, 100, \ldots, 300$. We assessed the bias and mean square error (MSE), respectively given by

$$\text{Bias}\,(\widehat{\xi}) = \frac{1}{N}\sum_{i=1}^{N}(\widehat{\xi}_i - \xi) \quad \text{and} \quad \text{MSE}\,(\widehat{\xi}) = \frac{1}{N}\sum_{i=1}^{N}(\widehat{\xi}$$

where ξ denotes the coefficient assessed in $N = 10000$ simulations.

We considered, as scenarios, the multiple regression models (with two covariates) over all the distribution parameters and over the conditional mean parameter, keeping the probabilities masses fixed. We chose the logit specification for the link functions, so that we are left with the systematic components

$$\begin{aligned}
\text{logit}(\mu_i) &= \beta_0 + \beta_1 u_{i1} + \beta_2 u_{i2} \\
\text{logit}(\sigma_i) &= \gamma_0 + \gamma_1 u_{i1} + \gamma_2 u_{i2}
\end{aligned}$$

and

$$\text{logit}(\rho_i) = \lambda_0 + \lambda_1 u_{i1} + \lambda_2 u_{i2}$$

where u_{i1} and u_{i2} are elements of the vector of variables generated by a uniform distribution, from -1 to 1, and the regression coefficients given by $\beta = (\beta_0, \beta_1, \beta_2) = \delta = (\delta_0, \delta_1, \delta_2) = \lambda = (\lambda_0, \lambda_1, \lambda_2) = (1, -2, 3)$.

6 and 7 refer to the ZINUL-RM, Figures 8, 9, 10 and 11 refer to the OINUL-RM, and Figures 12, 13, 14 and 15 relate to the ZOINUL-RM.

In a general context, the MSEs converge to zero for all models as the sample size increases, as shown in Figures 5, 7, 9, 11, 13 and 15. 7, 11 and 15 show that the MSE is influenced by the σ mixing parameter, since it decays along with the σ value. This result can be directly explained by the property of maximum likelihood estimators, since the decrease of σ implies a larger sub-sample resulting from the NUL distribution and, consequently, a smaller MSE associated with the estimate of μ.

10, 12 and 14 show an oscillation around zero with a tendency to zero as the sample size increases. In all scenarios, $\widehat{\beta}_0$ superestimates (is greater than) β_0 and, on average, has a bias value greater than those of $\widehat{\beta}_1$ and $\widehat{\beta}_2$.

The results of this study are presented as follows. The Figures 4, 5, 6 and 7 refer to the simulations performed in ZINUL- RM, the Figures 8, 9, 10 and 11 to OINUL-RM and the Figures 12, 13, 14 and 15 refer to ZOINUL-RM.

In this study it is observed that

- For all scenarios, the biases and MSE's converge to zero as the n sample size increases.

- In the regression models on the conditional mean parameter μ keeping the others fixed, Figures 6, 7, 10, 11, 14 and 15, the bias and MSE of $\widehat{\beta}$ increases as the blend parameter σ increases. This possibly occurs because increasing σ implies smaller subsamples of elements that follow a NUL distribution.

- In regression models over all parameters, the intercept usually has the lowest value for the MSE. See Figure 5, 9 and 13.

- The MSE of $\widehat{\beta}$ is generally smaller than $\widehat{\gamma}$ mainly for the smaller sample sizes considered in this study. See Figures 5, 9 and 13.

- In ZOINUL-RM, with σ and ρ fixed, the parameter ρ did not have a great influence on the bias and MSE of $\widehat{\beta}$. See Figures 14 and 15.

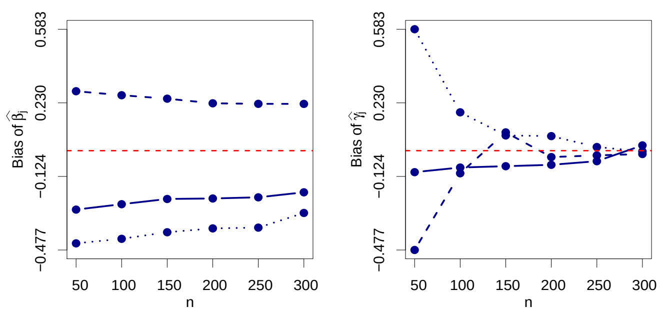

Figure 4 - Bias of the $\widehat{\beta}$ and $\widehat{\gamma}$ estimates of the regressors coefficients of the parameters vector in the ZINUL-RM (continuous line: j= 0, dashed lined: j= 1 and dotted line: j= 2).

9 Application

In this section, we adopt the ZINUL-RM to describe the behavior of the annual suicide rate in 2016 for countries in the African, Asian and European continents, taking into account information regarding sex, age group and socioeconomic indices of the respective countries, such as unemployment rate, gross domestic product (GDP) and human development index (HDI). The data used in this analysis are public domain data available on <https://www.who.int/>, <http://hdr.undp.org/en> and <https://databank.worldbank.org/source/world-development-indicators>.

In Table 1, we provide a brief description of the variables of numerical nature, and the GDP is presented through the logarithmic scale (neperian logarithm). Note

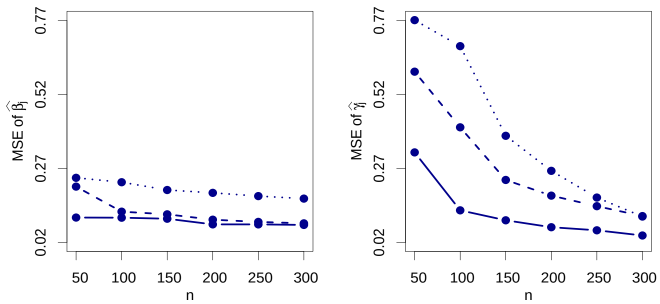

Figure 5 - MSE of the $\widehat{\beta}$ and $\widehat{\gamma}$ estimates of the regressors coefficients of the parameters vector in the ZINUL-RM (continuous line: j= 0, dashed lined: j= 1 and dotted line: j= 2).

