CHILEAN JOURNAL OF STATISTICS

Edited by Víctor Leiva

Volume 10 Number 2
December 2019
ISSN: 0718-7912 (print)
ISSN: 0718-7920 (online)

Published by the
Chilean Statistical Society
SOCHE
SOCIEDAD CHILENA DE ESTADÍSTICA

Aims

The Chilean Journal of Statistics (ChJS) is an official publication of the Chilean Statistical Society (`www.soche.cl`). The ChJS takes the place of *Revista de la Sociedad Chilena de Estadística*, which was published from 1984 to 2000.

The ChJS is an international scientific forum strongly committed to gender equality, open access of publications and data, and the new era of information. The ChJS covers a broad range of topics in statistics, data science, data mining, artificial intelligence, and big data, including research, survey and teaching articles, reviews, and material for statistical discussion. In particular, the ChJS considers timely articles organized into the following sections: Theory and methods, computation, simulation, applications and case studies, education and teaching, development, evaluation, review, and validation of statistical software and algorithms, review articles, letters to the editor.

The ChJS editorial board plans to publish one volume per year, with two issues in each volume. On some occasions, certain events or topics may be published in one or more special issues prepared by a guest editor.

Contents

Chilean Journal of Statistics
Vol. 10, No. 2, December 2019, 99–122

Statistical Modeling
Research Paper

On an asymmetric extension of the tobit model based on the tilted-normal distribution

Paulo H. Ferreira[1,*], Taciana K.O. Shimizu[2], Adriano K. Suzuki[2], and Francisco Louzada[2]

[1]Department of Statistics, Federal University of Bahia, Salvador, Brazil
[2]Department of Applied Mathematics and Statistics, University of São Paulo, São Carlos, Brazil

(Received: 25 February 2019 · Accepted in final form: 24 April 2019)

Abstract

In this paper, we introduce an asymmetric extension to the tobit model by assuming that the error term follows a tilted-normal distribution. The new model, namely tilted-normal tobit model, can be an useful alternative to other skewed tobit models, such as the skew-normal and power-normal tobit models. The method of maximum likelihood is used for estimating the model parameters. We provide some simulation studies for different sample sizes and parameter settings. In addition, we perform residual and influence diagnostic analysis. Finally, we use American food consumption data to illustrate the better performance of the model introduced.

Keywords: Censored regression model · Influence · Maximum likelihood estimation · Residual and influence diagnostic analysis · Tilted-normal distribution.

Mathematics Subject Classification: Primary 62J05 · Secondary 62N01.

1. Introduction

Tobit models are regression models whose range of the dependent variable is somehow constrained. They were first suggested in a pioneering work by Tobin (1958), to describe the relationship between a non-negative dependent variable (the ratio of total durable goods expenditure to total disposable income, per household) and a vector of independent variables (the age of the household head, and the ratio of liquid asset holdings to total disposable income). Tobin called his model the limited dependent variable model, however it and its various generalizations are popularly known among economists as tobit models, a phrase coined by Goldberger (1964) due to similarities with probit models (the term tobit aims to synthesize in one word Tobin's probit concept). Tobit models are also known as censored regression models. For discussion on properties, parameter estimation and asymptotic properties of estimators, see, e.g., Amemiya (1973, 1984, 1985) and Fair (1977).

*Corresponding author. Email: paulohenri@ufba.br

ISSN: 0718-7912 (print)/ISSN 0718-7920 (online)

http://www.soche.cl/chjs

The tobit specification is adequate for the situation where the sample proportion of zero observations is roughly equivalent to the left tail area of the assumed parametric distribution. The Cragg model (Cragg, 1971), which in the classical literature is known as the two-part model, is an alternative to tobit when the rate of zero observations is quite different from the probability of the left tail obtained with the assumed parametric model.

An interesting way of extending the tobit model is supposing that the probability distribution of the perturbations is no longer normal. For instance, Arellano-Valle et al. (2012) proposed an extension of the tobit model using the Student-t distribution, which is useful for statistical modeling of censored data sets involving observed variables with heavier tails than the normal distribution. Martínez-Flórez et al. (2013) assumed the power-normal distribution (Gupta and Gupta, 2008), thus providing an asymmetric alternative to tobit model. However, such a probability distribution is problematic, that is, of limited use, since it only accommodates low to moderate left-skewness. Moreover, Castro et al. (2014) extended the tobit model to the class of scale mixtures of normal distributions (Andrews and Mallows, 1974) from the Bayesian viewpoint. Other important contributions extending the tobit model by using asymmetric and/or heavy-tailed distributions are Garay et al. (2016, 2017), Mattos et al. (2018), Barros et al. (2018) and Desousa et al. (2018) among many others.

The main purpose of this paper is to focus on the study of the censored regression model, under the assumption that the error term follows the tilted-normal distribution (Maiti and Dey, 2012). Such probability distribution has received some attention in the recent literature, e.g. Louzada et al. (2018) applied the tilted-normal model to compositional data on percentages of players' points in the Brazilian men's volleyball super league 2014/2015. Parameter estimation is performed by using the maximum likelihood (ML) approach and its large sample properties. Application is implemented to American food consumption data set (USDA, 2000), where it is demonstrated that the proposed model can be very useful in fitting real data sets.

The paper is organized as follows. In Section 2, we define the tilted-normal distribution and discuss some of its properties. We present the tilted-normal tobit model and implement inference using the ML approach in Section 3. In Section 4, results of simulation studies reveal the good performance of the estimation approach and the appropriateness of some information criteria in distinguishing among candidate models. Section 5 presents an application to real data on consumption of tomato in the United States in 1994-1996 (USDA, 2000). Model fitting evaluation indicates that the data set in question is much better fitted by the tilted-normal tobit model than by the classic (standard or Type I) tobit model (Tobin, 1958), as well as by other asymmetric models, like the skew-normal tobit model (Hutton and Stanghellini, 2011) and the power-normal tobit model (Martínez-Flórez et al., 2013). Finally, some concluding remarks and directions for future work are given in Section 6. In the work of Hutton and Stanghellini (2011), the skew-normal tobit model was used to address the skewness and right-censoring problems in bounded health scores.

2. The Tilted-Normal Distribution

In this section, we present some basic properties of the tilted-normal distribution, including the probability density function (PDF) and the cumulative distribution function –CDF– (Subsection 2.1), the moments (Subsection 2.2), as well as other relevant issues (Subsection 2.3).

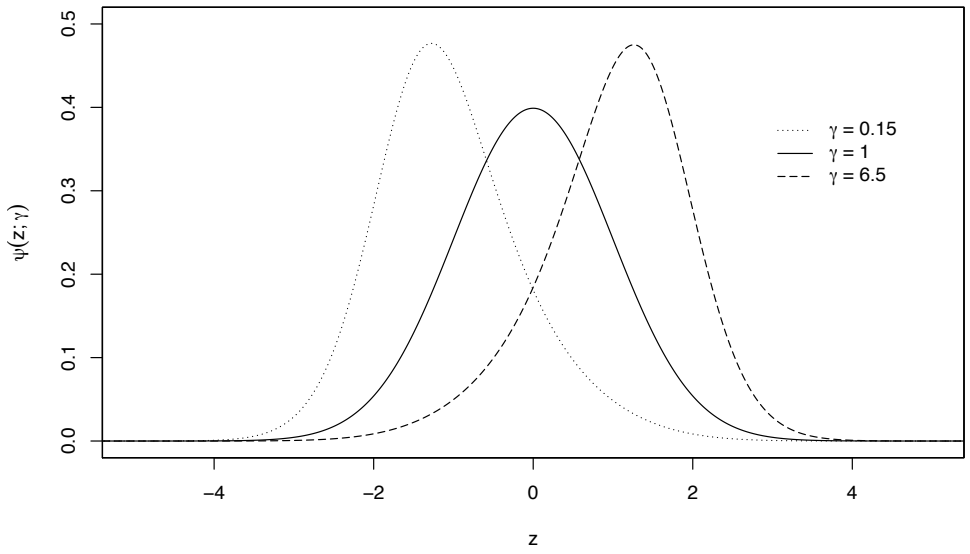

Figure 1. Tilted-normal PDF $\psi(z;\gamma)$ for some values of γ.

2.1 Probabilistic functions

Following the proposition of García et al. (2010) and Maiti and Dey (2012), the tilted-normal distribution is defined as follows. Let Z be a standard normal random variable, that is, $Z \sim \mathrm{N}(0,1)$. Following Marshall and Olkin (1997), the standard tilted-normal distribution, denoted by $\mathrm{TN}(0,1,\gamma)$, has PDF given by

$$\psi(z;\gamma) = \frac{\gamma\phi(z)}{\left[1-(1-\gamma)\left\{1-\Phi(z)\right\}\right]^2}, \quad z \in \mathrm{R},$$

where $\gamma > 0$ is a shape/skewness parameter, ϕ is the PDF of the standard normal distribution and Φ is the CDF of the standard normal distribution. The standard tilted-normal PDF is a unimodal function, which is skewed to the left if $\gamma > 1$ and to the right if $0 < \gamma < 1$, while $\gamma = 1$ indicates a standard normal PDF (Maiti and Dey, 2012). Figure 1 displays a few PDF graphs for different values of γ.

If Z is a random variable from a $\mathrm{TN}(0,1,\gamma)$ distribution, then the location-scale extension of Z, $Y = \mu + \sigma Z$, has PDF given by

$$\psi(y;\mu,\sigma,\gamma) = \frac{\frac{\gamma}{\sigma}\phi\left(\frac{y-\mu}{\sigma}\right)}{\left[1-(1-\gamma)\left\{1-\Phi\left(\frac{y-\mu}{\sigma}\right)\right\}\right]^2}, \tag{1}$$

as well as its CDF given by

$$\Psi(y;\mu,\sigma,\gamma) = \frac{\Phi\left(\frac{y-\mu}{\sigma}\right)}{1-(1-\gamma)\left\{1-\Phi\left(\frac{y-\mu}{\sigma}\right)\right\}}, \tag{2}$$

where $\mu \in \mathrm{R}$ and $\sigma > 0$. We will denote this extension by using the notation $Y \sim \mathrm{TN}(\mu,\sigma,\gamma)$.

2.2 Moments

For the model (1), García et al. (2010) showed that the k-th moment about the origin of the random variable Y is given by

$$\begin{aligned}\mu_k^{'} = \mathrm{E}\left[Y^k\right] &= \int_{-\infty}^{\infty} y^k \psi(y;\mu,\sigma,\gamma)\mathrm{d}y \\ &= \int_0^1 \left[\mu + \sigma\sqrt{2}\,\mathrm{erf}^{-1}\left(\frac{-u+\gamma-u\gamma}{u+\gamma-u\gamma}\right)\right]^k \mathrm{d}u,\end{aligned} \tag{3}$$

where $\mathrm{erf}^{-1}(w) = w\sqrt{\pi}/2 + \mathrm{O}\left(w^3\right) \simeq w\sqrt{\pi}/2$ is the inverse error function.

Although the expression (3) seems to be not available in compact form, the authors verified the following approximations:

$$\begin{aligned}\mu_1^{'} = \mathrm{E}\left[Y\right] &\simeq \frac{2(1-\gamma)^2\mu - \sigma\sqrt{2\pi}\left(1-\gamma^2+2\gamma\log(\gamma)\right)}{2\gamma(1-\gamma)}, \\ \mu_2^{'} = \mathrm{E}\left[Y^2\right] &\simeq \frac{\gamma}{2(1-\gamma)^3}\Big\{2(1-\gamma)^2\mu^2 + 2\left[\left(\gamma^2-1\right)\mu\sigma\sqrt{2\pi} + \left(1+6\gamma\left(1+\gamma\right)\pi\sigma^2\right)\right] \\ &\quad + 4\gamma\sigma\left[(1-\gamma)\mu\sqrt{2\pi} - (1+\gamma)\pi\sigma\right]\log(\gamma)\Big\}, \\ \mu_3^{'} = \mathrm{E}\left[Y^3\right] &\simeq \frac{-1}{4(1-\gamma)\gamma}\Big\{-4(1-\gamma)^3\mu^2\left[-1+\gamma(1+\gamma)\right] + 6(1-\gamma)^2\mu\pi\sigma^2\left[1+\gamma\left(2+\gamma^2\right)\right] \\ &\quad - \sqrt{2\pi}\pi\sigma^3\left[1+2\gamma-5\gamma^2+11\gamma^3+4\gamma^4-\gamma^5\right] - 6(1-\gamma)\gamma\sigma\sqrt{\pi}\Big[2\sqrt{2}(1-\gamma)\mu^2 \\ &\quad + 4\mu\sigma\sqrt{\pi}\left(1-\gamma^2\right) + \sqrt{2}(1+\gamma)^2\pi\sigma^2\Big]\log(\gamma)\Big\}.\end{aligned} \tag{4}$$

These quantities can be used to compute the approximate mean ($\mathrm{E}[Y] = \mu_1^{'}$), variance ($\mathrm{Var}[Y] = \mu_2^{'} - (\mu_1^{'})^2$) and skewness index ($\beta_1 = \mu_3^{'}/(\mu_2^{'})^{3/2}$) of the random variable Y, and are particularly useful for estimating the parameters by the method of moments.

2.3 Others

The model (1) can be extended by considering $\mu_i = \boldsymbol{x}_i^{\top}\boldsymbol{\beta}$, where $\boldsymbol{\beta}$ is an unknown vector of regression coefficients and $\boldsymbol{x}_i$ is a vector of known regressors correlated with the response vector, for $i = 1, \ldots, n$.

