



Article Modeling Proportion Data with Inflation by Using a Power-Skew-Normal/Logit Mixture Model

Guillermo Martínez-Flórez^{1,†}, Hector W. Gomez^{2,†} and Roger Tovar-Falón^{1,*,†}

- ¹ Departamento de Matemáticas y Estadística, Facultad de Ciencias Básicas, Universidad de Córdoba, Montería 230002, Colombia; guillermomartinez@correo.unicordoba.edu.co
- ² Departamento de Matemáticas, Facultad de Ciencias Básicas, Universidad de Antofagasta, Antofagasta 1240000, Chile; hector.gomez@uantof.cl
- * Correspondence: rjtovar@correo.unicordoba.edu.co
- + These authors contributed equally to this work.

Abstract: Rate or proportion data are modeled by using a regression model. The considered regression model can be used for studying phenomena with a response on the (0, 1), [0, 1), (0, 1], or [0, 1] intervals. To connect the response variable with the linear predictor in the regression model, we use a logit link function, which guarantees that the obtained prediction ranges between zero and one in the cases inflated at zero or one (or both). The model is complemented with the assumption that the errors follow a power-skew-normal distribution, resulting in a very flexible model, and with a non-singular information matrix, constituting an advantage over other existing models in the literature. To explain the probability of point mass at the values zero and/or one (inflated part), we used a polytomic logistic model with covariates. The results of two illustrations showed that the proposed model is a better alternative compared to widely known models in the literature.

Keywords: power-skew-normal distribution; unit power-skew-normal distribution; censorship; linear regression mixture model; maximum likelihood estimation

1. Introduction

Statistical modeling to explain variables, such as the concentration of sulfur in the tissue in 100 g of leaves of a certain genotype of bean (measured by turbidimetric methods), the proportion of children killed by unknown causes in the main cities of a country, the proportion of deaths caused by smoking, the prevalence rate of a certain disease in a community, the proportion of votes in favor of a presidential candidate for reelection, the proportion of income spent on education, and, in general, any response variable on the unit interval (0, 1) as proportions, rates, or indices, has been studied by several researchers, highlighting the works of Paolino [1], Cribari-Neto and Vasconcellos [2], Kieschnick and Mccullough [3], Ferrari and Cribari-Neto [4], and Vasconcellos and Cribari-Neto [5].

Among the most recent works, we emphasize Ospina and Ferrari [6,7], Bayes et al. [8] and Martínez-Flórez et al. [9,10] who have presented extensions of the works mentioned above, some of them by incorporating a set of covariates to the model. Other works in this same area are those of Mazucheli et al. [11–13] and Menezes et al. [14], which extend the Birbaum–Saunders, gamma, Weibull, and logistic models, respectively, to situations of models able to fit datasets whose variables are on a unity interval. These families have proven to be a good alternative to the beta model of Ferrari and Cribari-Neto [4] and the Kumaraswamy distribution by [15].

The previously mentioned distributions used for modeling proportions, rates, and indices as well as their respective extensions have special characteristics from which is possible to decide if is a favorable option for fitting a particular dataset; usually, the asymmetry and kurtosis coefficients are the most used. Unquestionably, these measures are associated with certain parameters of each model; generally, the parameters to which



Citation: Martínez-Flórez, G.; Gómez, H.W.; Tovar-Falón, R. Modeling Proportion Data with Inflation by Using a Power-Skew-Normal/Logit Mixture Model. *Mathematics* **2021**, *9*, 1989. https://doi.org/10.3390/math9161989

Academic Editor: Jiancang Zhuang

Received: 30 June 2021 Accepted: 17 August 2021 Published: 20 August 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). we refer are linked to characteristics of shape and/or asymmetry or the kurtosis of the distribution. Some works include the case of unit variables that contain an excessive amount of zeros and/or ones, and they are known in the literature as inflated distributions in the values of zero or one. Some works for dealing with these situations have been proposed by Ospina and Ferrari [7] and Martínez-Flórez et al. [10], among others.

The main objective of this article is to propose a new class of regression models based on the power-skew-normal distribution, which are useful for fitting data with response on the unit interval. The new models allow taking into account possible excesses of the zero and/or one values of the response variable and are also able to capture different forms of the response distribution, as well as high (or low) degrees of asymmetry and kurtosis present in the data.

The rest of this paper is organized as follows: Section 2 presents some asymmetric distributions and its main characteristics. In Section 3, the power-skew-normal/logit model is introduced, and its main properties are discussed. In addition, the statistical inference is carried out by using the maximum likelihood method. Section 4 presents the unit-power-skew-normal model for fitting data on the (0,1) interval. For this model, the maximum likelihood method is used to carry out the estimation of parameters. The score function and the elements of the observed information matrix are presented in detail. Section 5 presents the extension of the inflated unit-power-skew-normal model, which is an alternative to the inflated beta regression model. In particular, the log-UPSN model is studied. In Section 6, the doubly censored PSN model is presented, and the generalized two-part PSN model with covariates is studied as a particular case. Finally, in Section 7, two illustrative examples are reported and compared with several rival models.

2. Asymmetric Distributions

The study of families with flexible distributions capable of modeling different degrees of asymmetry and kurtosis has been of great interest in the recent statistical literature. Different works have been published, with initial works by Birnbaum [16], Lehmann [17], Roberts [18], and O'Hagan and Leonard [19] and more recently by Fernandez and Steel [20], Mudholkara and Hutson [21], Azzalini [22], Durrans [23], Gupta and Gupta [24], Arellano-Valle et al. [25,26], Gómez et al. [27], and Pewsey et al. [28]. Azzalini [22] introduces the skew-normal (SN) distribution by adding an extra parameter λ to the normal model. The inclusion of this new parameter allows for fitting data with high degrees of asymmetry. The probability density function (pdf) for the skew-normal model with location parameter μ and scale parameter σ is given by

$$f_{\rm SN}(x;\mu,\sigma,\lambda) = \frac{2}{\sigma} \phi\left(\frac{x-\mu}{\sigma}\right) \Phi\left(\lambda \frac{x-\mu}{\sigma}\right), \quad x \in \mathbb{R},$$
(1)

where $\mu, \lambda \in \mathbb{R}$ and $\sigma \in \mathbb{R}^+$. The functions $\phi(\cdot)$ and $\Phi(\cdot)$ denote the pdf and cumulative distribution function (cdf) of the standard normal distribution. The model in (1) is denoted by $X \sim SN(\mu, \sigma, \lambda)$, and the respective cdf is written as

$$\mathcal{F}_{\rm SN}\left(\frac{x-\mu}{\sigma},\lambda\right) = \Phi\left(\frac{x-\mu}{\sigma}\right) - 2T\left(\frac{x-\mu}{\sigma},\lambda\right), \quad x \in \mathbb{R},\tag{2}$$

where $T(\cdot, \lambda)$ is the Owen function (see [29]). Another asymmetric model widely studied in the statistical literature is the named power-normal (PN), which was initially introduced by Durrans [23]. The PN model is sometimes denominated the generalized Gaussian distribution. Later, Gupta and Gupta [24,30] studied some statistical properties of the PN model, and they called it the exponential distribution. On the other hand, Pewsey et al. [28] studied the statistical inference of the PN model by using the maximum likelihood method; here, the authors deduce the expected information matrix, and they show the non-singularity of the matrix. The pdf of the location-scale version of the PN model is given by

$$f_{\rm PN}(x;\mu,\sigma,\alpha) = \frac{\alpha}{\sigma} \phi\left(\frac{x-\mu}{\sigma}\right) \left\{ \Phi\left(\frac{x-\mu}{\sigma}\right) \right\}^{\alpha-1}, \quad x \in \mathbb{R},$$
(3)

where $\alpha \in \mathbb{R}^+$ is a shape parameter, which contributes strongly to the kurtosis of the model. The model in (3) is denoted by $X \sim PN(\mu, \sigma, \alpha)$, and the respective cdf is given by

$$\mathcal{F}_{\text{PN}}(x;\mu,\sigma,\alpha) = \left\{ \Phi\left(\frac{x-\mu}{\sigma}\right) \right\}^{\alpha}, \quad x \in \mathbb{R}.$$
(4)

An extension of the PN model which is capable of capturing greater ranges of asymmetry and kurtosis was proposed by Martínez-Flórez et al. [31]. This proposal, which is named the power-skew-normal (PSN) model, is originated by replacing the pdf and cdf of the normal distribution in the PN model by those of the SN model, that is, the PSN model contains both shape and asymmetry parameters. The pdf for the location-scale version of the PSN model is given by

$$f_{\rm PSN}(x;\mu,\sigma,\lambda,\alpha) = \alpha f_{\rm SN}(x,\mu,\sigma,\lambda) \left\{ \mathcal{F}_{\rm SN}\left(\frac{x-\mu}{\sigma},\lambda\right) \right\}^{\alpha-1}, \quad x \in \mathbb{R},$$
(5)

where $\mu, \lambda \in \mathbb{R}$ and $\sigma, \alpha \in \mathbb{R}^+$. This is denoted by $X \sim PSN(\mu, \sigma, \lambda, \alpha)$. One can observe that, for $\alpha = 1$, the SN model is obtained, while for $\lambda = 0$, the PN model is followed. The normal model is obtained when $\lambda = 0$ and $\alpha = 1$, that is, the PSN model is more flexible than the normal, SN and PN models. If $X \sim PSN(\mu, \sigma, \lambda, \alpha)$, then the cdf of X is given by

$$\mathcal{F}_{\text{PSN}}(x;\mu,\sigma,\lambda,\alpha) = \left\{\Phi\left(\frac{x-\mu}{\sigma}\right) - 2T\left(\frac{x-\mu}{\sigma},\lambda\right)\right\}^{\alpha}, \quad x \in \mathbb{R}.$$

where $T(\cdot, \lambda)$ is Owen's function. For λ and α values ranging in the (0.1;100) interval, the asymmetry and kurtosis coefficients for the PSN model are [-1.4676, 0.9953) and [1.4672, 5.4386], respectively; these intervals contain the respective asymmetry and kurtosis coefficients of the SN and PN models (see Pewsey et al. [28]). The extensions for the positive data of the random variable *X* following the SN, PN or PSN models are obtained by applying the transformation exp(*X*), and they are denominated as a log-skew-normal (LSN) distribution, log-power-normal (LPN) distribution and log-power-skew-normal (LPSN) distribution, respectively (see Martínez-Flórez et al. [9,32], Mateus-Figueras and Pawlosky-Glanh [33]).

Asymmetric models have become very useful statistical tools for modeling censored or truncated data using covariates: see, for example, the log-gamma model by Moulton and Halsey [34], the log-skew-normal model of Chai and Bailey [35], the power-normal model by Martínez-Flórez et al. [9], and log-power-normal models by Martínez-Flórez et al. [9,32]. In this work, we extend the PSN model to the case of proportions, rates or indices data. The proposed extension is useful for modeling data with a response on the unit interval with an excess of zeros and/or ones and covariates to explain the response and the excess of zeros and/or ones.

