UNIVERSIDADE FEDERAL DE GOIÁS
FACULDADE DE ADMINISTRAÇÃO, CIÊNCIAS CONTÁBEIS E CIÊNCIAS ECONÔMICAS - FACE/UFG
PROGRAMA DE PÓS-GRADUAÇÃO EM ECONOMIA

MÁRIO FERNANDO DE SOUSA

Two essays on Birnbaum-Saunders regression models for censored data

Durante o desenvolvimento deste trabalho o autor recebeu auxílio financeiro da CAPES

Goiânia

2016

sistema de bibliotecas ufg

UFG

TERMO DE CIÊNCIA E DE AUTORIZAÇÃO PARA DISPONIBILIZAR AS TESES E DISSERTAÇÕES ELETRÔNICAS NA BIBLIOTECA DIGITAL DA UFG

Na qualidade de titular dos direitos de autor, autorizo a Universidade Federal de Goiás (UFG) a disponibilizar, gratuitamente, por meio da Biblioteca Digital de Teses e Dissertações (BDTD/UFG), regulamentada pela Resolução CEPEC nº 832/2007, sem ressarcimento dos direitos autorais, de acordo com a Lei nº 9610/98, o documento conforme permissões assinaladas abaixo, para fins de leitura, impressão e/ou *download*, a título de divulgação da produção científica brasileira, a partir desta data.

1. Identificação do material bibliográfico: **[X] Dissertação** **[ ] Tese**

2. Identificação da Tese ou Dissertação

Nome completo do autor: Mário Fernando de Sousa

Título do trabalho: Two essays on Birnbaum-Saunders regression models for censored data

3. Informações de acesso ao documento:

Concorda com a liberação total do documento [X] SIM [ ] NÃO[1]

Havendo concordância com a disponibilização eletrônica, torna-se imprescindível o envio do(s) arquivo(s) em formato digital PDF da tese ou dissertação.

Mário Fernando de Sousa

Assinatura do (a) autor (a) [2]

Data: 26 / 04 / 2017

[1] Neste caso o documento será embargado por até um ano a partir da data de defesa. A extensão deste prazo suscita justificativa junto à coordenação do curso. Os dados do documento não serão disponibilizados durante o período de embargo.

[2] A assinatura deve ser escaneada.

MÁRIO FERNANDO DE SOUSA

Two essays on Birnbaum-Saunders regression models for censored data

Dissertação apresentada à comissão avaliadora do programa de Pós-Graduação em economia da Faculdade de Administração, Ciências Contábeis e Ciências Econômicas da Universidade Federal de Goiás, como requisito parcial para obtenção do título de mestre em economia.

Orientador: Prof. Dr. Helton Saulo Bezerra dos Santos.

Co-orientador: Prof. Dr. Victor Eliseo Leiva Sanchez.

Goiânia

2016

Ficha de identificação da obra elaborada pelo autor, através do
Programa de Geração Automática do Sistema de Bibliotecas da UFG.

Desousa, Mário Fernando

Two essays on Birnbaum-Saunders regression models for censored data [manuscrito]/ Mário Fernando Desousa. - 2016.

LIII, 53 f.

Orientador: Prof. Dr. Helton Saulo ; co-orientador Dr. Victor Leiva. Dissertação (Mestrado) - Universidade Federal de Goiás, Faculdade de Administração, Ciências Contábeis e Ciências Econômicas (FACE), Programa de Pós-Graduação em Economia (PPE), Goiânia, 2016.

Inclui lista de figuras e lista de tabelas.

1. Censored data. 2. Birnbaum-Saunders distribution. 3. Tobit models. 4. Maximum likelihood estimation. 5. Residuals. I. Saulo, Helton, orient. II. Título.

CDU 33

MINISTÉRIO DA EDUCAÇÃO
Universidade Federal de Goiás - UFG
Faculdade de Administração, Ciências Contábeis e Ciências Econômicas
Programa de Pós-Graduação em Econòmia

UFG

ATA DE DEFESA PÚBLICA DE DISSERTAÇÃO

Aos seis dias do mês de dezembro de 2016, no horário de 18.00 horas às 19.00 horas, foi realizada, em sessão pública, na sala 2110 da FACE, a defesa da dissertação *Two essays on Bimbaum-Saunders regression models for censored data.* de autoria do discente Mario Fernando de Sousa do Programa de Pós-Graduação em Economia- PPE da UFG. A Comissão Examinadora, constituída pelo Professor Helton Saulo Bezerra dos Santos, da Universidade Federal de Goiás (IMEUFG) (Membro Interno/Orientador), Professor Sandro Eduardo Monsueto, da Universidade Federal de Goiás (UFG)(Membro Interno) e pelo Professor Manoel Ferreira Santos-Neto, da Universidade Federal de Campina Grande (UFCG) (Membro Externo), emitiu o seguinte parecer:

Resultado Final:

☒ Aprovado

☐ Reprovado

Recomendações:

Escrita da banca

Eu, Helton Saulo Bezerra dos Santos, orientador do discente, lavrei a presente Ata, que segue assinada por mim e pelos demais membros da Comissão Examinadora.

Prof. Dr. Helton Saulo Bezerra dos Santos (orientador/Presidente/PPE/UFG)

Prof. Dr. Sandro Eduardo Monsueto (Examinador Interno/UFG/PPE)

Prof. Dr. Manoel Ferreira Santos-Neto (Membro Externo/UFCG)

Gilson Carlos de Assis Júnior
Assistente em Administração
FACE/UFG - SIAPE 1654190
SECRETÁRIO
PPE

Dedico este trabalho a Divina Maria Barbosa de Sousa, minha mãe, a Helena de Urzeda Ferro (*in memoriam*), minha avó e a Marizilda Nunes de Sousa, minha madrinha.

AGRADECIMENTOS

Ao fim de dois anos de mestrado faço questão de prestar meus agradecimentos a todas as pessoas que, em maior ou menor medida, contribuíram para o desenvolvimento deste trabalho:

À CAPES pelo apoio financeiro.

Ao professor Helton Saulo Bezerra dos Santos, pela orientação, dedicação e paciência durante todo o processo.

Ao professor Victor Eliseo Leiva Sanchez, pela dedicação e pelas valiosas contribuições enquanto co-orientador.

Aos professores Sandro Eduardo Monsueto e Manoel Ferreira dos Santos Neto pelas generosas e valiosas contribuições.

A todos os professores da Universidade Federal de Goiás (da graduação e do mestrado) que, de alguma forma, contribuíram para a minha formação por meio do ensino de suas respectivas disciplinas e da própria convivência. Aprendi com todos os discentes que o caráter, no final das contas, é o que realmente importa e esse é um aprendizado que ficará para sempre comigo. Não citarei nomes mas todos possuem os meus sinceros agradecimentos.

A todos os professores da Escola Corrossel (educação infantil), do Colégio Batista Goiano (ensino fundamental e dois primeiros anos do ensino médio) e do Colégio Atual (3º ano do ensino médio), pelo excelente tempo de aprendizado que passamos juntos. De modo especial agradeço à professora de matemática Deyse Gonçalves dos Reis que, ainda nas primeiras séries do ensino fundamental, notou uma aptidão ímpar para disciplina.

À Universidade Federal de Goiás pelo ambiente propício para o aprendizado e pela oportunidade de fazer o curso com excelentes professores.

A todos os técnicos administrativos, que trabalham nos bastidores de modo a operacionalizar as burocracias inerentes às atividades do setor público. Agradeço de modo especial ao Gilson Carlos de Assis Junior pela dedicação aos alunos do mestrado em economia.

Aos funcionários da limpeza e da segurança que trabalham arduamente para deixar o prédio sempre em boas condições.

A Divina Maria Barbosa de Sousa, minha mãe, Fernando Barbosa Ferro, meu avô, Helena de Urzeda Ferro (*in memoriam*), minha avó, e Marizilda Nunes de Sousa, minha madrinha, cujos esforços foram essenciais para que eu progredisse na vida acadêmica.

A todos os amigos, pela companhia, pelos estímulos positivos nos momentos necessários e pelos momentos de descontração. Agradeço, de forma especial, ao amigo Artur Barbosa Coelho cuja companhia diária me agrega muito, não só em termos de conhecimento, mas também na simplicidade e objetividade com que lida com as adversidades do cotidiano, mostrando que determinação e empenho são indispensáveis para atingirmos nossos objetivos.

"Only a Sith deals in absolutes."

(Star Wars episode III - Revenge of the Sith)

RESUMO

Este trabalho visa preencher uma lacuna existente na literatura pertinente à modelagem de dados assimétricos e censurados. O objetivo principal é oferecer uma contribuição via o desenvolvimento de dois modelos, os quais serão apresentados em dois artigos. No primeiro artigo é proposto o modelo tobit-Birnbaum-Saunders, ou seja, uma variação do modelo tobit clássico, com estimação baseada no método de máxima verossimilhança, resíduos, técnicas de diagnóstico e uma aplicação a dados reais. No segundo artigo é abordada a utilização de um modelo de mistura entre as distribuições Birnbaum-Saunders e Bernoulli, de modo a generalizar o modelo tobit-Birnbaum-Saunders e considerar a possibilidade de observações parciais abaixo do ponto de corte. Para o modelo de mistura são realizadas simulações de Monte Carlo e uma aplicação a dados reais. Os resultados mostram que, em ambos os casos, a distribuição Birnbaum-Saunders oferece os melhores resultados.

Palavras-chave: Dados censurados, Distribuição Birnbaum-Saunders, Modelos tobit, Máxima verossimilhança, Resíduos, Diagnósticos, Modelos de mistura.

ABSTRACT

This work aims to fill a gap in the literature on modeling asymmetric and censored data. The main objective is to provide a contribution by developing two models, which will be presented in two papers, respectively. In the first paper, we develop the tobit-Birnbaum-Saunders model, a variation of the standard tobit model. We discuss estimation based on the maximum likelihood method, residuals, diagnostic techniques and an empirical application. In the second paper, we propose the use of a mixture between the Birnbaum-Saunders and Bernoulli distributions. The objective is to generalize the tobit-Birnbaum-Saunders model in order to consider the possibility of partial observations below a cutoff point. For the mixture model, we carry out a Monte Carlo simulation study and an empirical application. The results show that, in both cases, the Birnbaum-Saunders distribution provides the best results.

Keywords: Censored data, Birnbaum-Saunders distribution, Tobit models, Maximum likelihood estimation, Residuals, Diagnostics, mixture models.

Contents

List of Figures **xi**

List of Tables **xii**

1 Introduction **1**

2 The tobit-BS model **3**

2.1 Introduction . . . 3
2.2 Background . . . 6
2.2.1 The tobit model . . . 6
2.2.2 The Birnbaum-Saunders distribution . . . 7
2.2.3 The log-Birnbaum-Saunders distribution . . . 8
2.3 The tobit-BS model . . . 10
2.3.1 Formulation . . . 10
2.3.2 Estimation . . . 10
2.3.3 Inference . . . 11
2.3.4 Model checking . . . 12
2.4 Diagnostic analysis . . . 13
2.4.1 Local Influence . . . 13
2.4.2 Global influence . . . 16
2.5 Application . . . 17
2.5.1 The data . . . 17
2.5.2 Software . . . 17
2.5.3 Monte Carlo simulations . . . 18
2.5.4 Data analysis . . . 18
2.5.5 Analysis under the normal tobit model . . . 20
2.5.6 Analysis under the tobit-BS model . . . 22
2.6 Concluding remarks . . . 23

3 A new mixture-based regression model for censored data **25**

3.1 Introduction . . . 25
3.2 Background . . . 27

3.2.1 The standard tobit model 27
3.2.2 Mixture models for censored data 27
3.2.3 The BS and log-BS distributions 28
3.3 The Bernoulli/BS mixture model 30
3.3.1 Formulation 30
3.3.2 Inference 32
3.4 Simulation study 33
3.5 Application 34
3.5.1 Exploratory data analysis 34
3.5.2 Results 35
3.6 Concluding remarks 36

4 Concluding remarks and future works **37**

Bibliography **39**

List of Figures

2.1 PDF of (left) BS$(\alpha, 1)$ and (right) log-BS$(\alpha, 0)$ distributions for the indicated value of α. 9

2.2 Histogram (left), TTT plot (center), and boxplots (right) for the measles vaccine data. . . . 20

2.3 Index plots of $C_i(\boldsymbol{\theta})$ (left), $C_i(\boldsymbol{\beta})$ (center) and $C_i(\sigma)$ (right) under the case-weight perturbation in the standard tobit model with the measles vaccine data. 21

2.4 QQ plot and its envelope for the GCS residual (left) and index plot of the GCD (right) for the standard tobit model with measles vaccine data. 21

2.5 QQ plot and its envelope for the GCS residual (left) and index plot of the GCD (right) for the tobit-BS model with measles vaccine data. 22

2.6 Index plots of $C_i(\boldsymbol{\theta})$ (left), $C_i(\boldsymbol{\beta})$ (center) and $C_i(\alpha)$ (right) under the case-weight perturbation in the tobit-BS model with the measles vaccine data. 23

3.1 PDF of (left) BS$(\alpha, 1)$ and (right) log-BS$(\alpha, 0)$ distributions for the indicated value of α. 30

3.2 Histogram (a), TTT plot (b), and boxplots (c) for the measles vaccine data. 35

List of Tables

2.1 Empirical bias and MSE (in parentheses) of the ML estimators for the tobit-BS model parameters by using the indicated sample sizes and parameter values with simulated data. 19
2.2 Descriptive summary for the measles vaccine data. 20
2.3 RCs (in %) in ML estimates for the indicated parameter (standard tobit model) and removed cases with the measles vaccine data. 21
2.4 ML estimates (with SE in parentheses) and model selection measures for the indicated models. 22
2.5 RCs (in %) in ML estimates for the indicated parameter (tobit-BS model) and removed cases with the measles vaccine data. 23

3.1 Monte-Carlo estimation of bias and MSE for the indicated sample sizes and parameter values 34
3.2 Descriptive summary for the measles vaccine data. 35
3.3 ML estimates (with SE in parentheses) and AIC values for the indicated models with the measles vaccine data 36

Chapter 1

Introduction

The linear regression approach with ordinary least squares (OLS) can not be applied to censored data because censoring, by definition, destroys the assumption of linearity required in the OLS framework, thus leading to biased estimators. Tobin (1958) proposed a different methodology, named the tobit model, that could fit the data appropriately by assuming a regression model whose response variable was censored to a prefixed limiting value. The censoring occurs when the response of the regression model is not directly observable, but its independent variables (or covariates) are.

Tobit models rely on the normality assumption. Proposals of tobit models that relax this assumption are extremely important, since it is common knowledge that most of the data available in the real world are often well modeled by non-normal distributions. A number of authors have noticed that the asymmetry of data of censored responses and their kurtosis usually are different from the expected for a normal distribution, so that more flexible tobit models are needed; see, for example, Martínez-Flores, Bolfarine and Gómez (2013a) and Rocha, Arellano and Loschi (2015).

The Birnbaum-Saunders (BS) distribution is a prominent model which has been widely studied; see Leiva (2016). This distribution was proposed by Birnbaum and Saunders (1969) and has two parameters that change its shape and scale. The BS distribution has been mainly considered to model failure times in engineering, but it has also been applied to model biological, business, economic, engineering, industry and environmental data. Some of its recent applications can be found in Qu and Xie (2011), Li, Chen and Xie (2012), Saulo et al. (2013), Leiva et al. (2017), Garcia-Papani et al. (2016), Wanke and Leiva (2015), Santos-Neto et al. (2016) and Leao et al. (2017).

The main objective of this dissertation is propose two asymmetric regression models for censored data by relaxing the assumption of normality and supposing that the errors of both models follow a BS distribution. The first model is called the tobit-Birnbaum-Saunders (tobit-BS) model, whereas the second is called the Bernoulli/Birnbaum-Saunders (Bernoulli/BS) model. For the tobit-BS model, in Chapter 2, we present: (i) inference based on the maximum likelihood (ML)

method; (ii) residuals and global and local influence tools for model checking and diagnostics; and (iii) an application of the proposed model by using the safety and immunogenecity case-study of measles vaccine in Haiti described by Moulton and Halsey (1995). The second model, which can be seen as a generalization of the first one, allows the possibility of a observation i, located below a threshold, to be either a partial observation from the assumed distribution or a realization from the point-mass distribution. For the Bernoulli/BS model, discussed in Chapter 3, we present: (i) inference based on the ML approach; (ii) a Monte Carlo (MC) simulation study; and (iii) an application of the proposed model to the same data considered in Chapter 2. In Chapter 4, we provide some conclusions and future works.