Regarding the other skewed distributions that could be used instead of the tilted-normal distribution, Gupta and Gupta (2008) observed that the estimation of the shape parameter of the skew-normal distribution (Azzalini, 1985) is problematic, among others, in the cases where the sample size is not large enough. Monti (2003) noticed that the estimate of the shape parameter is $\hat{\gamma} = \pm\infty$, even when the data are generated by a model with finite γ. Moreover, Pewsey et al. (2012) showed that the Fisher information matrix for the skew-normal distribution is singular under the symmetry hypothesis and, therefore, regularity conditions are not satisfied for the likelihood approach. The same authors also derived the Fisher information matrix for the location-scale version of the power-normal model (Gupta and Gupta, 2008) and have shown that, in addition to its several nice properties, it is not singular for the shape parameter $\gamma = 1$. However, as pointed out by Maiti and

Dey (2012), left-skewness is not so clear and modeling of left-skewed data will be misfit. This is due to the fact that such a distribution can only accommodate low to moderate left-skewness of the data distribution. Hence, the power-normal model is not appropriate for the cases where the data distribution exhibits strong left-skewness. This limitation also applies to the tilted-normal distribution, which can not capture high or moderate levels of skewness (when measured in an appropriate manner). In fact, Rubio and Steel (2012) and Jones (2015) discuss the restrictions of using the Marshall-Olkin transformation for inducing skewness in many symmetric models (including the normal one). Despite such limitation, we demonstrate here that the proposed tobit model based on the tilted-normal distribution can still be very useful in fitting real data sets as in Section 5.

3. The Tilted-Normal Tobit Model

In this section, we introduce the proposed extension of the tobit model using the tilted-normal distribution (Subsection 3.1) and discuss statistical inference based on the ML method (Subsection 3.2).

3.1 Formulation

Let $D_i = I(Y_i > 0)$, where $I(\cdot)$ is the indicator function. The tilted-normal tobit model can be defined by relating the observed dependent variable Y_i^{o} to the original (that is, of theoretical interest), but censored, dependent variable Y_i, as follows:

$$Y_i^{\mathrm{o}} = D_i Y_i \quad \text{and} \quad Y_i = \boldsymbol{x}_i^\top \boldsymbol{\beta} + \epsilon_i, \tag{5}$$

for $i = 1, \ldots, n$, where $\boldsymbol{\beta}$ is a $p \times 1$ unknown parameter vector, $\boldsymbol{x}_i$ is a $p \times 1$ vector of known independent variables, and the errors $\epsilon_i \sim \mathrm{TN}(0, \sigma, \gamma)$.

The value of the location parameter, 0, of ϵ_i implies, from the first expression of (4), that $\mathrm{E}[\epsilon_i] \simeq -\sigma\sqrt{2\pi}\left(1 - \gamma^2 + 2\gamma\log(\gamma)\right)/\left(2\gamma(1-\gamma)\right) < 0$, $\forall \sigma, \gamma > 0$ and $\gamma \neq 1$. Also, for $\sigma > 0$ fixed, $\mathrm{E}[\epsilon_i] \to -\infty$ when $\gamma \to 0^+$ and $\mathrm{E}[\epsilon_i] \to 0$ as $\gamma \to 1^-$. This location parameter choice follows from the work of Martínez-Flórez et al. (2013). However, it could also have been chosen in order to obtain $\mathrm{E}[\epsilon_i] = 0$, as in the normal model, and similarly as in the work of Mattos et al. (2018). Although, even in this case, the expected value of the observed dependent variable Y_i^{o} differs from the location parameter $\mu_i = \boldsymbol{x}_i^\top \boldsymbol{\beta}$, that is, $\mathrm{E}[Y_i^{\mathrm{o}} \mid \boldsymbol{x}_i] = \mathrm{E}[Y_i \mid Y_i > 0, \boldsymbol{x}_i]\, P(Y_i > 0 \mid \boldsymbol{x}_i)$, which, after some steps and considering $\epsilon_i \sim \mathrm{N}(0, \sigma^2)$, results in $\mathrm{E}[Y_i^{\mathrm{o}} \mid \boldsymbol{x}_i] = \Phi\left(\boldsymbol{x}_i^\top\boldsymbol{\beta}/\sigma\right)\left[\boldsymbol{x}_i^\top\boldsymbol{\beta} + \sigma\phi\left(\boldsymbol{x}_i^\top\boldsymbol{\beta}/\sigma\right)/\Phi\left(\boldsymbol{x}_i^\top\boldsymbol{\beta}/\sigma\right)\right] \neq \boldsymbol{x}_i^\top\boldsymbol{\beta}$ (see, e.g., Greene, 2012, Chapter 19).

Note, however, that for the case where $\epsilon_i \sim \mathrm{TN}(0, \sigma, \gamma)$, the main difficulty in obtaining $\mathrm{E}[Y_i^{\mathrm{o}} \mid \boldsymbol{x}_i]$, which would further allow us to analyze the effects of the inequality $\mathrm{E}[Y_i^{\mathrm{o}} \mid \boldsymbol{x}_i] \neq \boldsymbol{x}_i^\top\boldsymbol{\beta}$ on the intercept β_0 of the tilted-normal tobit model, is that there seems to be no explicit known expression for the conditional expectation $\mathrm{E}[Y_i \mid Y_i > 0, \boldsymbol{x}_i]$. Nevertheless, such expected value can be obtained numerically (as shown in Figure 4) or via approximations, e.g., by using some general results of the Marshall and Olkin (1997) family of distributions shown in Cordeiro et al. (2014), among others. We will leave this part of research for our future work.

The tilted-normal tobit model is basically a censored tilted-normal regression model with the tilted-normal distribution replacing the normal distribution for the error term. Thus, parameter estimation for the proposed model is related to parameter estimation for the censored tilted-normal distribution.

For the more general case, where the (known) left-censoring point is $c_i \in \mathrm{R}$, or even for the right-censoring case, we can obtain the estimation results by using the previous model (5), in the same way as stated in Martínez-Flórez et al. (2013).

The next subsection is devoted to implementation of parameter estimation by ML approach and discusses its properties in large samples.

3.2 Estimation

The ML estimators are the most commonly used in the literature. These estimators enjoy desirable properties and can be used for constructing confidence intervals for the model parameters. The normal approximation for the ML estimators in large sample distribution theory is easily handled either analytically or numerically.

In this work, we consider the ML estimation of the unknown parameters of the tilted-normal tobit model. The approach is described as follows.

Let $\boldsymbol{\theta} = \left(\boldsymbol{\beta}^\top, \sigma, \gamma\right)^\top$ be the vector of parameters of interest. Also suppose that the data consist of $n = n_0 + n_1$ observations $(\boldsymbol{x}_1, d_1 y_1), \ldots, (\boldsymbol{x}_n, d_n y_n)$, where n_0 and n_1 are the number of observations on the sets $N_0 = \{i : d_i = 0\} = \{i : y_i = 0\}$ and $N_1 = \{i : d_i = 1\} = \{i : y_i > 0\}$, respectively. Since the unobserved random variables $Y_1, \ldots, Y_n$ are independent, with $Y_i \sim \mathrm{TN}(\boldsymbol{x}_i^\top \boldsymbol{\beta}, \sigma, \gamma)$, we have $P\left(Y_i^{\mathrm{o}} = 0\right) = P\left(Y_i \leq 0\right) = \Phi\left(-\boldsymbol{x}_i^\top \boldsymbol{\beta}/\sigma\right) / \left(1 - (1-\gamma)\left\{1 - \Phi\left(-\boldsymbol{x}_i^\top \boldsymbol{\beta}/\sigma\right)\right\}\right)$, for $i \in N_0$, while for the non-nulls Y_i^{o}s we have that they are distributed as their respective Y_is, that is, $Y_i^{\mathrm{o}} \sim \mathrm{TN}(\boldsymbol{x}_i^\top \boldsymbol{\beta}, \sigma, \gamma)$, for $i \in N_1$. Thus, from the relations mentioned above, the likelihood function for the tilted-normal tobit model is given by

$$L(\boldsymbol{\theta}) = \prod_{i=1}^{n} \left[\frac{\Phi\left(\frac{-\boldsymbol{x}_i^\top \boldsymbol{\beta}}{\sigma}\right)}{1 - (1-\gamma)\left\{1 - \Phi\left(\frac{-\boldsymbol{x}_i^\top \boldsymbol{\beta}}{\sigma}\right)\right\}}\right]^{1-d_i} \left[\frac{\frac{\gamma}{\sigma}\phi\left(\frac{y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}}{\sigma}\right)}{\left(1 - (1-\gamma)\left\{1 - \Phi\left(\frac{y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}}{\sigma}\right)\right\}\right)^2}\right]^{d_i}.$$

Then, the corresponding log-likelihood function is expressed as

$$\begin{aligned}\ell(\boldsymbol{\theta}) = &\sum_{i=1}^{n}(1-d_i)\,\log\left(\Phi\left(\frac{-\boldsymbol{x}_i^\top \boldsymbol{\beta}}{\sigma}\right)\right) - \sum_{i=1}^{n}(1-d_i)\log\left(1 - (1-\gamma)\left\{1 - \Phi\left(\frac{-\boldsymbol{x}_i^\top \boldsymbol{\beta}}{\sigma}\right)\right\}\right) \\ &+ \log(\gamma)\sum_{i=1}^{n} d_i - \log(\sigma)\sum_{i=1}^{n} d_i + \sum_{i=1}^{n} d_i\,\log\left(\phi\left(\frac{y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}}{\sigma}\right)\right) \\ &- 2\sum_{i=1}^{n} d_i\,\log\left(1 - (1-\gamma)\left\{1 - \Phi\left(\frac{y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}}{\sigma}\right)\right\}\right). \end{aligned} \tag{6}$$

The ML estimator $\hat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ is obtained by directly solving the nonlinear equations: $U(\boldsymbol{\beta}) = \mathbf{0}$, $U(\sigma) = 0$ and $U(\gamma) = 0$, where $U(\cdot)$ denotes the score function (see Appendix for analytic description). Note that these equations can not be solved analytically, but we can use, for instance, the `optim` routine (method = "L-BFGS-B") of the R software to solve them numerically. Since regularity conditions are satisfied using the large sample distribution, the distribution of $\hat{\boldsymbol{\theta}}$ can be approximated by a multivariate normal distribution, that is, $\hat{\boldsymbol{\theta}} \sim \mathrm{N}_{p+2}(\boldsymbol{\theta}, [J_{p+2}(\hat{\boldsymbol{\theta}})]^{-1})$, to obtain confidence intervals and hypothesis testing for the parameters of the tilted-normal tobit model, where $J_{p+2}\hat{\boldsymbol{\theta}})$ is the $(p+2)\times(p+2)$ observed information matrix evaluated at $\hat{\boldsymbol{\theta}}$. The elements of the diagonal of $[J_{p+2}(\hat{\boldsymbol{\theta}})]^{-1}$ can be used to approximate the corresponding standard errors.

4. Simulation Studies

In this section, we present the main results obtained from Monte Carlo simulation studies aimed at verifying properties of the ML estimators of the tilted-normal tobit model parameters, with different sample sizes and censoring percentages (Subsection 4.1), as well as investigating the appropriateness of the chosen model selection criteria (Subsection 4.2).

4.1 Parameter recovery study

The first simulation study was based on $M = 2,000$ generated samples of sizes $n = 50$, 100, 300 and 500.

Without loss of generality, we took $\sigma = 1$ and $\beta_1 = 3.5$. It was considered a linear model with a single covariate X whose values were generated according to a $\mathrm{N}(0, 1)$ distribution. We assumed errors $\epsilon_i \sim \mathrm{TN}(0, \sigma, \gamma)$. To ensure a censoring percentage (that is, of zero y_i observations) of approximately 5%, 25%, 50% and 75%, we set the following true values for β_0, respectively (and also for different values of γ):

- For $\gamma = 0.5$: $\beta_0 = 6.4$, 2.8, 0.4 and -2.1;
- For $\gamma = 1$: $\beta_0 = 6$, 2.4, 0.05 and -2.5;
- For $\gamma = 2$: $\beta_0 = 5.5$, 2.1, -0.4 and -2.9;
- For $\gamma = 5$: $\beta_0 = 5$, 1.5, -0.9 and -3.4.

Observed data y_i were taken as $y_i = \max\{\beta_0 + \beta_1 x_i + \epsilon_i, 0\}$. In order to evaluate estimators performance for point estimates, the following quantities were considered: means, biases and mean squared errors (MSEs) of the parameter estimates, and estimated coverage lengths (CLs). We also assessed the performance of the proposed model through the coverage probabilities (CPs) of the 95% normal confidence intervals. ML estimates were computed by using the optim routine (method = "L-BFGS-B") of the R software.