3. The Power-Skew-Normal/Logit Mixture Model

The linear regression model with errors following a PSN($0, \sigma, \lambda, \alpha$) distribution was introduced and studied in detail by Martínez-Flórez et al. [32]. This model is expressed as

$$y_i = \delta_0 + \delta_1 z_{1i} + \dots + \delta_p z_{pi} + \varepsilon_i, \tag{6}$$

where $\delta = (\delta_0, \delta_1, \dots, \delta_p)^\top$ is an unknown vector of regression coefficients, $\mathbf{z}_i = (1, z_1, \dots, z_p)^\top$ is a vector containing *p* known explanatory variables with p < n, and $\varepsilon_i \sim \text{PSN}(0, \sigma, \lambda, \alpha)$ for $i = 1, 2, \dots, n$. It follows that

$$y_i \sim \text{PSN}(\mathbf{z}_i^{\top} \delta, \sigma, \lambda, \alpha), \quad i = 1, 2, \dots, n.$$

Since $\mathbb{E}(\varepsilon_i) = \alpha \sigma \int_0^1 \mathcal{F}_{SN}^{-1}(u) u^{\alpha-1} du \neq 0$, it follows that $\mathbb{E}(y_i) \neq \mathbf{z}_i^\top \delta$; therefore, it is necessary to correct the parameter δ_0 in the form $\delta_0^* = \delta_0 + \mu_{\varepsilon}$, where $\mu_{\varepsilon} = \mathbb{E}(\varepsilon_i)$. Thus, it is obtained that $\mathbb{E}(y_i) = \mathbf{z}_i^\top \delta^*$, where $\delta^* = (\delta_0^*, \delta_1, \dots, \delta_p)^\top$. The ordinary least squares method can be used to obtain an estimate of the parameter vector δ , which can be used as an initial value in the maximum likelihood estimation process.

The main interest in this paper is centered on the case where the measured variable has a response on the unit interval, and the expected response or predicted value falls outside of this unit interval, which could lead to negative estimates without any interpretation or meaning. To avoid these inconveniences, the assumption of response variable Y being a linear function of the vector of explanatory variables $\mathbf{z}_i^{\top} = (z_1, z_2, \dots, z_p)$ is replaced by the assumption of a non-linear transformation of this set of variables. This model is obtained by assuming that the location parameter of the y_i variable can be written as

$$g(\eta_i) = \sum_{i=1}^n \mathbf{z}_i^\top \boldsymbol{\delta}$$
(7)

where $g(\cdot)$ is a strictly monotonic link function whose second derivative exists. Two link functions $g(\cdot)$ widely used in practical situations that can be considered in (7) are the probit with $g(\eta_i) = \Phi(\eta_i)$, where $\Phi(\cdot)$ is the cdf of the standard normal distribution, and the logit function given by $g(\eta_i) = \log(\eta_i/(1 - \eta_i))$. These two options lead to very similar results in the predicted values, with some exceptions for extreme values. For the ease of handling deductions, in this work, we opt for the logit function. Thus, in this case, we have

$$\eta_i = \frac{\exp(\mathbf{z}_i^\top \boldsymbol{\delta})}{1 + \exp(\mathbf{z}_i^\top \boldsymbol{\delta})}, \quad i = 1, 2, \dots, n.$$
(8)

For the function in (8), the parameters are interpreted from the odds ratio between the odds of the prediction or mean when one of the variables is increased *m* units (keeping the rest of the explanatory variables fixed) and the odds without the increase. One can show that this quotient of odds ratios is given by $\exp(m\delta_k)$, where δ_k is the parameter associated with the explanatory variable increased by *m* units. It follows that the distribution of the study variable is

$$y_i \sim \text{PSN}(\eta_i, \sigma, \lambda, \alpha), \quad i = 1, 2, \dots, n.$$
 (9)

From the model in (9), some special cases can be obtained; for example, if $\alpha = 1$, the skew-normal-logit model is obtained, while, for $\lambda = 0$, the exponential or alpha-power-logit case is followed. If $\lambda = 0$ and $\alpha = 1$, it has the normal-logit model.

The parameter estimation of the PSN regression model on the unit interval (0,1) with logit link function can be obtained by using the maximum likelihood method. The loglikelihood function obtained from a random sample of size *n* is given by

$$\ell(\boldsymbol{\varphi}; \mathbf{y}) = n \log(\alpha) - n \log(\sigma) + \sum_{i=1}^{n} \log(f_{\mathrm{SN}}(w_i, \lambda)) + (\alpha - 1) \sum_{i=1}^{n} \log(\mathcal{F}_{\mathrm{SN}}(w_i, \lambda)), \quad (10)$$

where $w_i = (y_i - \eta_i)/\sigma$. for i = 1, 2, ..., n. To obtain the elements of the score function and the observed information matrix of the parameters $\boldsymbol{\varphi} = (\boldsymbol{\delta}^{\top}, \sigma)^{\top}$, we use the fact that

$$\frac{\partial \ell(\boldsymbol{\varphi}; \mathbf{y})}{\partial \delta_{i}} = \frac{\partial \ell(\boldsymbol{\varphi}; \mathbf{y})}{\partial w_{i}} \frac{\partial w_{i}}{\partial \eta_{i}} \frac{\partial \eta_{i}}{\partial \delta_{i}} \quad \text{and} \quad \frac{\partial \ell(\boldsymbol{\varphi}; \mathbf{y})}{\partial \sigma} = \frac{\partial \ell(\boldsymbol{\varphi}; \mathbf{y})}{\partial w_{i}} \frac{\partial w_{i}}{\partial \sigma},$$

where η_i is given in (8). Then, the elements of the score function can be written in the form

$$U(\delta_j) = \frac{1}{\sigma} \sum_{i=1}^n z_{ij} \eta_i w_i - \frac{\lambda}{\sigma} \sum_{i=1}^n z_{ij} \eta_i \xi(\lambda w_i) - \frac{\alpha - 1}{\sigma} \sum_{i=1}^n z_{ij} \eta_i \tau(\lambda w_i), \ j = 0, 1, 2, \dots, p, \quad (11)$$

$$U(\sigma) = -\frac{n}{\sigma} + \frac{1}{\sigma} \sum_{i=1}^{n} w_i^2 - \frac{\lambda}{\sigma} \sum_{i=1}^{n} w_i \xi(\lambda w_i) - \frac{\alpha - 1}{\sigma} \sum_{i=1}^{n} w_i \gamma(\lambda w_i),$$
(12)

$$U(\lambda) = \sum_{i=1}^{n} w_i \xi(\lambda w_i) - \frac{\alpha - 1}{1 + \lambda^2} \sum_{i=1}^{n} \nu(w_i), \quad U(\alpha) = \frac{n}{\alpha} + \sum_{i=1}^{n} \log(\mathcal{F}_{SN}(w_i, \lambda)).$$
(13)

where $\xi(w) = \frac{\phi(w)}{\Phi(w)}$, $\gamma(w) = \frac{f_{SN}(w,\lambda)}{F_{SN}(w,\lambda)}$ and $\nu(w) = \sqrt{\frac{2}{\pi}} \frac{\phi(\sqrt{1+\lambda^2}w)}{F_{SN}(w,\lambda)}$. The scores equations are obtained by setting the elements of the score function equal to zero, that is, the first derivative of $\ell(\varphi; \mathbf{y})$ with respect to the parameters $\delta_0, \delta_1, \ldots, \delta_p, \sigma, \lambda$, and α . By solving this system of equations, the maximum likelihood estimates are obtained. To maximize the log-likelihood function, it is necessary to use iterative numeric methods. Likewise, as in the standard case, the observed and expected Fisher information matrices are obtained as minus the Hessian matrix (the second derivative of $\ell(\varphi; \mathbf{y})$ with respect to the parameters) and the expected value of the elements of the observed information matrix, respectively. After some algebraic manipulations, it follows that the elements of the observed information matrix can be written in the form $\mathbf{H}(\varphi) = \sum_A h_{\varphi_j \varphi_{j'}}, h_{\varphi_j \varphi_{j'}} = -\partial^2 \ell(\varphi; \mathbf{y}) / \partial \varphi_j \partial \varphi_{j'}$ and $A = \{1, 2, 3, \ldots, n\}$ being the set of observations.

Letting $\zeta(w) = \xi(w)[\xi(w) + w]$ and $\tau(w) = \gamma(w)[\gamma(w) + w]$, the elements $h_{\varphi_j \varphi_{j'}}$ are given by

$$\begin{split} h_{\delta_{j}\delta_{l}} &= \frac{1}{\sigma^{2}} z_{ij} z_{il} \eta_{i}^{2} \Big[1 + \lambda^{2} \zeta(\lambda w_{i}) \Big] + \frac{\alpha - 1}{\sigma^{2}} z_{ij} z_{il} \eta_{i}^{2} [\tau(w_{i}) - \lambda \nu(w_{i}))] \\ &\quad + \frac{1}{\sigma} z_{ij} z_{il} \eta_{i} [-w_{i} + \lambda \xi(\lambda w_{i}) + (\alpha - 1)\gamma(w_{i})], \\ h_{\delta_{j}\eta} &= \frac{1}{\sigma^{2}} z_{ij} \eta_{i} \Big[2w_{i} - \lambda \xi(\lambda w_{i}) + \lambda^{2} w_{i} \zeta(\lambda w_{i}) \Big] \\ &\quad + \frac{\alpha - 1}{\sigma^{2}} z_{ij} \eta_{i} [-\gamma(w_{i}) + w_{i} \tau(w_{i}) - \lambda w_{i} \nu(w_{i})], \\ h_{\delta_{j}\lambda} &= \frac{1}{\sigma} z_{ij} \eta_{i} [\xi(\lambda w_{i}) + \lambda w_{i} \zeta(\lambda w_{i})] \\ &\quad + \frac{\alpha - 1}{\sigma} z_{ij} \eta_{i} \nu(w_{i}) \left[w_{i} + \frac{1}{1 + \lambda^{2}} \gamma(w_{i}) \right], \\ h_{\sigma\sigma} &= \frac{1}{\sigma^{2}} \Big[-1 + 3w_{i}^{2} - 2\lambda w_{i} \xi(\lambda w_{i}) + \lambda^{2} w_{i}^{2} \zeta(\lambda w_{i}) \Big] \\ &\quad + \frac{\alpha - 1}{\sigma^{2}} w_{i} [-2\gamma(w_{i}) + w_{i} \tau(w_{i}) - \lambda w_{i} \nu(w_{i})], \\ h_{\sigma\lambda} &= \frac{1}{\sigma} w_{i} [\xi(\lambda w_{i}) + \lambda w_{i} \zeta(\lambda w_{i})] + \frac{\alpha - 1}{\sigma} w_{i} \nu(w_{i}) \Big[w_{i} + \frac{1}{1 + \lambda^{2}} \gamma(w_{i}) \Big], \\ h_{\lambda\lambda} &= w_{i}^{2} \zeta(\lambda w_{i}) + \frac{\alpha - 1}{1 + \lambda^{2}} \nu(w_{i}) \Big[-\lambda \Big(\frac{2}{1 + \lambda^{2}} + w_{i}^{2} \Big) + \frac{1}{1 + \lambda^{2}} \nu(w_{i}) \Big], \\ h_{\delta_{j}\alpha} &= \frac{1}{\sigma} z_{ij} \eta_{i} \gamma(w_{i}), \qquad h_{\sigma\alpha} &= \frac{1}{\sigma} w_{i} \gamma(w_{i}), \qquad h_{\lambda\alpha} &= \frac{1}{1 + \lambda^{2}} \nu(w_{i}), \qquad h_{\alpha\alpha} &= \frac{1}{\alpha^{2}} \Big]. \end{split}$$