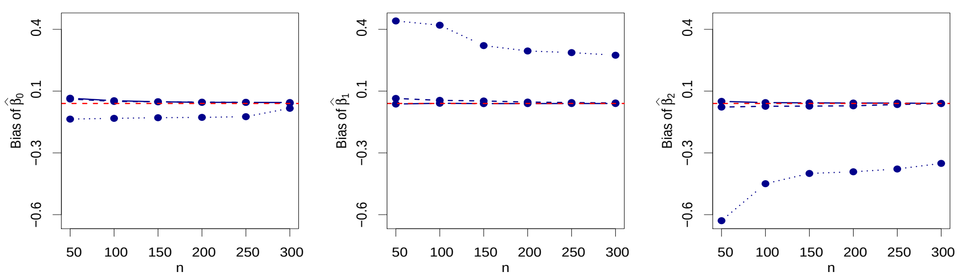

Figure 6 - Bias of the regressors coefficients estimates of μ, $\widehat{\beta} = (\widehat{\beta_0}, \widehat{\beta_1}, \widehat{\beta_2})$, in ZINUL-RM keeping σ fixed (continuous line: $\sigma = 0.2$, dashed line $\sigma = 0.5$ and dotted line: $\sigma = 0.8$).

that, compared to the other variables, the suicide rate is dimensionally much lower, which led us to work on the GDP in the logarithmic scale. In total, there were 140 observations (being 22 zeros) on the suicide rate in groups of individuals aged 15 to 24 (age 0); 25 to 34 (age 1); 35 to 54 (age 2); 55 to 74 (age 3), disaggregated by sex (female: 0 and male: 1). Mostly, Asian and European peoples constitute about 93% of the observations available on this database. For comparison purposes, we also consider the zero-inflated unit-Lindley (ZIUL-RM) and zero-inflated Beta (ZIBE-RM) regression models. The adjustments and related calculations were performed by using functions available on the `gamlss` (RIGBY; STASINOPOULOS, 2005) and `gamlss.inf` (ENEA et al., 2019) libraries, with the `R` software (R Core Team, 2019). We considered the logit specification for the binding functions. The regression models for the μ_i, σ_i and ϕ_i parameters are respectively given by the following

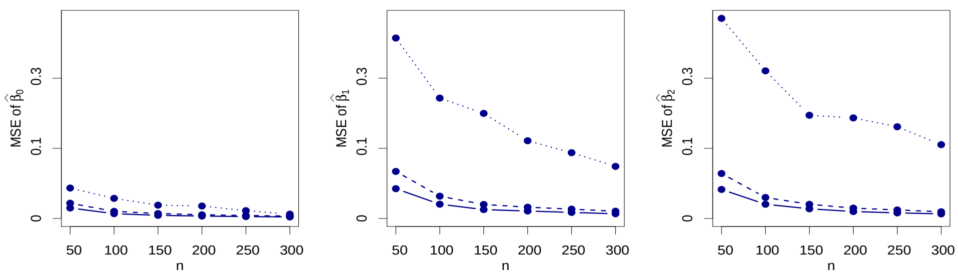

Figure 7 - MSE of the estimates of the regressors coefficients of μ, $\widehat{\beta} = (\widehat{\beta_0}, \widehat{\beta_1}, \widehat{\beta_2})$, in the ZINUL-RM keeping σ fixed (continuous line: $\sigma = 0.2$, dashed line $\sigma = 0.5$ and dotted line: $\sigma = 0.8$).

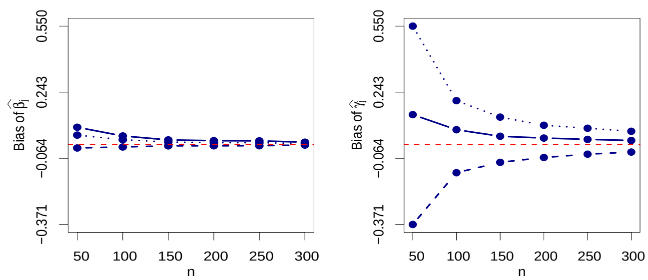

Figure 8 - Bias of the $\widehat{\beta}$, and $\widehat{\gamma}$ estimates of the regression coefficients of the parameters vector in the OINUL-RM (continuous line: j= 0, dashed lined: j= 1 and dotted line: j= 2).

components

$$\begin{aligned}\text{logit}(\mu_i) &= \beta_0 + \beta_1 \times \text{sex } 1_i + \beta_2 \times \text{age } 1_i + \beta_3 \times \text{age } 2_i + \beta_4 \times \text{age } 3_i \\ &+ \beta_5 \times \text{age } 4_i + \beta_6 \times \text{unemployment}_i + \beta_7 \times \text{HDI}_i\end{aligned}$$

and

$$\text{logit}(\sigma_i) = \gamma_0 + \gamma_1 \times \text{sex } 1_i + \gamma_2 \times \log(\text{GPD}_i) + \gamma_3 \times \text{unemployment}_i$$

Table 3 presents a detailed description with the estimates, standard errors and confidence intervals of the adjusted parameters. Note that the set of variables present in this analysis was not significant at the 5% level to explain the conditional mean of the ZIBE-RM. On the other hand, the variables sex, age group, unemployment rate, and HDI were significant at the 5% level for the ZINUL-RM and ZIUL-RM. Conversely, with regard to the probability mass at zero, the estimates

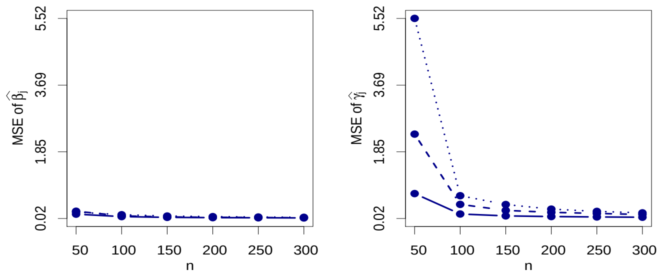

Figure 9 - MSE of the $\widehat{\beta}$ and $\widehat{\gamma}$ estimates of the regression coefficients of the parameters vector in the OINUL-RM (continuous line: j$=$ 0, dashed lined: j$=$ 1 and dotted line: j$=$ 2).