Let $\hat{\boldsymbol{\theta}} = (\widehat{\beta}_0, \widehat{\beta}_1, \widehat{\sigma}, \widehat{\gamma})^\top$ be the ML estimators of the tilted-normal tobit model parameters and $(s_{\widehat{\beta}_0}, s_{\widehat{\beta}_1}, s_{\widehat{\sigma}}, s_{\widehat{\gamma}})$ be their standard errors, which were computed by inverting the observed information matrix. The means, biases, MSEs, CLs and CPs can be estimated by the following equations:

$$\mathrm{Mean}\left(\hat{\theta}_j\right) = \frac{1}{M}\sum_{m=1}^{M}\widehat{\theta}_j^{(m)}, \quad \mathrm{Bias}\left(\hat{\theta}_j\right) = \frac{1}{M}\sum_{m=1}^{M}\left(\widehat{\theta}_j^{(m)} - \theta_j\right),$$

$$\mathrm{MSE}\left(\hat{\theta}_j\right) = \frac{1}{M}\sum_{m=1}^{M}\left(\widehat{\theta}_j^{(m)} - \theta_j\right)^2, \quad \mathrm{CL}\left(\hat{\theta}_j\right) = \frac{3.919928}{M}\sum_{m=1}^{M} s_{\widehat{\theta}_j^{(m)}}$$

and

$$\mathrm{CP}\left(\theta_j\right) = \frac{1}{M}\sum_{m=1}^{M} I\left(\widehat{\theta}_j^{(m)} - 1.959964\, s_{\widehat{\theta}_j^{(m)}} < \theta_j < \widehat{\theta}_j^{(m)} + 1.959964\, s_{\widehat{\theta}_j^{(m)}}\right),$$

for $j = 1, 2, 3, 4$, where $\widehat{\theta}_j^{(m)}$ is the ML estimate of θ_j obtained from the m^{th} replicated sample.

From Tables 1-4, it can be seen that the ML estimates of β_0 and β_1 are unstable, because these parameters are affected by the skewness parameter γ and the proportion of zero observations in the sample. However, the ML estimates become more stable as the

Table 1. Estimation results for the tilted-normal tobit model ($\gamma = 0.5$).

Sample size	Censoring percentage	Parameter	True value	Mean	Bias	MSE	CP	CL
50	5	β_0	6.4	6.4519	0.0519	0.5945	0.9800	3.0781
		β_1	3.5	3.5128	0.0128	0.0242	0.9415	0.5863
		σ	1	1.0012	0.0012	0.0165	0.9340	0.5361
		γ	0.5	0.8487	0.3487	0.9849	0.8565	4.2149
	25	β_0	2.8	2.8342	0.0342	0.6226	0.9840	3.3211
		β_1	3.5	3.5184	0.0184	0.0440	0.9220	0.7732
		σ	1	0.9980	-0.0020	0.0183	0.9400	0.5851
		γ	0.5	0.9119	0.4119	1.2976	0.8615	5.0279
	50	β_0	0.4	0.4397	0.0397	0.7172	0.9870	3.8817
		β_1	3.5	3.5226	0.0226	0.0935	0.9150	1.0962
		σ	1	0.9887	-0.0113	0.0246	0.9345	0.6950
		γ	0.5	0.9361	0.4361	1.3936	0.8435	6.3242
	75	β_0	-2.1	-2.0622	0.0378	1.0908	0.9850	5.1825
		β_1	3.5	3.5664	0.0664	0.3317	0.8945	1.9142
		σ	1	0.9630	-0.0370	0.0449	0.9180	0.9563
		γ	0.5	0.9565	0.4565	1.6456	0.8140	8.5444
100	5	β_0	6.4	6.4122	0.0122	0.3380	0.9770	2.1728
		β_1	3.5	3.5070	0.0070	0.0120	0.9365	0.4178
		σ	1	1.0034	0.0034	0.0083	0.9510	0.3611
		γ	0.5	0.7327	0.2327	0.5402	0.8945	2.5704
	25	β_0	2.8	2.8077	0.0077	0.3922	0.9845	2.4075
		β_1	3.5	3.5094	0.0094	0.0228	0.9325	0.5490
		σ	1	1.0060	0.0060	0.0096	0.9470	0.4015
		γ	0.5	0.7782	0.2782	0.6988	0.8825	3.1039
	50	β_0	0.4	0.4195	0.0195	0.5152	0.9870	2.8807
		β_1	3.5	3.5090	0.0090	0.0447	0.9300	0.7719
		σ	1	1.0072	0.0072	0.0138	0.9500	0.4878
		γ	0.5	0.8287	0.3287	0.8875	0.8645	4.0961
	75	β_0	-2.1	-2.0563	0.0437	0.7604	0.9820	4.0274
		β_1	3.5	3.5217	0.0217	0.1345	0.9295	1.3138
		σ	1	0.9945	-0.0055	0.0232	0.9475	0.6972
		γ	0.5	0.8941	0.3941	1.3086	0.8340	6.2404
300	5	β_0	6.4	6.3988	-0.0012	0.1036	0.9610	1.1943
		β_1	3.5	3.5026	0.0026	0.0039	0.9460	0.2414
		σ	1	1.0002	0.0002	0.0025	0.9555	0.1908
		γ	0.5	0.5756	0.0756	0.1010	0.9210	1.1327
	25	β_0	2.8	2.7921	-0.0079	0.1209	0.9655	1.3199
		β_1	3.5	3.5052	0.0052	0.0068	0.9430	0.3166
		σ	1	1.0012	0.0012	0.0028	0.9560	0.2090
		γ	0.5	0.5978	0.0978	0.1455	0.9180	1.3146
	50	β_0	0.4	0.3847	-0.0153	0.1704	0.9830	1.5937
		β_1	3.5	3.5063	0.0063	0.0135	0.9410	0.4446
		σ	1	1.0023	0.0023	0.0040	0.9530	0.2515
		γ	0.5	0.6397	0.1397	0.2584	0.9150	1.7361
	75	β_0	-2.1	-2.1111	-0.0111	0.3248	0.9865	2.3749
		β_1	3.5	3.5081	0.0081	0.0394	0.9335	0.7416
		σ	1	1.0041	0.0041	0.0083	0.9540	0.3754
		γ	0.5	0.7279	0.2279	0.5313	0.8890	2.9640
500	5	β_0	6.4	6.3938	-0.0062	0.0546	0.9630	0.9079
		β_1	3.5	3.5020	0.0020	0.0023	0.9470	0.1873
		σ	1	0.9999	-0.0001	0.0013	0.9465	0.1436
		γ	0.5	0.5463	0.0463	0.0498	0.9435	0.8300
	25	β_0	2.8	2.7915	-0.0085	0.0667	0.9670	1.0014
		β_1	3.5	3.5040	0.0040	0.0040	0.9445	0.2453
		σ	1	1.0004	0.0004	0.0016	0.9545	0.1566
		γ	0.5	0.5572	0.0572	0.0678	0.9305	0.9389
	50	β_0	0.4	0.3887	-0.0113	0.0900	0.9765	1.2022
		β_1	3.5	3.5031	0.0031	0.0082	0.9405	0.3438
		σ	1	1.0006	0.0006	0.0022	0.9540	0.1865
		γ	0.5	0.5770	0.0770	0.1084	0.9320	1.1725
	75	β_0	-2.1	-2.1190	-0.0190	0.1888	0.9845	1.8003
		β_1	3.5	3.5043	0.0043	0.0228	0.9335	0.5727
		σ	1	1.0035	0.0035	0.0047	0.9530	0.2768
		γ	0.5	0.6499	0.1499	0.2955	0.9050	2.0154

sample size increases. It can also be noted that the MSEs of the ML estimates of β_0, β_1, σ and γ decrease as the sample size increases, which is expected by us since ML estimators are consistent. As pointed out by Martínez-Flórez et al. (2013), bias correction methods, such as bootstrap or jackknife (Efron, 1982; Efron and Tibshirani, 1993), could be tried to improve small sample performance. The main conclusion here is that we are quite safe to work with the ML estimation method if sample sizes are large (that is, greater than 100).

Table 2. Estimation results for the tilted-normal tobit model ($\gamma = 1$).

Sample size	Censoring percentage	Parameter	True value	Mean	Bias	MSE	CP	CL
50	5	β_0	6	6.1417	0.1417	0.5485	0.9840	2.9170
		β_1	3.5	3.5138	0.0138	0.0248	0.9410	0.5952
		σ	1	0.9993	-0.0007	0.0143	0.9420	0.4983
		γ	1	1.3193	0.3193	1.6982	0.8365	6.4760
	25	β_0	2.4	2.5572	0.1572	0.5809	0.9840	3.2330
		β_1	3.5	3.5164	0.0164	0.0434	0.9300	0.7780
		σ	1	0.9934	-0.0066	0.0167	0.9435	0.5595
		γ	1	1.3130	0.3130	1.7376	0.8215	7.2424
	50	β_0	0.05	0.2379	0.1879	0.6702	0.9830	3.8073
		β_1	3.5	3.5213	0.0213	0.0895	0.9220	1.0849
		σ	1	0.9800	-0.0200	0.0232	0.9325	0.6784
		γ	1	1.3061	0.3061	1.8358	0.8025	8.6561
	75	β_0	-2.5	-2.2420	0.2580	1.0763	0.9775	5.1703
		β_1	3.5	3.5570	0.0570	0.3211	0.8985	1.9001
		σ	1	0.9463	-0.0537	0.0459	0.8975	0.9457
		γ	1	1.2107	0.2107	1.9302	0.7510	10.8050
100	5	β_0	6	6.0401	0.0401	0.2902	0.9905	2.0510
		β_1	3.5	3.5072	0.0072	0.0124	0.9390	0.4236
		σ	1	1.0026	0.0026	0.0068	0.9600	0.3323
		γ	1	1.2999	0.2999	1.2101	0.8715	4.5938
	25	β_0	2.4	2.4526	0.0526	0.3278	0.9920	2.3338
		β_1	3.5	3.5107	0.0107	0.0227	0.9350	0.5513
		σ	1	1.0035	0.0035	0.0082	0.9550	0.3836
		γ	1	1.3147	0.3147	1.3749	0.8605	5.3230
	50	β_0	0.05	0.1408	0.0908	0.4298	0.9890	2.8498
		β_1	3.5	3.5107	0.0107	0.0431	0.9285	0.7646
		σ	1	1.0006	0.0006	0.0119	0.9545	0.4777
		γ	1	1.3014	0.3014	1.4665	0.8425	6.4662
	75	β_0	-2.5	-2.3055	0.1945	0.7475	0.9720	4.0618
		β_1	3.5	3.5254	0.0254	0.1306	0.9290	1.3025
		σ	1	0.9843	-0.0157	0.0223	0.9380	0.7052
		γ	1	1.2642	0.2642	1.8509	0.7850	8.9251
300	5	β_0	6	5.9950	-0.0050	0.0855	0.9710	1.1322
		β_1	3.5	3.5029	0.0029	0.0040	0.9475	0.2447
		σ	1	1.0001	0.0001	0.0019	0.9535	0.1745
		γ	1	1.1421	0.1421	0.3988	0.9165	2.2529
	25	β_0	2.4	2.3905	-0.0095	0.1100	0.9760	1.2999
		β_1	3.5	3.5050	0.0050	0.0069	0.9445	0.3177
		σ	1	1.0016	0.0016	0.0025	0.9540	0.2016
		γ	1	1.1843	0.1843	0.5630	0.9055	2.6950
	50	β_0	0.05	0.0414	-0.0086	0.1468	0.9845	1.6296
		β_1	3.5	3.5061	0.0061	0.0132	0.9420	0.4390
		σ	1	1.0021	0.0021	0.0037	0.9625	0.2564
		γ	1	1.2282	0.2282	0.7680	0.9010	3.5501
	75	β_0	-2.5	-2.4638	0.0362	0.2994	0.9780	2.5066
		β_1	3.5	3.5097	0.0097	0.0377	0.9395	0.7342
		σ	1	1.0002	0.0002	0.0081	0.9505	0.4031
		γ	1	1.2854	0.2854	1.2776	0.8595	5.7356
500	5	β_0	6	5.9930	-0.0070	0.0487	0.9665	0.8641
		β_1	3.5	3.5022	0.0022	0.0024	0.9455	0.1899
		σ	1	1.0003	0.0003	0.0011	0.9520	0.1317
		γ	1	1.0855	0.0855	0.1869	0.9355	1.6243
	25	β_0	2.4	2.3894	-0.0106	0.0649	0.9695	0.9893
		β_1	3.5	3.5041	0.0041	0.0041	0.9450	0.2461
		σ	1	1.0012	0.0012	0.0015	0.9565	0.1515
		γ	1	1.1152	0.1152	0.2923	0.9235	1.9189
	50	β_0	0.05	0.0383	-0.0117	0.0913	0.9760	1.2347
		β_1	3.5	3.5035	0.0035	0.0079	0.9360	0.3394
		σ	1	1.0016	0.0016	0.0023	0.9520	0.1912
		γ	1	1.1501	0.1501	0.4136	0.9180	2.4745
	75	β_0	-2.5	-2.5073	-0.0073	0.2004	0.9780	1.9543
		β_1	3.5	3.5052	0.0052	0.0220	0.9375	0.5661
		σ	1	1.0031	0.0031	0.0053	0.9570	0.3070
		γ	1	1.2613	0.2613	0.9612	0.8865	4.4039

4.2 Misspecification study

The second simulation study was based on 3,000 generated samples of size $n = 500$. The main goal was to verify if we could distinguish between the proposed model and the candidate ones, in the light of the data set, based on the adopted model selection criteria: Akaike information criterion (AIC) (Akaike, 1977), corrected AIC (AICc) (Sugiura, 1978; Hurvich and Tsai, 1989), consistent AIC (CAIC) (Bozdogan, 1987; Anderson et al., 1998), Bayesian information criterion (BIC) (Schwarz, 1978), and Hannan-Quinn information criterion (HQIC) (Hannan and Quinn, 1979).