Now, by letting $\lambda = 0$ and $\alpha = 1$, and using numerical integration, the Fisher information matrix is given by

$$\mathbf{I}(\boldsymbol{\varphi}) = \begin{pmatrix} \mathbf{I}_{\mathrm{SN}}(\boldsymbol{\delta}, \sigma, \lambda) & \mathbf{I}_1 \\ & & \\ & \mathbf{I}_1^\top & \mathbf{1} \end{pmatrix},$$

where
$$\mathbf{I}_{1} = \left(\frac{0.9031}{\sigma}\mathbf{1}_{n}^{\top}\mathbf{M}\mathbf{Z}, -\frac{0.5956}{\sigma}, 0.7206\right)^{\top}$$
, with $\mathbf{M} = \text{Diag}(\eta_{1}, \eta_{2}, \dots, \eta_{n})$ and
 $\mathbf{I}_{\text{SN}}(\delta, \sigma, \lambda) = \left(\begin{array}{cc} \frac{1}{\sigma^{2}}(\mathbf{M}\mathbf{Z})^{\top}(\mathbf{M}\mathbf{Z}) & \mathbf{0}_{p+1} & \frac{1}{\sigma}b\mathbf{Z}^{\top}\mathbf{M}\mathbf{1}_{n} \\ \mathbf{0}_{p+1}^{\top} & \frac{2n}{\sigma^{2}} & \mathbf{0} \\ \frac{1}{\sigma}b\mathbf{1}_{n}^{\top}\mathbf{M}\mathbf{Z} & \mathbf{0} & nb^{2} \end{array}\right),$

with $b = \sqrt{\frac{2}{\pi}}$. The determinant of the **I**_{SN} matrix is given by

$$|\mathbf{I}_{\mathrm{SN}}(\delta,\sigma,\lambda)| = \frac{2nb^2}{\sigma^{2(p+1)}} \Big| (\mathbf{M}\mathbf{Z})^{\top} (\mathbf{M}\mathbf{Z}) \Big| \Big| n - \mathbf{1}_n^{\top} \mathbf{M}\mathbf{Z} \Big(\mathbf{Z}^{\top} \mathbf{M}\mathbf{M}\mathbf{Z}\Big)^{-1} \mathbf{Z}^{\top} \mathbf{M}\mathbf{1}_n \Big|$$

Since **Z** is of full-column rank, and **M** is a diagonal matrix, the rank of **MZ** is the same as the **Z**, that is, the matrix **MZ** is of full-column rank; therefore, $(\mathbf{MZ})^{\top}(\mathbf{MZ})$ will be full-rank and hence invertible, that is, its determinant is different from zero. Now, to find the last determinant, we write $\mathbf{Z} = (\mathbf{1}_n, \mathbf{Z}_1)$, where $\mathbf{Z}_1 = (Z_1, Z_2, \dots, Z_p)$; this partition leads to expressing the matrix $\mathbf{Z}^{\top}\mathbf{MMZ}$ as a partitioned matrix, for which we can use the existing expressions of the matrix algebra to find the inverse of a partitioned matrix, that is, we can determine $(\mathbf{Z}^{\top}\mathbf{MMZ})^{-1}$. With this result, it follows that, $\mathbf{1}_n^{\top}\mathbf{MZ}(\mathbf{Z}^{\top}\mathbf{MMZ})^{-1}\mathbf{Z}^{\top}\mathbf{M1}_n = \sum_{i=1}^n \eta_i^2$ and, as for $i = 1, 2, \dots, n, 0 < \eta_i < 1, \sum_{i=1}^n \eta_i^2 < n$, which leads to $|n - \mathbf{1}_n^{\top}\mathbf{MZ}(\mathbf{Z}^{\top}\mathbf{MMZ})^{-1}\mathbf{Z}^{\top}\mathbf{M1}_n| > 0$, concluding that $|\mathbf{I}_{SN}(\delta, \sigma, \lambda)| \neq 0$, that is, $\mathbf{I}_{SN}(\delta, \sigma, \lambda)$ is non-singular, and therefore its rows and/or columns are linearly independent. Thus, the rows and/or columns of the information matrix $\mathbf{I}(\boldsymbol{\varphi})$ are linearly independent. Therefore, the regularity conditions are satisfied, and the known \sqrt{n} -property for the maximum likelihood estimators is satisfied for all λ and α .

This important result further supports the hypothesis of singularity in the information matrix for the SN model for cases where the variable is a linear transformation of the location parameter. For other non-linear transformations, such as the asymmetric Birbaum–Saunders distributions studied by Vilca and Leiva-Sánchez [36], the asymmetric sinh-normal model of Leiva-Sánchez [37], and the asymmetric Birbaum–Saunders exponential distribution in Mattínez-Flórez et al. [38], the information matrices turned out to be non-singular.

4. Unit-Power-Skew-Normal Model

The unit-power-skew-normal (UPSN) model can be defined from the doubly truncated power-skew-normal (TPSN) distribution on the interval (0, 1), which has a pdf given by

$$f_{\text{TPSN}}(y_i; \mu, \sigma, \lambda, \alpha) = \frac{\frac{\alpha}{\sigma} f_{\text{SN}}(w_i, \lambda) \{\mathcal{F}_{\text{SN}}(w_i, \lambda)\}^{\alpha - 1}}{\{\mathcal{F}_{\text{SN}}(w_{1i}, \lambda)\}^{\alpha} - \{\mathcal{F}_{\text{SN}}(w_{0i}, \lambda)\}^{\alpha}}, \quad 0 < y_i < 1,$$
(14)

where

$$w_{0i} = -\frac{\mu}{\sigma}, \quad w_i = \frac{y_i - \mu}{\sigma} \quad \text{and} \quad w_{1i} = \frac{1 - \mu}{\sigma}.$$
 (15)

The properties of the doubly TPSN model can be studied from the properties of the truncated models. One can observe from Equation (14) that, if $\alpha = 1$, the standard unit skew-normal-logit model is obtained, while, for $\lambda = 0$, the standard unit exponential-logit or alpha-power-logit model is obtained. Finally, when $\lambda = 0$ and $\alpha = 1$, the standard unit normal-logit model is followed.

The cdf of the TPSN model is given by

$$\mathcal{F}_{\text{TPSN}}(y_i) = \frac{\left\{\mathcal{F}_{\text{SN}}(w_i, \lambda)\right\}^{\alpha} - \left\{\mathcal{F}_{\text{SN}}(w_{0i}, \lambda)\right\}^{\alpha}}{\left\{\mathcal{F}_{\text{SN}}(w_{1i}, \lambda)\right\}^{\alpha} - \left\{\mathcal{F}_{\text{SN}}(w_{0i}, \lambda)\right\}^{\alpha}}, \quad 0 < y_i < 1,$$
(16)

and the survival and Hazard functions are given by

$$\mathcal{S}_{\text{TPSN}}(y_i) = \frac{\{\mathcal{F}_{\text{SN}}(w_{1i},\lambda)\}^{\alpha} - \{\mathcal{F}_{\text{SN}}(w_i,\lambda)\}^{\alpha}}{\{\mathcal{F}_{\text{SN}}(w_{1i},\lambda)\}^{\alpha} - \{\mathcal{F}_{\text{SN}}(w_{0i},\lambda)\}^{\alpha}}, \quad 0 < y_i < 1,$$
(17)

$$r_{\text{TPSN}}(y_i) = \frac{\frac{\alpha}{\sigma} f_{\text{SN}}(w_i, \lambda) \{\mathcal{F}_{\text{SN}}(w_i, \lambda)\}^{\alpha - 1}}{\{\mathcal{F}_{\text{SN}}(w_i, \lambda)\}^{\alpha} - \{\mathcal{F}_{\text{SN}}(w_i, \lambda)\}^{\alpha}}, \quad R(t_i) = \frac{\frac{\alpha}{\sigma} f_{\text{SN}}(w_i, \lambda) \{\mathcal{F}_{\text{SN}}(w_i, \lambda)\}^{\alpha - 1}}{\{\mathcal{F}_{\text{SN}}(w_i, \lambda)\}^{\alpha} - \{\mathcal{F}_{\text{SN}}(w_{0i}, \lambda)\}^{\alpha}}, \tag{18}$$

respectively. The moments of the TPSN model can be calculated by the expression

$$\mathbb{E}(Y^{r}) = \frac{\alpha \sum_{j=1}^{r} \eta_{i}^{r-j} \sigma^{j} \varepsilon^{j}}{\left\{ \mathcal{F}_{SN}(w_{1i}, \lambda) \right\}^{\alpha} - \left\{ \mathcal{F}_{SN}(w_{0i}, \lambda) \right\}^{\alpha}},$$
(19)

where

$$\varepsilon = \int_{\mathcal{F}_{\mathrm{SN}}(w_{0i},\lambda)}^{\mathcal{F}_{\mathrm{SN}}(w_{1i},\lambda)} \mathcal{F}_{\mathrm{SN}}^{-1}(u,\lambda) u^{\alpha-1} du,$$

 $\mathcal{F}_{SN}^{-1}(\cdot,\lambda)$ being the inverse function of $\mathcal{F}_{SN}(\cdot,\lambda)$. The estimates of the parameters of the doubly TPSN-logit model (14) considering a set of covariates can be obtained by using the maximum likelihood method. The log-likelihood function for estimating $\boldsymbol{\varphi} = (\boldsymbol{\delta}^{\top}, \lambda, \alpha)^{\top}$ is given by

$$\ell(\boldsymbol{\varphi}, \mathbf{y}) = n \log(\alpha) - n \log(\sigma) + \sum_{y_i \in (0,1)} \log(f_{SN}(w_i, \lambda)) + (\alpha - 1) \sum_{y_i \in (0,1)} \log(\mathcal{F}_{SN}(w_i, \lambda)) - \sum_{y_i \in (0,1)} \log(\{\mathcal{F}_{SN}(w_{1i}, \lambda)\}^{\alpha} - \{\mathcal{F}_{SN}(w_{0i}, \lambda)\}^{\alpha}),$$
(20)

where w_{0i} , w_{1i} , w_i are defined in (15). The scores equations are obtained by setting the first derivative of $\ell(\boldsymbol{\varphi}, \mathbf{y})$ with respect to the parameters $\delta_0, \delta_1, \ldots, \delta_p, \sigma, \lambda$, and α , equal to zero. The solution of the resulting system of equations leads to maximum likelihood estimates, which is maximized by using iterative numeric methods.