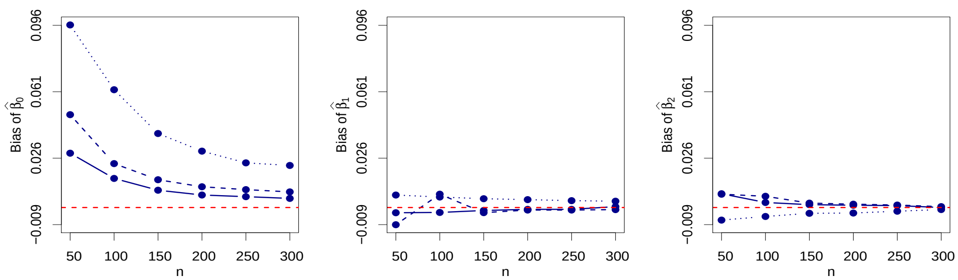

Figure 10 - Bias of the estimates of the regressors coefficients of μ, $\widehat{\beta} = (\widehat{\beta_0}, \widehat{\beta_1}, \widehat{\beta_2})$, in the OINUL-RM keeping σ fixed (continuous line: $\sigma = 0.2$, dashed line $\sigma = 0.5$ and dotted line: $\sigma = 0.8$).

are the same for the three models, since they have the same inflation structure. The significant variables at the 5% level, in this case, were sex and unemployment rate, and the latter was significant only in the presence of the GDP variable.

Table 2 shows the Akaike (AIC) and Schwarz Bayesian (BIC) information criteria used later to discriminate the models. Note that the ZINUL-RM has the lowest (best) value for the considered criteria. However, these values apparently do not show a considerable difference when compared to the AIC and BIC of the ZIUL-RM. Thus, a residual analysis of each model allowed us to measure and determine the model with the best fit for the data in question.

Figure 16 presents the construction of worm plot graphs that assess the deviation of residues from the quantiles of a standard normal distribution. In the ideal scenario, the deviations should follow, as close as possible, the red horizontal line within the bands bordered by the black dotted lines. In such conditions, the ZINUL-RM has the best fit, since its residues are distributed in a more satisfactory way within the area delimited by the bands.

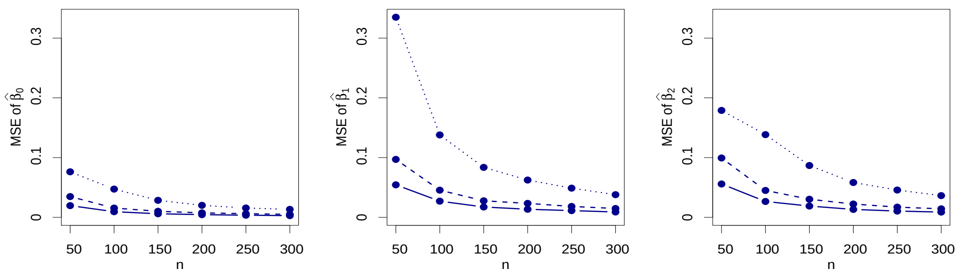

Figure 11 - MSE of the estimates of the regressors coefficients of μ, $\widehat{\beta} = (\widehat{\beta_0}, \widehat{\beta_1}, \widehat{\beta_2})$, in the OINUL-RM keeping σ fixed (continuous line: $\sigma = 0.2$, dashed line $\sigma = 0.5$ and dotted line: $\sigma = 0.8$).

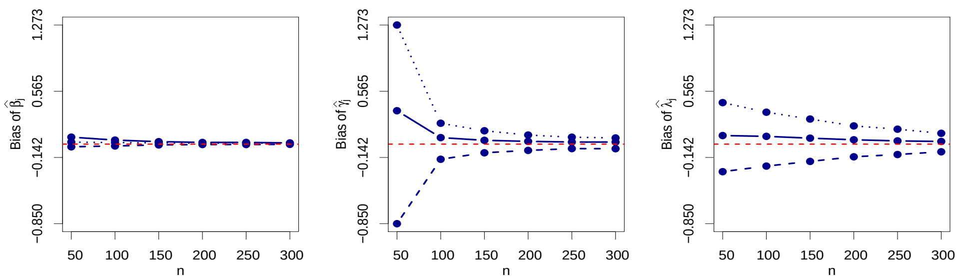

Figure 12 - Bias of the $\widehat{\beta}$, $\widehat{\gamma}$ and $\widehat{\lambda}$ estimates of the regressors coefficients of the parameters vector in the ZOINUL-RM (continuous line: j= 0, dashed lined: j= 1 and dotted line: j= 2).

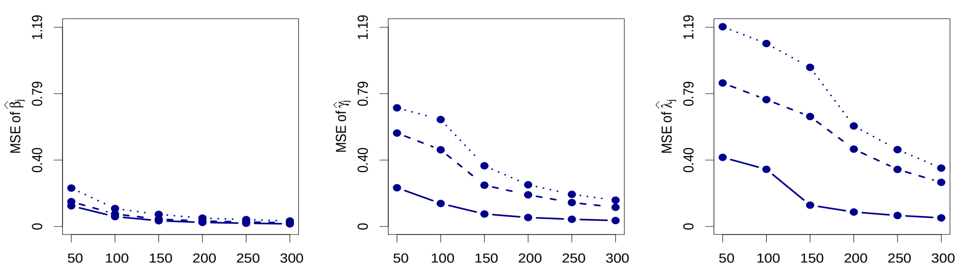

Figure 13 - MSE of the $\widehat{\beta}$, $\widehat{\gamma}$ and $\widehat{\lambda}$ estimates of the regressors coefficients of the parameters vector in the ZOINUL-RM (continuous line: j= 0, dashed lined: j= 1 and dotted line: j= 2).