Chapter 2

On a new tobit-Birnbaum-Saunders model with an application to antibody response to vaccine

The tobit model allows a censored response variable to be described by covariates in different areas. For example, its applications cover economics, engineering, environment and medicine. A strong assumption of the standard tobit model is that its errors follow a normal distribution. However, not all applications are well modeled by this distribution. Some efforts have relaxed the normality assumption by considering more flexible distributions. Nevertheless, presence of asymmetry could not be described by these flexible distributions. We explore a real-world data example of measles vaccine in Haiti which confirms this asymmetry problem. Then, to solve such a problem, we propose a tobit model with errors following a Birnbaum-Saunders distribution, which has shown to be a good alternative to describe medical data. We develop inference based on the maximum likelihood method and derive a type of residual. We perform global and local influence diagnostics to assess the sensitivity of the maximum likelihood estimators to atypical cases. We conduct a real-world data example with the proposed methodology with the help of the R software. The results show the good performance of the tobit-Birnbaum-Saunders model.

Keywords Birnbaum-Saunders distribution; Diagnostics; Maximum Likelihood Estimation; R software; Residuals; Tobit models.

2.1 Introduction

A relevant topic of study is the determination of antibody concentration by quantitative assays. Such a topic is of relevance because there is always a concentration value (threshold) below which an exact measurement cannot be obtained regardless of the employed technique. However, this antibody concentration value is a function of the associated assay. When dealing with data from an assay where left-censoring is present, the lower detection limit (LDL) can be used to substitute this value with a censored case. In special, this substitution is applied in a safety and immunogenecity case-study of measles vaccine in Haiti presented by Moulton and Halsey (1995), which corresponds to the example explored in this paper. The data associated with this case-study

have left-censoring. Then, standard tobit models could be used to provide estimates of interest (BARROS et al., 2010). However, the standard tobit model has a strong assumption stating that the model errors follow a normal distribution, which is symmetric. In addition to left-censoring detected in the mentioned case-study, presence of asymmetry is also identified. Ignoring the effect of asymmetry can be harmful and conduct to estimates significantly different from the true values. We analyze the antibody concentration data in four stages. First, we consider the standard (normal or Gaussian) tobit model. Second, we show its inadequacy by verifying the assumptions on which this model relies and the evidence of asymmetry in the data. Third, we correct this inadequacy by deriving a new tobit model as described below. Fourth, we demonstrate empirically how the new tobit model describes these data very well.

The tobit model was proposed by Tobin (1958) for a limited (censored) dependent variable (or response) and named tobit by Goldberger (1964), because of its similarity with the probit model. According to Amemiya (1984), Tobin (1958) was motivated to develop his model by a case-study where he needed to analyze the relationship between household expenditure on a durable good and household incomes. The common regression approach with ordinary least squares could not be used because there were many cases where the expenditure was zero, which destroyed the assumption of linearity. To solve the problem, Tobin (1958) proposed a model that could fit the data appropriately formulating a regression model whose response was censored to a prefixed limiting value.

Tobit models have been used extensively to describe censored responses in several areas. The censoring occurs when the response of the regression model is not directly observable, but its independent variables (or covariates). At the beginning, tobit models were not widely applied due to the computational complexity needed in the parameter estimation. However, nowadays this is no longer a problem because of the computational development with different softwares having available these models, in particular, the R software (R-TEAM, 2016, <www.R-project.org>). Although tobit models were born in economics, they have been applied in biology and engineering as well. The biology research relies most in analyzing the survival time of a patient (LEIVA et al., 2007); the engineers use it to model time to failure of various types of materials or machines (VILLEGAS; PAULA; LEIVA, 2011); whereas the sociologists employ it to describe the duration of marriage, residing in a particular region or unemployment, as well the time between births, but the tobit models have also been used to describe inheritance, ratio of unemployed hours to employed hours and expected age of retirement (AMEMIYA, 1984). Tobit models can also be used in environmental sciences, where censored data also are present (THORARINSDOTTIR; GNEITING, 2010; HELSEL, 2011).

As mentioned, tobit models have a strong normality assumption. Proposals of tobit models that relax the assumption of normality are of importance, because most of the data available in the real-world are often well modeled by non-normal distributions. A number of authors have noticed that the asymmetry of data of censored responses and their kurtosis can be, and usually are, different from the expected for a normal distribution, so that more flexible tobit models are

needed. The interested reader is referred to Barros et al. (2010), Barros et al. (2017), Arellano et al. (2012), Martínez-Flores, Bolfarine and Gómez (2013a), Martínez-Flores, Bolfarine and Gómez (2013b), Garay et al. (2015), Massuia et al. (2015) and Rocha, Arellano and Loschi (2015) for some works related to non-normal tobit models.

The Birnbaum-Saunders (BS) distribution has been widely studied. It is positively skewed, has a failure rate with upside-down bathtub shape, and a close relation with the normal model. For more details about the BS distribution, the interested reader is referred to the seminal paper by Birnbaum and Saunders (1969) and the books by (JOHNSON; KOTZ; BALAKRISHNAN, 1995, pp. 651-663) and Leiva (2016). The BS distribution was derived in terms of shape and scale parameters, but the latter is also its median. Thus, the BS distribution can be seen as an analogue to the normal model, but in an asymmetrical setting, where the median can be a better measure of central tendency than the mean. The BS distribution has been applied to model biological, business, economic, engineering, industry and environmental data, which have been conducted by international, transdisciplinary groups of researchers. Some of its recent applications are attributed to Qu and Xie (2011), Ferreira, Gomes and Leiva (2012), Li, Chen and Xie (2012), Saulo et al. (2013), Leiva et al. (2014a), Leiva et al. (2014b), Leiva et al. (2014), Leiva et al. (2015), Leiva et al. (2015), Leiva et al. (2016), Leiva et al. (2016), Leiva et al. (2017), Garcia-Papani et al. (2016), Wanke and Leiva (2015), Marchant, Leiva and Cysneiros (2016), Marchant et al. (2016), Santos-Neto et al. (2016) and Leao et al. (2017). The BS distribution has shown to be a good alternative to describe medical data (BARROS; PAULA; LEIVA, 2008; LEIVA et al., 2007; LEAO et al., 2017).

The main objective of this paper is propose a tobit model by relaxing the assumption of normality and supposing that its errors follow a BS distribution, which we denote as the tobit-BS model. According to the best of our knowledge, this topic has not been studied before. The secondary objectives of this paper are: (i) to develop inference for the tobit-BS model based on the maximum likelihood (ML) method; (ii) to derive residuals and global and local influence tools for model checking and diagnostics; and (iii) to carry out an application of the proposed model by using the safety and immunogenecity case-study of measles vaccine in Haiti described by Moulton and Halsey (1995). The results indicate that the tobit-BS model presents an excellent performance compared with the standard tobit model.

The paper is constituted, along with this introduction, of five sections. Section 2.2 presents a background of standard tobit models and of the BS distribution and its logarithmic version (log-BS). Section 2.3 formulates the tobit-BS model and provides estimation, inference and residual analysis based on the ML method. Section 2.4 derives diagnostics for the proposed model including global/local influence tools. Section 2.5 carries out the mentioned empirical application with real-world data. Finally, Section 2.6 discusses some concluding remarks about this research.

2.2 Background

2.2.1 The tobit model

Let $\boldsymbol{Y} = (Y_1, \ldots, Y_m, Y_{m+1}, \ldots, Y_n)^\top$ be a sample of size n, that is, independent (IND) random variables but not necessarily independent identically distributed (IID). Assume that this sample includes m censored data to the left and $n - m$ observed (complete or uncensored) data. Thus, such censoring scheme can be visualized under a regression setting with a censored response Y^*, which is a (unobserved) latent variable. Hence, the m censored data (unobserved)correspond to the values of Y^* less than or equal to a threshold point y_0 (censoring to the left), so that all of these data take the value y_0. The other $n - m$ data (observed) are related to values of Y^* greater than y_0, which can be described by a linear regression structure of the type $\boldsymbol{x}_i^\top \boldsymbol{\beta}$. All this modeling environment may be formulated by the normal tobit model with censored response to the left as

$$Y_i = \begin{cases} y_o, & Y_i^* \leq y_0, \quad i = 1, \ldots, m; \\ \boldsymbol{x}_i^\top \boldsymbol{\beta} + \varepsilon_i, & Y_i^* > y_0, \quad i = m+1, \ldots, n; \end{cases} \tag{2.1}$$

where $\varepsilon_i \overset{\text{IID}}{\sim} \text{N}(0, \sigma^2)$ is the model error term, $\boldsymbol{\beta}$ is a vector of regression coefficients corresponding to unknown parameters to be estimated, and $\boldsymbol{x}_i$ is a vector containing the values of covariates. Observe that y_0 given in (2.1) is a prefixed limiting value that does the response of the regression model defined in (2.1) to be limited (or censored), as mentioned by Tobin (1958).

Note the similarity between the normal probit model and the normal tobit model defined in (2.1). In the normal probit model, the response is a latent (unobserved) variable described by

$$Y_i^* = \boldsymbol{x}_i^\top \boldsymbol{\beta} + \varepsilon_i, \quad i = 1, \ldots, n, \tag{2.2}$$

where $\boldsymbol{x}_i$, $\boldsymbol{\beta}$ and ε_i are defined analogously as in (2.1). As it is not possible to observe the latent variable Y_i^*, the indicator variable

$$Y_i = \begin{cases} 0, & Y_i^* \leq y_0, \quad i = 1, \ldots, m; \\ 1, & Y_i^* > y_0, \quad i = m+1, \ldots, n; \end{cases} \tag{2.3}$$

is defined, However, instead of using $\boldsymbol{x}^\top \boldsymbol{\beta}$, from (2.2) and (2.3), the expected response is formulated as

$$\text{E}[Y|\boldsymbol{x}] = \text{P}(Y = 1) = \text{P}(Y^* > y_0) = \text{P}(\boldsymbol{x}^\top \boldsymbol{\beta} + \varepsilon > y_0) = 1 - \Phi(y_0 - \boldsymbol{x}^\top \boldsymbol{\beta}), \tag{2.4}$$

where Φ is the cumulative distribution function (CDF) of the standard normal distribution. Note that other CDFs might be assumed for Φ in (2.4), so expanding the covering of the probit model.

Probit and tobit models are the same for the latent variable (Y^*), but models for the ob-

served response (Y) are different. In the tobit model, we know the value of Y^* when $Y^* > y_0$, whereas in the probit model we just know that $Y^* > y_0$, but we do not know its value. Thus, there is more information in the tobit model than in the probit model. Also, the estimates of $\boldsymbol{\beta}$ in the tobit model are more efficient than in the probit model. Moreover, for the censored cases of the probit model, it is not possible to estimate the variance of Y^*. However, in the tobit model, this variance can be estimated from the data. The interested reader is referred to Scott (1997, p. 199) for details about logit, probit and tobit models.

2.2.2 The Birnbaum-Saunders distribution

If a random variable T follows a Birnbaum-Saunders distribution with shape parameter α and scale parameter σ, we use the notation $T \sim \text{BS}(\alpha, \sigma)$. This distribution can be defined by its CDF given by

$$F_T(t;\alpha,\sigma) = \Phi\left(\frac{1}{\alpha}\left(\sqrt{t/\sigma} - \sqrt{\sigma/t}\right)\right), \quad t > 0, \alpha > 0, \sigma > 0. \tag{2.5}$$

Then, the probability density function (PDF) of T obtained from (2.5) is expressed as

$$f_T(t;\alpha,\sigma) = \frac{1}{2\alpha}\left(\sqrt{1/\sigma t} + \sqrt{\sigma/t^{\frac{3}{2}}}\right)\phi\left(\frac{1}{\alpha}\left(\sqrt{t/\sigma} - \sqrt{\sigma/t}\right)\right), \quad t > 0, \alpha > 0, \sigma > 0, \tag{2.6}$$

where ϕ is the standard normal PDF. Thus, the PDF in (2.6) can be rewritten as

$$f_T\left(t;\alpha,\sigma\right) = \frac{\exp(\alpha^{-2})}{2\alpha\sqrt{2\pi\sigma}}\exp\left(-\frac{1}{2\alpha^2}\left(\frac{t}{\sigma} + \frac{\sigma}{t}\right)\right)t^{-\frac{3}{2}}\left(t+\sigma\right), \quad t > 0, \alpha > 0, \sigma > 0. \tag{2.7}$$

Note that the results provided in (2.5) and (2.6) may be obtained from the transformation theorem of random variables by using

$$T = \sigma\left(\alpha Z/2 + \sqrt{(\alpha Z/2)^2 + 1}\right)^2, \tag{2.8}$$

where $Z \sim \text{N}(0,1)$. Also from (2.8), it may be verified that a continuous random variable T has a BS distribution with parameters $\alpha > 0$ and $\sigma > 0$, if and only if $Z = (1/\alpha)(\sqrt{T/\sigma} - \sqrt{\sigma/T}) \sim \text{N}(0,1)$. Some properties of the BS distribution are presented as follows. If $T \sim \text{BS}(\alpha,\sigma)$, then: (i) $\text{E}(T) = \sigma(1+\alpha^2/2)$ and $\text{Var}(T) = (\alpha\sigma)^2(1+5\alpha^2/4)$; (ii) for $b > 0$, $bT \sim \text{BS}(\alpha, b\sigma)$, which means that the BS distribution is closed under scalar multiplication (proportionality); (iii) $1/T \sim \text{BS}(\alpha, 1/\sigma)$, implying that the BS distribution is closed under reciprocation; (iv) the median of the distribution of T is σ, which can be directly obtained when $q = 0.5$ from its quantile function

given by

$$t(q;\alpha,\sigma)=F_T^{-1}(q;\alpha,\sigma)=\sigma\left(\frac{\alpha z(q)}{2}+\sqrt{\left(\frac{\alpha z(q)}{2}\right)^2+1}\right)^2,\quad 0<q<1,$$

where $z(q)$ is the standard normal quantile function; and (v) the BS distribution is positively skewed as α increases, being strongly skewed as an exponential (EXP) distribution when $\alpha>2$, and approximately symmetrical around σ as α goes to zero; see Figure 2.1(left). Properties of proportionality and reciprocation given above in (ii) and (iii) are easily verified by using once again the mentioned transformation theorem. In addition to the five properties above mentioned, the BS distribution possesses the following interesting features: (vi) by using its genesis, an analogy in the modeling of medical data can be obtained (DESMOND, 1985; CRAMÉR, 1946, p. 219); (vii) in the context of medical data, it has shown to be a very competitive model in terms of fitting (BARROS; PAULA; LEIVA, 2008; LEIVA et al., 2007; LEAO et al., 2017); (viii) it is a member of the family of log-symmetric distributions, as the log-normal (LN) distribution (MARSHALL; OLKIN, 2007; VANEGAS; PAULA, 2016a; VANEGAS; PAULA, 2016b). Note that the log-symmetric family corresponds to a random variable with the same distribution as its reciprocal or as ordinary symmetry of the distribution of the logged random variable (JONES, 2008). More details about log-symmetric distributions and, in particular of the log-BS distribution, are discussed in Section 2.2.3. Furthermore of the eight properties above mentioned, the BS distribution has a survival function (SF) and a hazard rate (HR) given respectively by

$$S_T(t;\alpha,\sigma)=\Phi\left(-\frac{1}{\alpha}\left(\sqrt{\frac{t}{\sigma}}-\sqrt{\frac{\sigma}{t}}\right)\right),\quad h_T(t;\alpha,\sigma)=\frac{\phi\left(\frac{1}{\alpha}\left(\sqrt{\frac{t}{\sigma}}-\sqrt{\frac{\sigma}{t}}\right)\right)\frac{t^{-3/2}(t+\sigma)}{2\alpha\sigma^{1/2}}}{\Phi\left(-\frac{1}{\alpha}\left(\sqrt{\frac{t}{\sigma}}-\sqrt{\frac{\sigma}{t}}\right)\right)},\quad t>0.$$

Note that the HR of the BS distribution has the following characteristics: (ix) it is unimodal for any α, increasing for $t<t_c$, and decreasing for $t>t_c$, where t_c denotes its change-point; (x) it goes to $1/(2\alpha^2\beta)$ as $t\to\infty$; (xi) it tends to be decreasing for $\alpha>2$, as $\alpha\to\infty$; and (xii) it tends to be increasing as $\alpha\to 0$; see Kundu, Kannan and Balakrishnan (2008).