The covariance matrix and standard errors for the TPSN model can be obtained from the inverse of the observed information matrix, given by minus the second derivative of the log-likelihood function in Equation (20) with respect to the parameters of the model, δ , σ , λ , and α . Then, the observed information matrix of the truncated unit PSN-logit model can be found from the elements of the matrix of the unit PSN-logit model; these elements can be written as

$$\begin{split} \kappa_{\delta_{j}\delta_{l}} &= \sum_{y_{i}\in(0,1)} h_{\delta_{j}\delta_{l}} + \sum_{y_{i}\in(0,1)} \left[\frac{\psi_{\delta_{j}\delta_{l}}^{''}(1-0)}{\psi(1-0)} - \frac{\psi_{\delta_{j}}^{'}(1-0)\psi_{\delta_{l}}^{'}(1-0)}{\psi^{2}(1-0)} \right] \\ \kappa_{\delta_{j}\sigma} &= \sum_{y_{i}\in(0,1)} h_{\delta_{j}\sigma} + \sum_{y_{i}\in(0,1)} \left[\frac{\psi_{\delta_{j}\sigma}^{''}(1-0)}{\psi(1-0)} - \frac{\psi_{\delta_{j}}^{'}(1-0)\psi_{\sigma}^{'}(1-0)}{\psi^{2}(1-0)} \right] \\ \kappa_{\delta_{j}\lambda} &= \sum_{y_{i}\in(0,1)} h_{\delta_{j}\sigma} + \sum_{y_{i}\in(0,1)} \left[\frac{\psi_{\delta_{j}\lambda}^{''}(1-0)}{\psi(1-0)} - \frac{\psi_{\delta_{j}}^{'}(1-0)\psi_{\lambda}^{'}(1-0)}{\psi^{2}(1-0)} \right] \\ \kappa_{\delta_{j}\alpha} &= \sum_{y_{i}\in(0,1)} h_{\delta_{j}\alpha} + \sum_{y_{i}\in(0,1)} \left[\frac{\psi_{\delta_{j}\lambda}^{''}(1-0)}{\psi(1-0)} - \frac{\psi_{\delta_{j}}^{'}(1-0)\psi_{\lambda}^{'}(1-0)}{\psi^{2}(1-0)} \right] \end{split}$$

$$\begin{split} \kappa_{\sigma\sigma} &= \sum_{y \in (0,1)} h_{\sigma\sigma} + \sum_{y_i \in (0,1)} \left[\frac{\psi_{\sigma\sigma}'(1-0)}{\psi(1-0)} - \frac{\left(\psi_{\sigma}'(1-0)\right)^2}{\psi^2(1-0)} \right] \\ \kappa_{\sigma\lambda} &= \sum_{y_i \in (0,1)} h_{\sigma\lambda} + \sum_{y_i \in (0,1)} \left[\frac{\psi_{\sigma\lambda}'(1-0)}{\psi(1-0)} - \frac{\psi_{\sigma}'(1-0)\psi_{\lambda}'(1-0)}{\psi^2(1-0)} \right] \\ \kappa_{\sigma\alpha} &= \sum_{y_i \in (0,1)} h_{\sigma\alpha} + \sum_{y_i \in (0,1)} \left[\frac{\psi_{\sigma\alpha}'(1-0)}{\psi(1-0)} - \frac{\psi_{\sigma}'(1-0)\psi_{\alpha}'(1-0)}{\psi^2(1-0)} \right] \\ \kappa_{\lambda\lambda} &= \sum_{y_i \in (0,1)} h_{\lambda\lambda} + \sum_{y_i \in (0,1)} \left[\frac{\psi_{\lambda\lambda}'(1-0)}{\psi(1-0)} - \frac{\left(\psi_{\lambda}'(1-0)\right)^2}{\psi^2(1-0)} \right] \\ \kappa_{\lambda\alpha} &= \sum_{y_i \in (0,1)} h_{\lambda\alpha} + \sum_{y \in (0,1)} \left[\frac{\psi_{\lambda\alpha}'(1-0)}{\psi(1-0)} - \frac{\psi_{\lambda}'(1-0)\psi_{\alpha}'(1-0)}{\psi^2(1-0)} \right] \\ \kappa_{\alpha\alpha} &= \frac{n}{\alpha^2} + \sum_{y_i \in (0,1)} \left[\frac{\psi_{\alpha\alpha}'(1-0)}{\psi(1-0)} - \frac{\left(\psi_{\alpha}'(1-0)\right)^2}{\psi^2(1-0)} \right] \end{split}$$

where

$$\begin{split} \psi(1-0) &= \psi(1) - \psi(0), \quad \text{with} \quad \psi(y) = (\mathcal{F}_{\text{SN}}(w;\lambda))^{\alpha}, \\ \psi_{\varphi}^{'}(1-0) &= \psi_{\varphi}^{'}(1) - \psi_{\varphi}^{'}(0), \quad \text{with} \quad \psi_{\varphi}^{'}(y) = \frac{\partial \psi(y)}{\partial \varphi}, \\ \psi_{\varphi_{j}\varphi_{l}}^{''}(1-0) &= \psi_{\varphi_{j}\varphi_{l}}^{''}(1) - \psi_{\varphi_{j}\varphi_{l}}^{''}(0), \quad \text{with} \quad \psi_{\varphi_{j}\varphi_{l}}^{''}(y) = \frac{\partial^{2}\psi(y)}{\partial \varphi_{j}\partial \varphi_{l}}. \end{split}$$

In addition, for

$$\iota(w) = \sqrt{\frac{2}{\pi}} \frac{\phi(\sqrt{1+\lambda^2}w)}{f_{\rm SN}(w;\lambda)},$$

it follows that

$$\begin{split} \psi_{\delta_{j}}^{'}(y_{i}) &= -z_{ij}\eta_{i}f_{\mathrm{PSN}}(y_{i};\eta_{i},\sigma,\lambda,\alpha), \\ \psi_{\lambda}^{'}(y_{i}) &= -\sqrt{\frac{\pi}{2}}\frac{\alpha}{1+\lambda^{2}}\{\mathcal{F}_{\mathrm{PSN}}(y_{i};\eta_{i},\sigma,\lambda,\alpha)\}^{\alpha-1}, \\ \psi_{\sigma}^{'}(y_{i}) &= -w_{i}f_{\mathrm{PSN}}(y_{i};\eta_{i},\sigma,\lambda,\alpha), \\ \psi_{\alpha}^{'}(y_{i}) &= \log(\mathcal{F}_{\mathrm{PSN}}(y_{i};\eta_{i},\sigma,\lambda,\alpha))\{\mathcal{F}_{\mathrm{PSN}}(y_{i};\eta_{i},\sigma,\lambda,\alpha)\}^{\alpha}, \\ \psi_{\delta_{j}\delta_{l}}^{''}(y_{i}) &= -\frac{z_{ij}z_{il}}{\sigma}f_{\mathrm{PSN}}(y_{i};\eta_{i},\sigma,\lambda,\alpha)\left[\eta_{i}^{2}(w_{i}-\lambda\iota(w_{i})-(\alpha-1)\gamma(w_{i}))+\eta_{i}\right], \\ \psi_{\delta_{j}\sigma}^{''}(y_{i}) &= -\frac{z_{ij}}{\sigma}w_{i}f_{\mathrm{PSN}}(y_{i};\eta_{i},\sigma,\lambda,\alpha)\left[w_{i}-\lambda\iota(w_{i})-(\alpha-1)\gamma(w_{i})\right], \\ \psi_{\delta_{j}\delta_{l}}^{''}(y_{i}) &= -z_{ij}\eta_{i}\gamma(w_{i})\left[\frac{\alpha}{\sigma}w_{i}\{\mathcal{F}_{\mathrm{SN}}(w_{i};\lambda)\}^{\alpha}-\frac{\alpha-1}{1+\lambda^{2}}f_{\mathrm{PSN}}(y_{i};\eta_{i},\sigma,\lambda,\alpha)\right], \\ \psi_{\delta_{j}\alpha}^{''}(y_{i}) &= -z_{ij}\eta_{i}f_{\mathrm{PSN}}(y_{i};\eta_{i},\sigma,\lambda,\alpha)\left[\frac{1}{\alpha}+\log(\mathcal{F}_{\mathrm{SN}}(w_{i};\lambda))\right], \\ \psi_{\sigma\sigma}^{''}(y_{i}) &= \frac{z_{i}}{\sigma}f_{\mathrm{PSN}}(y_{i};\eta_{i},\sigma,\lambda,\alpha)\left[2-w_{i}^{2}+(\alpha-1)w_{i}\gamma(w_{i})\right] \\ &\quad +\frac{\alpha\lambda}{\sigma^{2}}w_{i}^{2}\nu(w_{i})\{\mathcal{F}_{\mathrm{SN}}(w_{i};\lambda)\}^{\alpha}, \end{split}$$

$$\begin{split} \psi_{\sigma\lambda}^{''}(y_i) &= -w_i \gamma(w_i) \left[\frac{\alpha}{\sigma} w_i \{ \mathcal{F}_{\mathrm{SN}}(w_i;\lambda) \}^{\alpha} - \frac{\alpha - 1}{1 + \lambda^2} f_{\mathrm{PSN}}(y_i;\eta_i,\sigma,\lambda,\alpha) \right], \\ \psi_{\sigma\alpha}^{''}(y_i) &= -w_i f_{\mathrm{PSN}}(y_i;\eta_i,\sigma,\lambda,\alpha) \left[\frac{1}{\alpha} + \log(\mathcal{F}_{\mathrm{SN}}(w_i;\lambda)) \right], \\ \psi_{\lambda\lambda}^{''}(y_i) &= -\frac{\alpha}{1 + \lambda^2} \gamma(w_i) \{ \mathcal{F}_{\mathrm{SN}}(w_i;\lambda) \}^{\alpha} \left[\lambda(w_i^2 - 2) - \frac{\alpha - 1}{1 + \lambda^2} \gamma(w_i) \right], \\ \psi_{\lambda\alpha}^{''}(y_i) &= -\frac{1}{1 + \lambda^2} \gamma(w_i) \{ \mathcal{F}_{\mathrm{SN}}(w_i;\lambda) \}^{\alpha} [1 + \alpha \log(\mathcal{F}_{\mathrm{SN}}(w_i;\lambda))], \\ \psi_{\alpha\alpha}^{''}(y_i) &= (\log(\mathcal{F}_{\mathrm{SN}}(w_i;\lambda)))^2 \{ \mathcal{F}_{\mathrm{SN}}(w_i;\lambda) \}^{\alpha}. \end{split}$$

According to the results found for the the PSN-logit model, the information matrix of the model is non-singular; therefore, for large sample sizes, we have

$$\hat{\boldsymbol{\varphi}} \stackrel{D}{\longrightarrow} N_{p+4} \Big(\boldsymbol{\varphi}, \mathbf{I}(\boldsymbol{\varphi})^{-1} \Big).$$

That is, the vector of the estimators is consistent and has a normal asymptotic distribution, with covariance matrix being the inverse of the Fisher information matrix. In practice, since the matrix $\mathbf{H}(\boldsymbol{\varphi})$ is consistent for $\mathbf{I}(\boldsymbol{\varphi})$, then we can take $\boldsymbol{\Sigma} = \mathbf{H}^{-1}(\boldsymbol{\varphi})$ as the covariance matrix of the vector of estimators of the standard unit PSN-logit regression model.

5. Inflated Unit-Power-Skew-Normal Model

Ospina and Ferrari [6] introduced the zero-one inflated beta (BIZU) model, which is a mixture between a random variable with Bernoulli distribution with parameter γ , for $0 < \gamma < 1$, and a reparameterized beta distribution of parameters μ and σ . Particular cases of this model follow for the situations of a unique inflated extreme value (zero or one) called BIZ and BIU, respectively. These ideas can be extended to the truncated unit-PSN model.