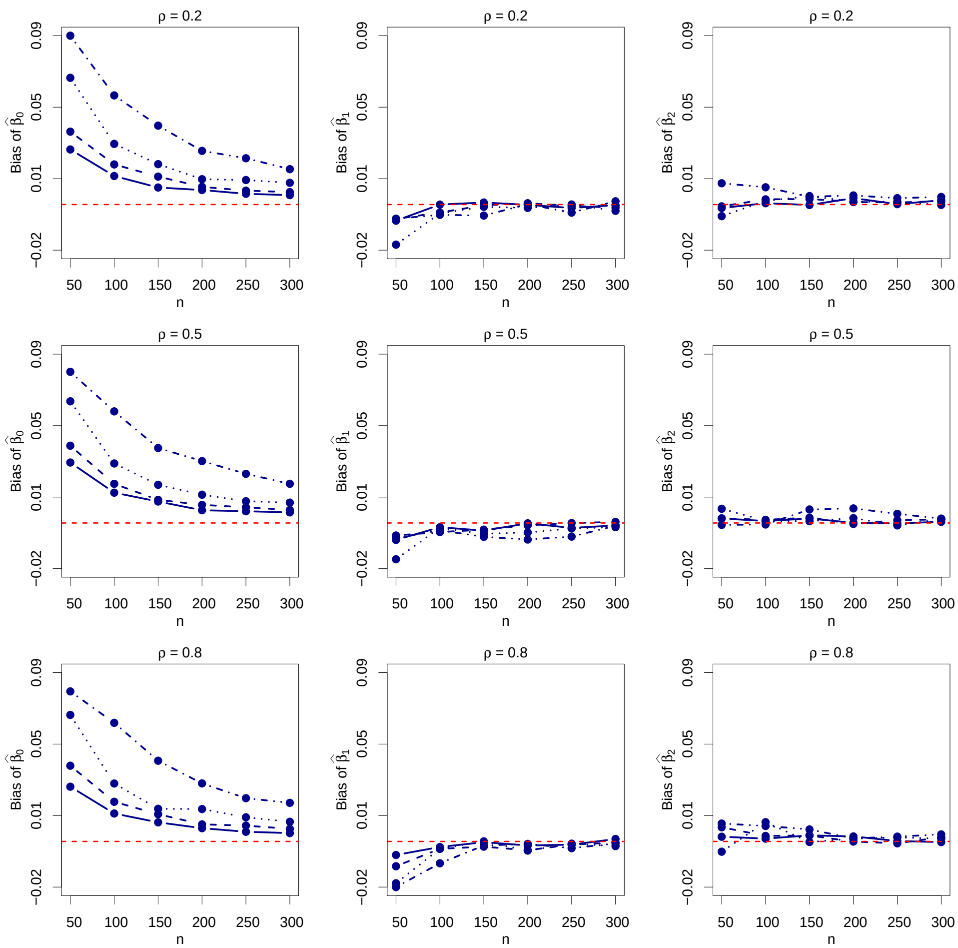

Figure 14 - Bias of the estimates of the regressors coefficients of μ, $\widehat{\beta} = (\widehat{\beta_0}, \widehat{\beta_1}, \widehat{\beta_2})$, in the ZOINUL-RM keeping σ and ρ fixed (continuous line: $\sigma = 0.2$, dashed line $\sigma = 0.4$, dotted line: $\sigma = 0.6$ and dash-dotted line: $\sigma = 0.8$).