2.2.3 The log-Birnbaum-Saunders distribution

When modeling data with the BS distribution, the log-BS distribution is needed. A random variable Y has a log-BS distribution with shape ($\alpha>0$) and location ($\mu\in\mathbb{R}$) parameters, which is denoted by log-BS(α,μ), if and only if $Z=(2/\alpha)\text{sinh}((Y-\mu)/2)\sim \text{N}(0,1)$. Then, the CDF of Y is given by

$$F_Y(y;\alpha,\mu)=\Phi\left(\frac{2}{\alpha}\text{sinh}\left(\frac{y-\mu}{2}\right)\right),\quad y\in\mathbb{R},\mu\in\mathbb{R},\alpha>0. \tag{2.9}$$

Consequently, from (2.9), the PDF of Y is obtained as

$$f_Y(y;\alpha,\mu) = \frac{1}{\alpha\sqrt{2\pi}}\cosh\left(\frac{y-\mu}{2}\right)\exp\left(-\frac{2}{\alpha^2}\sinh^2\left(\frac{y-\mu}{2}\right)\right), \quad y\in\mathbb{R}, \mu\in\mathbb{R}, \alpha>0, \tag{2.10}$$

whereas the logarithm of the PDF given in (2.10) is expressed as

$$\log(f_Y(y;\alpha,\mu)) = -\log(2) - \frac{\log(2\pi)}{2} + \log\left(\frac{2}{\alpha}\cosh\left(\frac{y-\mu}{2}\right)\right) - \frac{2}{\alpha^2}\sinh^2\left(\frac{y-\mu}{2}\right), \quad y\in\mathbb{R}. \tag{2.11}$$

Some properties of the log-BS distribution are presented as follows. If $Y \sim$ log-BS(α,μ), then: (i) $T = \exp(Y) \sim$ BS(α,σ), which means that the log-BS PDF given in (2.10) can be obtained from the standard normal PDF as in (2.9) or from the BS PDF defined in (2.7); (ii) $\mathrm{E}(Y) = \mu$; (iii) there is no closed form for the variance of Y, but based upon an asymptotic approximation for the moment generating function of the log-BS distribution, it follows that, if $\alpha \to 0$, then $\mathrm{Var}(T) = \alpha^2 - \alpha^4/4$, whereas that, if $\alpha \to \infty$, then $\mathrm{Var}(T) = 4(\log^2(\sqrt{2}\alpha) + 2 - 2\log(\sqrt{2}\alpha))$; (iv) if $X = \pm Y + d$, then $X \sim$ log-BS$(\alpha, \pm\mu + d)$; and (v) the log-BS distribution is symmetric around μ, unimodal for $\alpha \leq 2$ and bimodal for $\alpha > 2$; see Figure 2.1(right). Therefore, a further property of the BS distribution is that its logarithmic version has a flexible bimodality, which is not shared by other natural competitors. Note that: (vi) if $T \sim$ BS(α,σ), then $Y = \log(T) \sim$ log-BS$(\alpha, \log(\sigma))$ (RIECK; NEDELMAN, 1991); (vii) if $T \sim$ LN(μ,σ), then $Y = \log(T) \sim$ N(μ,σ) (CROW; SHIMIZU, 1988); (viii) if $T \sim$ IG(μ,σ), then $Y = \log(T) \sim$ log-IG$(\sigma, \log(\mu)$ (KOTZ; LEIVA; SANHUEZA, 2010); and (ix) if $T \sim$ GA(α,σ), then $Y = \log(T) \sim$ log-GA(α,σ) (JOHNSON; KOTZ; BALAKRISHNAN, 1995), where IG and GA are acronyms for the inverse Gaussian and gamma distributions, respectively.

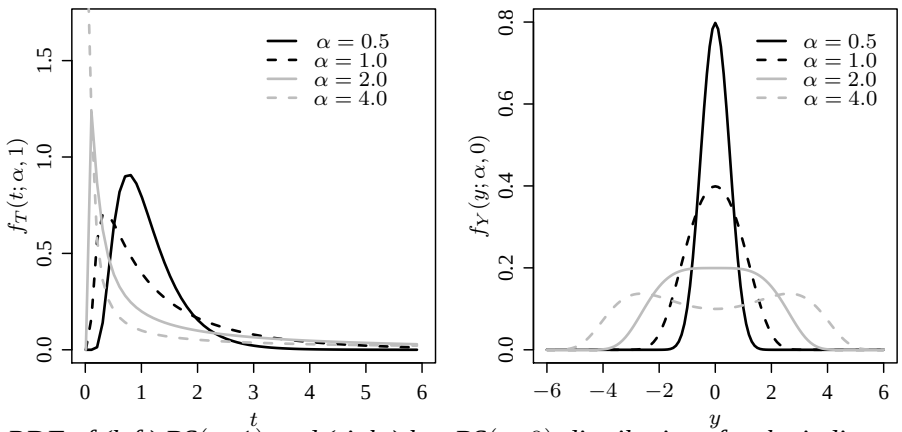

Figure 2.1: *PDF of (left) BS$(\alpha,1)$ and (right) log-BS$(\alpha,0)$ distributions for the indicated value of α.*

2.3 The tobit-BS model

2.3.1 Formulation

Consider the BS regression model

$$T_i = \exp(\boldsymbol{x}_i^\top \boldsymbol{\beta})\delta_i, \quad i = 1, \ldots, n, \tag{2.12}$$

originally proposed by Rieck and Nedelman (1991), where T_i is the response, $\boldsymbol{x}_i$ and $\boldsymbol{\beta}$ are analogously defined as in (2.2), and $\delta_i \sim \mathrm{BS}(\mathrm{ff}, 1)$ is the model error. By applying logarithm in (2.12), we obtain

$$Y_i = \boldsymbol{x}_i^\top \boldsymbol{\beta} + \varepsilon_i, \quad i = 1, \ldots, n, \tag{2.13}$$

where $Y_i = \log(T_i)$ is the log-response, $\boldsymbol{x}_i$ and $\boldsymbol{\beta}$ are similar to (2.1), and $\varepsilon_i = \log(\delta_i) \sim$ log-BS$(\alpha, 0)$ is the error term of the model. Then, based on (2.1) and (2.13), we propose a tobit-BS model as

$$Y_i = \begin{cases} y_0, & Y_i^* \leq y_0, \quad i = 1, \ldots, m; \\ \boldsymbol{x}_i^\top \boldsymbol{\beta} + \varepsilon_i, & Y_i^* > y_0, \quad i = m+1, \ldots, n; \end{cases} \tag{2.14}$$

where $Y_i^* = \log(T_i^*)$, and $\boldsymbol{\beta}$, $\boldsymbol{x}_i$ and ε_i are given in (2.13). Note that the BS distribution have a positive support. Thus, the corresponding left-censored response must be strictly positive.

2.3.2 Estimation

We estimate the parameters of the tobit-BS model defined in (2.14) with the ML method, in which case the log-likelihood function for $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}^\top)^\top$ obtained from (2.11) takes the form

$$\ell(\boldsymbol{\theta}) = -(n-m)\log(2) - (n-m)\frac{\log(2\pi)}{2} + \sum_{i=1}^{m}(\log \Phi(\zeta_{i2}^{\mathrm{c}})) + \sum_{i=m+1}^{n}\left(\log(\zeta_{i1}) - \frac{\zeta_{i2}^2}{2}\right), \tag{2.15}$$

where

$$\zeta_{i2}^{\mathrm{c}} = \frac{2}{\alpha}\sinh\left(\frac{y_0 - \boldsymbol{x}_i^\top \boldsymbol{\beta}}{2}\right), \quad \zeta_{i1} = \frac{2}{\alpha}\cosh\left(\frac{y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}}{2}\right), \quad \zeta_{i2} = \frac{2}{\alpha}\sinh\left(\frac{y_i - \boldsymbol{x}_i^\top \boldsymbol{\beta}}{2}\right). \tag{2.16}$$

To obtain the ML estimator of $\boldsymbol{\theta}$, it is necessary to maximize the log-likelihood function given in (2.15). The corresponding score vector is $\dot{\boldsymbol{\ell}} = \partial\ell(\boldsymbol{\theta})/\partial\boldsymbol{\theta} = (\dot{\ell}_\alpha, \dot{\boldsymbol{\ell}}_{\boldsymbol{\beta}}^\top)^\top$, which contains the first

partial derivatives of (2.15), where

$$\dot{\ell}_{\alpha}=\begin{cases}-\frac{\lambda(\zeta_{i2}^{\mathrm{c}})\zeta_{i2}^{\mathrm{c}}}{\alpha}, & i=1,\ldots,m;\\ \frac{\zeta_{i2}^{2}-1}{\alpha}, & i=m+1,\ldots,n;\end{cases}\quad \dot{\boldsymbol{\ell}}_{\boldsymbol{\beta}}=\begin{cases}-\frac{x_{ij}\cosh(\delta)\lambda(\zeta_{i2}^{\mathrm{c}})}{\alpha}, & i=1,\ldots,m;\\ \frac{x_{ij}\sinh(2\delta_i)}{\alpha^2}-\frac{x_{ij}\tanh(\delta_i)}{2}, & i=m+1,\ldots,n;\end{cases} \tag{2.17}$$

with $\delta_i = y_0 - x_i^{\top}\beta/2$, if $i = 1,\ldots,m$, and $\delta_i = y_i - x_i^{\top}\boldsymbol{\beta}/2$, if $i = m+1,\ldots,n$, and $\lambda(\zeta_{i2}^{\mathrm{c}}) = \phi(\zeta_{i2}^{\mathrm{c}})/\Phi(\zeta_{i2}^{\mathrm{c}})$. The ML estimate of $\boldsymbol{\theta}$ is obtained equating (2.17) to zero. However, the system of equations defined by these two equations does not have an analytic solution. Leiva et al. (2007) suggested to use the Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-Newton algorithm, employing as starting values for the numerical procedure $\widehat{\alpha}^2 = 4(\sinh((y_i - \boldsymbol{x}_i^{\top}\widehat{\boldsymbol{\beta}})/2))^2/(n-m)$ and $\widehat{\boldsymbol{\beta}} = (\boldsymbol{X}^{\top}\boldsymbol{X})^{-1}\boldsymbol{X}^{\top}\boldsymbol{y}$, where $\boldsymbol{X}$ is a matrix composed by the rows $\boldsymbol{x}_i$. Recall that the log-BS distribution can be bimodal if $\alpha > 2$. It implicates that the log-likelihood function has more than one maximum value. However, Rieck and Nedelman (1991) and Leiva (2016) argued that $\alpha > 2$ is unusual in practice, which means that the maximum point is often unique.

2.3.3 Inference

Assuming that some regularity conditions defined in Cox and Hinkley (1974) are satisfied, the ML estimators $\widehat{\alpha}$ and $\widehat{\boldsymbol{\beta}}$ are consistent and follow a multivariate normal joint asymptotic distribution. This distribution has an asymptotic mean vector with elements α and $\boldsymbol{\beta}$ and have an asymptotic covariance matrix equal to $\mathcal{J}(\boldsymbol{\theta})^{-1}$, which can be approximated by the expected Fisher information matrix. Therefore, as $n \to \infty$, we have that

$$\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \quad \overset{\mathrm{d}}{\to} \quad \mathrm{N}_{p+1}\left(\mathbf{0}_{p+1}, \mathcal{J}(\boldsymbol{\theta})^{-1}\right), \tag{2.18}$$

where $\mathcal{J}(\boldsymbol{\theta}) = \lim_{n\to\infty}(1/n)\mathcal{I}(\boldsymbol{\theta})$, with $\mathcal{I}(\boldsymbol{\theta})$ being the expected Fisher information matrix. In addition, $\overset{\mathrm{d}}{\to}$ means convergence in distribution to and $\mathbf{0}_{p+1}$ is a $p \times 1$ vector of zeros. Note that $\widehat{\mathcal{I}}(\boldsymbol{\theta})^{-1}$ is a consistent estimator of the asymptotic variance-covariance matrix of $\widehat{\boldsymbol{\theta}}$, $\mathcal{J}(\boldsymbol{\theta})^{-1}$ say. In practice, one may approximate the expected Fisher information matrix by its observed version (EFRON; HINKLEY, 1978), whereas the diagonal elements of the inverse observed information matrix may be used to approximate the corresponding standard errors (SEs). The observed Fisher information matrix is obtained from the Hessian matrix, which contains the second partial derivatives of (2.15) and it is given by

$$\ddot{\boldsymbol{\ell}} = \begin{pmatrix} \mathrm{tr}(\boldsymbol{G}) & \boldsymbol{k}^{\top}\boldsymbol{X} \\ \boldsymbol{X}^{\top}\boldsymbol{k} & \boldsymbol{X}^{\top}\boldsymbol{V}\boldsymbol{X} \end{pmatrix}, \tag{2.19}$$

where $\boldsymbol{V} = \text{diag}\{v_1(\boldsymbol{\theta}), v_2(\boldsymbol{\theta}), v_3(\boldsymbol{\theta}), \ldots, v_n(\boldsymbol{\theta})\}$, $\boldsymbol{k} = (k_1(\boldsymbol{\theta}), k_2(\boldsymbol{\theta}), k_3(\boldsymbol{\theta}), \ldots, k_n(\boldsymbol{\theta}))^\top$ and $\boldsymbol{G} = \text{diag}\{g_1(\boldsymbol{\theta}), g_2(\boldsymbol{\theta}), g_3(\boldsymbol{\theta}), \ldots, g_n(\boldsymbol{\theta})\}$, with

$$v_i(\boldsymbol{\theta}) = \begin{cases} \dfrac{\sinh(\delta_i)\lambda(\zeta_{i2}^{\text{c}})}{2\alpha} - \dfrac{\zeta_{i1}^{\text{c}}\zeta_{i2}^{\text{c}}\cosh(\delta_i)\lambda(\zeta_{i2}^{\text{c}})}{2\alpha} - \dfrac{\cosh^2(\delta_i)\lambda^2(\zeta_{i2}^{\text{c}})}{\alpha^2}, & i = 1, \ldots, m; \\ \dfrac{1}{4}(\text{sech}(\delta_i))^2) - \dfrac{1}{\alpha^2}\cosh(2\delta_i), & i = m+1, \ldots, n; \end{cases}$$

$$k_i(\boldsymbol{\theta}) = \begin{cases} -\dfrac{1}{2\alpha}\lambda(\zeta_{i2}^{\text{c}})(\zeta_{i2}^{\text{c}})^2\zeta_{i1}^{\text{c}} + \dfrac{\lambda(\zeta_{i2}^{\text{c}})\cosh(\delta_i)}{\alpha^2} - \dfrac{\zeta_{i2}^{\text{c}}}{\alpha^2}\lambda^2(\zeta_{i2}^{\text{c}})\cosh(\delta_i), & i = 1, \ldots, m; \\ -\dfrac{2}{\alpha^3}\sinh(2\delta_i), & i = m+1, \ldots, n; \end{cases}$$

$$g_i(\boldsymbol{\theta}) = \begin{cases} -\dfrac{\zeta_{i2}^{\text{c}}}{\alpha^2}(((\zeta_{i2}^{\text{c}})^2\lambda(\zeta_{i2}^{\text{c}}) + \zeta_{i2}^{\text{c}}\lambda^2(\zeta_{i2}^{\text{c}})) - 2\lambda(\zeta_{i2}^{\text{c}})), & i = 1, \ldots, m; \\ \dfrac{1}{\alpha^2} - \dfrac{3\zeta_{i2}}{\alpha^2}, & i = m+1, \ldots, n, \end{cases}$$

and ζ_{i1}, ζ_{i2} and ζ_{i2}^{c} presented in (2.16), whereas $\zeta_{i1}^{\text{c}} = (2/\alpha)\cosh((y_0 - \boldsymbol{x}_i^\top\boldsymbol{\beta})/2)$, for $i = 1, \ldots, m$. Asymptotic inference for the tobit-BS model parameters can be made by using the results given in (2.18) and (2.19).

2.3.4 Model checking

The objective of the residual analysis is to assess if the errors hold the distributional, heteroscedasticity and autocorrelation assumptions, as well the presence of atypical data. In classic regression models, Pearson and studentized residuals are often used. However, in tobit models, generally, even under normality, these types of residuals are inadequate (BARROS et al., 2010). For the tobit-BS model, we propose the generalized Cox-Snell (GCS) residual, which is often used in generalized linear models and survival analysis. The GCS residual is given by

$$r_i^{\text{GCS}} = -\log(\widehat{S}_Y(y_i; \alpha, \boldsymbol{\beta}, \boldsymbol{x})), \quad i = 1, \ldots, n,$$

where $\widehat{S}_Y$ is the estimated SF of the model defined in (2.14). Note that the estimate SF for the log-BS model evaluated at the case i is given by

$$S_Y(y_i; \alpha, \boldsymbol{\beta}, \boldsymbol{x}) = \Phi\left(-\frac{2}{\widehat{\alpha}}\sinh\left(\frac{y_i - \widehat{\mu}_i}{2}\right)\right), \quad i = 1, \ldots, n.$$

Here, regardless the specification of the model, if it is correct, then the GCS residual follows an EXP(1) distribution (BHATTI, 2010; LEIVA et al., 2014).