By considering that the mass point at value zero can be modeled by a Bernoulli random variable with parameter γ , namely Ber($y; \gamma$), and the responses between zero and one can be modeled by the truncated centered unit-power-skew-normal distribution, f(y) with parameter $\boldsymbol{\varphi} = (\mu, \sigma, \lambda, \alpha)^{\top}$, the random variable on the unit interval [0, 1] then follows a truncated unit distribution inflated at zero and one, with parameters ($\vartheta, \gamma, \mu, \sigma, \lambda, \alpha$) if its pdf is represented by the mixture

$$g(y) = egin{cases} artheta(1-\gamma), & ext{if } y = 0, \ (1-artheta)f(y), & ext{if } 0 < y < 1, \ artheta\gamma, & ext{if } y = 1. \end{cases}$$

where $0 < \vartheta, \gamma < 1$. By the construction shown in the previous pdf, it holds that $\operatorname{Prob}[y = 0] = p(1 - \gamma)$ and $\operatorname{Prob}[y = 1] = \vartheta \gamma, \vartheta$ being the mixture parameter. For w, w_0 , and w_1 defined as in Equation (15), the cdf of Y can be written as

$$\mathcal{F}_{Y}(y;\mu,\sigma,\alpha) = \begin{cases} \vartheta(1-\gamma), & \text{if } y \leq 0, \\ \vartheta(1-\gamma) + (1-\vartheta) \frac{\{\mathcal{F}_{SN}(w;\lambda)\}^{\alpha} - \{\mathcal{F}_{SN}(w_{0};\lambda)\}^{\alpha}}{\{\mathcal{F}_{SN}(w_{1};\lambda)\}^{\alpha} - \{\mathcal{F}_{SN}(w_{0};\lambda)\}^{\alpha}}, & \text{if } 0 < y < 1, \\ 1, & \text{if } y \geq 1. \end{cases}$$

Considering the parameterization $\pi_1 = \vartheta \gamma$ and $\pi_0 = \vartheta - \pi_1$, where $0 < \pi_0, \pi_1, \pi_0 + \pi_1 < 1$, the above model can be written in the form

$$g(y) = \begin{cases} \pi_0, & \text{if } y = 0, \\ (1 - \pi_0 - \pi_1)f(y), & \text{if } 0 < y < 1, \\ \pi_1, & \text{if } y = 1. \end{cases}$$

From this model, inflation at zero is obtained by taking $\pi_1 = 0$, and inflation at one follows by taking $\pi_0 = 0$. Now, we introduce covariates in the model. For the discrete part, we assume that the responses in zero and one can be explained by the covariate vectors $\mathbf{z}_{(0)i} = (1, z_{0i1}, \dots, z_{0iq})^{\top}$ and $\mathbf{z}_{(1)i} = (1, z_{1i1}, \dots, z_{1ir})^{\top}$, respectively. Then, following the construction of a logistic model with polytomous response, it is obtained that

$$\pi_{0i} = \operatorname{Prob}(y_i = 0) = \frac{\exp\left(\mathbf{z}_{(0)i}^\top \delta_{(0)}\right)}{1 + \exp\left(\mathbf{z}_{(0)i}^\top \delta_{(0)}\right) + \exp\left(\mathbf{z}_{(1)i}^\top \delta_{(1)}\right)},$$
(21)

$$\pi_{1i} = \operatorname{Prob}(y_i = 1) = \frac{\exp\left(\mathbf{z}_{(1)i}^{\top} \delta_{(1)}\right)}{1 + \exp\left(\mathbf{z}_{(0)i}^{\top} \delta_{(0)}\right) + \exp\left(\mathbf{z}_{(1)i}^{\top} \delta_{(1)}\right)},\tag{22}$$

$$\pi_{01i} = 1 - \pi_{0i} - \pi_{1i} = \operatorname{Prob}(y_i \in (0, 1)) = \frac{1}{1 + \exp(\mathbf{z}_{(0)i}^\top \delta_{(0)}) + \exp(\mathbf{z}_{(1)i}^\top \delta_{(1)})}, \quad (23)$$

where $\delta_{(0)} = (\delta_{00}, \delta_{01}, \dots, \delta_{0q})^{\top}$ and $\delta_{(1)} = (\delta_{10}, \delta_{11}, \dots, \delta_{1r})^{\top}$ are vectors of unknown parameters associated with the covariate vectors $z_{(0)}$ and $z_{(1)}$, respectively. For the continuous part, we continue assuming a truncated centered unit-PSN model, with parameters $(\delta, \sigma, \lambda, \alpha)^{\top}$, defined in (14). One can show that the log-likelihood function for the parameters vector $\boldsymbol{\varphi} = (\delta_{(0)}^{\top}, \delta_{(1)}^{\top}, \delta^{\top}, \sigma, \lambda, \alpha)^{\top}$, given $\mathbf{z}_{(0)}, \mathbf{z}, \mathbf{z}_{(1)}$, and \mathbf{Y} , can be written in the form

$$\ell(\boldsymbol{\varphi}) = \ell(\boldsymbol{\delta}_{(0)}, \boldsymbol{\delta}_{(1)}) + \ell(\boldsymbol{\delta}, \boldsymbol{\sigma}, \boldsymbol{\lambda}, \boldsymbol{\alpha}),$$

where

$$\ell(\delta_{(0)},\delta_{(1)}) = \sum_{0} \mathbf{z}_{(0)i}\delta_{(0)} + \sum_{1} \mathbf{z}_{(1)i}\delta_{(1)} - \sum_{i=1}^{n} \log\Big[1 + \exp\big(\mathbf{z}_{(0)i}^{\top}\delta_{(0)}\big) + \exp\big(\mathbf{z}_{(1)i}^{\top}\delta_{(1)}\big)\Big].$$

and $\ell(\delta, \sigma, \lambda, \alpha)$ is defined in Equation (20).

This guarantees that the parameter estimates can be obtained in separate forms. The score functions and the observed information matrix are obtained by differentiating the log-likelihood function once and twice, with respect to the parameters, respectively. The fact that the log-likelihood function can be broken down into two independent components implies that the Fisher information matrix is a diagonal block, that is, it can be written as

$$\mathbf{I}(\boldsymbol{\varphi}) = \text{Diag}\Big\{\mathbf{I}(\boldsymbol{\delta}_{(0)}, \boldsymbol{\delta}_{(1)}), \mathbf{I}(\boldsymbol{\delta}, \sigma, \lambda, \alpha)\Big\}$$

where $I(\delta_{(0)}, \delta_{(1)})$ is related to the discrete part and $I(\delta, \sigma, \lambda, \alpha)$ to the set of parameters of the continuous part. This matrix coincides with the respective matrix for the previous case of the model for the standard on interval (0, 1).

The elements of the observed information matrix for the discrete part are presented in Appendix A. Taking the expected value to these elements, the Fisher information matrix is obtained. Likewise, given the properties of the inverse of a diagonal matrix, one can conclude that the covariance matrix of the estimators vector can be written as

$$\boldsymbol{\Sigma} = \mathrm{Diag} \Big\{ \mathbf{I}^{-1} \Big(\boldsymbol{\delta}_{(0)}, \boldsymbol{\delta}_{(1)} \Big), \mathbf{I}^{-1} \big(\boldsymbol{\delta}, \boldsymbol{\sigma}, \boldsymbol{\lambda}, \boldsymbol{\alpha} \big) \Big\}.$$

Confidence intervals for φ_r with confidence coefficient $\omega = 100(1 - \psi)\%$ can be obtained as $\hat{\varphi}_r \mp z_{1-\omega/2}\sqrt{\hat{\sigma}(\hat{\varphi}_r)}$. Taking $\rho_{1i} = 0$, the inflated model at zero is followed, UPSNIZ($\delta_{(0)}^{\top}, \delta^{\top}, \sigma, \lambda, \alpha$), and taking $\rho_{0i} = 0$, the inflated model at one is followed, UPSNIU($\delta_{(1)}^{\top}, \delta^{\top}, \sigma, \lambda, \alpha$).

The Log-UPSN Model

In some cases, the random variable Y_i does not follow a UPSN($\mu, \sigma, \lambda, \alpha$) distribution; however, the random variable log(Y_i) can have a UPSN ($\mu, \sigma, \lambda, \alpha$) distribution. In those cases, it is said that the random variable Y_i follows a truncated log-unit-PSN model, and its pdf is given by

$$f(y_i) = \frac{\alpha f_{\rm SN}(w_i;\lambda) \{\mathcal{F}_{\rm SN}(w_i;\lambda)\}^{\alpha-1}}{\sigma(y_i+1) \left(\{\mathcal{F}_{\rm SN}(w_1;\lambda)\}^{\alpha} - \{\mathcal{F}_{\rm SN}(w_0;\lambda)\}^{\alpha}\right)}, \quad 0 < y_i < 1,$$
(24)

where

$$w_0 = -\frac{\mu}{\sigma}, \ w_i = \frac{\log(y_i + 1) - \mu}{\sigma}, \ w_1 = \frac{\log(2) - \mu}{\sigma},$$
 (25)

for i = 1, 2, ..., n. $y_i + 1$ is used instead of y_i due to the non-existence of the logarithm at the point $y_i = 0$. In the cases with covariates, it holds that

$$\log(y_i) \sim \text{UPSN}(\delta, \sigma, \lambda, \alpha), \quad y_i \in (0, 1),$$

where $\mathbf{z}_i = (1, z_{i1}, \dots, z_{ip})^{\top}$. This new model is denoted by LUPSN($\delta, \sigma, \lambda, \alpha$). For $i = 1, 2, \dots, n$ in (25), μ is replaced by η_i which is defined in (8).

The estimation of the parameters follows the same routine as in the case of the UPSN model; likewise, the information matrix of this model can be obtained from the information matrix of the UPSN model. It is enough to change y_i to $\log(y_i + 1)$ in the respective expressions. For this model, in the case of inflation at zero and/or one, that is, in the intervals, [0, 1], [0, 1), and (0, 1] are used for the discrete part—a random variable binomial under a logit link function, similar to the case of the UPSN model.

6. Doubly Censored PSN Model

In this section, the model given by Moulton and Halsey [39] is generalized to the case of a mixture model for two limit points, lower and upper. One of the first models for the fit of the mixture between a discrete and a continuous random variable was proposed by Cragg [40], often called *the two-part model*. Under the Cragg model, the pdf of y_i can be formally written as $g(y_i) = p_i I_i + (1 - p_i) f(y_i)(1 - I_i)$, where p_i is the probability that determines the relative contribution made by the point mass to the general mixture distribution, f is a density function with positive support, $I_i = 0$ if $y_i > 0$ and $I_i = 1$ if $y_i \leq 0$. In this model, the two components are determined by different stochastic processes, so a positive response is necessarily reached from f. On the other hand, a zero comes from the point mass distribution. This model, however, does not consider the situation of a lower limit and that part of the observations may be below the lower limit.