Table 3. Estimation results for the tilted-normal tobit model ($\gamma = 2$).

Sample size	Censoring percentage	Parameter	True value	Mean	Bias	MSE	CP	CL
50	5	β_0	5.5	5.7563	0.2563	0.5080	0.9745	2.8682
		β_1	3.5	3.5112	0.1548	0.0241	0.9395	0.5863
		σ	1	0.9814	0.1141	0.0134	0.9280	0.4887
		γ	2	1.9730	1.6222	2.6308	0.7890	9.9064
	25	β_0	2.1	2.3960	0.2960	0.5682	0.9750	3.1662
		β_1	3.5	3.5123	0.0123	0.0399	0.9335	0.7515
		σ	1	0.9745	-0.0255	0.0161	0.9355	0.5509
		γ	2	1.9013	-0.0987	2.6386	0.7630	10.5477
	50	β_0	-0.4	-0.0271	0.3729	0.7246	0.9690	3.7854
		β_1	3.5	3.5204	0.0204	0.0869	0.9210	1.0689
		σ	1	0.9564	-0.0436	0.0237	0.9165	0.6806
		γ	2	1.7980	-0.2020	2.8972	0.7410	12.1170
	75	β_0	-2.9	-2.4003	0.4997	1.2122	0.9485	5.1567
		β_1	3.5	3.5541	0.0541	0.3009	0.8955	1.8605
		σ	1	0.9143	-0.0857	0.0494	0.8740	0.9397
		γ	2	1.5492	-0.4508	3.0574	0.6790	14.1499
100	5	β_0	5.5	5.6038	0.1038	0.2478	0.9710	2.0787
		β_1	3.5	3.5049	0.0049	0.0118	0.9445	0.4156
		σ	1	0.9935	-0.0065	0.0064	0.9520	0.3417
		γ	2	2.1771	0.1771	2.3208	0.8435	8.0449
	25	β_0	2.1	2.2390	0.1390	0.2746	0.9730	2.3670
		β_1	3.5	3.5076	0.0076	0.0202	0.9405	0.5298
		σ	1	0.9887	-0.0113	0.0077	0.9485	0.3940
		γ	2	2.1057	0.1057	2.4177	0.8260	8.8352
	50	β_0	-0.4	-0.1877	0.2123	0.4311	0.9595	2.9385
		β_1	3.5	3.5110	0.0110	0.0396	0.9325	0.7483
		σ	1	0.9825	-0.0175	0.0121	0.9420	0.5042
		γ	2	2.0225	0.0225	2.6717	0.8000	10.6324
	75	β_0	-2.9	-2.5234	0.3766	0.7913	0.9530	4.0722
		β_1	3.5	3.5261	0.0261	0.1224	0.9285	1.2666
		σ	1	0.9539	-0.0461	0.0237	0.9125	0.7083
		γ	2	1.7361	-0.2639	2.8601	0.7240	12.4367
300	5	β_0	5.5	5.5060	0.0060	0.0872	0.9630	1.2106
		β_1	3.5	3.5026	0.0026	0.0038	0.9455	0.2393
		σ	1	0.9989	-0.0011	0.0023	0.9490	0.1945
		γ	2	2.2276	0.2276	1.3580	0.9025	4.7415
	25	β_0	2.1	2.1138	0.0138	0.1114	0.9630	1.4184
		β_1	3.5	3.5040	0.0040	0.0061	0.9445	0.3043
		σ	1	0.9991	-0.0009	0.0030	0.9475	0.2310
		γ	2	2.2532	0.2532	1.7165	0.8855	5.6553
	50	β_0	-0.4	-0.3504	0.0496	0.1487	0.9525	1.8175
		β_1	3.5	3.5068	0.0068	0.0124	0.9450	0.4281
		σ	1	0.9951	-0.0049	0.0046	0.9470	0.3008
		γ	1	2.1821	0.1821	1.9111	0.8760	7.1213
	75	β_0	-2.9	-2.7586	0.1414	0.2937	0.9545	2.6569
		β_1	3.5	3.5126	0.0126	0.0349	0.9390	0.7116
		σ	1	0.9831	-0.0169	0.0094	0.9260	0.4415
		γ	2	2.0438	0.0438	2.3756	0.8085	9.5972
500	5	β_0	5.5	5.4934	-0.0066	0.0549	0.9615	0.9335
		β_1	3.5	3.5023	0.0023	0.0023	0.9455	0.1857
		σ	1	1.0006	0.0006	0.0014	0.9530	0.1492
		γ	2	2.1799	0.1799	0.8502	0.9230	3.5385
	25	β_0	2.1	2.0939	-0.0061	0.0728	0.9660	1.0957
		β_1	3.5	3.5040	0.0040	0.0037	0.9460	0.2358
		σ	1	1.0012	0.0012	0.0019	0.9525	0.1777
		γ	2	2.2187	0.2187	1.1618	0.9085	4.2419
	50	β_0	-0.4	-0.3887	0.0113	0.1047	0.9570	1.4216
		β_1	3.5	3.5032	0.0032	0.0075	0.9405	0.3305
		σ	1	0.9994	-0.0006	0.0031	0.9510	0.2342
		γ	2	2.2231	0.2231	1.5191	0.8950	5.5508
	75	β_0	-2.9	-2.8224	0.0776	0.2014	0.9575	2.1558
		β_1	3.5	3.5070	0.0070	0.0203	0.9375	0.5472
		σ	1	0.9914	-0.0086	0.0063	0.9375	0.3557
		γ	2	2.1280	0.1280	2.0353	0.8575	8.0622

As in the simulation study presented in the previous subsection, we considered a linear model with a single covariate $X \sim \mathrm{N}(0, 1)$ and set $\beta_1 = 3.5$. We also assumed the following distributions for the errors:

- Normal: that is, $\epsilon_i \sim \mathrm{N}(0, 1)$. To ensure a censoring percentage of about 5%, 25%, 50% and 75%, we took the following true values for β_0, respectively: 6, 2.4, 0.1 and -2.4;
- Skew-normal: that is, $\epsilon_i \sim \mathrm{SN}(0, 1, \gamma)$ (for details on the skew-normal distribution, see Azzalini, 1985). To consider the two kinds of skewness this distribution has (left-skewed if $\gamma < 0$ and right-skewed if $\gamma > 0$, while for $\gamma = 0$ the distribution reduces to the

Table 4. Estimation results for the tilted-normal tobit model ($\gamma = 5$).

Sample size	Censoring percentage	Parameter	True value	Mean	Bias	MSE	CP	CL
50	5	β_0	5	5.3540	0.3540	0.3342	0.9460	2.8536
		β_1	3.5	3.5111	0.0111	0.0201	0.9475	0.5493
		σ	1	0.9412	-0.0588	0.0162	0.8680	0.5013
		γ	5	3.5575	-1.4425	8.1214	0.7440	19.0512
	25	β_0	1.5	1.8797	0.3797	0.3676	0.9520	3.2018
		β_1	3.5	3.5134	0.0134	0.0358	0.9400	0.7148
		σ	1	0.9292	-0.0708	0.0202	0.8700	0.5740
		γ	5	3.4673	-1.5327	8.3702	0.7460	21.0448
	50	β_0	-0.9	-0.4860	0.4140	0.4478	0.9570	3.8100
		β_1	3.5	3.5140	0.0140	0.0727	0.9260	1.0123
		σ	1	0.9083	-0.0917	0.0307	0.8520	0.7005
		γ	5	3.3800	-1.6200	8.8377	0.7575	24.6978
	75	β_0	-3.4	-2.9325	0.4675	0.8327	0.9380	5.1829
		β_1	3.5	3.5171	0.0171	0.2475	0.9065	1.7592
		σ	1	0.8662	-0.1338	0.0603	0.8165	0.9525
		γ	5	3.2562	-1.7438	9.4535	0.7985	31.3246
100	5	β_0	5	5.2705	0.2705	0.2364	0.9070	2.2265
		β_1	3.5	3.5057	0.0057	0.0098	0.9495	0.3863
		σ	1	0.9601	-0.0399	0.0088	0.8905	0.3841
		γ	5	3.8660	-1.1340	6.6349	0.7725	15.6729
	25	β_0	1.5	1.8164	0.3164	0.2865	0.9030	2.5228
		β_1	3.5	3.5080	0.0080	0.0169	0.9440	0.4980
		σ	1	0.9500	-0.0500	0.0111	0.8795	0.4426
		γ	5	3.6951	-1.3049	7.4738	0.7565	17.3047
	50	β_0	-0.9	-0.5265	0.3735	0.3687	0.9255	2.9950
		β_1	3.5	3.5063	0.0063	0.0335	0.9410	0.7008
		σ	1	0.9354	-0.0646	0.0160	0.8755	0.5327
		γ	5	3.4990	-1.5010	8.3924	0.7285	19.4424
	75	β_0	-3.4	-2.9793	0.4207	0.5580	0.9445	4.1795
		β_1	3.5	3.5194	0.0194	0.1060	0.9355	1.2020
		σ	1	0.9051	0.1512	0.0319	0.8505	0.7477
		γ	5	3.3094	-1.6906	9.1418	0.7580	25.4297
300	5	β_0	5	5.1046	0.1046	0.0970	0.9320	1.4321
		β_1	3.5	3.5023	0.0023	0.0031	0.9510	0.2211
		σ	1	0.9848	-0.0152	0.0035	0.9125	0.2501
		γ	5	4.6549	-0.3451	4.4836	0.8570	11.5500
	25	β_0	1.5	1.6473	0.1473	0.1269	0.9185	1.6924
		β_1	3.5	3.5036	0.0036	0.0052	0.9505	0.2844
		σ	1	0.9781	-0.0219	0.0045	0.8970	0.2970
		γ	5	4.4234	-0.5766	5.0965	0.8290	13.1186
	50	β_0	-0.9	-0.6767	0.2233	0.1929	0.8930	2.0705
		β_1	3.5	3.5073	0.0073	0.0105	0.9445	0.3966
		σ	1	0.9638	-0.0362	0.0071	0.8735	0.3650
		γ	5	4.0483	-0.9517	6.0402	0.7815	14.9564
	75	β_0	-3.4	-3.0605	0.3395	0.3464	0.9010	2.8726
		β_1	3.5	3.5092	0.0092	0.0295	0.9345	0.6632
		σ	1	0.9402	-0.0598	0.0142	0.8540	0.5025
		γ	5	3.5831	-1.4169	7.9578	0.7305	18.6356
500	5	β_0	5	5.0518	0.0518	0.0607	0.9445	1.1470
		β_1	3.5	3.5014	0.0014	0.0019	0.9450	0.1711
		σ	1	0.9928	-0.0072	0.0021	0.9460	0.2015
		γ	5	4.9160	-0.0840	3.4303	0.8955	9.5602
	25	β_0	1.5	1.5847	0.0847	0.0826	0.9315	1.3724
		β_1	3.5	3.5032	0.0032	0.0032	0.9470	0.2198
		σ	1	0.9880	-0.0120	0.0029	0.9295	0.2425
		γ	5	4.7424	-0.2576	4.1078	0.8560	11.1365
	50	β_0	-0.9	-0.7555	0.1445	0.1282	0.9100	1.7211
		β_1	3.5	3.5030	0.0030	0.0062	0.9480	0.3060
		σ	1	0.9774	-0.0226	0.0047	0.9030	0.3051
		γ	5	4.4380	-0.5620	5.0809	0.8275	13.2592
	75	β_0	-3.4	-3.1236	0.2764	0.2595	0.8970	2.4194
		β_1	3.5	3.5058	0.0058	0.0173	0.9395	0.5098
		σ	1	0.9553	-0.0447	0.0094	0.8605	0.4235
		γ	5	3.8121	-1.1879	6.8930	0.7495	16.4207

normal one), and ensure a censoring percentage of approximately 5%, 25%, 50% and 75%, we set the following true values for β_0, respectively (and also for different values of shape/skewness parameter γ):

- For $\gamma = -2.2$: $\beta_0 = 6.6$, 3.1, 0.9 and -1.7;
- For $\gamma = -1.2$: $\beta_0 = 6.5$, 3, 0.2 and -1.8;
- For $\gamma = 1.2$: $\beta_0 = 5.4$, 1.8, -0.7 and -3.1;
- For $\gamma = 2.2$: $\beta_0 = 5$, 3, -0.7 and -3.3.

- Power-normal: that is, $\epsilon_i \sim \text{PN}(0, 1, \gamma)$ (for details on the power-normal distribution, see Gupta and Gupta, 2008). To consider the two kinds of skewness this distribution has (left-skewed if $0 < \gamma < 1$ and right-skewed if $\gamma > 1$, while for $\gamma = 1$ the distribution reduces to the normal one), and ensure a censoring percentage of approximately 5%, 25%, 50% and 75%, we assumed the following true values for β_0, respectively (and also for different values of shape/skewness parameter γ):
  - For $\gamma = 0.35$: $\beta_0 = 7.2$, 3.6, 1.1 and -1.5;
  - For $\gamma = 2.8$: $\beta_0 = 5$, 1.5, -1 and -3.2;
  - For $\gamma = 10$: $\beta_0 = 4.2$, 0.7, -1.5 and -4.3.
- Tilted-normal: that is, $\epsilon_i \sim \text{TN}(0, 1, \gamma)$. In order to consider the two kinds of skewness this distribution has, and ensure a censoring percentage of approximately 5%, 25%, 50% and 75%, we set the following true values for β_0, respectively (and also for different values of γ):
  - For $\gamma = 6.5$: $\beta_0 = 5$, 1.5, -1 and -3.5;
  - For $\gamma = 2$: $\beta_0 = 5.5$, 2, -0.5 and -2.8;
  - For $\gamma = 0.5$: $\beta_0 = 6.5$, 2.7, 0.3 and -2.1;
  - For $\gamma = 0.15$: $\beta_0 = 7$, 3.5, 1 and -1.4.