2.4 Diagnostic analysis

2.4.1 Local Influence

One way to assess the effect produced by one case on the ML estimates is the deletion of each of them from the data set and then evaluate if it exercises influence on the estimates or not. This approach is known as global influence and is presented in Section 2.4.2. However, the local influence method relies on a geometric differentiation, taking the curvature of the plane of the log-likelihood function. This method does not require any deletion. Often the differential comparison is made before and after a perturbation. There are many ways to conduct a local influence analysis. We use case-weight, response and covariate perturbation schemes. Recalling that $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}^\top)^\top$ is the vector of parameters, let $\ell(\boldsymbol{\theta}|\boldsymbol{\omega})$ be the log-likelihood function of the model defined in (2.14) perturbed by $\boldsymbol{\omega}$, where $\boldsymbol{\omega}$ is a subset of $\Omega \in \mathbb{R}^n$. One way to evaluate the influence of a perturbation over the estimates of $\boldsymbol{\theta}$ is to use the likelihood distance (LD), which is given by

$$\text{LD}(\boldsymbol{\omega}) = 2(\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}})), \tag{2.20}$$

where $\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$ is the ML estimate of $\boldsymbol{\theta}$ based on $\ell(\boldsymbol{\theta}|\boldsymbol{\omega})$. To detect what case exercises influence on $\text{LD}(\boldsymbol{\omega})$ defined in (2.20), we study the normal curvature in the neighbourhood of the point of no perturbation $\boldsymbol{\omega}_0$, in the direction of a unit vector $\boldsymbol{l}$, with $||\boldsymbol{l}||= 1$. The normal curvature (BARROS et al., 2010) can be expressed as

$$C_{\boldsymbol{l}}(\boldsymbol{\theta}) = 2|\boldsymbol{l}^\top \boldsymbol{\Delta}^\top \ddot{\boldsymbol{\ell}}^{-1} \boldsymbol{\Delta} \boldsymbol{l}|, \tag{2.21}$$

where $\boldsymbol{\Delta}$ is a matrix of perturbations and $\ddot{\boldsymbol{\ell}}$ is the Hessian matrix given in (2.19). The matrix $\boldsymbol{\Delta}$ has elements $\partial^2 \ell(\theta|\boldsymbol{\omega})/\partial\theta_j \partial\omega_i$, for $j = 1, \ldots, p+1, i = 1, \ldots, n$, and must be evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega} = \boldsymbol{\omega}_0$. To determine what case is influential under small perturbations, Barros et al. (2010) proposed an index plot based on the eigenvector of $\boldsymbol{l}_{\max}$, which can be constructed by using the maximum eigenvalue of

$$\boldsymbol{B}(\boldsymbol{\theta}) = |\boldsymbol{\Delta}^\top \ddot{\boldsymbol{\ell}}^{-1} \boldsymbol{\Delta}|, \tag{2.22}$$

evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega} = \boldsymbol{\omega}_0$. On the one hand, if the interest relies just on the vector of parameter $\boldsymbol{\beta}$, (2.21) becomes

$$C_{\boldsymbol{l}}(\boldsymbol{\theta}) = 2|\boldsymbol{l}^\top \boldsymbol{\Delta}^\top (\ddot{\boldsymbol{\ell}}^{-1} - \boldsymbol{B_1}) \boldsymbol{\Delta} \boldsymbol{l}|. \tag{2.23}$$

Note that (2.23) removes α from the analysis. Thus, the influence detection is only made on $\boldsymbol{\beta}$, where $\boldsymbol{B_1}$ under a tobit-BS model takes the form

$$\boldsymbol{B_1} = \begin{pmatrix} \text{tr}(\boldsymbol{G})^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}.$$

On the other hand, if the interest is just on α, (2.21) becomes

$$C_{\boldsymbol{l}}(\boldsymbol{\theta}) = 2|\boldsymbol{l}^\top \boldsymbol{\Delta}^\top (\ddot{\boldsymbol{\ell}}^{-1} - \boldsymbol{B_2})\boldsymbol{\Delta}\boldsymbol{l}|,$$

where $\boldsymbol{B}_2$ for the tobit-BS model is given by

$$\boldsymbol{B_2} = \begin{pmatrix} 0 & \mathbf{0} \\ \mathbf{0} & (\boldsymbol{X}^\top \boldsymbol{V}\boldsymbol{X})^{-1} \end{pmatrix},$$

with tr$(\boldsymbol{G})$ and $(\boldsymbol{X}^\top \boldsymbol{V}\boldsymbol{X})^{-1}$ being obtained from the Hessian matrix of the tobit-BS model given in (2.19). The maximum normal curvature vector, denoted by $\boldsymbol{l}_{\max}$, is an important direction to assess the local influence on $\widehat{\boldsymbol{\theta}}$. The vector $\boldsymbol{l}_i = \boldsymbol{e}_{in}$ denotes the direction of the case i, where $\boldsymbol{e}_{in}$ is the canonical basis of $\mathbb{R}^n$. It assumes zero for every case except the case i, which assumes the value one. The normal curvature of each case is given by $C_i(\theta) = 2|b_{ii}|$, where b_{ii} represents the ith element of the matrix defined in (2.22) for each case. An observation is potentially influential on $\boldsymbol{\theta}$ if $C_i(\widehat{\boldsymbol{\theta}}) > 2C(\widehat{\boldsymbol{\theta}})$, where $C(\widehat{\boldsymbol{\theta}})$ is the mean of the C_is for $i = 1, \ldots, n$ (LESAFFRE; VERBEKE, 1998). Next, we present three perturbation schemes and their corresponding perturbation matrices $\boldsymbol{\Delta}$.

Case-weight perturbation This scheme permits us to detect how cases with different weights affect the ML estimates of $\boldsymbol{\theta}$. Consider the log-likelihood function $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{=1}^{n} \boldsymbol{\omega}_i \ell_i$, with ℓ_i given in (2.15) and $\boldsymbol{\omega}_i \in [0, 1]$ being a perturbation vector. Taking its partial derivative with respect to $\boldsymbol{\omega}^\top$, we obtain

$$\frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \boldsymbol{\omega}^\top} = \sum_{i=1}^{m} \ell_i(\boldsymbol{\theta}) \boldsymbol{e}_{in}^\top,$$

with $\boldsymbol{e}_{in}^\top$ denoting an $n \times 1$ vector. After evaluating $\boldsymbol{\theta}$ at $\widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega}$ at $\boldsymbol{\omega}_0$, we obtain a perturbation matrix

$$\boldsymbol{\Delta} = \sum_{i=1}^{n} \boldsymbol{h}_i \boldsymbol{e}_{in}^\top, \tag{2.24}$$

where $\boldsymbol{h}_i$ is given by

$$\boldsymbol{h}_i = \begin{pmatrix} \frac{\partial \ell_i(\boldsymbol{\theta})}{\partial \alpha} \\ \frac{\partial \ell_i(\boldsymbol{\theta})}{\partial \beta} \end{pmatrix}.$$

From the log-likelihood function defined in (2.15), we obtain an explicit expression for $\boldsymbol{h}_i$, which is

$$\begin{aligned} \dot{\boldsymbol{\ell}}_{\alpha_i}(\boldsymbol{\theta}|\boldsymbol{\omega}) &= a_i = \begin{cases} -\frac{\lambda(\zeta_{i2}^{\mathsf{C}})\zeta_{i2}^{\mathsf{C}}}{\alpha} & i = 1, \ldots, m; \\ \frac{1}{\alpha}(\zeta_{i2}^2 - 1) & i = m+1, \ldots, n; \end{cases} \\ \dot{\boldsymbol{\ell}}_{\boldsymbol{\beta}_i}(\boldsymbol{\theta}|\boldsymbol{\omega}) &= b_i = \begin{cases} -\frac{\lambda(\zeta_{i2}^{\mathsf{C}})\zeta_{i1}^{\mathsf{C}}}{2} & i = 1, \ldots, m; \\ \frac{1}{2}\left(\zeta_{i1}\zeta_{i2} - \frac{\zeta_{i2}}{\zeta_{i1}}\right) & i = m+1, \ldots, n. \end{cases} \end{aligned}$$

The case-weight perturbation matrix given by $\boldsymbol{\Delta}$ in (2.24) is decomposed in $\boldsymbol{\Delta}_\alpha = (a_1, \ldots, a_n)$ and $\boldsymbol{\Delta}_\beta = \boldsymbol{X}^\top \text{diag}\{b_1, \ldots, b_n\}$.

Response perturbation There are different scenarios that one can consider for the response perturbation. We consider here a response perturbation scheme with an additive perturbation, which is defined by $Y_{i\omega} = Y_i + \omega_i S_Y$, for $i = m+1, \ldots, n$, where S_Y is a scale component that can be the standard deviation (SD) of the response. Let us consider that the log-likelihood function of the tobit-BS model is given by $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^n \ell_i(\boldsymbol{\theta}|\boldsymbol{\omega})$, with the relevant part being

$$\ell_i(\boldsymbol{\theta}) = \begin{cases} \sum_{i=1}^m (\log \Phi(\zeta_{i2\omega 1}^{\text{c}})), & i = 1, \ldots, m; \\ \sum_{i=m+1}^n \left(\log(\zeta_{i1\omega 1}) - \frac{1}{2}(\zeta_{i2\omega 1}^2)\right), & i = m+1, \ldots, n; \end{cases} \tag{2.25}$$

where $\zeta_{i1\omega 1}, \zeta_{i2\omega 1}, \zeta_{i2\omega 1}^{\text{c}}$ are as defined in (2.16) after changing Y for $Y_{i\omega}$. The process of obtaining the perturbation matrix in this scheme is composed by two steps. First, we take the derivative of (2.25) with respect to $\boldsymbol{\omega}^\top$, that is,

$$\frac{\partial \ell_i(\boldsymbol{\theta}|\boldsymbol{\omega}_i)}{\partial \boldsymbol{\omega}^\top} = \begin{cases} \mathbf{0}_m^\top, & i = 1, \ldots, m, \\ \frac{S_Y}{2}\left(\frac{\zeta_{i2\omega 1}}{\zeta_{i1\omega 1}} - \zeta_{i2\omega 1}\zeta_{i1\omega 1}\right), & i = m+1, \ldots, n. \end{cases} \tag{2.26}$$

Then, the partial derivative of (2.26) with respect to $\boldsymbol{\theta}$ is computed and evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega} = \boldsymbol{\omega}_0$, to obtain the response perturbation matrix $\boldsymbol{\Delta}$ with elements

$$\boldsymbol{\Delta}_\alpha = (c_{m+1}, \ldots, c_n), \quad \boldsymbol{\Delta}_{\boldsymbol{\beta}} = \boldsymbol{X}^\top \text{diag}\{d_{m+1}, \ldots, d_n\},$$

where

$$c_i = \frac{S_Y \zeta_{i1}\zeta_{i2}}{\alpha}, \quad d_i = S_Y\left(\frac{1}{\alpha^2}\cosh(2\delta_i) - \frac{1}{4}\left(\text{sech}\,(\delta_i)\right)^2\right).$$

Note that the response perturbation scheme, in a tobit model, makes sense only for the non-censored part of the data. This occurs because the censored part of the data is either unobservable or below the threshold y_0, otherwise, the case i receives the same value y_0. Then, there is no perturbation in this part of the data.

Covariate perturbation As in the response scheme, there are several ways to perturb a continuous covariate. Here, we insert an additive perturbation that takes the form

$$x_{it\omega} = x_{it} + \omega_i S_X, \quad i = 1, \ldots, n,$$

where S_X can be the SD of the corresponding covariate X_t. Considering the log-likelihood function for the tobit-BS model to be $\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = \sum_{i=1}^n \ell_i(\boldsymbol{\theta}|\boldsymbol{\omega})$ from (2.15), we obtain as the relevant

part for the covariate perturbation scheme as

$$\ell_i(\boldsymbol{\theta}|\boldsymbol{\omega}) = \begin{cases} \sum_{i=1}^{m}(\log \Phi(\zeta^{\mathrm{c}}_{i2\boldsymbol{\omega}2})) & i = 1,\ldots,m; \\ \sum_{i=m+1}^{n}\left(\log(\zeta_{i1\boldsymbol{\omega}2}) - \frac{1}{2}(\zeta^2_{i2\boldsymbol{\omega}2})\right) & i = m+1,\ldots,n. \end{cases} \tag{2.27}$$

In order to obtain the perturbation matrix of (2.27), we take the derivative with respect to $\boldsymbol{\theta}$ and then with respect to $\boldsymbol{\omega}$, evaluating it at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega} = \boldsymbol{\omega}_0$ to obtain

$$\boldsymbol{\Delta}_\alpha = \begin{cases} -\frac{S_x\boldsymbol{\beta}_t}{\alpha}\left(\frac{\zeta^{\mathrm{c}}_{i1}\lambda(\zeta^{\mathrm{c}}_{i2})}{2} - \zeta^{\mathrm{c}}_{i2}\left(\frac{2\lambda(\zeta^{\mathrm{c}}_{i2})\sinh(2\delta_i)}{\alpha^2} + \frac{\lambda^2\zeta^{\mathrm{c}}_{i2}}{2}\right)\right), & i = 1,\ldots,m; \\ -\frac{2}{\alpha^3}S_x\boldsymbol{\beta}_t\sinh(2\delta_i), & i = m+1,\ldots,n; \end{cases}$$

and $\boldsymbol{\Delta}_{\boldsymbol{\beta}}$, which is a $p \times n$ matrix with elements

$$\Delta_{\boldsymbol{\beta}_{ij}} = \begin{cases} -\frac{S_x\boldsymbol{\beta}_t x_{ij}}{4}\left(\frac{\lambda(\zeta^{\mathrm{c}}_{i2})(\zeta^{\mathrm{c}}_{i1})^2\zeta^{\mathrm{c}}_{i2}}{\Phi(\zeta^{\mathrm{c}}_{i2})} + (\zeta^{\mathrm{c}}_{i2})^2\lambda^2(\zeta^{\mathrm{c}}_{i2}) - \lambda(\zeta^{\mathrm{c}}_{i2})\zeta^{\mathrm{c}}_{i2}\right) & \\ -\frac{S_x\zeta^{\mathrm{c}}_{i2}\lambda(\zeta^{\mathrm{c}}_{i2})}{2}, & i = 1,\ldots,m; \\ S_x\boldsymbol{\beta}_t x_{ij}\left(\frac{1}{4}\left(\operatorname{sech}(\delta_i)\right)^2 - \frac{1}{\alpha^2}\cosh(2\delta_i)\right), & i = m+1,\ldots,n, \end{cases}$$

if $j \neq t$, or

$$\Delta_{\boldsymbol{\beta}_{ij}} = \begin{cases} -\frac{S_x\boldsymbol{\beta}_t x_{ij}}{4}\left(\frac{\lambda(\zeta^{\mathrm{c}}_{i2})(\zeta^{\mathrm{c}}_{i1})^2\zeta^{\mathrm{c}}_{i2}}{\Phi(\zeta^{\mathrm{c}}_{i2})} + (\zeta^{\mathrm{c}}_{i2})^2\lambda^2(\zeta^{\mathrm{c}}_{i2}) - \lambda(\zeta^{\mathrm{c}}_{i2})\zeta^{\mathrm{c}}_{i2}\right) - \frac{S_x\zeta^{\mathrm{c}}_{i2}\lambda(\zeta^{\mathrm{c}}_{i2})}{2} & \\ -\frac{S_x x_{ij}}{\alpha}\left(\cosh(2\delta_i)\lambda(\zeta^{\mathrm{c}}_{i2})\right), & i = 1,\ldots,m; \\ S_x\boldsymbol{\beta}_t x_{ij}\left(\frac{1}{4}\left(\operatorname{sech}(\delta_i)\right)^2 - \frac{1}{\alpha^2}\cosh(2\delta_i)\right) & \\ +S_x\left(\frac{1}{\alpha^2}\sinh(2\delta_i) - \frac{1}{2}\tanh(\delta_i)\right), & i = m+1,\ldots,n, \end{cases}$$

if $j = t$.