We extend Cragg's model [40] to the case of the doubly censored and centered powerskew-normal model. A random variable is said to be doubly censored when measurements above the upper limit of detection and below the lower limit of detection are taken as those values. The lower and upper detection limits are specified by the researcher and generally depend on the measuring device used to produce the measurements. For our particular case, the lower and upper detection limits are given by $y_i = 0$ and $y_i = 1$, respectively. For $(y_1^*, y_2^*, \dots, y_n^*)$, a random sample where, for $i = 1, 2, \dots, n$, $y_i^* \sim \text{PSN}(\mu, \sigma, \lambda, \alpha)$, the doubly censored random variable PSN between zero and one is defined as

$$y_i = egin{cases} 0, & ext{if } y_i^* \leq 0, \ y_i^*, & ext{if } 0 < y_i^* < 1, \ 1, & ext{if } y_i^* \geq 1 \end{cases}$$

We use the notation $\Upsilon \sim \text{DCPSN}(\mu, \sigma, \lambda, \alpha)$. The contribution of the uncensored observations to the likelihood function, $0 < y_i < 1$, is given by the density function

$$\frac{\alpha}{\sigma} f_{\rm SN}(w_i;\lambda) \{\mathcal{F}_{\rm SN}(w_i;\lambda)\}^{\alpha-1}.$$
(26)

On the other hand, the contribution of the censored observations at y = 0 is given by

$$\mathcal{F}(0) = \{\mathcal{F}_{\mathrm{SN}}(w_{0i};\lambda)\}^{\alpha},\tag{27}$$

while the contribution of the censored observations at y = 1 is given by

$$1 - \mathcal{F}(1) = 1 - \{\mathcal{F}_{SN}(w_{1i};\lambda)\}^{\alpha}.$$
(28)

Then, from (26)–(28), the DCPSN($\mu, \sigma, \lambda, \alpha$) model has a pdf given by

$$f(y_i) = \begin{cases} \{\mathcal{F}_{\mathrm{SN}}(w_{0i};\lambda)\}^{\alpha}, & \text{if } y_i^* \le 0, \\\\ \frac{\alpha}{\sigma} f_{\mathrm{SN}}(w_i;\lambda) \{\mathcal{F}_{\mathrm{SN}}(w_i;\lambda)\}^{\alpha-1}, & \text{if } 0 < y_i^* < 1 \\\\ 1 - \{\mathcal{F}_{\mathrm{SN}}(w_{1i};\lambda)\}^{\alpha}, & \text{if } y_i^* \ge 1 \end{cases}$$

where w_{0i} , w_{1i} and w_i are defined in Equation (15).

The parameters estimation of the DCPSN model can be achieved by maximizing the log-likelihood function given by

$$\begin{split} \ell(\boldsymbol{\varphi}; \mathbf{y}) &= \alpha \sum_{0} \log[\mathcal{F}_{\mathrm{SN}}(w_{0}; \lambda)] + \sum_{1} \log\left[1 - \left\{\mathcal{F}_{\mathrm{SN}}(w_{1}; \lambda)\right\}^{\alpha}\right] \\ &+ \sum_{y_{i} \in (0, 1)} \left\{\log(\alpha) - \log(\sigma) + \log(f_{\mathrm{SN}}(w_{i}; \lambda)) + (\alpha - 1)\log(\mathcal{F}_{\mathrm{SN}}(w_{i}; \lambda))\right\}. \end{split}$$

Generalized Two-Part PSN Model with Covariates

Moulton and Halsey [39] generalize the two-part model by explicitly allowing the possibility that some limited responses are the result of the censoring interval of f. This means that an observed zero can be a realization from the point mass distribution or partial observation of f with a critical value not precisely known, but close to (0, T) for a small prespecified constant T, the lower detection limit.

Formally,

$$g(y_i) = [\pi_i + (1 - \pi_i)F(T)]I_i + (1 - \pi_i)f(y_i)(1 - I_i),$$

where *F* is the cdf associated with the *f* density function. In many studies, T = 0. Therefore, a large family of mixed models can be generated by varying the basic density *f* and the corresponding link function π_i . One can see that if $\pi_i = 0$, for i = 1, ..., n, the Moulton and Halsey [39] model is reduced to the Tobit model.

The two-part model by Moulton and Halsey [39] is extended to the situations of doubly censored responses. If π_0 denotes the proportion of observations below the lower detection limit, $y_i = 0$, and π_1 denotes the proportion of observations above the upper detection limit, $y_i = 1$, then the doubly censored model can be defined from the pdf.

$$g(y_i) = \pi_{0i} + (1 - \pi_{0i} - \pi_{1i})F(0)I_{(-\infty,0])} + (1 - \pi_{0i} - \pi_{1i})f(y_i)I_{(0,1)} + \pi_{1i} + (1 - \pi_{0i} - \pi_{1i})(1 - F(1))I_{[1,\infty)}.$$

with F(0), F(1), as in the generalized doubly censored model of Cragg [40], and $f(y_i)$ is the distribution of the truncated PSN model defined on the (0, 1) interval. Some mixture models have been used in practical applications in different fields such as biology, econ-

omy, agriculture, etc.—to mention a few, the probit/truncated-normal, logit/log-normal, logit/log gamma and probit/log-skew-normal mixture models (see Chai and Bailey [35] and Martínez-Flórez et al. [9]).

We consider an extension of the two-part generalized model for the situations of logit/doubly censored power-skew-normal model, together with covariates in each part of the model. Denoting $\mathbf{z}_{(0)i} = (1, z_{0i1}, \dots, z_{0iq})^{\top}$ and $\mathbf{z}_{(1)i} = (1, z_{1i1}, \dots, z_{1ir})^{\top}$ as auxiliary covariates for the discrete part at zero and one, respectively; denoting a set of covariates $\mathbf{z}_i = (1, z_{i1}, \dots, z_{ip})^{\top}$ for the continuous part at (0, 1); and letting π_0 be the proportion of observations below zero, with $y_i = 0$ being the lower detection limit and π_1 the proportion of observations above one, with $y_i = 1$ as the upper detection limit, then the extension of the Moulton and Halsey [39] model for the case of the doubly censored PSN model is represented by the density function

$$g(y_i) = \begin{cases} \pi_{0i} + (1 - \pi_{0i} - \pi_{1i}) \{ \mathcal{F}_{SN}(w_{0i}; \lambda) \}^{\alpha}, & \text{if } y_i \leq 0, \\ (1 - \pi_{0i} - \pi_{1i}) \frac{\alpha}{\sigma} f_{SN}(w_i; \lambda) \{ \mathcal{F}_{SN}(w_i; \lambda) \}^{\alpha - 1}, & \text{if } 0 < y_i < 1, \\ \pi_{1i} + (1 - \pi_{0i} - \pi_{1i}) (1 - \{ \mathcal{F}_{SN}(w_{1i}; \lambda) \}^{\alpha}), & \text{if } y_i \geq 1. \end{cases}$$

$$(29)$$

where π_{0i} , and π_{1i} are the point mass probabilities at the values zero and one, respectively, and w_{0i} , w_{1i} and w_i are defined as in the equations given in (15). For modeling the responses at the mass points y = 0 and y = 1, we define a binomial random variable with logit link function and polytomous response as defined in Equations (21) to (23).

A more general model, where only a proportion, $\pi = 1 - \pi_0 - \pi_1$, with $0 < \pi_0, \pi_1, \pi < 1$, of censored observations come from the censored PSN model and the rest of the censored observations, say π_0 100%, are located below or at the point y = 0, and π_1 100% are located above or at the point y = 1, can be obtained from the model in Expression (29).

The log-likelihood function for estimating the parameter vector of the model, $\boldsymbol{\varphi} = (\boldsymbol{\delta}_{(0)}^{\top}, \boldsymbol{\delta}_{(1)}^{\top}, \boldsymbol{\delta}^{\top}, \sigma, \lambda, \alpha)^{\top}$ given $\mathbf{z}_{(0)}, \mathbf{z}, \mathbf{z}_{(1)}$, is

$$\ell(\boldsymbol{\varphi}) = \sum_{0} \log \left[\frac{\exp(\mathbf{z}_{(0)i}\delta_{(0)}) + \{\mathcal{F}_{SN}(w_{0i},\lambda)\}^{\alpha}}{1 + \exp(\mathbf{z}_{(0)i}^{\top}\delta_{(0)}) + \exp(\mathbf{z}_{(1)i}^{\top}\delta_{(1)})} \right] \\ + \sum_{1} \log \left[\frac{\exp(\mathbf{z}_{(1)i}\delta_{(1)}) + 1 - \{\mathcal{F}_{SN}(w_{1i},\lambda)\}^{\alpha}}{1 + \exp(\mathbf{z}_{(0)i}^{\top}\delta_{(0)}) + \exp(\mathbf{z}_{(1)i}^{\top}\delta_{(1)})} \right] \\ + n_{01} (\log(\alpha) - \log(\sigma)) \\ + \sum_{i \in (0,1)} \left\{ \log(f_{SN}(w_{i},\lambda)) + (\alpha - 1)\log(\mathcal{F}_{SN}(w_{i},\lambda)) \right\} \\ - \sum_{i \in (0,1)} \log \left[1 + \exp(\mathbf{z}_{(0)i}^{\top}\delta_{(0)}) + \exp(\mathbf{z}_{(1)i}^{\top}\delta_{(1)}) \right].$$
(30)

To obtain the information matrix, we proceed as in the case of the truncated UPSN model in the interval [0, 1]. Again, the right-censored or left-censored cases will be special cases of this model for $\pi_0 = 0$ and $\pi_1 = 0$, respectively. The log-doubly censored case is constructed in the same way as was done for the truncated UPSN model, that is, by taking

$$h(y_i) = \begin{cases} \{\mathcal{F}_{\mathrm{SN}}(w_0;\lambda)\}^{\alpha}, & \text{if } y_i^* \leq 0, \\ \\ \frac{\alpha}{\sigma(y_i+1)} f_{\mathrm{SN}}(w_i;\lambda) \{\mathcal{F}_{\mathrm{SN}}(w_i;\lambda)\}^{\alpha-1}, & \text{if } 0 < y_i^* < 1, \\ \\ 1 - \{\mathcal{F}_{\mathrm{SN}}(w_1;\lambda)\}^{\alpha}, & \text{if } y_i^* \geq 1 \end{cases}$$

with w_0 , w_1 , and w_i defined as in (25).

7. Examples

In this section, we present two examples which allow us to illustrate the applicability of the proposed models.

7.1. Example 1

The first example is related to the household expenditure on food of 38 households taken from Griffiths et al. [41]; this dataset is available in the *betareg* library of the R Development Core Team [42]. The response variable is the relationship or rate of food/income, that is, the proportion of the family income spent on food, while the explanatory variables are: the family income mentioned above and the number of people living in the household. Ferrari and Cribari-Neto [4] modeled this set of variables through the beta regression model; therefore, we will implement the fit of the proportion of family income spent on food, explained through the covariates family income and number of people living at home, using PSN, SN, PN, and normal families of distributions, by using a logit link function. Likewise, we will fit the truncated PSN model with a logit link function. The estimation of the parameters for the previously mentioned models was carried out via maximum likelihood by using the *optim* function of R Development Core Team [42]. To compare the distributions in question, the AIC criteria by Akaike [43] and the corrected AIC (AICC) of Cavanaugh [44] were used. The criteria are defined by

$$AIC = -2\ell(\varphi) + 2p$$
 and $AICC = -2\ell(\varphi) + \frac{2n(p+1)}{n-p-2}$

where *p* is the number of parameters of the model in question. The maximum likelihood estimates, with standard errors in parentheses, are presented in Table 1. According to the results shown by the AIC and AICC criteria, the best fit is the truncated PSN-logit (TPSNL), followed by the PSN and SN models with logit link function.