It is important to note that the shape/skewness parameter values presented above, were chosen in order to ensure a skewness measure of approximately -0.5, -0.2, 0.2 and 0.5, respectively (in the order that such values appear), for each error distribution (with the exception of the power-normal distribution for the first case, since -0.5 is less than ≈ -0.48, which is the lowest skewness measure that can be accommodated by such a model). The observed data y_i were taken as $y_i = \max\{\beta_0 + \beta_1 x_i + \epsilon_i, 0\}$, for $i = 1, \ldots, n$.

For each obtained sample and for each situation described above, we applied the following procedures: all four models (tobit-N, tobit-SN, tobit-PN and tobit-TN, where tobit-N stands for the normal tobit model, tobit-SN is the skew-normal tobit model, tobit-PN is the power-normal tobit model, and tobit-TN is the tilted-normal tobit model) were fitted to the data set and then the best one was selected according to the AIC, AICc, CAIC, BIC and HQIC criteria. The proportion of times each model was chosen is shown in Tables 5-9. The results in these tables indicate that the true model from which the sample was generated shows a higher proportion, except for the cases where the degree of asymmetry is weak.

5. Application

In this section, we illustrate the applicability of our proposed tobit-TN model (Subsection 5.2) and its diagnostics (Subsection 5.3) using an American food consumption data set (Subsection 5.1) extracted from the 1994-1996 Continuing Survey of Food Intakes by Individuals (CSFII) (USDA, 2000).

5.1 Data

In the CSFII, two nonconsecutive days of dietary data for individuals of all ages residing in the United States were collected via in-person interviews using 24 hours recall. Each sample person reported the amount of each food item consumed. Where two days were reported, there is also a third record regarding daily averages. Socioeconomic and demographic data for the sample households and their members were also collected in the survey. Here, the size of the extracted sample is $n = 304$ adults aged 20 or older (we only consider one member per household). In our application, presented in detail in this section, we select the amount of tomatoes consumed (in 400 grams) by them as the response variable.

Table 5. The proportion of times each tobit model is selected as the best one according to the AIC criterion.

True model	Fitted model			
	tobit-N	tobit-SN	tobit-PN	tobit-TN
tobit-N 5%	0.8000	0.0130	0.0950	0.0920
tobit-N 25%	0.8037	0.0147	0.0933	0.0883
tobit-N 50%	0.7863	0.0130	0.1047	0.0960
tobit-N 75%	0.7817	0.0280	0.1010	0.0893
tobit-SN 5% ($\gamma = -2.2$)	0.0007	0.5393	0.2753	0.1847
tobit-SN 25%	0.0033	0.4817	0.3367	0.1783
tobit-SN 50%	0.0260	0.4233	0.3490	0.2017
tobit-SN 75%	0.1693	0.2967	0.3347	0.1993
tobit-SN 5% ($\gamma = -1.2$)	0.3130	0.1893	0.2683	0.2293
tobit-SN 25%	0.3860	0.1657	0.2447	0.2037
tobit-SN 50%	0.5240	0.1067	0.2010	0.1683
tobit-SN 75%	0.6037	0.0773	0.1683	0.1507
tobit-SN 5% ($\gamma = 1.2$)	0.3410	0.1273	0.2773	0.2543
tobit-SN 25%	0.4187	0.1143	0.2287	0.2383
tobit-SN 50%	0.5077	0.1027	0.1863	0.2033
tobit-SN 75%	0.6460	0.1057	0.1033	0.1450
tobit-SN 5% ($\gamma = 2.2$)	0.0017	0.5117	0.3120	0.1747
tobit-SN 25%	0.0033	0.5120	0.2967	0.1880
tobit-SN 50%	0.0297	0.4517	0.2837	0.2350
tobit-SN 75%	0.2013	0.3620	0.1860	0.2507
tobit-PN 5% ($\gamma = 0.35$)	0.2917	0.0000	0.5177	0.1907
tobit-PN 25%	0.3367	0.0000	0.4820	0.1813
tobit-PN 50%	0.4227	0.0010	0.4180	0.1583
tobit-PN 75%	0.5593	0.0030	0.3080	0.1297
tobit-PN 5% ($\gamma = 2.8$)	0.3530	0.1263	0.2850	0.2357
tobit-PN 25%	0.4067	0.1150	0.2477	0.2307
tobit-PN 50%	0.5173	0.1073	0.1760	0.1993
tobit-PN 75%	0.6273	0.1013	0.1257	0.1457
tobit-PN 5% ($\gamma = 10$)	0.0070	0.3173	0.4350	0.2407
tobit-PN 25%	0.0587	0.2870	0.3987	0.2557
tobit-PN 50%	0.0967	0.2727	0.3657	0.2650
tobit-PN 75%	0.3153	0.2593	0.1913	0.2340
tobit-TN 5% ($\gamma = 6.5$)	0.0030	0.0000	0.1137	0.8833
tobit-TN 25%	0.0147	0.0027	0.1460	0.8367
tobit-TN 50%	0.0607	0.0013	0.1970	0.7410
tobit-TN 75%	0.2023	0.0007	0.2613	0.5357
tobit-TN 5% ($\gamma = 2$)	0.3023	0.0003	0.2733	0.4240
tobit-TN 25%	0.3733	0.0000	0.2660	0.3607
tobit-TN 50%	0.4643	0.0007	0.2420	0.2930
tobit-TN 75%	0.5910	0.0057	0.2103	0.1930
tobit-TN 5% ($\gamma = 0.5$)	0.3283	0.1277	0.1737	0.3703
tobit-TN 25%	0.4010	0.1273	0.1477	0.3240
tobit-TN 50%	0.4893	0.1123	0.1257	0.2727
tobit-TN 75%	0.6457	0.1200	0.0810	0.1533
tobit-TN 5% ($\gamma = 0.15$)	0.0033	0.1537	0.1577	0.6853
tobit-TN 25%	0.0047	0.1690	0.1887	0.6377
tobit-TN 50%	0.0237	0.1720	0.2080	0.5963
tobit-TN 75%	0.1053	0.2527	0.1787	0.4633

Table 10 presents the definitions and sample statistics for all considered variables, where we see that the proportion of tomato-consuming individuals in the data set is around 70%. Among those consuming, an individual on average consumes 66.12 grams of tomatoes per day. The histogram and boxplots of tomato consumption are presented in Figures 2 and 3, respectively. Proposed by Hubert and Vandervieren (2008) and used when the data are skewed distributed, the adjusted boxplot (see Figure 3 right panel) indicates that some potential outliers identified by the usual boxplot (see Figure 3 left panel) are not outliers.

Table 11 shows asymmetry and kurtosis coefficients for complete data and also for positive ys. Note that values for the asymmetry and kurtosis coefficients justify using the skewed alternatives to the tobit-N model, e.g. the proposed tobit-TN model.

5.2 Model results

Following Martínez-Flórez et al. (2013), a more emphatic indication that an asymmetric model should be considered comes from testing the hypothesis of a tobit-N model against

Table 6. The proportion of times each tobit model is selected as the best one according to the AICc criterion.

True model	Fitted model			
	tobit-N	tobit-SN	tobit-PN	tobit-TN
tobit-N 5%	0.8050	0.0123	0.0933	0.0893
tobit-N 25%	0.8093	0.0147	0.0900	0.0860
tobit-N 50%	0.7923	0.0123	0.1013	0.0940
tobit-N 75%	0.7887	0.0273	0.0977	0.0863
tobit-SN 5% ($\gamma = -2.2$)	0.0007	0.5393	0.2753	0.1847
tobit-SN 25%	0.0033	0.4817	0.3367	0.1783
tobit-SN 50%	0.0273	0.4233	0.3480	0.2013
tobit-SN 75%	0.1720	0.2967	0.3323	0.1990
tobit-SN 5% ($\gamma = -1.2$)	0.3177	0.1887	0.2660	0.2277
tobit-SN 25%	0.3907	0.1633	0.2433	0.2027
tobit-SN 50%	0.5317	0.1057	0.1977	0.1650
tobit-SN 75%	0.6087	0.0773	0.1660	0.1480
tobit-SN 5% ($\gamma = 1.2$)	0.3493	0.1267	0.2737	0.2503
tobit-SN 25%	0.4240	0.1137	0.2263	0.2360
tobit-SN 50%	0.5147	0.1013	0.1833	0.2007
tobit-SN 75%	0.6517	0.1050	0.1010	0.1423
tobit-SN 5% ($\gamma = 2.2$)	0.0020	0.5113	0.3120	0.1747
tobit-SN 25%	0.0033	0.5120	0.2967	0.1880
tobit-SN 50%	0.0303	0.4513	0.2833	0.2350
tobit-SN 75%	0.2037	0.3617	0.1860	0.2487
tobit-PN 5% ($\gamma = 0.35$)	0.2963	0.0000	0.5150	0.1887
tobit-PN 25%	0.3440	0.0000	0.4767	0.1793
tobit-PN 50%	0.4310	0.0010	0.4133	0.1547
tobit-PN 75%	0.5660	0.0030	0.3030	0.1280
tobit-PN 5% ($\gamma = 2.8$)	0.3597	0.1257	0.2817	0.2330
tobit-PN 25%	0.4123	0.1140	0.2453	0.2283
tobit-PN 50%	0.5233	0.1070	0.1737	0.1960
tobit-PN 75%	0.6340	0.1007	0.1233	0.1420
tobit-PN 5% ($\gamma = 10$)	0.0077	0.3173	0.4347	0.2403
tobit-PN 25%	0.0603	0.2867	0.3977	0.2553
tobit-PN 50%	0.0983	0.2723	0.3653	0.2640
tobit-PN 75%	0.3210	0.2583	0.1883	0.2323
tobit-TN 5% ($\gamma = 6.5$)	0.0033	0.0000	0.1137	0.8830
tobit-TN 25%	0.0153	0.0027	0.1457	0.8363
tobit-TN 50%	0.0627	0.0013	0.1963	0.7397
tobit-TN 75%	0.2070	0.0007	0.2600	0.5323
tobit-TN 5% ($\gamma = 2$)	0.3073	0.0003	0.2717	0.4207
tobit-TN 25%	0.3783	0.0000	0.2643	0.3573
tobit-TN 50%	0.4707	0.0007	0.2393	0.2893
tobit-TN 75%	0.5957	0.0057	0.2073	0.1913
tobit-TN 5% ($\gamma = 0.5$)	0.3333	0.1273	0.1720	0.3673
tobit-TN 25%	0.4053	0.1273	0.1460	0.3213
tobit-TN 50%	0.4957	0.1110	0.1237	0.2697
tobit-TN 75%	0.6537	0.1183	0.0787	0.1493
tobit-TN 5% ($\gamma = 0.15$)	0.0037	0.1537	0.1577	0.6850
tobit-TN 25%	0.0050	0.1690	0.1887	0.6373
tobit-TN 50%	0.0243	0.1717	0.2080	0.5960
tobit-TN 75%	0.1083	0.2517	0.1777	0.4623

an asymmetric tobit model (e.g. the tobit-TN model), that is,

$$\mathrm{H}_0 : \gamma = 1 \quad \text{versus} \quad \mathrm{H}_1 : \gamma \neq 1,$$

using the likelihood ratio statistic:

$$\Lambda = \frac{L_{\text{tobit-N}}(\boldsymbol{\theta})}{L_{\text{tobit-TN}}(\boldsymbol{\theta})}.$$

This leads to the observed value: $-2\log(\Lambda) = 50.5177$, which is greater than the 5% critical value of the Chi-square distribution with one degree of freedom, given by $\chi^2_{1;0.95} = 3.8415$. Therefore, we can conclude that the tobit-TN model fits the American food consumption data set (tomato consumption) better than the standard tobit model (that is, the tobit-N model).

Table 12 presents the parameter estimates for the tobit-N and tobit-TN models, as well as for the other asymmetric alternatives, such as the tobit-SN and tobit-PN models. Notice that all the information criteria choose the tobit-TN model as the best one.

Table 7. The proportion of times each tobit model is selected as the best one according to the CAIC criterion.