2.4.2 Global influence

A global influence method commonly used in statistical modeling is the Cook distance, which allows us to assess the effect of each case on the estimated parameters. The procedure proposed by Cook (1977) suggests the deletion of each case and the evaluation of the log-likelihood function without the case i. The generalized Cook distance (GCD) for a standard tobit model was given by Barros et al. (2010) and it takes the form

$$\mathrm{GCD}_i(\boldsymbol{\theta}) = \frac{1}{p+1}(\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_{(i)})\ddot{\boldsymbol{\ell}}^{-1}(\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_{(i)}), \quad i = 1,\ldots,n, \tag{2.28}$$

where p is the number of model coefficients and $\widehat{\boldsymbol{\theta}}_{(i)}$ is the ML estimate of $\boldsymbol{\theta}$ without the case i. In order to facilitate the calculations, a first order approximation $\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_i \approx \ddot{\boldsymbol{\ell}}^{-1}_{(i)}\dot{\boldsymbol{\ell}}_{(i)}$ is used in (2.28)

and it becomes

$$\text{GCD}_i(\boldsymbol{\theta}) = \frac{1}{p+1}(\dot{\boldsymbol{\ell}}_{(i)}^{\top}\ddot{\boldsymbol{\ell}}_{(i)}^{-1}(-\ddot{\boldsymbol{\ell}})\ddot{\boldsymbol{\ell}}_{(i)}^{-1}\dot{\boldsymbol{\ell}}_{(i)}), \quad i = 1, \ldots, n, \tag{2.29}$$

where $\dot{\boldsymbol{\ell}}_{(i)}$ and $\ddot{\boldsymbol{\ell}}_{(i)}$ are the score vector and the Hessian matrix from the tobit-BS model defined in (2.17) and (2.19), respectively, without considering the case i and evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$. Note that, if the approximation described in (2.29) is not used, we must calculate the GCD eliminating each case. However, with this approximation, we need to calculate the GCD just once. Usually, the diagnostics analysis relies on the vector $\boldsymbol{\beta}$ of coefficients. Then, in that situation, (2.28) becomes

$$\text{GCD}_i(\boldsymbol{\beta}) = \frac{1}{p}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)})\ddot{\boldsymbol{\ell}}^{-1}(\widehat{\boldsymbol{\beta}} - \widehat{\boldsymbol{\beta}}_{(i)}), \quad i = 1, \ldots, n.$$

To determine whether the case i is potentially influential on $\boldsymbol{\beta}$, we use the same benchmark used by Zhu and Zhang (2004) and Barros et al. (2010), which is $2/n$. If the value of the GCD for the case i is greater than $2/n$, it is potentially influential on the estimated vector of parameters.

2.5 Application

2.5.1 The data

The standard tobit and tobit-BS models are now used to analyze a real-world data set data provided by Moulton and Halsey (1995) from a case-study of measles vaccines. Neutralization antibody levels were collected from 330 children at 12 months of age. The LDL was 0.1 international units (IU) or -2.306 in logarithm scale. Around 86 (26.1%) of the cases are below the LDL and then recorded as 0.1. The following covariates were considered: X_1 indicates the type of vaccine used (0 if Schwartz and 1 if Edmonston-Zagreb); X_2 is the level of the dosage (0 if medium and 1 if high); and X_3 is the children's gender (0 if male and 1 if female). The data set can be obtained from the authors under request or in Moulton and Halsey (1995). In this particular data set all covariables are dummy but it is interesting to note that the proposed methodology does not impose any restriction on the independent variables.

2.5.2 Software

The estimation of parameters for the standard tobit model was performed by using the `tobit()` function of the `AER` package of the software `R` to fit tobit regression models (KLEIBER; ZEILEIS, 2008; KLEIBER; ZEILEIS, 2015), whereas influence diagnostics for generalized tobit models, including the normal tobit model, can be produced by the `tobitdiag` package (SANTOS-NETO, 2016; BARROS et al., 2017) to be installed with the command

```
devtools::install_github("tobitproject/tobitdiag")
```

Estimation and diagnostic results regarding the tobit-BS model were implemented by functions developed in R. The codes are available under request from the authors. In addition, we use the `robustbase` package to construct box-plots adjusted for asymmetrical data (ROUSSEEUW et al., 2015).

2.5.3 Monte Carlo simulations

We present a Monte Carlo simulation study with 5000 replications to evaluate the performance of the ML estimators of the tobit-BS model parameters. The sample sizes considered are $n = 100, 300, 500$, with parameters $\alpha = 0.25, 0.5, 1.5, 2.5, \boldsymbol{\beta} = (0.2, 0.5)^\top$ and censoring proportions equal to $\varrho = 0.1, 0.25, 0.4, 0.5$. One covariate X is considered, where $X \sim \text{Uniform}(0, 1)$. We compute the empirical bias and mean squared error (MSE) in order to do the performance evaluation. All numerical calculations were done in the R software. Table 3.1 presents the results obtained for the indicated sample sizes, parameters values and censoring proportions. This table shows that, for $\alpha = 0.25, 0.5, 1.5, 2.5$, and $\varrho = 0.1, 0.25, 0.4$, the empirical bias and MSE decrease when n increases, as expected. Table 3.1 indicates that the model is not adequate to deal with $\alpha = 1.5, 2.5$, for a censoring proportion $\varrho = 0.5$. In these cases, the estimators are not consistent. In general, the results provide a good performance of the tobit-BS model.

2.5.4 Data analysis

Table 2.2 reports the descriptive statistics of the observed neutralization antibody levels from the measles vaccines case-study. This table includes the median, mean, SD and coefficients of variation (CV), skewness (CS) and kurtosis (CK) values. The CK and CS indicate the positively skewed nature and high kurtosis level of the data distribution. Figure 2.2 shows the scaled total time on test (TTT) plot, histogram and boxplots for the measles vaccine data.

The TTT plot is a tool to assess whether a particular distribution is suitable or not for a data set. This is done by characterizing the shape of an HR. We can detect the type of HR that the data have and then choose a suitable distribution. Let $h_T(t) = f_T(t)/(1 - F_T(t))$ be the HR of a random variable T, where f_T and F_T are the PDF and CDF of T, respectively. Then, the TTT function is $W(u) = H^{-1}(u)/H^{-1}(1)$, for $0 \leq u \leq 1$, where $H^{-1}(u) = \int_0^{F_T^{-1}(u)} (1 - F_T(z))\mathrm{d}z$, with F_T^{-1} denoting the inverse function of the CDF of T. A plot of the points $(k/n, W_n(k/n))$ can approximate W, with $W_n(k/n) = (\sum_{i=1}^{k} t_{(i)} + (n-k)t_k)/\sum_{i=1}^{n} t_{(i)}$, for $k = 1, \ldots, n$, and $t_{(i)}$ denoting the ith observed order statistic; see, for example, Figure 1 in Azevedo et al. (2012) for some theoretical shapes of scaled TTT curves.

Note that the skewed nature is confirmed by the histogram of Figure 2.2(left), whereas the TTT plot displayed in Figure 2.2(center) suggests a decreasing HR. Moreover, Figure 2.2(right) indicates that some outliers considered by the usual boxplot are not outliers when we consider the

Table 2.1: *Empirical bias and MSE (in parentheses) of the ML estimators for the tobit-BS model parameters by using the indicated sample sizes and parameter values with simulated data.*

n		$\varrho = 0.10$			$\varrho = 0.25$		
	α	$\widehat{\alpha}$	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{\alpha}$	$\widehat{\beta}_0$	$\widehat{\beta}_1$
100	0.25	0.00080(0.00036)	0.01419(0.00207)	−0.01754(0.00686)	0.01817(0.00062)	−0.03534(0.00313)	0.02691(0.00847)
	0.50	0.00947(0.00304)	0.10164(0.01724)	−0.07389(0.02747)	−0.03035(0.00250)	0.11585(0.01866)	−0.08166(0.02755)
	1.50	−0.05657(0.02271)	0.24971(0.17299)	−0.10247(0.15517)	−0.48643(0.24338)	0.39486(0.19059)	−0.18133(0.13820)
	2.50	−0.39635(0.23887)	0.40309(0.26798)	−0.16064(0.34102)	−0.93813(0.89697)	0.59277(0.42502)	−0.21088(0.25953)
300	0.25	0.00068(0.00013)	0.01497(0.00084)	−0.01895(0.00250)	0.01811(0.00044)	−0.03438(0.00180)	0.02583(0.00318)
	0.50	0.00349(0.00188)	0.10003(0.01303)	−0.07169(0.01297)	−0.02981(0.00175)	0.11350(0.01471)	−0.07978(0.01380)
	1.50	−0.11450(0.01684)	0.15481(0.12207)	−0.10074(0.06064)	−0.48247(0.23481)	0.39321(0.16613)	−0.16377(0.06822)
	2.50	−0.38176(0.18983)	0.37443(0.21699)	−0.14910(0.12536)	−0.92855(0.86751)	0.57346(0.37655)	−0.20073(0.11413)
500	0.25	0.00026(0.00008)	0.01246(0.00061)	−0.01556(0.00168)	0.01302(0.00041)	−0.03389(0.00152)	0.02549(0.00219)
	0.50	0.00301(0.00050)	0.09919(0.01121)	−0.07020(0.01050)	−0.01660(0.00053)	0.07255(0.01102)	−0.04193(0.01088)
	1.50	−0.00455(0.00212)	0.09372(0.08997)	−0.09511(0.03865)	−0.38160(0.23317)	0.39169(0.16357)	−0.11684(0.05626)
	2.50	−0.30617(0.17775)	0.32750(0.10479)	−0.11983(0.08525)	−0.92740(0.86323)	0.51755(0.37194)	−0.11511(0.08992)
n		$\varrho = 0.40$			$\varrho = 0.50$		
	α	$\widehat{\alpha}$	$\widehat{\beta}_0$	$\widehat{\beta}_1$	$\widehat{\alpha}$	$\widehat{\beta}_0$	$\widehat{\beta}_1$
100	0.25	0.05175(0.00303)	−0.10352(0.01300)	0.06820(0.01463)	0.07072(0.00539)	−0.14971(0.02494)	0.07805(0.01724)
	0.50	−0.02178(0.00312)	0.05843(0.00950)	−0.07950(0.02995)	−0.01307(0.00830)	0.00899(0.01016)	−0.08380(0.03368)
	1.50	−0.62570(0.62582)	0.49902(0.41086)	−0.15857(0.13771)	−2.58339(13.00365)	1.55824(4.19966)	0.15169(0.37660)
	2.50	−1.74532(5.89450)	1.05179(2.08633)	−0.11833(0.30030)	−4.08696(35.66735)	1.85482(5.92291)	0.22605(0.70380)
300	0.25	0.05044(0.00297)	−0.10259(0.01129)	0.06722(0.00781)	0.07055(0.00509)	−0.14899(0.02306)	0.07746(0.00974)
	0.50	0.00124(0.00107)	0.05811(0.00519)	−0.05327(0.01469)	−0.01041(0.00144)	0.00530(0.00096)	−0.08279(0.01563)
	1.50	−0.55699(0.39553)	0.45312(0.26916)	−0.11652(0.06998)	−3.97332(34.04745)	2.04444(6.37383)	0.14853(0.33840)
	2.50	−1.28276(2.64321)	0.80957(1.01003)	−0.10019(0.12096)	−7.82872(221.02820)	2.28197(9.16912)	0.14196(0.71841)
500	0.25	0.04374(0.00296)	−0.10229(0.01092)	0.06718(0.00648)	0.06293(0.00504)	−0.14860(0.02259)	0.07712(0.00817)
	0.50	−0.00098(0.00067)	0.05804(0.00457)	−0.02318(0.01166)	−0.00965(0.00083)	0.00421(0.00152)	−0.08128(0.01183)
	1.50	−0.52299(0.27465)	0.43081(0.19146)	−0.10058(0.05614)	−3.85664(30.10071)	2.02818(6.64127)	0.14783(0.27190)
	2.50	−1.53321(3.64631)	0.99536(1.59385)	−0.09344(0.11624)	−6.85616(177.49570)	2.10903(9.75583)	0.10385(0.60329)

adjusted boxplot.

Table 2.2: *Descriptive summary for the measles vaccine data.*

n	Min	Max	Mean	Median	SD	CV	CS	CK
330	0.10	15.47	1.20	0.40	2.10	174.74%	3.46	14.37

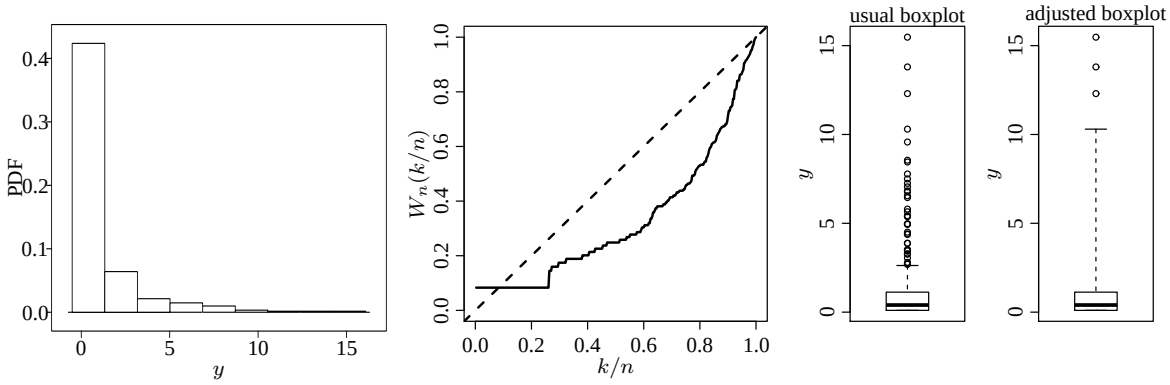

Figure 2.2: *Histogram (left), TTT plot (center), and boxplots (right) for the measles vaccine data.*

2.5.5 Analysis under the normal tobit model

Let us consider a standard tobit model defined by

$$Y_i = \begin{cases} 0.1, & Y_i^* \leq 0.1,\ i = 1, \ldots, 85, \\ Y_i^* = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{2i} + \beta_3 x_{3i} + \varepsilon_i, & Y_i^* > 0.1,\ i = 86, \ldots, 330, \end{cases}$$

where $\varepsilon_i \overset{\text{IID}}{\sim} \text{N}(0, \sigma^2)$. The response Y_i is a latent variable that is observed for values greater than 0.1 and censored otherwise. The ML estimates, computed by the BFGS method, along with the SEs and p-values of the t-test for the standard tobit model, are presented in Table 2.4. Note that the parameters β_1, β_2 and β_3 are not statistically significant at a 5% level. In Table 2.4, we also show the values for Akaike (AIC) and Bayes (BIC) information criteria. Figure 2.4 (left) shows the quantile versus quantile (QQ) plot with simulated envelope for the GCS residual based on the standard tobit model. This figure indicates that the GCS residual does not present an agreement with the EXP(1) distribution.

Next, we carry out a diagnostic analysis based on global and local influence. Figure 2.3 displays index plots of local influence under the case-weight perturbation, which shows the cases #328, #329 and #330 as potentially influential. In order to assess the global influence on the ML estimates when a case is removed, we analyze the index plot of the GCD shown in Figure 2.4 (right). It indicates the cases #328 and #328 as potentially influential.

The impact of the detected influential cases on the model inference can be checked by computing the relative change (RC), which is obtained by removing influential cases and re-estimating the parameters as $\text{RC}_{\theta_{j(i)}} = |[\widehat{\theta}_j - \widehat{\theta}_{j(i)}]/\widehat{\theta}_j| \times 100\%$, where $\widehat{\theta}_{j(i)}$ is the ML estimate of θ_j after

removing the case i, for $j = 1, \ldots, 4$ and $i = 1, \ldots, n$, with $\theta_1 = \beta_0$, $\theta_2 = \beta_1$, $\theta_3 = \beta_2$, and $\theta_4 = \beta_3$. Table 2.3 reports the RCs in the parameter estimates obtained by considering the data with dropped cases. Note that, in general, the RCs are large. Note also that after the exclusion of the cases #328 and #330, β_3 becomes significant at 10%, and after the exclusion of the group that contains the cases #328, #329 and #330, β_3 becomes significant at 5%; see values detached in gray. Therefore, we have inferential changes after removing potentially influential cases. We expect that the tobit-BS model does not present these inference problems.