Table 1. Parameter estimates (standard errors) with logit link function for the normal, Beta, SN, PN PSN, and TPSN regression models.

| Estimator | Beta | Normal-Logit | SN-Logit | PN-Logit | PSN-Logit | TPSNL |
|------------------|----------|--------------|----------|----------|-----------|----------|
| $\hat{\delta}_0$ | -0.6225 | -1.0665 | -1.2348 | -0.8717 | -1.1259 | -1.2810 |
| | (0.2238) | (0.2061) | (0.1123) | (0.1363) | (0.1995) | (0.1663) |
| $\hat{\delta}_1$ | -0.0122 | 0.0997 | 0.1376 | 0.0995 | 0.1246 | 0.1376 |
| | (0.0030) | (0.0086) | (0.0114) | (0.0077) | (0.0179) | (0.0167) |
| $\hat{\delta}_2$ | 0.1184 | -0.0258 | -0.0513 | -0.0266 | -0.0441 | -0.0488 |
| | (0.0353) | (0.0025) | (0.0050) | (0.0020) | (0.0071) | (0.0067) |
| $\hat{\sigma}$ | 35.610 | 0.0374 | 0.0232 | 0.0271 | 0.0189 | 0.0202 |
| | (8.080) | (0.0083) | (0.0021) | (0.0053) | (0.0031) | (0.0029) |
| $\hat{\lambda}$ | | | 6.4714 | | 7.4719 | 7.8010 |
| | | | (0.9057) | | (2.1724) | (1.7758) |
| â | | | | 0.3771 | 0.2814 | 0.3891 |
| | | | | (0.1575) | (0.1463) | (0.1867) |
| AIC | -82.66 | -132.42 | -150.30 | -129.77 | -152.54 | -153.73 |
| AICC | -78.78 | -128.55 | -145.59 | -125.06 | -146.81 | -148.00 |

We now compare the normal-logit (NL) model with the PSN-logit (PSNL) model through a hypothesis test.

$$H_0: (\lambda, \alpha) = (0, 1)$$
 versus $H_1: (\lambda, \alpha) \neq (0, 1)$.

Using the likelihood ratio statistic,

$$\Lambda = \frac{\mathcal{L}_{\rm NL}(\hat{\boldsymbol{\varphi}})}{\mathcal{L}_{\rm PSNL}(\hat{\boldsymbol{\varphi}})}$$

where $\mathcal{L}(\cdot)$ denotes the likelihood function, we obtain

$$-2\log(\Lambda) = 24.1226$$

which is greater than the value of the $\chi^2_{2,95\%} = 5.99$. Thus, the PSN-logit model is a good alternative for fitting the dataset. The PSN-logit model is also compared to the PN-logit (PNL) model and the SN-logit (SNL) model by the hypothesis tests

$$H_{01}: \lambda = 0$$
 versus $H_{11}: \lambda \neq 0$, and $H_{02}: \alpha = 1$ versus $H_{12}: \alpha \neq 1$,

respectively, using the likelihood ratio statistics

$$\Lambda_1 = \frac{\mathcal{L}_{\text{PNL}}(\hat{\boldsymbol{\varphi}})}{\mathcal{L}_{\text{TPSNL}}(\hat{\boldsymbol{\varphi}})} \quad \text{and} \quad \Lambda_2 = \frac{\mathcal{L}_{\text{SNL}}(\hat{\boldsymbol{\varphi}})}{\mathcal{L}_{\text{TPSNL}}(\hat{\boldsymbol{\varphi}})}.$$

The numerical results were

$$-2\log(\Lambda_1) = 24.7694$$
 and $-2\log(\Lambda_2) = 4.2394$

which is greater than $\chi^2_{1,95\%} = 3.84$. The TPSNL model showed a better fit to the data compared to the other considered models.

The transformed martingale residuals r_{MT_i} , introduced by Barros et al. [45], were considered with the goal to identify atypical observations and/or model misspecification. The transformed martingale residuals are defined by

$$r_{MT_i} = \operatorname{sgn}(r_{M_i}) \sqrt{-2[r_{M_i} + v_i \log(v_i - r_{M_i})]}; i = 1, 2, \dots, n_i$$

where $r_{M_i} = v_i + \log(S(e_i, \hat{\varphi}))$ is the martingale residual introduced by Ortega et al. [46]; v_i is an indicator function of the censorship of the *i*th observation with $v_i = 0$ if the *i*th observation is censored and $v_i = 1$ if the *i*th observation is uncensored; sgn(·) is the sign function; and $S(e_i; \hat{\varphi})$ represents the survival function evaluated at e_i , where $\hat{\varphi}$ are the MLE for φ .

The r_{MT_i} plots with a generated envelope for the SN, PSN, and PSNT models are presented in Figure 1a–c. The graphs show that the PSN and PSNT regression models with logit link function present good fits, compared to the rest of the fitted regression models.



Figure 1. Envelope graphs for r_{MT_i} : (a) SN-logit model, (b) PSN-logit, and (c) PSNT-Logit.

7.2. *Example* 2

In the second example, we consider a dataset referring to the broadcasting of cable television in the USA. The data correspond to 282 communities that are essentially individual franchise areas with cable television allocation. The data were taken from The Federal Communications Commission (FCC) and are described in detail in Appendix E of FCC 93–177 [47]. The variable of interest, *Y*, is the proportion of households with cable television that purchase additional services.

The explanatory variables are: Z_1 = the logarithm of the average income in the franchise (lin) given in thousands of dollars; Z_2 = the percentage of children in the franchise (child); Z_3 = the number of channels with local signal (ltv); and (Z_4 =) the age in years of the cable television system (agehe). This dataset is inflated to zero, 68 zeros, which corresponds to 21.98% of the observations, that is, the dataset is left censored. A graph of the response variable Y = proportion of households with cable television that purchase additional services can be seen in Figure 2a. For this set of variables, the beta zero inflated (BIZ) linear regression model, the truncated PSN inflated at zero (PSNIZ) linear regression model, and the generalized two-part PSN model were fitted with detection limit at $y_i = 0$, that is, zero-censored (CGPSN), these last two had a logit link function between zero and one, $y_i \in (0, 1)$. After fitting each of these models, it was found that the significant variables were the logarithm of income (Z_1), for the censored part at $y_i = 0$. The estimates of the parameters and the fitted models are found in Table 2. According to the AIC and AICC criteria, the CGPSN and PSNIZ models present a good fit compared to the BIZ model.

| Ŷ | BIZ | CGPSN | PSNIZ |
|---|--|---|---|
| $\hat{\pi}_{0i}$ | $\frac{\exp\left(-0.7147-0.090z_{4i}\right)}{1\!+\!\exp\left(-0.7147-0.0904z_{4i}\right)}$ | $\frac{\exp\left(-0.7992 - 0.0948 z_{4i}\right)}{1 + \exp\left(-0.7992 - 0.0948 z_{4i}\right)}$ | $\frac{\exp\left(-0.7147 - 0.0904 z_{4i}\right)}{1 + \exp\left(-0.7147 - 0.0904 z_{4i}\right)}$ |
| $\hat{\eta}_i$ | $\frac{\exp(-8.295+0.736z_{1i})}{1+\exp(-8.295+0.736z_{1i})}$ | $\frac{\exp(-8.9465+0.7862z_{1i})}{1+\exp(-8.9465+0.7862z_{1i})}$ | $\frac{\exp(-8.6806+0.7749z_{1i})}{1+\exp(-8.6806+0.7749z_{1i})}$ |
| $\hat{\sigma} \ \hat{\lambda} \ \hat{lpha}$ | 1.7830 (0.0902) | 0.2444 (0.0293) 3.3627 (1.4804) 0.2803 (0.2849) | 0.2402 (0.0194) 2.1795 (1.0192) 0.3592 (0.2804) |
| AIC AICC | 121.95 124.26 | 112.75 115.43 | 111.91 114.57 |

Table 2. Maximum likelihood estimates of the parameters for fitted models.

The r_{MT_i} graphs with envelopes generated for the BIZ, PSNIZ, and GCPSN models are found in Figures 2b and 3a,b, which show that the PSNIZ and CGPSN regression models with logit link present a good fit, compared to the BIZ model—that is to say that these models are new alternatives to fit variables of rate and proportions, such as the proportion of households with cable TV that acquire additional services.



Figure 2. (a) Histogram for the variable γ and (b) envelope graphics for r_{MT_i} , BIZ model.



Figure 3. Envelope graphics for r_{MT_i} , (a) GCPSN-logit model, and (b) PSNIZ-logit model.

8. Conclusions

In this paper, new regression models for fitting data on the intervals (0, 1), [0, 1), (0, 1], or [0, 1] were proposed. The main statistical properties of the proposed models and the problem of the parameters' estimation are studied in detail by using the maximum likelihood method. For the fitting regression model, which can explain the phenomena under study, such as rates or proportions, a logit link function was implemented, with which it is guaranteed that the prediction obtained by the model is between zero and one. The results show that the models present a non-singular information matrix, and the applications show great potential in the proposed models, are more flexible than certain rival models, and fit better to some real datasets.

Author Contributions: Conceptualization, G.M.-F., H.W.G. and R.T.-F.; Methodology, G.M.-F. and H.W.G.; Data curation, G.M.-F., H.W.G. and R.T.-F.; Formal analysis, G.M.-F., H.W.G. and R.T.-F.; Funding acquisition, H.W.G. and R.T.-F.; Investigation, G.M.-F., H.W.G. and R.T.-F.; Resources, G.M.-F., H.W.G. and R.T.-F.; Software, G.M.-F. and H.W.G.; Supervision, G.M.-F.; Validation, G.M.-F. and R.T.-F.; Visualization, G.M.-F., H.W.G. and R.T.-F.; Writing—original draft, G.M.-F., H.W.G. and R.T.-F.; Visualization, G.M.-F.; Visualization, G.M.-F.; Vi

Writing—review & editing, G.M.-F., H.W.G. and R.T.-F. All authors have read and agreed to the published version of the manuscript.

Funding: The research of H.W. Gómez was supported by Grant SEMILLERO UA-2021 (Chile). The research of R. Tovar-Falón and G. Martínez-Flórez was supported by the project: Resolución de Problemas de Situaciones Reales Usando Análisis Estadístico a través del Modelamiento Multidimensional de Tasas y Proporciones; Esquemas de Monitoreamiento para Datos Asimétricos no Normales y una Estrategia Didáctica para el Desarrollo del Pensamiento Lógico-Matemático, Universidad de Córdoba, (Colombia). Code FCB-05-19.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Details about data available are given in Section 7.