True model	Fitted model			
	tobit-N	tobit-SN	tobit-PN	tobit-TN
tobit-N 5%	0.9897	0.0017	0.0063	0.0023
tobit-N 25%	0.9900	0.0013	0.0037	0.0050
tobit-N 50%	0.9873	0.0020	0.0057	0.0050
tobit-N 75%	0.9787	0.0050	0.0087	0.0077
tobit-SN 5% ($\gamma = -2.2$)	0.0220	0.5363	0.2640	0.1777
tobit-SN 25%	0.0670	0.4707	0.2997	0.1627
tobit-SN 50%	0.2410	0.3657	0.2390	0.1543
tobit-SN 75%	0.6440	0.1720	0.1020	0.0820
tobit-SN 5% ($\gamma = -1.2$)	0.7730	0.0703	0.0833	0.0733
tobit-SN 25%	0.8320	0.0507	0.0633	0.0540
tobit-SN 50%	0.9110	0.0287	0.0313	0.0290
tobit-SN 75%	0.9373	0.0177	0.0200	0.0250
tobit-SN 5% ($\gamma = 1.2$)	0.7913	0.0487	0.0840	0.0760
tobit-SN 25%	0.8450	0.0377	0.0563	0.0610
tobit-SN 50%	0.8960	0.0280	0.0340	0.0420
tobit-SN 75%	0.9457	0.0220	0.0107	0.0217
tobit-SN 5% ($\gamma = 2.2$)	0.0373	0.4980	0.3020	0.1627
tobit-SN 25%	0.0667	0.4823	0.2793	0.1717
tobit-SN 50%	0.2513	0.3603	0.2193	0.1690
tobit-SN 75%	0.6440	0.1723	0.0793	0.1043
tobit-PN 5% ($\gamma = 0.35$)	0.7610	0.0000	0.1753	0.0637
tobit-PN 25%	0.8137	0.0000	0.1330	0.0533
tobit-PN 50%	0.8637	0.0003	0.0983	0.0377
tobit-PN 75%	0.9437	0.0010	0.0430	0.0123
tobit-PN 5% ($\gamma = 2.8$)	0.7900	0.0453	0.0907	0.0740
tobit-PN 25%	0.8363	0.0380	0.0677	0.0580
tobit-PN 50%	0.8940	0.0333	0.0367	0.0360
tobit-PN 75%	0.9440	0.0203	0.0147	0.0210
tobit-PN 5% ($\gamma = 10$)	0.0947	0.2923	0.3963	0.2167
tobit-PN 25%	0.3253	0.2150	0.2793	0.1803
tobit-PN 50%	0.4680	0.1657	0.2073	0.1590
tobit-PN 75%	0.7557	0.0950	0.0627	0.0867
tobit-TN 5% ($\gamma = 6.5$)	0.0570	0.0000	0.1033	0.8397
tobit-TN 25%	0.1447	0.0027	0.1180	0.7347
tobit-TN 50%	0.3437	0.0013	0.1237	0.5313
tobit-TN 75%	0.6633	0.0003	0.0950	0.2413
tobit-TN 5% ($\gamma = 2$)	0.7707	0.0000	0.0977	0.1317
tobit-TN 25%	0.8223	0.0000	0.0757	0.1020
tobit-TN 50%	0.8717	0.0000	0.0593	0.0690
tobit-TN 75%	0.9400	0.0007	0.0310	0.0283
tobit-TN 5% ($\gamma = 0.5$)	0.7747	0.0567	0.0530	0.1157
tobit-TN 25%	0.8467	0.0467	0.0307	0.0760
tobit-TN 50%	0.8917	0.0330	0.0277	0.0477
tobit-TN 75%	0.9503	0.0270	0.0070	0.0157
tobit-TN 5% ($\gamma = 0.15$)	0.0503	0.1487	0.1463	0.6547
tobit-TN 25%	0.0860	0.1570	0.1680	0.5890
tobit-TN 50%	0.2110	0.1473	0.1557	0.4860
tobit-TN 75%	0.4707	0.1560	0.0993	0.2740

In Figure 4, we show a scatter plot of $\hat{\mathrm{E}}\left[Y_i^{\mathrm{o}} \mid \boldsymbol{x}_i\right]$ (calculated numerically using adaptive quadrature implemented by the integrate function in R) versus $\boldsymbol{x}_i\hat{\boldsymbol{\beta}}$, for $i = 1, 2, \ldots, 304$. Besides the fact that $\hat{\mathrm{E}}\left[Y_i^{\mathrm{o}} \mid \boldsymbol{x}_i\right] \neq \boldsymbol{x}_i\hat{\boldsymbol{\beta}}$, there seems to be a slightly quadratic relationship between these two quantities.

5.3 Residual and influence diagnostic analysis

Next, we perform a residual analysis to detect atypical observations and/or model misspecification. We can generate envelopes as suggested by Atkinson (1981), based on the generalized Cox-Snell (GCS) residuals, which for the case of tilted-normal distribution are defined as $r_i^{\mathrm{GCS}} = -\log\left(1 - \hat{\Psi}(y_i; \hat{\mu}_i, \hat{\sigma}, \hat{\gamma})\right)$, $i = 1, \ldots, n$, where $\hat{\Psi}$ denotes the CDF (2) fitted to the data. The results (half-normal plots with simulated envelopes) are shown in Figure 5, from which we can see that the tobit-TN model fits better the American food consumption data set.

In order to identify influential observations, we can generate graphs of the generalized Cook distance (Cook, 1977, 1986), where a high value of this measure indicates that the

Table 8. The proportion of times each tobit model is selected as the best one according to the BIC criterion.

True model	Fitted model			
	tobit-N	tobit-SN	tobit-PN	tobit-TN
tobit-N 5%	0.9857	0.0020	0.0090	0.0033
tobit-N 25%	0.9827	0.0017	0.0080	0.0077
tobit-N 50%	0.9807	0.0023	0.0093	0.0077
tobit-N 75%	0.9703	0.0057	0.0120	0.0120
tobit-SN 5% ($\gamma = -2.2$)	0.0120	0.5380	0.2697	0.1803
tobit-SN 25%	0.0460	0.4760	0.3110	0.1670
tobit-SN 50%	0.1800	0.3850	0.2680	0.1670
tobit-SN 75%	0.5503	0.2033	0.1380	0.1083
tobit-SN 5% ($\gamma = -1.2$)	0.7137	0.0830	0.1097	0.0937
tobit-SN 25%	0.7873	0.0637	0.0820	0.0670
tobit-SN 50%	0.8737	0.0343	0.0477	0.0443
tobit-SN 75%	0.9067	0.0247	0.0317	0.0370
tobit-SN 5% ($\gamma = 1.2$)	0.7330	0.0613	0.1060	0.0997
tobit-SN 25%	0.8017	0.0487	0.0750	0.0747
tobit-SN 50%	0.8623	0.0370	0.0457	0.0550
tobit-SN 75%	0.9270	0.0310	0.0147	0.0273
tobit-SN 5% ($\gamma = 2.2$)	0.0250	0.5037	0.3057	0.1657
tobit-SN 25%	0.0470	0.4937	0.2840	0.1753
tobit-SN 50%	0.2010	0.3850	0.2330	0.1810
tobit-SN 75%	0.5657	0.2083	0.0987	0.1273
tobit-PN 5% ($\gamma = 0.35$)	0.6983	0.0000	0.2197	0.0820
tobit-PN 25%	0.7567	0.0000	0.1760	0.0673
tobit-PN 50%	0.8130	0.0003	0.1373	0.0493
tobit-PN 75%	0.9107	0.0010	0.0660	0.0223
tobit-PN 5% ($\gamma = 2.8$)	0.7363	0.0577	0.1117	0.0943
tobit-PN 25%	0.7937	0.0453	0.0850	0.0760
tobit-PN 50%	0.8547	0.0437	0.0527	0.0490
tobit-PN 75%	0.9147	0.0330	0.0227	0.0297
tobit-PN 5% ($\gamma = 10$)	0.0657	0.3030	0.4070	0.2243
tobit-PN 25%	0.2693	0.2340	0.3020	0.1947
tobit-PN 50%	0.3963	0.1863	0.2377	0.1797
tobit-PN 75%	0.6933	0.1153	0.0810	0.1103
tobit-TN 5% ($\gamma = 6.5$)	0.0417	0.0000	0.1053	0.8530
tobit-TN 25%	0.1083	0.0027	0.1240	0.7650
tobit-TN 50%	0.2810	0.0013	0.1377	0.5800
tobit-TN 75%	0.5893	0.0003	0.1200	0.2903
tobit-TN 5% ($\gamma = 2$)	0.7033	0.0000	0.1223	0.1743
tobit-TN 25%	0.7773	0.0000	0.0973	0.1253
tobit-TN 50%	0.8260	0.0000	0.0777	0.0963
tobit-TN 75%	0.9147	0.0007	0.0407	0.0440
tobit-TN 5% ($\gamma = 0.5$)	0.7193	0.0670	0.0670	0.1467
tobit-TN 25%	0.7950	0.0567	0.0457	0.1027
tobit-TN 50%	0.8497	0.0443	0.0350	0.0710
tobit-TN 75%	0.9313	0.0340	0.0100	0.0247
tobit-TN 5% ($\gamma = 0.15$)	0.0350	0.1513	0.1497	0.6640
tobit-TN 25%	0.0627	0.1597	0.1743	0.6033
tobit-TN 50%	0.1647	0.1553	0.1673	0.5127
tobit-TN 75%	0.4023	0.1803	0.1110	0.3063

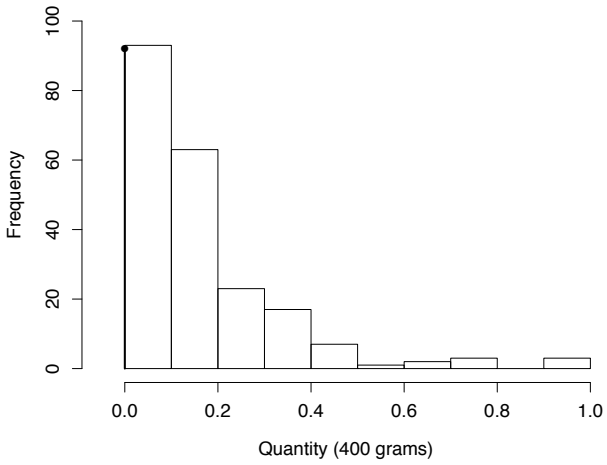

Figure 2. Distribution of the tomato consumption. The vertical line at zero on x axis represents individuals that did not consume tomatoes during the survey period.

Table 9. The proportion of times each tobit model is selected as the best one according to the HQIC criterion.

True model	Fitted model			
	tobit-N	tobit-SN	tobit-PN	tobit-TN
tobit-N 5%	0.9300	0.0047	0.0327	0.0327
tobit-N 25%	0.9330	0.0070	0.0337	0.0263
tobit-N 50%	0.9260	0.0053	0.0380	0.0307
tobit-N 75%	0.9037	0.0140	0.0433	0.0390
tobit-SN 5% ($\gamma = -2.2$)	0.0033	0.5390	0.2743	0.1833
tobit-SN 25%	0.0117	0.4813	0.3313	0.1757
tobit-SN 50%	0.0653	0.4167	0.3250	0.1930
tobit-SN 75%	0.3270	0.2653	0.2460	0.1617
tobit-SN 5% ($\gamma = -1.2$)	0.4983	0.1393	0.1913	0.1710
tobit-SN 25%	0.5780	0.1143	0.1730	0.1347
tobit-SN 50%	0.7083	0.0653	0.1263	0.1000
tobit-SN 75%	0.7707	0.0520	0.0893	0.0880
tobit-SN 5% ($\gamma = 1.2$)	0.5360	0.1010	0.1877	0.1753
tobit-SN 25%	0.6053	0.0833	0.1557	0.1557
tobit-SN 50%	0.6940	0.0680	0.1147	0.1233
tobit-SN 75%	0.8113	0.0643	0.0510	0.0733
tobit-SN 5% ($\gamma = 2.2$)	0.0053	0.5110	0.3110	0.1727
tobit-SN 25%	0.0143	0.5080	0.2933	0.1843
tobit-SN 50%	0.0820	0.4347	0.2687	0.2147
tobit-SN 75%	0.3580	0.3020	0.1463	0.1937
tobit-PN 5% ($\gamma = 0.35$)	0.4790	0.0000	0.3800	0.1410
tobit-PN 25%	0.5470	0.0000	0.3297	0.1233
tobit-PN 50%	0.6297	0.0010	0.2707	0.0987
tobit-PN 75%	0.7597	0.0023	0.1723	0.0657
tobit-PN 5% ($\gamma = 2.8$)	0.5430	0.0960	0.1943	0.1667
tobit-PN 25%	0.6030	0.0817	0.1657	0.1497
tobit-PN 50%	0.7023	0.0790	0.1057	0.1130
tobit-PN 75%	0.8033	0.0663	0.0593	0.0710
tobit-PN 5% ($\gamma = 10$)	0.0193	0.3153	0.4290	0.2363
tobit-PN 25%	0.1337	0.2703	0.3667	0.2293
tobit-PN 50%	0.2130	0.2457	0.3123	0.2290
tobit-PN 75%	0.4873	0.1940	0.1417	0.1770
tobit-TN 5% ($\gamma = 6.5$)	0.0097	0.0000	0.1120	0.8783
tobit-TN 25%	0.0387	0.0027	0.1400	0.8187
tobit-TN 50%	0.1307	0.0013	0.1770	0.6910
tobit-TN 75%	0.3650	0.0007	0.2043	0.4300
tobit-TN 5% ($\gamma = 2$)	0.4937	0.0000	0.2083	0.2980
tobit-TN 25%	0.5740	0.0000	0.1870	0.2390
tobit-TN 50%	0.6540	0.0000	0.1583	0.1877
tobit-TN 75%	0.7703	0.0027	0.1127	0.1143
tobit-TN 5% ($\gamma = 0.5$)	0.5143	0.1037	0.1197	0.2623
tobit-TN 25%	0.6013	0.0977	0.0900	0.2110
tobit-TN 50%	0.6953	0.0793	0.0710	0.1543
tobit-TN 75%	0.8103	0.0773	0.0330	0.0793
tobit-TN 5% ($\gamma = 0.15$)	0.0073	0.1533	0.1573	0.6820
tobit-TN 25%	0.0177	0.1663	0.1863	0.6297
tobit-TN 50%	0.0637	0.1673	0.1970	0.5720
tobit-TN 75%	0.0637	0.1673	0.1970	0.5720

Table 10. Variable definitions and sample statistics ($n = 304$).