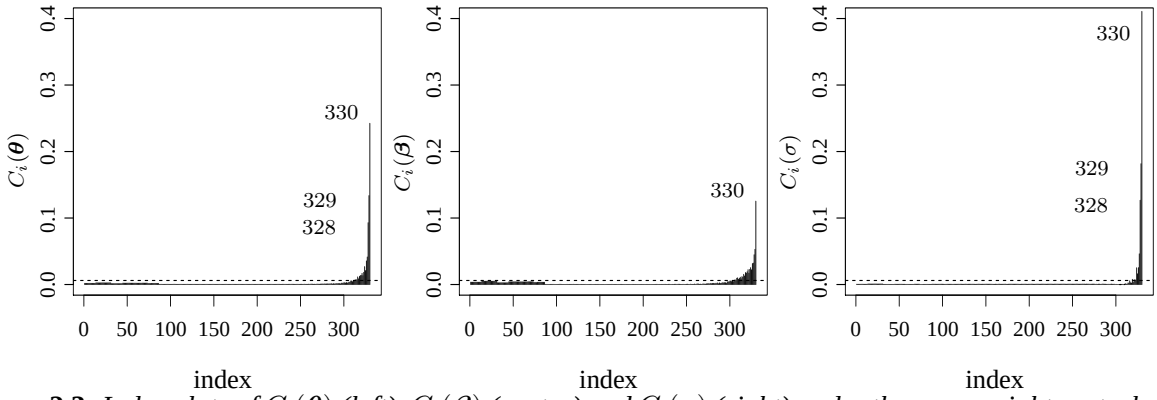

Figure 2.3: *Index plots of $C_i(\boldsymbol{\theta})$ (left), $C_i(\boldsymbol{\beta})$ (center) and $C_i(\sigma)$ (right) under the case-weight perturbation in the standard tobit model with the measles vaccine data.*

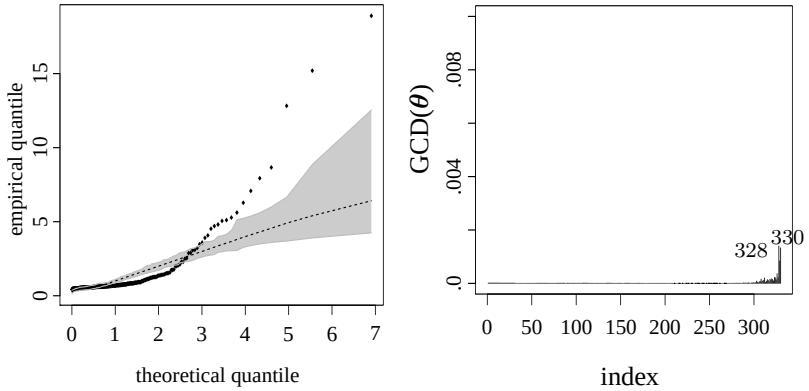

Figure 2.4: *QQ plot and its envelope for the GCS residual (left) and index plot of the GCD (right) for the standard tobit model with measles vaccine data.*

Table 2.3: *RCs (in %) in ML estimates for the indicated parameter (standard tobit model) and removed cases with the measles vaccine data.*

Removed case(s)	Coefficients			
	β_0	β_1	β_2	β_3
{328}	81.95	83.74	10.17	42.84
{329}	91.43	76.11	7.46	10.42
{330}	88.40	154.93	0.59	54.34
{328,329}	70.08	49.86	22.52	12.92
{328,330}	67.01	129.03	29.47	78.04
{329,330}	76.75	121.38	31.72	24.42
{328,329,330}	55.15	95.15	62.13	48.13

2.5.6 Analysis under the tobit-BS model

We consider now an analysis under the tobit-BS model defined by equation (2.14) where $\varepsilon_i \overset{\text{IID}}{\sim}$ log-BS$(\alpha, 0)$ for the measles vaccine data. Table 2.4 reports the ML estimates, SEs and p-values of the t-test for tobit-BS model parameters, computed by the BFGS method. In addition, the AIC and BIC criteria are reported. From Table 2.4, we observe that the parameters β_1, β_2 and β_3 are not statistically significant at a 5% level. We also observe that the tobit-BS model provides a better fit compared to the standard tobit model based on the values of AIC and BIC. Figure 2.5 (left) shows the QQ plot with simulated envelope of the GCS residual for the tobit-BS model. From this figure, note that the GCS residual presents an excellent agreement with the EXP(1) distribution.

Table 2.4: *ML estimates (with SE in parentheses) and model selection measures for the indicated models.*

Model	θ	Estimates	p-value	AIC	BIC
Normal	$\log(\sigma)$	0.94512(0.04688)	<0.0001	1299.27	1318.27
	β_0	0.59700(0.28878)	0.0387		
	β_1	0.22507(0.29733)	0.4491		
	β_2	−0.22827(0.29580)	0.4403		
	β_3	0.27128(0.29675)	0.3606		
BS	α	1.54569(0.04819)	<0.0001	1168.60	1187.55
	β_0	−0.91057(0.10487)	<0.0001		
	β_1	0.18889(0.11112)	0.08915		
	β_2	0.07371(0.10979)	0.50203		
	β_3	0.12135(0.11030)	0.27125		

Note, from table 2.4 that the normal model suggests that all covariables are not signifficant, however the tobit-BS model suggests that the type of vaccine used is slightly significant, the results indicate that the receivers of the Edmonston-Zagreb vaccine have higher concentration levels of the strain $\exp(0.22507) - 1 = 0.25241\%$.

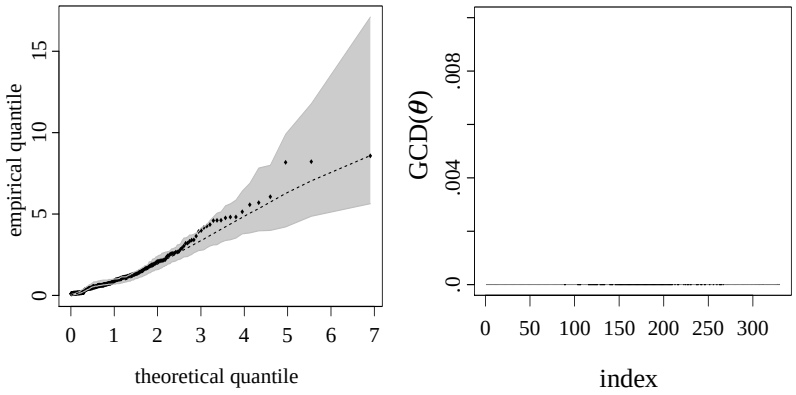

Figure 2.5: *QQ plot and its envelope for the GCS residual (left) and index plot of the GCD (right) for the tobit-BS model with measles vaccine data.*

Index plots of C_i under the case-weight perturbation are displayed in Figure 2.6 (left). This figure detects the cases #328 and #330 as potential influential observations for the tobit-BS model. Figure 2.6 (right) presents the GCD index plot, which shows no evidence of influential cases.

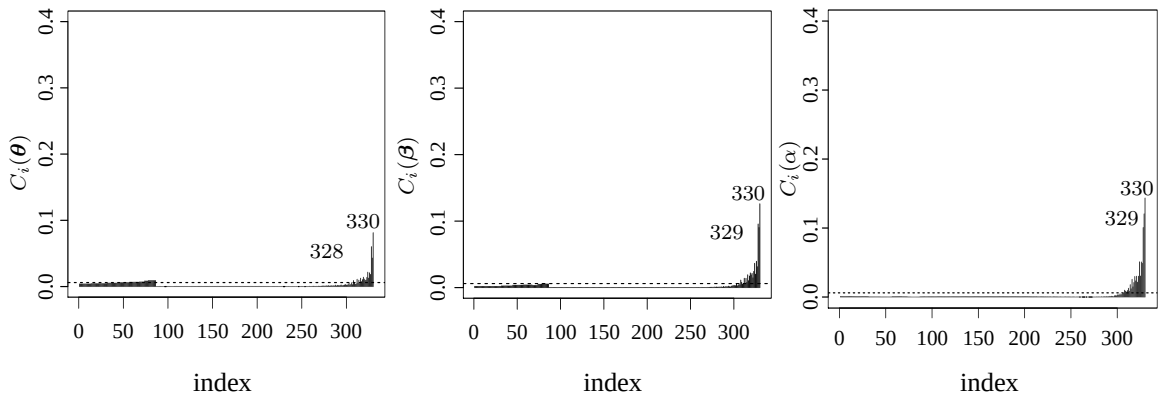

Figure 2.6: *Index plots of $C_i(\boldsymbol{\theta})$ (left), $C_i(\boldsymbol{\beta})$ (center) and $C_i(\alpha)$ (right) under the case-weight perturbation in the tobit-BS model with the measles vaccine data.*

Table 2.5 reports the RCs in the parameter estimates of the tobit-BS model, where the removed cases include observations detected as potential influential ones for both of the standard tobit and tobit-BS models. In terms of model sensitivity, we observe a less pronounced influence of the removed cases on the ML estimates of the tobit-BS model parameters (Table 2.5) than on the ML estimates of the standard tobit model parameters (Table 2.3). This result can be interpreted as a robustness of the tobit-BS model to atypical observations. We observe that β_1 is not significant at 10% after eliminating the observation #330. However, in general, the diagnostic measures identify potentially influential cases do not alter the inference of the model. Thus, we can conclude that the tobit-BS model is better than the standard tobit model for the data set considered in this application.

Table 2.5: *RCs (in %) in ML estimates for the indicated parameter (tobit-BS model) and removed cases with the measles vaccine data.*

Removed case(s)	Coefficients			
	β_0	β_1	β_2	β_3
{328}	4.28	4.23	28.31	16.11
{329}	2.13	5.83	27.89	18.61
{330}	2.34	18.40	29.52	18.93
{328,329}	6.45	15.47	56.51	2.52
{328,330}	6.66	8.90	58.18	35.25
{329,330}	4.47	7.23	57.34	0.32
{328,329,330}	8.83	2.34	86.36	16.62

2.6 Concluding remarks

It is widely known that the normal distribution does not always model data adequately in different areas. This also occurs with the tobit model, which describes a censored response variable by covariates in diverse fields, such as economy and medicine. In this work, we have proposed a new tobit model based on the Birnbaum-Saunders distribution. This distribution has shown to be a good alternative to describe medical data. We have estimated the parameters of the tobit-Birnbaum-Saunders model with the maximum likelihood method and have taken advantage of its properties to propose asymptotic inference for such parameters. We also have developed global and local

influence tools for the tobit-BS model. As a global influence tool we have derived the generalized Cook distance and for local influence procedures we have derived the normal curvatures under some schemes of perturbation. We also have carried out a residual analysis. As a final task we have carried out an application of the new tobit-Birnbaum-Saunders model with real-world of measles vaccine data in Haiti, due to the presence of asymmetry and high kurtosis present in these data. The application has shown the superiority of the tobit-Birnbaum-Saunders model over the standard tobit model, providing a strong evidence that the Birnbaum-Saunders distribution is a good model for accommodating outliers. We carried to out this application with the help of the R software and of some its existing packages and of codes developed by us.

Chapter 3

A new mixture-based regression model for censored data

Most of the models designed for modeling censored data rely on the assumption of normality for the error distribution. It is well known that not all applications are well modeled by this distribution. Some efforts have relaxed the normality assumption by considering more flexible distributions such as t and log-alpha-power. Nevertheless, these models do not consider partial observations from the assumed distribution which potentially conducts to biased inference. We have explored a real data example of measles vaccine in Haiti and confirmed both the possibility of partial observation and asymmetry problems. Then, to solve such problems, we propose a mixture model consisting of the Birnbaum-Saunders and Bernoulli distributions. We discuss estimation of the model parameters based on the maximum likelihood method. We then carry out a Monte Carlo simulation study to evaluate the performance of the maximum likelihood estimators. We use the R software in all computations and the results favor the proposed methodology.

Keywords Censoring; Birnbaum-Saunders distribution; Maximum Likelihood Estimation; R software; mixture model; Monte Carlo simulation.

3.1 Introduction

The determination of antibody concentration by quantitative assays is a very important topic of research, because there is always a concentration value ($\mathcal{T}$) below which an exact measurement cannot be obtained regardless of the employed technique. Nevertheless, this antibody concentration value ($\mathcal{T}$) is a function of the associated assay. When left-censoring is present in data from an assay, the lower detection limit (LDL) can be used to substitute a value for the censored observation, namely, the value $\mathcal{T}$. In special, this substitution is applied in a safety and immunogenecity study of measles vaccine in Haiti presented by Moulton and Halsey (1995), an example explored in this paper. The data associated with the case-study described by Moulton and Halsey (1995) have both left-censoring and asymmetry, and tobit models can be used to provide estimates of interest; see Barros et al. (2010). However, the standard tobit model (TOBIN, 1958) has a strong assumption stating that the model errors follow a normal distribution, which is symmetric. Ignor-

ing the effect of asymmetry can be harmful and conducts to significantly biased estimates; some flexible tobit models, in terms of asymmetry, can be seen in Martínez-Flores, Bolfarine and Gómez (2013a) and Rocha, Arellano and Loschi (2015). Moreover, the standard tobit model, regardless the distribution, has a drawback concerning the fact that this model does not take into account a lower limit and that some observations may be below this lower limit, which is the case in the measles vaccine data. Finally, the standard tobit model can not cope with extreme heaviness of the censored part of the distribution, namely, the number of zeros surpasses the expected amount. In the measles vaccine data, 87 out of 370 values are below the LDL.

A two-part model for the situation of zero excess was considered by Cragg (1971). The Cragg model considers the possibility of having observations from the assumed distribution f and from the point mass distribution. In this model, the log-normal distribution was considered for the independent variable. The Cragg model, however, does not consider the existence of both a lower limit and some observations below this limit. Moulton and Halsey (1995) proposed a straightforward generalization of the two-part model, called Bernoulli/lognormal model, by considering the possibility of some limiting responses resulting from interval censoring associated with f. The generalized two-part model allows the possibility of a observation i, if located below $\mathcal{T}$ to be either a partial observation from f or a realization of the point mass distribution.

The Birnbaum-Saunders (BS) distribution is positively skewed and has a failure rate with upside-down bathtub shape and a close relation with the normal distribution; see Birnbaum and Saunders (1969), Johnson, Kotz and Balakrishnan (1995) and Leiva (2016). This distribution has two parameters, i.e., the shape and scale parameters, where the latter is also its median. Thus, the BS distribution can be seen as an analogue to the normal model, but in an asymmetrical setting, where the median is generally considered to be a better measure of central tendency than the mean. The BS distribution has been applied to biological, economic, engineering and environmental data, among others, and some of its recent applications are discussed in Qu and Xie (2011), Ferreira, Gomes and Leiva (2012), Li, Chen and Xie (2012), Saulo et al. (2013), Leiva et al. (2014b), Leiva et al. (2015), Leiva et al. (2015), Garcia-Papani et al. (2016) and Santos-Neto et al. (2016).

The main objective of this paper is to propose a regression model for censored data based on the mixture between the BS and Bernoulli distributions, that is, a censored continuous distribution and a point mass distribution located below the detection limit. The proposed model extends to the BS case the Moulton and Halsey (1995)'s Bernoulli/lognormal model. The secondary objectives of this paper are: (i) to develop inference for the Bernoulli/BS model based on the maximum likelihood (ML) method; (ii) to perform a Monte Carlo (MC) simulation study to evaluate the performance of the ML estimators; and (iii) to carry out an application of the proposed model by using the safety and immunogenecity study of measles vaccine in Haiti described by Moulton and Halsey (1995). The results indicate that the Bernoulli/BS model has the best performance in terms of fit.

The rest of the chapter proceeds as follows. Section 3.2 provides a background of mixture

models for censored data and of the BS distribution and its logarithmic transformation. Section 3.3 formulates the Bernoulli/BS model along with inference and estimation based on the ML method. Section 3.4 presents a Monte Carlo simulation study. Section 3.5 carries out an empirical application with real world data. Finally, Section 3.6 presents some concluding remarks.

3.2 Background

3.2.1 The standard tobit model

Consider a sample of size n, $\boldsymbol{Y} = (Y_1, \ldots, Y_m, Y_{m+1}, \ldots, Y_n)^\top$, namely, independent (IND) random variables but not independent identically distributed (IID) necessarily. Consider also that this sample includes m censored data to the left and $n - m$ observed (complete or uncensored) data. Thus, such censoring scheme can be visualized under a regression setting with a censored response Y^*, which is a (unobserved) latent variable. Hence, the m censored data (unobserved) correspond to the values of Y^* less than or equal to a threshold point y_0 (censoring to the left), so that all of these data take the value y_0. The other $n - m$ data (observed) are related to values of Y^* greater than y_0, which can be described by a linear regression structure of the type $\boldsymbol{x}_i^\top \boldsymbol{\beta}$. Then, the normal tobit model with censored response to the left can be formulated as

$$Y_i = \begin{cases} y_o, & \text{if } Y_i^* \leq y_0, \quad i = 1, \ldots, m; \\ \boldsymbol{x}_i^\top \boldsymbol{\beta} + \varepsilon_i, & \text{if } Y_i^* > y_0, \quad i = m+1, \ldots, n; \end{cases} \tag{3.1}$$

where $\varepsilon_i \overset{\text{IID}}{\sim} \text{N}(0, \sigma^2)$ is the model error term, $\boldsymbol{\beta}$ is a vector of regression coefficients corresponding to unknown parameters to be estimated, and $\boldsymbol{x}_i$ is a vector containing the values of explanatory variables (covariates). Note that y_0 given in (3.1) is a prefixed limiting value that does the response of the regression model defined in (3.1) to be limited (or censored); see Tobin (1958).

3.2.2 Mixture models for censored data

Cragg (1971) proposed a simple zero excess model described by

$$g(y_i) = p_i I_i + (1 - p_i) f(y_i)(1 - I_i), \tag{3.2}$$

where p_i is a probability that determines the relative contribution made by one point distribution to the mixture distribution as a whole, $f(\cdot)$ is the probability density function (PDF) of Y and I_i receives the value 0 if $y_i > 0$ and 1 if $y_i \leq 0$. Note that the PDF in (3.2) is not defined, thus the model can be freely modified. This model however does not consider the possibility of a lower limit.