Acknowledgments: G. Martínez-Flórez and R. Tovar-Falón acknowledge the support given by Universidad de Córdoba, Montería, Colombia.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Elements of the Observed Information Matrix for the Discrete Part of the Inflated UPSN Model

In this section, expressions for the elements of the observed information matrix for the discrete part of the inflated unit-power-skew-normal model are presented.

$$J_{\delta_{(0)r}\delta_{(0)r'}} = \sum_{i=1}^{n} \frac{z_{(0)ip} z_{(0)ip'} \exp(\mathbf{z}_{(0)i}^{\top} \delta_{(0)})}{\left(1 + \exp(\mathbf{z}_{(0)i}^{\top} \delta_{(0)}) + \exp(\mathbf{z}_{(1)i}^{\top} \delta_{(1)})\right)^{2}} \\ + \sum_{i=1}^{n} \frac{z_{(0)ip} z_{(0)ip'} \exp(\mathbf{z}_{(0)i}^{\top} \delta_{(0)}) \exp(\mathbf{z}_{(1)i}^{\top} \delta_{(1)})}{\left(1 + \exp(\mathbf{z}_{(0)i}^{\top} \delta_{(0)}) + \exp(\mathbf{z}_{(1)i}^{\top} \delta_{(1)})\right)^{2}},$$

$$J_{\delta_{(1)q}\delta_{(0)r}} = -\sum_{i=1}^{n} \frac{z_{(0)ip} z_{(1)iq} \exp(\mathbf{z}_{(0)i}^{\top} \delta_{(0)}) \exp(\mathbf{z}_{(1)i}^{\top} \delta_{(1)})}{\left(1 + \exp(\mathbf{z}_{(0)i}^{\top} \delta_{(0)}) + \exp(\mathbf{z}_{(1)i}^{\top} \delta_{(1)})\right)^{2}},$$

$$\begin{split} J_{\delta_{(1)q}\delta_{(1)q'}} &= \sum_{i=1}^{n} \frac{z_{(1)iq} z_{(1)iq'} \exp(\mathbf{z}_{(1)i}^{\top} \delta_{(1)})}{\left(1 + \exp(\mathbf{z}_{(0)i}^{\top} \delta_{(0)}) + \exp(\mathbf{z}_{(1)i}^{\top} \delta_{(1)})\right)^{2}} \\ &+ \sum_{i=1}^{n} \frac{z_{(1)iq} z_{(1)iq'} \exp(\mathbf{z}_{(1)i}^{\top} \delta_{(1)}) \exp(\mathbf{z}_{(0)i}^{\top} \delta_{(0)})}{\left(1 + \exp(\mathbf{z}_{(0)i}^{\top} \delta_{(0)}) + \exp(\mathbf{z}_{(1)i}^{\top} \delta_{(1)})\right)^{2}} \end{split}$$

References

- Paolino, P. Maximum likelihood estimation of models with beta-distributed dependent variables. *Political Anal.* 2001, *9*, 325–346. [CrossRef]
- 2. Cribari-Neto, F.; Vasconcellos, K.L.P. Nearly unbiased maximum likelhood estimation for the beta distribution. *J. Stat. Comput. Simul.* **2002**, *72*, 107–118. [CrossRef]
- 3. Kieschnick, R.; Mccullough, B.D. Regression analysis of variates observed on (0,1). Stat. Model. 2003, 3, 193–213. [CrossRef]
- 4. Ferrari, S.; Cribari-Neto, F. Beta regression for modelling rates and proportions. J. Appl. Stat. 2004, 31, 799–815. [CrossRef]
- 5. Vasconcellos, K.L.P.; Cribari-Neto, F. Improved maximum likelihood estimation in a new class of beta regression models. *Braz. J. Probab. Stat.* **2005**, *19*, 13–31.
- 6. Ospina, R.; Ferrari, S.L.P. Inflated beta distributions. *Stat. Pap.* **2010**, *51*, 111–126. [CrossRef]
- Ospina, R.; Ferrari, S.L.P. A general class of zero-or-one inflated beta regression models. *Comput. Stat. Data Anal.* 2012, 56, 1609–1623. [CrossRef]

- 8. Bayes, C.; Bazan, J.; García, C. A new robust regression model for proportions. Bayesian Anal. 2012, 7, 841–866. [CrossRef]
- 9. Martínez-Flórez, G.; Bolfarine, H.; Gómez, H.W. Asymmetric regression models with limited responses with an application to antibody response to vaccine. *Biom. J.* 2013, *55*, 156–172. [CrossRef]
- 10. Martínez-Flórez, G.; Bolfarine, H.; Gómez, H.W. Power-models for proportions with zero/one excess. *Appl. Math. Inf. Sci.* 2018, 12, 293–303. [CrossRef]
- 11. Mazucheli, J.; Menezes, A.F.B.; Dey, S. The unit-Birnbaum-Saunders distribution with applications. Chil. J. Stat. 2018, 9, 47–57.
- 12. Mazucheli, J.; Menezes, A.F.B.; Dey, S. Improved maximum-likelihood estimators for the parameters of the unit-gamma distribution. *Commun. Stat.-Theory Methods* **2018**, 47, 3767–3778. [CrossRef]
- 13. Mazucheli, J.; Menezes, A.F.B.; Ghitany, M.E. The unit-Weibull distribution and associated inference. *J. Appl. Probab. Stat.* **2018**, 13, 1–22.
- 14. Menezes, A.F.B.; Mazucheli, J.; Dey, S. The unit-logistic distribution: Different methods of estimation. *Pesqui. Oper.* **2018**, *38*, 555–578. [CrossRef]
- 15. Kumaraswamy, P. A generalized probability density function for double-bounded random processes. *J. Hydrol.* **1980**, *46*, 79–88. [CrossRef]
- 16. Birnbaum, Z.W. Effect of linear truncation on a multinormal population. Ann. Math. Statist. 1950, 21, 272–279. [CrossRef]
- 17. Lehmann, E.L. The power of rank Tests. Ann. Math. Statist. 1953, 24, 23–243. [CrossRef]
- 18. Roberts, C. A Correlation Model Useful in the Study of Twins. J. Am. Stat. Assoc. 1966, 61, 1184–1190. [CrossRef]
- 19. O'Hagan, A.; Leonard, T. Bayes estimation subject to uncertainty about parameter constraints. *Biometrika* **1976**, *63*, 201–202. [CrossRef]
- 20. Fernandez, C.; Steel, M. On bayesian modeling of fat tails and skewness. J. Am. Stat. Assoc. 1998, 93, 359-371.
- 21. Mudholkara, G.S.; Hutson, A. The epsilon–skew–normal distribution for analyzing near-normal data. *J. Stat. Plan. Inference* 2000, 83, 291–309. [CrossRef]
- 22. Azzalini, A. A class of distributions which includes the normal ones. *Scand. J. Stat.* **1985**, *12*, 171–178.
- 23. Durrans, S.R. Distributions of fractional order statistics in hydrology. Water Resour. Res. 1992, 28, 1649–1655. [CrossRef]
- 24. Gupta, R.C.; Gupta, R.D. Generalized skew normal model. Test 2004, 13, 501–524. [CrossRef]
- 25. Arellano-Valle, R.B.; Gómez, H.W.; Quintana, F.A. A new class of skew-normal distributions. *Commun. Stat.-Theory Methods* 2004, 33, 1465–1480. [CrossRef]
- 26. Arellano-Valle, R.B.; Gómez, H.W.; Quintana, F.A. Statistical inference for a general class of asymmetric distributions. *J. Stat. Plan. Inference* **2005**, 128, 427–443. [CrossRef]
- 27. Gómez, H.W.; Venegas, O.; Bolfarine, H. Skew-symmetric distributions generated by the distribution function of the normal distribution. *Environmetrics* 2007, *18*, 395–407. [CrossRef]
- 28. Pewsey, A.; Gómez, H.W.; Bolfarine, H. Likelihood-based inference for power distributions. Test 2012, 21, 775–789. [CrossRef]
- 29. Owen, D.B. Tables for computing bi-variate normal probabilities. Ann. Math. Stat. 1956, 27,1075–1090. [CrossRef]
- 30. Gupta, R.D.; Gupta, R.C. Analyzing skewed data by power normal model. Test 2008, 17, 197–210. [CrossRef]
- 31. Martínez-Flórez, G.; Bolfarine, H.; Gómez, H.W. Skew-normal alpha-power model. Statistics 2014, 48, 1414–1428. [CrossRef]
- 32. Martínez-Flórez, G.; Bolfarine, H.; Gómez, H.W. The log alpha-power asymmetric distribution with application to air pollution. *Environmetrics* **2014**, *25*, 44–56. [CrossRef]
- Mateus-Figueras, G.; Pawlosky-Glanh, V. Una Alternativa a la Distribución Lognormal; Actas del XXVII Congreso Nacional de Estadística e Investigación Operativa (SEIO); Sociedad de Estadística e Investigación Operativa: Barcelona, Spain, 2003; pp. 1849–1858.
- 34. Moulton, L.; Halsey, N.H. A mixed gamma model for regression analyses of quantitative assay data. *Vaccine* **1996**, *14*, 1154–1158. [CrossRef]
- 35. Chai, H.S.; Bailey, K.R. Use of log-skew-normal distribution in analysis of continuous data with a discrete component at zero. *Stat. Med.* **2008**, *27*, 3643–3655. [CrossRef]
- 36. Vilca-Labra, F.; Leiva-Sánchez, V. A new fatigue life model based on the family of skew-elliptical distributions. *Commun. Stat.*-*Theory Methods* **2006**, *35*, 229–244. [CrossRef]
- 37. Leiva-Sánchez, V.; Vilca-Labra, F.; Balakrishnan, N.; Sanhueza, A. A skewed sinh-normal distribution and its properties and application to air pollution. *Commun. Stat.-Theory Methods* **2010**, *36*, 426–443. [CrossRef]
- Martínez-Flórez, G.; Bolfarine, H.; Gómez, Y.M.; Gómez, H.W. A unification of families of Birnbaum-Saunders distributions with applications. REVSTAT 2020, 18, 637–660.
- 39. Moulton, L.H.; Halsey, N.H. A mixture model with detection limits for regression analyses of antibody response to vaccine. *Biometrics* **1995**, *51*, 1570–1578. [CrossRef] [PubMed]
- 40. Cragg, J. Some statistical models for limited dependent variables with application to the demand for durable goods. *Econometrica* **1971**, *39*, 829–844. [CrossRef]
- 41. Griffiths, W.E.; Hill, R.C.; Judge, G.G. Learning and Practicing Econometrics, 1st ed.; Wiley: Hoboken, NJ, USA, 1993.
- 42. R Development Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2017. Available online: http://www.R-project.org (accessed on 12 February 2021).
- 43. Akaike, H. A new look at statistical model identification. IEEE Trans. Autom. Control 1995, AU-19, 716–722.

- 44. Cavanaugh, J.E. Unifying the derivations for the Akaike and corrected Akaike information criteria. *Stat. Probab. Lett.* **1997**, *33*, 201–208. [CrossRef]
- 45. Barros, M.; Paula, G.A.; Leiva, V. A new class of survival regression models with heavy-tailed errors: Robustness and diagnostics. *Lifetime Data Anal.* 2010, 14, 316–332. [CrossRef] [PubMed]
- 46. Ortega, E.M.; Bolfarine, H.; Paula, G.A. Influence diagnostics in generalized log-gamma regression models. *Comput. Stat. Data Anal.* **2003**, *42*, 165–186. [CrossRef]
- 47. Federal Communications Commission. *Report and Order and Further Notice of Proposed Rulemaking*; MM Docker 92-266 (3 May 1993); Federal Communications Commission: Washington, DC, USA, 1993; Volume FCC 93–177, p. 6134.