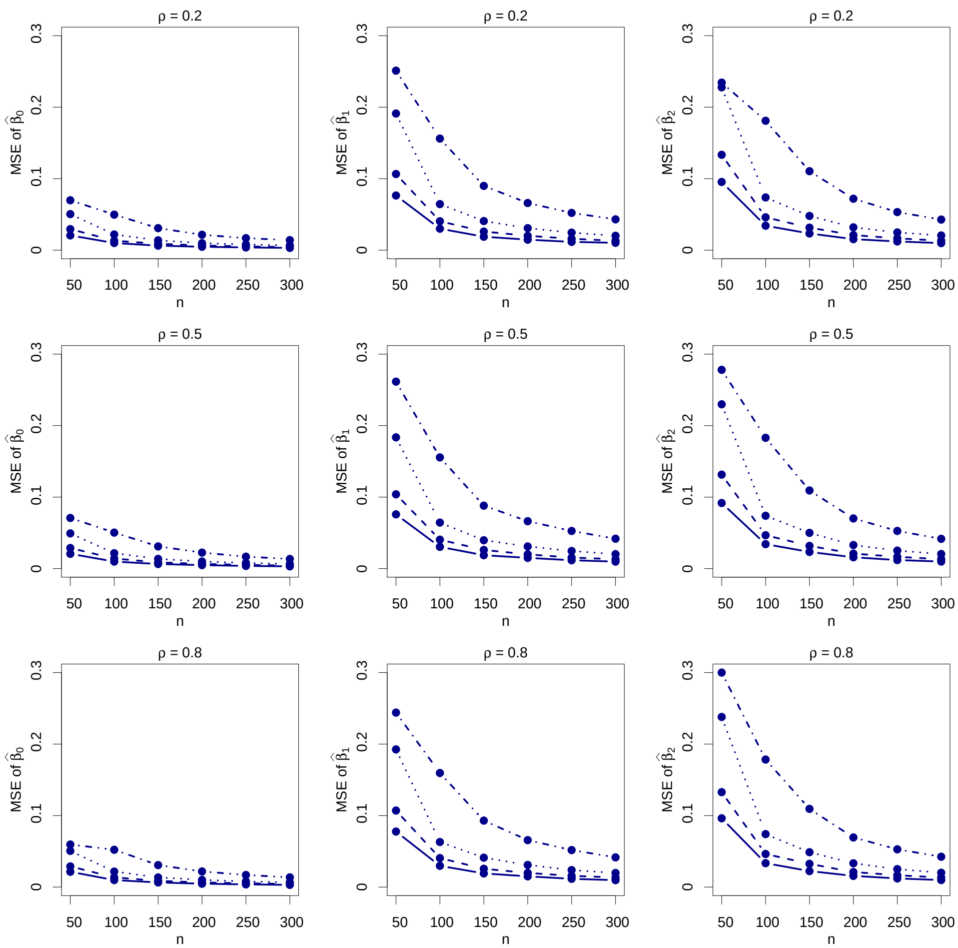

Figure 15 - MSE of the estimates of the regressors coefficients of μ, $\widehat{\beta} = (\widehat{\beta_0}, \widehat{\beta_1}, \widehat{\beta_2})$, in the ZOINUL-RM keeping σ and ρ fixed (continuous line: $\sigma = 0.2$, dashed line $\sigma = 0.4$, dotted line: $\sigma = 0.6$ and dash-dotted line: $\sigma = 0.8$).

Table 1 - Statistical summary of numerical variables

	Suicide Rate	log(GDP)	Unemployment Rate	HDI
Minimum	0	23.08	0.15	0.73
1º Quartile	2.5×10^{-5}	23.23	6.01	0.77
Median	7.3×10^{-5}	24.57	7.12	0.84
Mean	1.2×10^{-4}	24.90	8.71	0.84
3º Quartile	1.5×10^{-4}	26.69	12.95	0.87
Maximum	9.7×10^{-4}	27.38	17.62	0.94

Table 2 - Models Discrimination Statistics

Model	AIC	BIC
ZINUL-RM	-2119.50	-2084.20
ZIUL-RM	-2094.20	-2058.90
ZIBE-RM	-1817.03	-1778.79

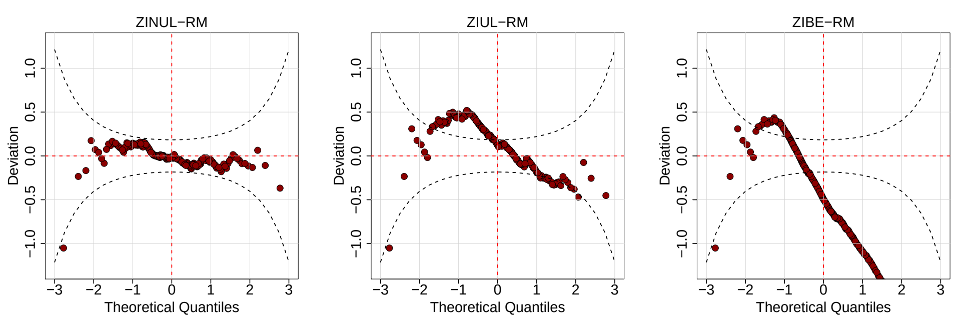

Figure 16 - Worm plot of residues for ZINUL-RM, ZIUL-RM and ZIBE-RM models.

Table 3 - Summary of Adjusted Models

	ZINUL-RM			ZIUL-RM			ZIBE-RM		
	Estimate	SE	CI of 95%	Estimate	SE	CI of 95%	Estimate	SE	CI of 95%
β_0	-13.841	0.886	$(-15.578\,;\,-12.104)$	-12.144	1.143	$(-14.384\,;\,-9.904)$	-7.529	1.108	$(-9.701\,;\,-5.357)$
β_1	1.216	0.126	$(0.969\,;\,1.463)$	1.472	0.176	$(1.127\,;\,1.817)$	0.310	0.169	$(-0.021\,;\,0.641)$
β_2	-0.111	0.195	$(-0.493\,;\,0.271)$	0.073	0.275	$(-0.466\,;\,0.612)$	0.002	0.267	$(-0.521\,;\,0.525)$
β_3	0.425	0.195	$(0.043\,;\,0.807)$	0.457	0.277	$(-0.086\,;\,0.100)$	0.112	0.266	$(-0.409\,;\,0.633)$
β_4	0.247	0.197	$(-0.139\,;\,0.633)$	0.561	0.281	$(0.010\,;\,1.112)$	0.110	0.268	$(-0.415\,;\,0.635)$
β_5	1.079	0.205	$(0.677\,;\,1.481)$	1.069	0.292	$(0.497\,;\,1.641)$	0.261	0.278	$(-0.284\,;\,0.806)$
β_6	-0.037	0.011	$(-0.059\,;\,-0.015)$	-0.060	0.022	$(-0.103\,;\,-0.017)$	-0.010	0.017	$(-0.043\,;\,0.023)$
β_7	4.858	1.020	$(2.859\,;\,6.857)$	2.621	1.304	$(0.065\,;\,5.177)$	0.849	1.262	$(-1.624\,;\,3.322)$
γ_0	14.390	8.191	$(-1.664\,;\,30.444)$	14.390	8.191	$(-1.664\,;\,30.444)$	14.390	8.191	$(-1.664\,;\,30.444)$
γ_1	-2.270	1.089	$(-4.404\,;\,-0.136)$	-2.270	1.089	$(-4.404\,;\,-0.136)$	-2.270	1.089	$(-4.404\,;\,-0.136)$
γ_2	-0.599	0.315	$(-1.216\,;\,0.018)$	-0.599	0.315	$(-1.216\,;\,0.018)$	-0.599	0.315	$(-1.216\,;\,0.018)$
γ_3	-0.207	0.097	$(-0.397\,;\,-0.0179)$	-0.207	0.097	$(-0.397\,;\,-0.017)$	-0.207	0.097	$(-0.397\,;\,-0.0179)$
logit(ϕ)	—	—	—	—	—	—	-2.740	0.052	$(2.638\,;\,2.842)$

10 Conclusion

In this paper, we propose inflated versions of the NUL distribution to deal with continuous unit variables in the presence of zeros, ones, as well as zeros and ones at the same time. Our distributions are viable alternatives to literary distributions that have the same support, since they have a number of important properties (such as closed forms for moments for the quantile function). In addition, they are members of the exponential family and describe, within the subinterval $(0, 1)$, a unimodal asymmetric behavior. The simulation study showed that the MLE of the μ conditional mean was influenced by the σ mixing parameter, with a lower MSE for values below σ. When applied to real data on the suicide rates in the year of 2016, which is characterized by a strong concentration very close to zero, the ZINUL-RM was the one that had the best fit when compared to the ZIUL-RM and ZIBE-RM. Finally, we believe that our models can be useful for researchers and professionals looking for probability distributions with similar characteristics.