Moulton and Halsey (1995) extended the model in Cragg (1971) to a generalized two-

part model by considering the possibility of limiting responses coming from interval censoring in $f(\cdot)$. In the extended model, a zero can be either a realization of a point mass distribution or the result of a censoring in $f(\cdot)$. The generalized model incorporates an intermediary possibility that an observed zero can be a realization of the point mass or a partial observation of $f(\cdot)$ with an unknown critical value lying in some point between $(0, \mathcal{T})$. The PDF for the extended two-part model is given by

$$g(y_i) = [p_i + (1 - p_i)\, F(\mathcal{T})]\, \mathrm{I}_i + (1 - p_i) f(y_i)(1 - \mathrm{I}_i), \tag{3.3}$$

where $F(\cdot)$ is the cumulative distribution function associated with $f(\cdot)$. A large family of mixture model can be created by changing $f(\cdot)$ and the link function. Martínez-Flores, Bolfarine and Gómez (2013b) noted that, in Equation (3.3), it is assumed that just a proportion τ of censored observations comes from the assumed distribution, with the rest $(1-\tau)$ coming from a population of low responders located at, or below $\mathcal{T}$. This mixture is modeled by assuming a Bernoulli random variable D with

$$pr(D = 1) = \tau = 1 - p. \tag{3.4}$$

Chai and Bailey (2008) listed the use of many mixture models such as probit/truncated-normal, logit/lognormal, logit/log-gamma and probit/log-skew-normal with applications in biology, economics and agriculture. Martínez-Flores, Bolfarine and Gómez (2013b) considered a new approach for asymmetric data with positive support based on Equation (3.3) called the Bernoulli/log-power-normal (Bernoulli/LPN) model. Note that if $\varepsilon_i \overset{\text{IID}}{\sim} \mathrm{N}(0, \sigma^2)$ and $p_i = 0$, $i = 1, \ldots, n$, in Equation (3.3) then the generalized two-part model becomes the standard tobit model .

3.2.3 The BS and log-BS distributions

Let T be a random variable following a BS distribution with shape parameter α and scale parameter σ, with the following notation $T \sim \mathrm{BS}(\alpha, \sigma)$. Then,

- the cumulative distribution function (CDF) of T is given by

$$F_T(t; \alpha, \sigma) = \Phi\left(\frac{1}{\alpha}\left(\sqrt{t/\sigma} - \sqrt{\sigma/t}\right)\right), \quad t > 0, \alpha > 0, \sigma > 0; \tag{3.5}$$

- the PDF of T obtained from (3.5) is expressed as

$$f_T(t; \alpha, \sigma) = \frac{1}{2\alpha}\left(\sqrt{1/\sigma t} + \sqrt{\sigma/t^{\frac{3}{2}}}\right)\phi\left(\frac{1}{\alpha}\left(\sqrt{t/\sigma} - \sqrt{\sigma/t}\right)\right), \quad t > 0, \alpha > 0, \sigma > 0, \tag{3.6}$$

where ϕ is the standard normal PDF. Alternatively, (3.5) can be expressed as

$$f_T(t;\alpha,\sigma) = \frac{\exp(\alpha^{-2})}{2\alpha\sqrt{2\pi\sigma}} \exp\left(-\frac{1}{2\alpha^2}\left(\frac{t}{\sigma}+\frac{\sigma}{t}\right)\right) t^{-\frac{3}{2}}(t+\sigma), \quad t>0, \alpha>0, \sigma>0; \tag{3.7}$$

- the results provided in (3.5) and (3.6) may be obtained from the transformation theorem of random variables by using

$$T = \sigma\left(\alpha Z/2 + \sqrt{(\alpha Z/2)^2+1}\right)^2, \tag{3.8}$$

where $Z \sim \mathrm{N}(0,1)$.

- from (3.8), it may be verified that a continuous random variable T has a BS distribution with parameters $\alpha>0$ and $\sigma>0$, if and only if $Z = (1/\alpha)(\sqrt{T/\sigma} - \sqrt{\sigma/T}) \sim \mathrm{N}(0,1)$. Some properties of the BS distribution are presented as follows. If $T \sim \mathrm{BS}(\alpha,\sigma)$, then: (i) $\mathrm{E}(T) = \sigma(1+\alpha^2/2)$ and $\mathrm{Var}(T) = (\alpha\sigma)^2(1+5\alpha^2/4)$; (ii) for $b>0$, $bT \sim \mathrm{BS}(\alpha, b\sigma)$, which means that the BS distribution is closed under scalar multiplication (proportionality); (iii) $1/T \sim \mathrm{BS}(\alpha, 1/\sigma)$, implying that the BS distribution is closed under reciprocation; (iv) the median of the distribution of T is σ, which can be directly obtained when $q=0.5$ from its quantile function given by

$$t(q;\alpha,\sigma) = F_T^{-1}(q;\alpha,\sigma) = \sigma\left(\frac{\alpha z(q)}{2} + \sqrt{\left(\frac{\alpha z(q)}{2}\right)^2+1}\right)^2, \quad 0<q<1,$$

where $z(q)$ is the standard normal quantile function; and (v) the BS distribution is positively skewed as α increases and approximately symmetrical around σ as α goes to zero; see Figure 3.1(left). Properties of proportionality and reciprocation given above in (ii) and (iii) are easily verified by using once again the mentioned transformation theorem;

- the survival function (SF) and hazard rate (HR) of $T \sim \mathrm{BS}(\alpha,\sigma)$ are given respectively by

$$\begin{aligned} S_T(t;\alpha,\sigma) &= \Phi\left(-\frac{1}{\alpha}\left(\sqrt{\frac{t}{\sigma}} - \sqrt{\frac{\sigma}{t}}\right)\right), \\ h_T(t;\alpha,\sigma) &= \frac{\phi\left(\frac{1}{\alpha}\left(\sqrt{\frac{t}{\sigma}} - \sqrt{\frac{\sigma}{t}}\right)\right)\frac{t^{-3/2}(t+\sigma)}{2\alpha\sigma^{1/2}}}{\Phi\left(-\frac{1}{\alpha}\left(\sqrt{\frac{t}{\sigma}} - \sqrt{\frac{\sigma}{t}}\right)\right)}, \quad t>0. \end{aligned}$$

The logarithmic version (log-BS) of the BS distribution is used when data modeling is needed. A random variable Y has a log-BS distribution with shape ($\alpha>0$) and location ($\mu\in\mathbb{R}$) parameter, which is denoted by log-BS(α,μ), if and only if $Z = (2/\alpha)\sinh((Y-\mu)/2) \sim \mathrm{N}(0,1)$.

Then, the CDF of Y is given by

$$F_Y(y;\alpha,\mu) = \Phi\left(\frac{2}{\alpha}\text{sinh}\left(\frac{y-\mu}{2}\right)\right), \quad y \in \mathbb{R}, \mu \in \mathbb{R}, \alpha > 0. \tag{3.9}$$

Consequently, from (3.9), the PDF of Y is obtained as

$$f_Y(y;\alpha,\mu) = \frac{1}{\alpha\sqrt{2\pi}}\cosh\left(\frac{y-\mu}{2}\right)\exp\left(-\frac{2}{\alpha^2}\text{sinh}^2\left(\frac{y-\mu}{2}\right)\right), \quad y \in \mathbb{R}, \mu \in \mathbb{R}, \alpha > 0, \tag{3.10}$$

whereas that the logarithm of the PDF given in (3.10) is expressed as

$$\log(f_Y(y;\alpha,\mu)) = -\log(2) - \frac{\log(2\pi)}{2} + \log\left(\frac{2}{\alpha}\cosh\left(\frac{y-\mu}{2}\right)\right) - \frac{2}{\alpha^2}\text{sinh}^2\left(\frac{y-\mu}{2}\right), \quad y \in \mathbb{R}.$$

Important properties of the log-BS distribution are as follows. If $Y \sim$ log-BS(α,μ), then: (i) $T = \exp(Y) \sim$ BS(α,σ), which means that the log-BS PDF given in (3.10) can be obtained from the standard normal PDF as in (3.9) or from the BS PDF defined in (3.7); (ii) $\text{E}(Y) = \mu$; (iii) there is no closed form for the variance of Y, but based upon an asymptotic approximation for the moment generating function of the log-BS distribution, it follows that, if $\alpha \to 0$, then $\text{Var}(T) = \alpha^2 - \alpha^4/4$, whereas that, if $\alpha \to \infty$, then $\text{Var}(T) = 4(\log^2(\sqrt{2}\alpha) + 2 - 2\log(\sqrt{2}\alpha))$; (iv) if $X = \pm Y + d$, then $X \sim$ log-BS$(\alpha, \pm\mu + d)$; and (v) the log-BS distribution is symmetric around μ, unimodal for $\alpha \leq 2$ and bimodal for $\alpha > 2$; see Figure 3.1(right).

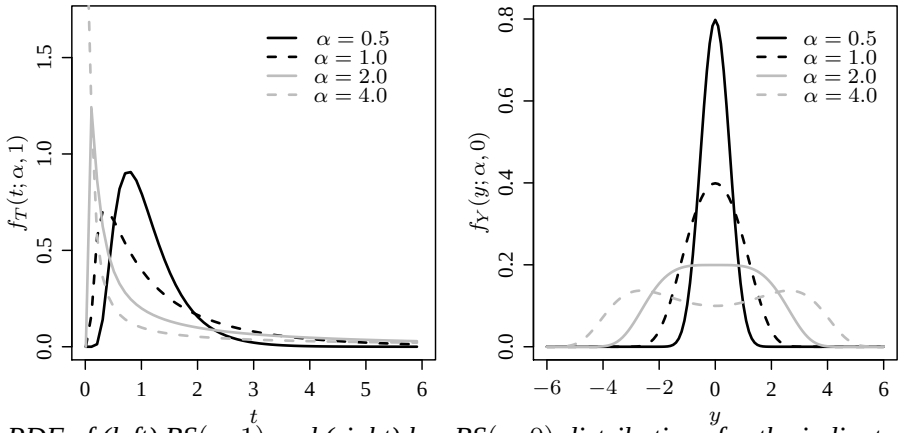

Figure 3.1: *PDF of (left) BS$(\alpha,1)$ and (right) log-BS$(\alpha,0)$ distributions for the indicated value of α.*

3.3 The Bernoulli/BS mixture model

3.3.1 Formulation

We propose a mixture model between the Bernoulli and BS distributions (Bernoulli/BS) by assuming that $f(\cdot)$ in Equation (3.3) is a log-BS PDF as in (3.10), with a logit link for p_i. Then,

Equation (3.3) can be rewritten as follows

$$g(y_i) = [p + (1-p)\,\Phi\left(\zeta_{i2}^{\mathrm{c}}\right)]\,\mathrm{I_i} + (1-p)\left[\frac{\mathrm{c}_1}{\alpha}\cosh\left(\frac{y-\mu_1}{2}\right)\exp\left(-\frac{2}{\alpha^2}\sinh^2\left(\frac{y-\mu_1}{2}\right)\right)\right](1-\mathrm{I_i}), \tag{3.11}$$

where $\mathrm{c}_1 = 1/\sqrt{2\pi}$, $\mu_1 = \boldsymbol{x}_{(1)i}^{\top}\boldsymbol{\beta}_{(1)}$, $\zeta_{i2}^{\mathrm{c}} = (2/\alpha)\sinh(y_0 - \boldsymbol{x}_{(1)i}^{\top}\boldsymbol{\beta}_{(1)})/2$,

$$\mathrm{I}_i = \begin{cases} 1 & \text{if } y \le y_0 \\ 0 & \text{if } y > y_0, \end{cases} \tag{3.12}$$

and $\Phi(\cdot)$ is the CDF of the standard normal distribution, with $\boldsymbol{x}_{(1)}$ being the covariates associated with $\boldsymbol{\beta}_{(1)}$. We assume a logit link for the random variable D defined in Equation (3.4), thus it is possible to include covariates as follows

$$\operatorname{logit}\left[pr\left(D = 1|\boldsymbol{x}_{(2)}\right)\right] = \mathbf{x}_{(2)}{}^{\top}\boldsymbol{\beta}_{(2)}, \tag{3.13}$$

where $\boldsymbol{x}_{(2)}$ are the covariates associated with $\boldsymbol{\beta}_{(2)}$. The formulation of the logit link defined in Equation (3.13) becomes,

$$\tau_i = 1 - p_i = \frac{\exp\left(\mathbf{x}_{(2)}{}^{\top}\boldsymbol{\beta}_{(2)}\right)}{1 + \exp\left(\mathbf{x}_{(2)}{}^{\top}\boldsymbol{\beta}_{(2)}\right)}, \tag{3.14}$$

where the vector $\mathbf{x}_{(2)}$ has a dimension q, which can be, and usually is, different from the dimension of the vector $\mathbf{x}_{(1)}$.

Combining the Equations (3.11) and (3.14) we obtain the individual contribution to the likelihood function of the mixture Bernoulli/BS model that is given in Equation (3.15),

$$\begin{aligned} L_i(\boldsymbol{\theta}) = &\left[1 + \frac{\exp(\mathbf{x}_{(2)}^{\top}\boldsymbol{\beta}_{(2)})}{1 + \exp(\mathbf{x}_{(2)}^{\top}\boldsymbol{\beta}_{(2)})}\left[\Phi\left(\zeta_{i2}^{\mathrm{c}}\right) - 1\right]\right]^{\mathrm{I_i}} \times \\ &\left[\frac{\exp(\mathbf{x}_{(2)}^{\top}\boldsymbol{\beta}_{(2)})}{1 + \exp(\mathbf{x}_{(2)}^{\top}\boldsymbol{\beta}_{(2)})}\left(\frac{\mathrm{c}_1}{\alpha}\cosh\left(\frac{y-\mu_1}{2}\right)\exp\left(-\frac{2}{\alpha^2}\sinh^2\left(\frac{y-\mu_1}{2}\right)\right)\right)\right]^{1-\mathrm{I_i}}, \end{aligned} \tag{3.15}$$

where c_1, μ_1 and ζ_{i2}^{c} are defined in Equation (3.11), $\boldsymbol{\theta} = (\alpha, \boldsymbol{\beta}_{(1)}^{\top}, \boldsymbol{\beta}_{(2)}^{\top})^{\top}$ and $\mathrm{I_i}$ is defined in Equation (3.12).

The log-likelihood function for the mixture Bernoulli/BS is obtained by taking the logarithm of Equation (3.15), which gives

$$\begin{aligned} \ell(\boldsymbol{\theta}) = &-(n-m)\log(2) - (n-m)\frac{\log(2\pi)}{2} \\ &+ \sum_{i=1}^{m}\mathrm{I_i}\left[\log\left(1 + \exp(\mathbf{x}_{(2)}^{\top}\boldsymbol{\beta}_{(2)})\left[\Phi\left(\zeta_{i2}^{\mathrm{c}}\right) - 1\right]\right) - \log\left(1 + \exp\left(\mathbf{x}_{(2)}^{\top}\boldsymbol{\beta}_{(2)}\right)\right)\right] \\ &+ \sum_{m+1}^{n}(1 - \mathrm{I_i})\left[\mathbf{x}_{(2)}^{\top}\boldsymbol{\beta}_{(2)} + \log\left(\zeta_{i1}\right) - \frac{1}{2}\zeta_{i2}^2 - \log\left(1 + \exp(\mathbf{x}_{(2)}^{\top}\boldsymbol{\beta}_{(2)})\right)\right], \end{aligned} \tag{3.16}$$

where ζ_{i2}^{c} is defined in Equation (3.11) and ζ_{i1} and ζ_{i2} are defined in Equation (3.17),

$$\zeta_{i1} = \tfrac{2}{\alpha}\cosh\left(\tfrac{y_i - \boldsymbol{x}_{(1)i}^{\top}\boldsymbol{\beta}_{(1)}}{2}\right), \quad \zeta_{i2} = \tfrac{2}{\alpha}\sinh\left(\tfrac{y_i - \boldsymbol{x}_{(1)i}^{\top}\boldsymbol{\beta}_{(1)}}{2}\right). \tag{3.17}$$

To obtain the ML estimators, it is necessary to maximize the log-likelihood function given in (3.16). The score vector is $\dot{\boldsymbol{\ell}} = \partial\ell(\boldsymbol{\theta})/\partial\boldsymbol{\theta} = \left(\dot{\ell}_{\alpha}, \dot{\boldsymbol{\ell}}_{\boldsymbol{\beta}_{(1)}}^{\top}, \dot{\boldsymbol{\ell}}_{\boldsymbol{\beta}_{(2)}}^{\top}\right)^{\top}$, which contains the first partial derivatives of (3.16), where

$$\begin{aligned}
\dot{\ell}_{\alpha} &= \begin{cases} -\frac{1}{\alpha}\left[\frac{\exp(x_{(2)}\boldsymbol{\beta}_{(2)})\phi(\zeta_{i2}^{\mathrm{c}})\zeta_{i2}^{\mathrm{c}}}{1+\exp(x_{(2)}\boldsymbol{\beta}_{(2)})\left[\Phi(\zeta_{i2}^{\mathrm{c}})-1\right]}\right], & i = 1,\ldots,m; \\ \frac{1}{\alpha}\left[\zeta_{i2}^{2} - 1\right], & i = m+1,\ldots,n; \end{cases} \\
\dot{\boldsymbol{\ell}}_{\boldsymbol{\beta}_{(1)}} &= \begin{cases} -\frac{x_{(1)}}{2}\left[\frac{\exp(x_{(2)}\boldsymbol{\beta}_{(2)})\phi(\zeta_{i2}^{\mathrm{c}})\zeta_{i1}^{\mathrm{c}}}{1+\exp(x_{(2)}\boldsymbol{\beta}_{(2)})\left[\Phi(\zeta_{i2}^{\mathrm{c}})-1\right]}\right], & i = 1,\ldots,m; \\ \frac{x_{(1)}}{2}\left[\zeta_{i1}\zeta_{i2} - \frac{\zeta_{i2}}{\zeta_{i1}}\right], & i = m+1,\ldots,n; \end{cases} \\
\dot{\boldsymbol{\ell}}_{\boldsymbol{\beta}_{(2)}} &= \begin{cases} x_{(2)}\left[\frac{\exp(x_{(2)}\boldsymbol{\beta}_{(2)})\left[\Phi(\zeta_{i2}^{\mathrm{c}})-1\right]}{1+\exp(x_{(2)}\boldsymbol{\beta}_{(2)})\left[\Phi(\zeta_{i2}^{\mathrm{c}})-1\right]} - \tau\right], & i = 1,\ldots,m. \\ x_{(2)}\left[1-\tau\right], & i = m+1,\ldots,n; \end{cases}
\end{aligned} \tag{3.18}$$

The ML estimator of $\boldsymbol{\theta}$ is obtained equating (3.18) to zero. Note that $\dot{\boldsymbol{\ell}}_{\boldsymbol{\beta}_{(2)}}$ can be solved analytically however, the system of equations defined by $\dot{\ell}_{\alpha} = 0$, $\dot{\boldsymbol{\ell}}_{\boldsymbol{\beta}_{(1)}} = 0$ and $\dot{\boldsymbol{\ell}}_{\boldsymbol{\beta}_{(2)}} = 0$ does not have an analytic solution. In this paper, we solve them by an iterative procedure for non-linear optimization, i.e., the Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-Newton method; see Leiva et al. (2007).