Variable	Definition	Mean	Standard Deviation
Dependent variable: amount consumed			
Tomato (in 400 grams)	Quantity of tomatoes consumed	0.1153	0.1598
	Among the consuming ($n = 212$; 69.74%)	0.1653	0.1684
Continuous explanatory variable			
Income	Household income as the proportion of poverty threshold	2.3730	0.8489
Binary explanatory variables (yes = 1; no = 0)			
Age 20-30	Age is 20-30	0.1480	
Age 31-40	Age is 31-40	0.1776	
Age 41-50	Age is 41-50	0.1974	
Age 51-60	Age is 51-60	0.1743	
Age > 60	Age > 60 (reference)	0.3026	
Northeast	Resides in the Northeastern states	0.1579	
Midwest	Resides in the Midwestern states	0.2336	
West	Resides in the Western states	0.2204	
South	Resides in the Southern states (reference)	0.3882	

Source: Compiled from the CSFII, USDA, 1994-1996.

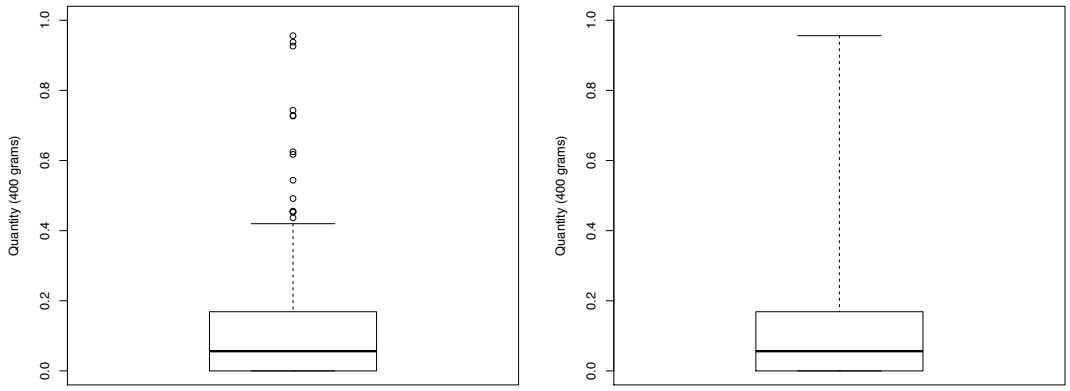

Figure 3. Usual boxplot (left panel) and adjusted boxplot (right panel) for the tomato consumption data.

Table 11. Descriptive statistics for GCS residuals of the tobit-N model.

n	Mean	Standard deviation	Skewness	Kurtosis
304	1.0970	1.3159	4.4515	26.8052
212	1.3603	1.4983	3.8351	20.0321

Table 12. Parameter estimates (standard errors in parenthesis) for tobit-N, tobit-SN, tobit-PN and tobit-TN models.

Parameter	Fitted model			
	tobit-N	tobit-SN	tobit-PN	tobit-TN
β_0 (Intercept)	-0.0025	-0.1541	-0.9440	0.5662
	(0.0446)	(0.0378)	(0.1849)	(0.0852)
β_{11} (Age 20-30)	-0.0419	-0.0164	-0.0228	-0.0222
	(0.0397)	(0.0328)	(0.0323)	(0.0284)
β_{12} (Age 31-40)	-0.0744	-0.0439	-0.0503	-0.0503
	(0.0371)	(0.0317)	(0.0306)	(0.0274)
β_{13} (Age 41-50)	-0.0142	0.0053	-0.0032	-0.0081
	(0.0353)	(0.0277)	(0.0283)	(0.0254)
β_{14} (Age 51-60)	-0.0152	0.0094	0.0017	-0.0029
	(0.0369)	(0.0293)	(0.0296)	(0.0264)
β_{21} (Northeast)	0.0845	0.0511	0.0516	0.0344
	(0.0368)	(0.0281)	(0.0296)	(0.0268)
β_{22} (Midwest)	0.0499	0.0240	0.0261	0.0173
	(0.0326)	(0.0259)	(0.0262)	(0.0234)
β_{23} (West)	0.0253	0.0166	0.0191	0.0160
	(0.0328)	(0.0269)	(0.0267)	(0.0235)
β_3 (Income)	0.0292	0.0151	0.0171	0.0133
	(0.0147)	(0.0120)	(0.0118)	(0.0103)
σ	0.2024	0.2675	0.3930	0.2325
	(0.0104)	(0.0147)	(0.0354)	(0.0185)
γ	-	4.3851	99.9963	0.0114
		(2.1513)	(75.3681)	(0.0059)
Log-likelihood	-37.3939	-22.7356	-20.2673	-12.1351
AIC	94.7879	67.4711	62.5347	46.2702
AICc	95.6920	68.5433	63.6068	47.3424
CAIC	141.9582	119.3584	114.4220	98.1576
BIC	131.9582	108.3584	103.4220	87.1576
HQIC	109.6569	83.8270	78.8905	62.6261

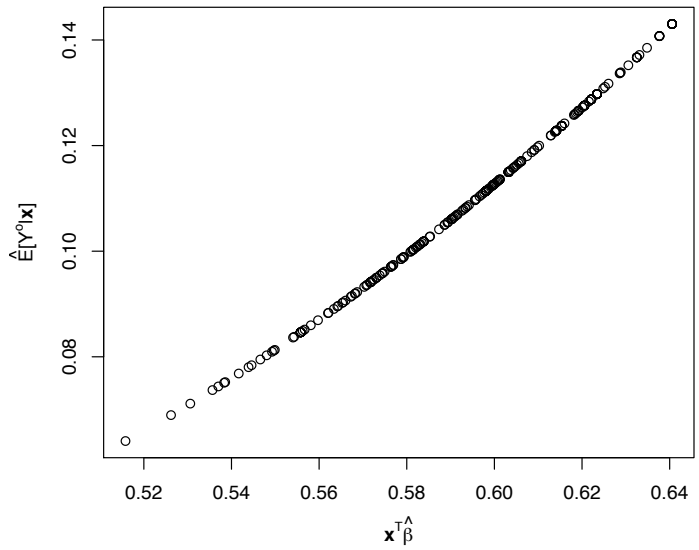

Figure 4. Scatter plot of $\hat{\mathrm{E}}\left[Y_i^{\mathrm{o}} \mid \boldsymbol{x}_i\right]$ versus $\boldsymbol{x}_i\hat{\boldsymbol{\beta}}$, $i = 1, 2, \ldots, 304$, for tobit-TN model.

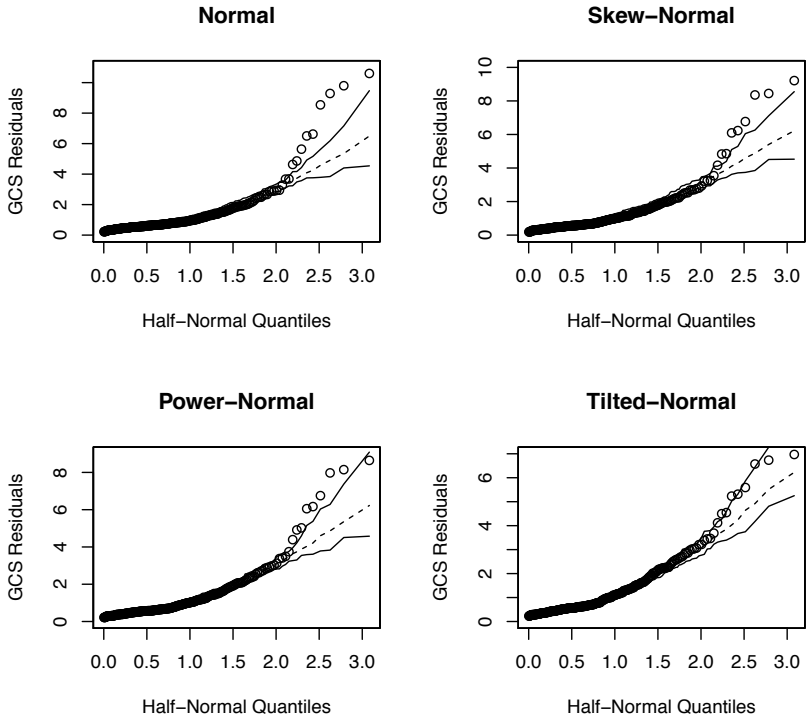

Figure 5. Half-normal plots with simulated envelopes for the GCS residuals.

corresponding observation has a high impact on the ML estimates of the parameters. We can use 1.0 as the cut-off value, as employed by some authors, like Imon (2005). From Figure 6, we note that, under the tobit-N model fitting, the observations 187 and 237 are influential on the ML estimates. However, with the tobit-SN, tobit-PN and tobit-TN models fitted, the scenario has changed: no observation is considered influential on the

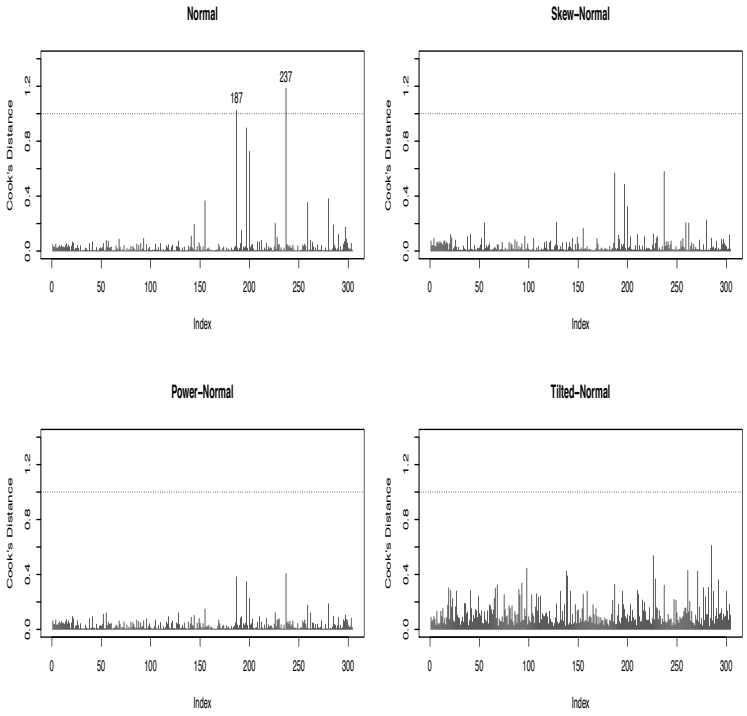

Figure 6. Generalized Cook distance. The influential observations are numbered.

parameter estimates, showing that these models are more robust.

6. Conclusions and Further Research

This paper discussed an asymmetric alternative for the standard tobit model (Tobin, 1958). It was based on the tilted-normal distribution (Maiti and Dey, 2012). The standard tobit model is a special case of the proposed model, which can also be seen as an alternative for the tobit-SN model (Hutton and Stanghellini, 2011) and tobit-PN model (Martínez-Flórez et al., 2013). Parameter estimates were obtained by using the ML method, which was also used for deriving large sample properties for the estimators. All the simulations and statistical analyses were performed using the programming language R version 3.3.1 (R Core Team, 2016). The computational code is available from the authors upon request. Simulation studies indicated good parameter recovery with the estimation approach developed, and appropriateness of the chosen model selection criteria. Since the standard tobit model is a special case of the tobit-TN model, the likelihood ratio statistic can be used for testing the standard tobit model null hypothesis. Application to an American food consumption data set (tomato consumption) indicated that the tobit-TN model can be an useful alternative to the standard tobit model, as well as to some of its asymmetric versions (tobit-SN and tobit-PN models). However, although the tobit-TN model was valid, that is, it has shown an adequate fitting to the data set at hand, we could also have considered a mixture of normal or tilted-normal distributions, for instance, as well as skewed heavy-tailed distributions for the error term, since the tomato consumption data seemed to have a long right tail. More study in this direction is desired.

Future work may also include to consider the use of other flexible distributions with better inferential properties and higher flexibility (e.g. the Families 1 to 4 considered in Jones, 2015) in the tobit framework. Other possible extension of the tobit model considers that the error term follows the centered skew-normal Birnbaum-Saunders distribution proposed by Chaves et al. (2019). Despite of being straightforward, our proposed ML estimation approach performs well, as demonstrated in the simulation results shown in Section 4. However, an interesting alternative to the direct maximization of the log-likelihood function, a procedure that sometimes can be quite cumbersome, is to use the Expectation-Maximization (EM) algorithm (Dempster et al., 1977) or some other extensions like the Monte Carlo EM (MCEM) (Wei and Tanner, 1990), Expectation Conditional Maximization (ECM) (Meng and Rubin, 1993), ECM Either (ECME) (Liu and Rubin, 1994) or the Stochastic Approximation of EM (SAEM) algorithm (Delyon et al., 1999). As stated in Mattos et al. (2018), the EM algorithm is a very popular iterative optimization strategy in models with non-observed or incomplete data, and has many attractive features such as numerical stability, simplicity of implementation and quite reasonable memory requirements. Thus, the EM algorithm provides an interesting setting for the ML estimation of tobit models, including for instance the estimation or prediction of the censored observations. Arellano-Valle et al. (2012), Garay et al. (2016, 2017) and Mattos et al. (2018) developed efficient EM-type algorithms for the ML estimation of their proposed extensions of the standard tobit model (Tobin, 1958). The derivation of an EM-type approach for our proposed tobit-TN model, e.g., by using some general mathematical properties of the Marshall-Olkin family of distributions shown in Cordeiro et al. (2014), will be the subject to our future work. We also intend to develop a Bayesian framework for the tobit-TN model, as similarly as in the works of Garay et al. (2015) and Massuia et al. (2017).