Acknowledgments

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior – Brasil (CAPES) – Finance Code 001

FERREIRA, A. B.; MAZUCHELI, J. The Zero, One and Zero-and-one-inflated New unit-Lindley Distributions. *Braz. J. Biom.,* Lavras, v.40, n.3, p.291-326, 2022.

▪ RESUMO: Neste artigo propomos as distribuições New unit-Lindley inflacionada em zero, em um e em zero e um como extensões naturais da distribuição New unit-Lindley para modelar respostas contínuas medidas nos intervalos $[0, 1)$, $(0, 1]$ e $[0, 1]$. Estas distribuições foram construídas a partir de combinações convexas entre a distribuição New unit-Lindley e as distribuições degenerada em zero, em um e Bernoulli. Elas também contam com uma série de propriedades interessantes tais como serem membros da família exponencial além de contar com formas as funções de distribuição acumulada, quantil e para os momentos. Aspectos inferenciais e estruturas de regressão são discutidas neste trabalho bem como um estudo de simulação Monte Carlo para avaliar a performance dos coeficientes regressores. Por fim, trazemos uma aplicação a dados reais sobre a taxa de suicídio no ano de 2016.

Palavras Chaves: Modelos Inflacionados, Estimador de Máxima Verossimilhança, Simulação Monte Carlo, Modelos de Regressão.

References

ABRAMOWITZ, M.; STEGUN, I. A. *Handbook of mathematical functions with formulas, graphs, and mathematical tables.* New York: United States Department of Commerce, National Bureau of Standards, 1974.

BAPAT, S. R.; BHARDWAJ, R. On an Inflated Unit-Lindley Distribution. *arXiv preprint arXiv:2102.04687*, 2021.

CASELLA, G.; BERGER, R. L. *Statistical inference*. Califórnia: Duxbury Pacific Grove, 2002. v. 2.

CHAI, H. et al. A marginalized two-part beta regression model for microbiome compositional data. *PLoS computational biology*, v. 14, n. 7, p. e1006329, 2018.

CORLESS, R. M. et al. On the Lambert W function. *Advances in Computational Mathematics*, v. 5, n. 1, p. 329–359, 1996.

CRIBARI-NETO, F.; SANTOS, J. Inflated Kumaraswamy distributions. *Anais da Academia Brasileira de Ciências*, v. 91, n. 2, 2019.

ENEA, M. et al. *gamlss.inf: Fitting Mixed (Inflated and Adjusted) Distributions*. [S.l.], 2019. R package version 1.0-1. Disponível em: <https://CRAN.R-project.org/package=gamlss.inf>.

GHITANY, M. E. et al. The unit-inverse Gaussian distribution: A new alternative to two-parameter distributions on the unit interval. *Communications in Statistics-Theory and Methods*, v. 48, n. 14, p. 3423–3438, 2019.

JODRÁ, P. Computer generation of random variables with Lindley or Poisson-Lindley distribution via the Lambert W function. *Mathematics and Computers in Simulation*, v. 81, n. 4, p. 851–859, 2010.

LINDLEY, D. V. Fiducial distributions and bayes' theorem. *Journal of the Royal Statistical Society: Series B (Methodological)*, v. 20, n. 1, p. 102–107, 1958.

LIU, P. et al. Zero-one-inflated simplex regression models for the analysis of continuous proportion data. *Statistics and Its Interface*, v. 13, n. 2, p. 193–208, 2020.

MAZUCHELI, J.; BAPAT, S. R.; MENEZES, A. F. B. A new one-parameter unit-lindley distribution. *Chilean Journal of Statistics (ChJS)*, v. 11, n. 1, 2020.

MAZUCHELI, J.; MENEZES, A. F. B.; CHAKRABORTY, S. On the one parameter unit-Lindley distribution and its associated regression model for proportion data. *Journal of Applied Statistics*, v. 46, n. 4, p. 700–714, 2019.

MAZUCHELI, J.; MENEZES, A. F. B.; DEY, S. Unit-Gompertz distribution with applications. *Statistica*, v. 79, n. 1, p. 25–43, 2019.

MAZUCHELI, J.; MENEZES, A. F. B.; GHITANY, M. E. The unit-Weibull distribution and associated inference. *Journal of Applied Probability and Statistics*, v. 13, p. 1–22, 2018.

MENEZES, A. F.; MAZUCHELI, J.; BOURGUIGNON, M. A parametric quantile regression approach for modelling zero-or-one inflated double bounded data. *Biometrical Journal*, v. 63, n. 4, p. 841–858, 2021.

NELDER, J. A.; WEDDERBURN, R. W. Generalized linear models. *Journal of the Royal Statistical Society: Series A (General)*, Wiley Online Library, v. 135, n. 3, p. 370–384, 1972.

OSPINA, R.; FERRARI, S. L. A general class of zero-or-one inflated beta regression models. *Computational Statistics & Data Analysis*, v. 56, n. 6, p. 1609–1623, 2012.

OSPINA, R.; FERRARI, S. L. P. Inflated beta distributions. *Statistical Papers*, v. 51, n. 1, p. 111, 2010.

QUEIROZ, F. F.; LEMONTE, A. J. A broad class of zero-or-one inflated regression models for rates and proportions. *Canadian Journal of Statistics*, v. 49, n. 2, p. 566–590, 2021.

R Core Team. *R: A Language and Environment for Statistical Computing*. Vienna, Austria, 2019. Disponível em: <https://www.R-project.org/>.