Appendix: Score functions

In this appendix, we show the score functions of the log-likelihood function (6). These quantities are obtained as follows:

$$U(\boldsymbol{\beta}) = \frac{\partial \ell\,(\boldsymbol{\theta})}{\partial \boldsymbol{\beta}} = \frac{1}{\sigma}\sum_{i=1}^{n} d_i \left[z_i + 2(1-\gamma)k_i\Phi(z_i)\right]\boldsymbol{x}_i^{\top} - \frac{1}{\sigma}\sum_{i=1}^{n}(1-d_i)\left[w_{0i} - (1-\gamma)k_{0i}\Phi(z_{0i})\right]\boldsymbol{x}_i^{\top},$$

$$U(\sigma) = \frac{\partial \ell\,(\boldsymbol{\theta})}{\partial \sigma} = -\frac{1}{\sigma}\sum_{i=1}^{n} d_i \left[1 - z_i^2 - 2(1-\gamma)k_i\Phi(z_i)z_i\right] - \frac{1}{\sigma}\sum_{i=1}^{n}(1-d_i)z_{0i}\left[w_{0i} - (1-\gamma)k_{0i}\Phi(z_{0i})\right]$$

and

$$U(\gamma) = \frac{\partial \ell\,(\boldsymbol{\theta})}{\partial \gamma} = \frac{1}{\gamma}\sum_{i=1}^{n} d_i - 2\sum_{i=1}^{n} d_i k_i\left[1-\Phi(z_i)\right] - \sum_{i=1}^{n}(1-d_i)k_{0i}\left[1-\Phi(z_{0i})\right],$$

where $z_{0i} = -\boldsymbol{x}_i^{\top}\boldsymbol{\beta}/\sigma$, $z_i = \left(y_i - \boldsymbol{x}_i^{\top}\boldsymbol{\beta}\right)/\sigma$, $k_{0i} = \left[1-(1-\gamma)\left\{1-\Phi(z_{0i})\right\}\right]^{-1}$, $k_i = \left[1-(1-\gamma)\left\{1-\Phi(z_i)\right\}\right]^{-1}$ and $w_{0i} = \phi(z_{0i})/\Phi(z_{0i})$.

Acknowledgements

The research is supported by the Brazilian organization FAPESP.

References

Akaike, H., 1977. On entropy maximization principle. In Krishnaiah, P.R. (Ed.), Applications of Statistics. North-Holland, Amsterdam, pp. 27-41.

Amemiya, T., 1973. Regression analysis when the dependent variable is truncated normal. Econometrica, 41, 997-1016.

Amemiya, T., 1984. Tobit models: A survey. Journal of Econometrics, 24, 3-61.

Amemiya, T., 1985. Advanced Econometrics. Harvard University Press, Cambridge.

Anderson, D.R., Burnham, K.P., and White, G.C., 1998. Comparison of Akaike information criterion and consistent Akaike information criterion for model selection and statistical inference from capture-recapture studies. Journal of Applied Statistics, 25, 263-282.

Andrews, D.F., and Mallows, C.L., 1974. Scale mixtures of normal distributions. Journal of the Royal Statistical Society B, 36, 99-102.

Arellano-Valle, R.B., Castro, L.M., González-Farías, G., and Muñoz-Gajardo, K.A., 2012. Student-t censored regression model: Properties and inference. Statistical Methods and Applications, 21, 453-473.

Atkinson, A.C., 1981. Two graphical displays for outlying and influential observations in regression. Biometrika, 68, 13-20.

Azzalini, A., 1985. A class of distributions which includes the normal ones. Scandinavian Journal of Statistics, 12, 171-178.

Barros, M., Galea, M., Leiva, V., and Santos-Neto, M., 2018. Generalized tobit models: Diagnostics and application in econometrics. Journal of Applied Statistics, 45, 145-167.

Bozdogan, H., 1987. Model selection and Akaike's information criterion (AIC): The general theory and its analytical extensions. Psychometrika, 52, 345-370.

Castro, L.M., Lachos, V.H., Ferreira, G.P., and Arellano-Valle, R.B., 2014. Partially linear censored regression models using heavy-tailed distributions: A Bayesian approach. Statistical Methodology, 18, 14-31.

Chaves, N.L., Azevedo, C.L.N., Vilca-Labra, F., and Nobre, J.S., 2019. A new Birnbaum-Saunders type distribution based on the skew-normal model under a centered parameterization. Chilean Journal of Statistics, 10, 55-76.

Cook, R.D., 1977. Detection of influential observation in linear regression. Technometrics, 19, 15-18.

Cook, R.D., 1986. Assessment of local influence. Journal of the Royal Statistical Society B, 48, 133-169.

Cordeiro, G.M., Lemonte, A.J., and Ortega, E.M.M., 2014. The Marshall-Olkin family of distributions: Mathematical properties and new models. Journal of Statistical Theory and Practice, 8, 343-366.

Cragg, J.G., 1971. Some statistical models for limited dependent variables with application to the demand for durable goods. Econometrica, 39, 829-844.

Delyon, B., Lavielle, M., and Moulines, E., 1999. Convergence of a stochastic approximation version of the EM algorithm. The Annals of Statistics, 27, 94-128.

Dempster, A., Laird, N., and Rubin, D., 1977. Maximum likelihood from incomplete data via the EM algorithm. Journal of the Royal Statistical Society B, 39, 1-38.

Desousa, M.F., Saulo, H., Leiva, V., and Scalco, P., 2018. On a tobit-Birnbaum-Saunders model with an application to medical data. Journal of Applied Statistics, 45, 932-955.

Efron, B., 1982. The Jackknife, the Bootstrap, and Other Resampling Plans. CBNS-NSF Regional Conference Series in Applied Mathematics, Monograph 38. Philadelphia, PA: SIAM.

Efron, B., and Tibshirani, R.J., 1993. An Introduction to the Bootstrap. Chapman and Hall, New York.

Fair, R.C., 1977. A note on the computation of the tobit estimator. Econometrica, 45, 1723-1727.

Garay, A.M., Bolfarine, H., Lachos, V.H., and Cabral, C.R.B., 2015. Bayesian analysis of censored linear regression models with scale mixtures of normal distributions. Journal of Applied Statistics, 42, 2694-2714.

Garay, A.M., Lachos, V.H., Bolfarine, H., and Cabral, C.R.B., 2017. Linear censored regression models with scale mixtures of normal distributions. Statistical Papers, 58, 247-278.

Garay, A.M., Lachos, V.H., and Lin, T.-I. 2016. Nonlinear censored regression models with heavy-tailed distributions. Statistics and its Interface, 9, 281-293.

García, V.J., Goméz-Déniz, E., and Vázquez-Polo, F.J., 2010. A new skew generalization of the normal distribution: Properties and applications. Computational Statistics and Data Analysis, 54, 2021-2034.

Goldberger, A.S., 1964. Econometric Theory. Wiley, New York.

Greene, W.H., (2012). Econometric Analysis. Pearson, Boston.

Gupta, D. and Gupta, R.C., 2008. Analyzing skewed data by power normal model. TEST, 17, 197-210.

Hannan, E.J., and Quinn, B.G., 1979. The determination of the order of an autoregression. Journal of the Royal Statistical Society B, 41, 190-195.

Hubert, M., and Vandervieren, E., 2008. An adjusted boxplot for skewed distributions. Computational Statistics and Data Analysis, 52, 5186-5201.

Hurvich, C.M., and Tsai, C., 1989. Regression and time series model selection in small samples. Biometrika, 76, 297-307.

Hutton, J.L. and Stanghellini, E., 2011. Modelling bounded health scores with censored skew-normal distributions. Statistics in Medicine, 30, 368-376.

Imon, A.H.M.R., 2005. Identifying multiple influential observations in linear regression. Journal of Applied Statistics, 32, 929-946.

Jones, M.C., 2015. On families of distributions with shape parameters. International Statistical Review, 83, 175-192.

Liu, C., and Rubin, D.B., 1994. The ECME algorithm: A simple extension of EM and ECM with faster monotone convergence. Biometrika, 81, 633-648.

Louzada, F., Shimizu, T.K.O., Suzuki, A.K., Mazucheli, J., and Ferreira, P.H., 2018. Compositional regression modeling under tilted normal errors: An application to a Brazilian super league volleyball data set. Chilean Journal of Statistics, 9, 33-53.

Maiti, S.S., and Dey, M., (2012). Tilted normal distribution and its survival properties. Journal of Data Science, 10, 225-240.

Marshall, A.W. and Olkin, I., 1997. A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. Biometrika, 84, 641-652.

Martínez-Flórez, G., Bolfarine, H., and Gómez, H.W., 2013. The alpha-power tobit model. Communications in Statistics: Theory and Methods, 42, 633-643.

Massuia, M.B., Garay, A.M., Cabral, C.R.B., and Lachos, V.H., 2017. Bayesian analysis of censored linear regression models with scale mixtures of skew-normal distributions. Statistics and its Interface, 10, 425-439.

Mattos, T.B., Garay, A.M., and Lachos, V.H., 2018. Likelihood-based inference for censored linear regression models with scale mixtures of skew-normal distributions. Journal of Applied Statistics, 45, 2039-2066.

Meng, X.L., and Rubin, B.D., 1993. Maximum likelihood estimation via the ECM algorithm: A general framework. Biometrika, 80, 267-278.

Monti, A.C., 2003. A note on the estimation of the skew normal and the skew exponential power distributions. Metron, 61, 205-219.

Pewsey, A., Gómez, H.W., and Bolfarine, H., 2012. Likelihood-based inference for power distributions. TEST, 21, 775-789.

R Core Team, 2016. R: A Language and Environment for Statistical Computing. (Version 3.3.1) R Foundation for Statistical Computing, Vienna, Austria.

Rubio, F.J., and Steel, M.F.J., 2012. On the Marshall-Olkin transformation as a skewing mechanism. Computational Statistics and Data Analysis, 56, 2251-2257.

Schwarz, G.E., 1978. Estimating the dimension of a model. Annals of Statistics 6, 461-464.

Sugiura, N., 1978. Further analysis of the data by Akaike's information criterion and the finite corrections. Communications in Statistics: Theory and Methods, 7, 13-26.

Tobin, J., 1958. Estimation of relationships for limited dependent variables. Econometrica, 26, 24-36.

USDA, 2000. Continuing survey of food intakes by individuals 1994-1996. CD-ROM. Agricultural Research Service, Washington, DC.

Wei, G.C., and Tanner, M.A., 1990. A Monte Carlo implementation of the EM algorithm and the poor man's data augmentation algorithms. Journal of the American Statistical Association, 85, 699-704.

Information for authors

The editorial board of the Chilean Journal of Statistics (ChJS) is seeking papers, which will be refereed. We encourage the authors to submit a PDF electronic version of the manuscript in a free format to Víctor Leiva, Editor-in-Chief of the ChJS (E-mail: `chilean.journal.of.statistics@gmail.com`). Submitted manuscripts must be written in English and contain the name and affiliation of each author followed by a leading abstract and keywords. The authors must include a"cover letter" presenting their manuscript and mentioning: "We confirm that this manuscript has been read and approved by all named authors. In addition, we declare that the manuscript is original and it is not being published or submitted for publication elsewhere".

Preparation of accepted manuscripts

Manuscripts accepted in the ChJS must be prepared in Latex using the ChJS format. The Latex template and ChJS class files for preparation of accepted manuscripts are available at `http://chjs.mat.utfsm.cl/files/ChJS.zip`. Such as its submitted version, manuscripts accepted in the ChJS must be written in English and contain the name and affiliation of each author, followed by a leading abstract and keywords, but now mathematics subject classification (primary and secondary) are required. AMS classification is available at `http://www.ams.org/mathscinet/msc/`. Sections must be numbered 1, 2, etc., where Section 1 is the introduction part. References must be collected at the end of the manuscript in alphabetical order as in the following examples:

Arellano-Valle, R., 1994. Elliptical Distributions: Properties, Inference and Applications in Regression Models. Unpublished Ph.D. Thesis. Department of Statistics, University of São Paulo, Brazil.

Cook, R.D., 1997. Local influence. In Kotz, S., Read, C.B., and Banks, D.L. (Eds.), Encyclopedia of Statistical Sciences, Vol. 1., Wiley, New York, pp. 380-385.

Rukhin, A.L., 2009. Identities for negative moments of quadratic forms in normal variables. Statistics and Probability Letters, 79, 1004-1007.

Stein, M.L., 1999. Statistical Interpolation of Spatial Data: Some Theory for Kriging. Springer, New York.

Tsay, R.S., Peña, D., and Pankratz, A.E., 2000. Outliers in multivariate time series. Biometrika, 87, 789-804.

References in the text must be given by the author's name and year of publication, e.g., Gelfand and Smith (1990). In the case of more than two authors, the citation must be written as Tsay et al. (2000).

Copyright