RIGBY, R. A.; STASINOPOULOS, D. M. Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, v. 54, n. 3, p. 507–554, 2005.

RIVAS, L.; CAMPOS, F. Zero inflated waring distribution. *Communications in Statistics-Simulation and Computation*, p. 1–16, 2021.

SANTOS, B.; BOLFARINE, H. Bayesian analysis for zero-or-one inflated proportion data using quantile regression. *Journal of Statistical Computation and Simulation*, v. 85, n. 17, p. 3579–3593, 2015.

SANTOS, B.; BOLFARINE, H. Bayesian quantile regression analysis for continuous data with a discrete component at zero. *Statistical Modelling*, v. 18, n. 1, p. 73–93, 2018.

SILVA, A. R. et al. Augmented-limited regression models with an application to the study of the risk perceived using continuous scales. *Journal of Applied Statistics*, v. 48, n. 11, p. 1998–2021, 2021.

TOMARCHIO, S. D.; PUNZO, A. Modelling the loss given default distribution via a family of zero-and-one inflated mixture models. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, v. 182, n. 4, p. 1247–1266, 2019.

XIE, M.; HE, B.; GOH, T. Zero-inflated poisson model in statistical process control. *Computational Statistics & data Analysis*, v. 38, n. 2, p. 191–201, 2001.

Appendix

```
# Implementation of models for R software

rm(list = ls()) graphics.off()

library('lamW');
library('expint');
library('gamlss');
library('gamlss.inf');

# ---- New unit-Lindley distribution (Nul)

# if X ~ Nul(mu), then E[X] = mu.

#---- probability density function

dNul <- function(y, mu, log = FALSE){

  theta = mu/(1-mu);

  t1 = theta * theta;
  t5 = y * y;
  t12 = exp(-theta * (0.1e1 - y) / y);
  fy = t1 / (0.1e1 + theta) / t5 / y * t12;

  if(log){fy <- log(fy)}

  return(fy)
}

#---- cumulative distribution function

pNul <- function(q, mu, lower.tail = TRUE, log.p = FALSE){

  theta = mu/(1-mu);

  t4 = 0.1e1 / q;
  t6 = exp((-0.1e1 + q) * theta * t4);
  cdf = (q + theta) * t6 * t4 / (0.1e1 + theta);

  if(!lower.tail){cdf <- 1-cdf}
  if(log.p){cdf <- log(cdf)}
```

```
  return(cdf)
}

#---- quantile function

qNul <- function(p, mu, lower.tail = TRUE, log.p = FALSE){

  theta = mu/(1-mu);

  if(log.p){p <- exp(p)}
  if(!lower.tail){p <- 1 - p}

  t1 = 0.1e1 + theta;
  t3 = exp(-t1);
  t5 = lamW::lambertWm1(-p * t1 * t3);
  qtf = -theta/(t5 + 1);

  return(qtf)
}

#---- random values

rNul <- function(n, mu){

  stopifnot( n > 0)

  n <- ceiling(n);
  u <- runif(n);
  qNul(u, mu);
}



# ---------------------------- gamlss implantation


pow <- function(x, y) x^y

Nul <- function (mu.link = "logit"){
  mstats <- checklink("mu.link", "Nulindley",
                      substitute(mu.link),c("logit",
                      "probit", "cloglog", "cauchit",
                      "own"))
```

```
  structure(list(family = c("Nul", "New unit-Lindley
  distribution"),
                 parameters = list(mu = TRUE),
                 nopar = 1,
                 type = "Continuous",
                 mu.link = as.character(substitute(mu.link)
                 ),
                 mu.linkfun = mstats$linkfun,
                 mu.linkinv = mstats$linkinv,
                 mu.dr = mstats$mu.eta,
                 dldm = function(y, mu){t1 = mu * mu; t3 =
                 0.2e1 * y; t10 = pow(-0.1e1 + mu, 0.2e1);
                 (t1 * y + (-t3 - 0.1e1) * mu + t3) / mu /
                 t10 / y },
                 d2ldm2 = function(y, mu){t1 = mu * mu;
                 t11 = -0.1e1 + mu; t12 = t11 * t11;
                 (-y * t1 * mu + (0.3e1 * y + 0.2e1) *
                 t1 - 0.6e1 * mu * y + 0.2e1 * y) /
                 t11 / t12 / y / t1 },
                 G.dev.incr = function(y, mu, ...){-2*
                 dNul(y, mu,  log = TRUE)},
                 rqres = expression(rqres(pfun = "pNul",
                 type = "Continuous", y, mu)),
                 mu.initial = expression(mu <- mean(y)),
                 mu.valid = function(mu) all(mu>0 & mu<1),
                 y.valid = function(y) all( y>0 &  y<1),
                 mean = function(mu){mu},
                 variance = function(mu){theta = -mu/(-1+
                 mu);
                 t1 = expint(theta, 1); t2 = theta * theta;
                 t6 = exp(theta); return( t1 * t2 / (0.1e1
                 + theta) * t6 )}),
            class = c("gamlss.family", "family"))
}

# ---- inflated models

# The inflated models can be easily obtained through the
# gamlss.inf library which uses logit specification in the # fits.

library('gamlss.inf')

gen.Inf0to1(family = "Nul", type.of.Inflation =  "Zero")
```

```
gen.Inf0to1(family = "Nul", type.of.Inflation = "One")

# In the parameterization adopted in this work for the
# ZOINUL distribution, we need to implement it directly as
# an object of the gamlss library. So we have


#---- The Zero-and-One inflated New unit-Lindley distribution

#---- probability density function

dZOINUL <- function(y, mu, sigma, nu, log = FALSE){

  fy <- ifelse( y > 0 & y < 1, (1 - sigma)*dNUL(y, mu),
  sigma*(1-nu))
  fy <- ifelse( y==1, sigma*nu, fy)

  if(log){fy <- log(fy)}
  return(fy)
}


#---- cumulative distribution function

pZOINUL <- function(q, mu, sigma, nu, lower.tail = TRUE,
log.p = FALSE){

  cdf <- ifelse( q > 0 & q < 1, sigma*(1-nu) + (1-sigma)*
  pNUL(q = q, mu, lower.tail = TRUE, log.p = FALSE),
  sigma*(1-nu))
  cdf <- ifelse( q>=1, 1, cdf)

  if(!lower.tail){cdf <- 1-cdf}
  if(log.p){cdf <- log(cdf)}

  return(cdf)
}


#---- quantile function


qZOINUL <- function(p, mu, sigma, nu, lower.tail = TRUE,
log.p = FALSE){
```

```
  if(log.p){p <- exp(p)}
  if(!lower.tail){p <- 1 - p}

  p0 <- sigma*(1-nu);

  q <- ifelse( p <= sigma*(1-nu), 0, qNUL( (p - p0)/
  (1 - sigma) , mu, lower.tail = TRUE, log.p = FALSE))
  q <- ifelse( p >= 1- sigma*nu, 1, q)
  q
}


#---- random values


rZOINUL <- function(n, mu, sigma, nu){

  n <- ceiling(n)
  p <- runif(n)
  r <- qZOINUL(p, mu, sigma, nu = nu)
  r
}


# ---------------------------- gamlss implantation


ZOINUL <- function(mu.link ="logit", sigma.link ="logit",
   nu.link ="logit"){
  mstats <- checklink("mu.link","ZOINUL",
  substitute(mu.link),
                      c("logit", "probit", "cloglog", "log"
                      ,"own"))
  dstats <- checklink("sigma.link","ZOINUL",
  substitute(sigma.link),
                      c("logit", "probit", "cloglog", "log"
                      ,"own"))
  vstats <- checklink("nu.link","ZOINUL",
  substitute(nu.link),
                      c("logit", "probit", "cloglog", "log"
                      ,"own"))
  structure(list(family = c("ZOINUL", "The Zero-and-One
  inflated New unit-Lindley distribution"),
```

```
parameters = list(mu =TRUE, sigma =TRUE, nu =TRUE),
nopar = 3,
type = "Mixed",
mu.link = as.character(substitute(mu.link)),
sigma.link = as.character(substitute(sigma.link)),
nu.link = as.character(substitute(nu.link)),
mu.linkfun = mstats$linkfun,
sigma.linkfun = dstats$linkfun,
nu.linkfun = vstats$linkfun,
mu.linkinv = mstats$linkinv,
sigma.linkinv = dstats$linkinv,
nu.linkinv = vstats$linkinv,
mu.dr = mstats$mu.eta,
sigma.dr = dstats$mu.eta,
nu.dr = vstats$mu.eta,
dldm = function(y, mu){
  t1 = mu * mu;
  t3 = 0.2e1 * y;
  t10 = pow(mu - 0.1e1, 0.2e1);
  dldm = ifelse( ((y == 0) | (y == 1)), 0,  0.1e1 /
  y / t10 / mu*(y * t1 + mu * (-t3 - 0.1e1) + t3))
  dldm
},
d2ldm2 = function(y, mu){
  t1 = mu * mu;
  t11 = mu - 0.1e1;
  t12 = t11 * t11;
  d2ldm2 = ifelse( (y == 0) | (y == 1), 0, 0.1e1 /
  t1 / y / t12 / t11 * (-t1 * mu * y + t1 * (0.3e1
  * y + 0.2e1) - 0.6e1 * mu * y + 0.2e1 * y) )
     d2ldm2
  },
  dldd = function(y, sigma){
    dldd = ifelse( ((y == 0) | (y == 1)), 0.1e1 /
    sigma, 0.1e1 / (-0.1e1 + sigma))
    dldd
  },
  d2ldd2 = function(y, sigma){
    t1 = sigma * sigma;
    t2 = pow(-0.1e1 + sigma, 0.2e1);
    d2ldd2 = ifelse( (y == 0) | (y == 1), -0.1e1 /
    t1, -0.1e1 / t2)
    d2ldd2
  },
```

```
        dldv = function(y, nu){
          dldv = ifelse( y==0,  -0.1e1 / (0.1e1 - nu), 0)
          dldv = ifelse( y==1, 0.1e1 / nu, dldv)
          dldv
        },
        d2ldv2 = function(y, nu){
          t1 = nu * nu;
          t2 = pow(0.1e1 - nu, 0.2e1);
          d2ldv2 = ifelse( y==0, -0.1e1 / t2, 0)
          d2ldv2 = ifelse( y==1,  -0.1e1 / t1, d2ldv2)
          d2ldv2
        },
        d2ldmdd = function(y){
          d2ldmdd = rep(0, length(y))
          d2ldmdd
        },
        d2ldmdv = function(y){
          d2ldmdv = rep(0, length(y))
          d2ldmdv
        },
        d2ldddv = function(y, sigma, nu){
          d2ldddv = rep(0, length(y))
          d2ldddv
        },
        G.dev.incr = function(y, mu, sigma, nu, ...){-2 *
        dZOINUL(y, mu, sigma, nu, log = TRUE)},
        rqres = expression(rqres(pfun = "pZOINUL", type =
        "Mixed",mass.p = c(0,1), prob.mp=cbind(sigma, nu)
        , y, mu, sigma, nu = nu)),
        mu.initial =expression(mu <- mean(y[y>0 & y<1])),
        sigma.initial =expression(sigma <- mean(y==0) ),
        nu.initial =expression(nu <- mean(y==1) ),
        mu.valid =function(mu) all(mu > 0 & mu < 1),
        sigma.valid =function(sigma) all(sigma>0 &
        sigma<1),
        nu.valid =function(nu) all(nu > 0 & nu < 1),
        y.valid =function(y) all(y >= 0 & y <= 1)),
    class = c("gamlss.family", "family"))
}
```

Received in 14.07.2021.

Approved after revised in 07.10.2021.