3.3.2 Inference

Assuming that the regularity conditions defined in Cox and Hinkley (1974) are satisfied, the ML estimators $\widehat{\alpha}$, $\widehat{\boldsymbol{\beta}}_{(1)}$ and $\widehat{\boldsymbol{\beta}}_{(2)}$ are consistent and follow a multivariate normal joint asymptotic distribution. This distribution has an asymptotic mean vector with elements α, $\boldsymbol{\beta}_{(1)}$, and $\boldsymbol{\beta}_{(2)}$ and have an asymptotic covariance matrix equal to $\mathcal{J}(\boldsymbol{\theta})^{-1}$, which can be approximated by the expected Fisher information matrix. Therefore, as $n \to \infty$, we have that

$$\sqrt{n}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \quad \xrightarrow{\mathrm{d}} \quad \mathrm{N}_{p+1}\left(\mathbf{0}_{p+1}, \mathcal{J}(\boldsymbol{\theta})^{-1}\right), \tag{3.19}$$

where $\mathcal{J}(\boldsymbol{\theta}) = \lim_{n\to\infty}(1/n)\mathcal{I}(\boldsymbol{\theta})$, with $\mathcal{I}(\boldsymbol{\theta})$ being the expected Fisher information matrix. In addition, $\xrightarrow{\mathrm{d}}$ means convergence in distribution to and $\mathbf{0}_{p+1}$ is a $p \times 1$ vector of zeros. Note that $\widehat{\mathcal{I}}(\boldsymbol{\theta})^{-1}$ is a consistent estimator of the asymptotic variance-covariance matrix of $\widehat{\boldsymbol{\theta}}$, $\mathcal{J}(\boldsymbol{\theta})^{-1}$ say. In practice, one may approximate the expected Fisher information matrix by its observed version Efron and Hinkley (1978), whereas the elements of the diagonal of the inverse of the observed information matrix can be used to approximate the corresponding standard errors (SEs). The ob-

served Fisher information matrix is obtained from the Hessian matrix, which contains the second partial derivatives of (3.19) and given by

$$\ddot{\boldsymbol{\ell}} = \begin{pmatrix} \operatorname{tr}(\boldsymbol{G}) & \boldsymbol{k}_{(1)}^{\top}\boldsymbol{X}_{(1)} & \boldsymbol{k}_{(2)}{}^{\top}\boldsymbol{X}_{(2)} \\ \boldsymbol{X}_{(1)}^{\top}\boldsymbol{k}_{(1)} & \boldsymbol{X}_{(1)}^{\top}\boldsymbol{V}_{(1)}\boldsymbol{X}_{(1)} & \boldsymbol{X}_{(1)}^{\top}\boldsymbol{D}\boldsymbol{X}_{(2)} \\ \boldsymbol{X}_{(2)}^{\top}\boldsymbol{k}_{(2)} & \boldsymbol{X}_{(2)}^{\top}\boldsymbol{D}\boldsymbol{X}_{(1)} & \boldsymbol{X}_{(2)}^{\top}\boldsymbol{V}_{(2)}\boldsymbol{X}_{(2)} \end{pmatrix}, \tag{3.20}$$

where $\boldsymbol{V}_{(1)} = \operatorname{diag}\{v_{(1)1}(\boldsymbol{\theta}), v_{(1)2}(\boldsymbol{\theta}), v_{(1)3}(\boldsymbol{\theta}), \ldots, v_{(1)n}(\boldsymbol{\theta})\}$, $\boldsymbol{V}_{(2)} = \operatorname{diag}\{v_{(2)1}(\boldsymbol{\theta}), v_{(2)2}(\boldsymbol{\theta}), v_{(2)3}(\boldsymbol{\theta}), \ldots, v_{(2)n}(\boldsymbol{\theta})\}$, $\boldsymbol{k}_{(1)} = (k_{(1)1}(\boldsymbol{\theta}), k_{(1)2}(\boldsymbol{\theta}), k_{(1)3}(\boldsymbol{\theta}), \ldots, k_{(1)n}(\boldsymbol{\theta}))^{\top}$, $\boldsymbol{k}_{(2)} = (k_{(2)1}(\boldsymbol{\theta}), k_{(2)2}(\boldsymbol{\theta}), k_{(2)3}(\boldsymbol{\theta}), \ldots, k_{(2)n}(\boldsymbol{\theta}))^{\top}$, $\boldsymbol{D} = \operatorname{diag}\{d_1(\boldsymbol{\theta}), d_2(\boldsymbol{\theta}), d_3(\boldsymbol{\theta}), \ldots, d_n(\boldsymbol{\theta})\}$ and $\boldsymbol{G} = \operatorname{diag}\{g_1(\boldsymbol{\theta}), g_2(\boldsymbol{\theta}), g_3(\boldsymbol{\theta}), \ldots, g_n(\boldsymbol{\theta})\}$, with

$$g_i(\boldsymbol{\theta}) = \begin{cases} \frac{1}{\alpha^2}\left[\frac{\exp(x_{(2)}\boldsymbol{\beta}_{(2)})\phi(\zeta_{i2}^{c})\zeta_{i1}^{c}}{\left(1+\exp(x_{(2)}\boldsymbol{\beta}_{(2)})\left[\Phi(\zeta_{i2}^{c})-1\right]\right)} + \frac{\zeta_{i2}^{2c}\phi(\zeta_{i2}^{c})\exp(x_{(2)}\boldsymbol{\beta}_{(2)})}{\left(1+\exp(x_{(2)}\boldsymbol{\beta}_{(2)})\left[\Phi(\zeta_{i2}^{c})-1\right]\right)}\right] \\ -\frac{1}{\alpha^2}\left[\frac{\phi^2(\zeta_{i2}^{c})\zeta_{i2}^{2c}\exp(2x_{(2)}\boldsymbol{\beta}_{(2)})}{\left(1+\exp(x_{(2)}\boldsymbol{\beta}_{(2)})\left[\Phi(\zeta_{i2}^{c})-1\right]\right)}\right], & i = 1,\ldots,m; \\ -\frac{1}{\alpha^2}\left[3\zeta_{i2}^{2}-1\right], & i = m+1,\ldots,n; \end{cases}$$

$$k_{(1)i}(\boldsymbol{\theta}) = \begin{cases} \frac{\exp(x_{(2)}\boldsymbol{\beta}_{(2)})}{2\alpha}\left[\frac{\zeta_{i1}^{c}\phi\left(\zeta_{i2}^{c}\right)+\zeta_{i1}^{c}\phi\left(\zeta_{i2}^{c}\right)\zeta_{i2}^{2c}}{\left(1+\exp(x_{(2)}\boldsymbol{\beta}_{(2)})\left[\Phi\left(\zeta_{i2}^{c}\right)-1\right]\right)} + \frac{\exp(x_{(2)}\boldsymbol{\beta}_{(2)})\phi^2\left(\zeta_{i2}^{c}\right)\zeta_{i1}^{c}\zeta_{i2}^{c}}{\left(1+\exp(x_{(2)}\boldsymbol{\beta}_{(2)})\left[\Phi(\zeta_{i2}^{c})-1\right]\right)^2}\right], & i = 1,\ldots,m; \\ \frac{1}{\alpha}\left[\zeta_{i1}\zeta_{i2}\right], & i = m+1,\ldots,n; \end{cases}$$

$$k_{(2)i}(\boldsymbol{\theta}) = \begin{cases} -\frac{\exp(x_{(2)}\boldsymbol{\beta}_{(2)})\phi(\zeta_{i2}^{c})\zeta_{i2}^{c}}{\alpha}\left[\frac{\left(1+\exp(x_{(2)}\boldsymbol{\beta}_{(2)})\left[\Phi(\zeta_{i2}^{c})-1\right]\right)-1}{\left(1+\exp(x_{(2)}\boldsymbol{\beta}_{(2)})\left[\Phi(\zeta_{i2}^{c})-1\right]\right)^2}\right], & i = 1,\ldots,m; \\ 0, & i = m+1,\ldots,n; \end{cases}$$

$$d_i(\boldsymbol{\theta}) = \begin{cases} -\frac{1}{2}\left[\frac{\exp(x_{(2)}\boldsymbol{\beta}_{(2)})\phi(\zeta_{i2}^{c})\zeta_{i1}^{c}}{\left(1+\exp(x_{(2)}\boldsymbol{\beta}_{(2)})\left[\Phi(\zeta_{i2}^{c})-1\right]\right)} - \frac{\exp(2x_{(2)}\boldsymbol{\beta}_{(2)})\phi(\zeta_{i2}^{c})\zeta_{i1}^{c}\left[\Phi(\zeta_{i2}^{c})-1\right]}{\left(1+\exp(x_{(2)}\boldsymbol{\beta}_{(2)})\left[\Phi(\zeta_{i2}^{c})-1\right]\right)^2}\right], & i = 1,\ldots,m; \\ 0, & i = m+1,\ldots,n; \end{cases}$$

$$v_{(1)i}(\boldsymbol{\theta}) = \begin{cases} \frac{\exp(x_{(2)}\boldsymbol{\beta}_{(2)})}{4}\left[\frac{-\phi(\zeta_{i2}^{c})\zeta_{i2}^{c}+\zeta_{i1}^{2c}\zeta_{i2}^{c}\phi(\zeta_{i2}^{c})}{\left(1+\exp(x_{(2)}\boldsymbol{\beta}_{(2)})\left[\Phi(\zeta_{i2}^{c})-1\right]\right)} + \frac{\phi^2(\zeta_{i2}^{c})\zeta_{i1}^{2c}\exp(x_{(2)}\boldsymbol{\beta}_{(2)})}{\left(1+\exp(x_{(2)}\boldsymbol{\beta}_{(2)})\left[\Phi(\zeta_{i2}^{c})-1\right]\right)^2}\right], & i = 1,\ldots,m; \\ \frac{1}{4}\left[1-(\zeta_{i2}^{2}/\zeta_{i1}^{2})-\zeta_{i1}^{2}-\zeta_{i2}^{2}\right], & i = m+1,\ldots,n; \end{cases}$$

$$v_{(2)i}(\boldsymbol{\theta}) = \begin{cases} \frac{\exp\left(\mathbf{x}_{(2)}^{\top}\boldsymbol{\beta}_{(2)}\right)\left[\Phi\left(\zeta_{i2}^{c}\right)-1\right]}{\left(1+\exp(\mathbf{x}_{(2)}^{\top}\boldsymbol{\beta}_{(2)})\left[\Phi\left(\zeta_{i2}^{c}\right)-1\right]\right)} - \frac{\left(\exp\left(\mathbf{x}_{(2)}^{\top}\boldsymbol{\beta}_{(2)}\right)\left[\Phi\left(\zeta_{i2}^{c}\right)-1\right]\right)^2}{\left(1+\exp\left(\mathbf{x}_{(2)}^{\top}\boldsymbol{\beta}_{(2)}\right)\left[\Phi\left(\zeta_{i2}^{c}\right)-1\right]\right)^2} - \tau+\tau^2, & i = 1,\ldots,m; \\ \tau-\tau^2, & i = m+1,\ldots,n. \end{cases}$$

3.4 Simulation study

We present a Monte Carlo simulation study with 5000 replications that intends to reveal the performance of the maximum likelihood estimators for the parameters of the Bernoulli/BS model. The sample sizes considered are $n = 100$, $n = 300$ and $n = 500$, with parameters $\alpha = 0.1, 0.5, 1, 2, 4$, $\beta_{(1)} = (0.2, 0.5)^{\top}$ and $\beta_{(2)} = (1, 2)^{\top}$. We consider one covariate X, where $X \sim$

Uniform$(0,1)$. The generated values for the response variable were obtained as follows

$$t_i = \begin{cases} 1, & \text{with probability } 1 - \frac{\exp(\beta_{(2)0}+\beta_{(2)1}x_i)}{1+\exp(\beta_{(2)0}+\beta_{(2)1}x_i)}, \\ \exp\left(\beta_{(1)0} + \beta_{(1)1}x_i\right)\delta_i, & \text{with probability } \frac{\exp(\beta_{(2)0}+\beta_{(2)1}x_i)}{1+\exp(\beta_{(2)0}+\beta_{(2)1}x_i)}, \end{cases}$$

where $\delta \sim \text{BS}(\alpha, 1)$. In order to obtain y_i we take the natural logarithm of t_i.

We compute the bias and mean squared error (MSE) in order to evaluate the estimators' performance. All numerical evaluations were done in the R software. Table 3.1 presents the results obtained for the indicated sample sizes and parameters values. This table shows that for $\alpha = 0.1, 0.5, 1$ and 2, the bias and MSE decrease as the size of the sample increases, which is expected. In general, the results show the good performance of the Bernoulli/BS model.

Table 3.1: *Monte-Carlo estimation of bias and MSE for the indicated sample sizes and parameter values*

		n=100			n=300			n=500		
α	θ	Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
0.1	α	0.21103	0.11103	0.62099	0.14891	0.04891	0.00328	0.10005	4.69350e−5	0.00039
	$\beta_{(1)0} = 0.2$	0.21215	0.01215	1.03058	0.20326	0.00326	0.08643	0.20309	0.00309	0.01018
	$\beta_{(1)1} = 0.5$	0.48652	−0.01348	0.98195	0.50374	0.00374	0.09117	0.50057	0.00057	0.00977
	$\beta_{(2)0} = 1$	1.00140	0.00140	1.17457	1.00115	0.00115	0.08867	0.99797	−0.00203	0.00997
	$\beta_{(2)1} = 2$	1.98399	−0.01600	1.04944	2.00469	0.00469	0.09159	2.00104	0.00104	0.01007
0.5	α	0.50211	0.00211	0.01549	0.49848	−0.00152	0.00343	0.50004	4.22419e−5	0.00032
	$\beta_{(1)0} = 0.2$	0.22839	0.02839	0.81110	0.21216	0.01216	0.14482	0.20293	0.00293	0.00919
	$\beta_{(1)1} = 0.5$	0.49694	−0.00306	0.79049	0.50993	0.00993	0.15961	0.50054	0.00054	0.00881
	$\beta_{(2)0} = 1$	1.00613	0.00613	0.83217	1.01332	0.01332	0.17506	0.99807	−0.00192	0.00899
	$\beta_{(2)1} = 2$	1.98798	−0.01201	0.81395	2.01255	0.01255	0.16855	2.00098	0.00098	0.00908
1	α	0.99855	−0.00145	0.06463	0.99941	−0.00059	0.00279	1.00004	4.69354e−5	0.00039
	$\beta_{(1)0} = 0.2$	0.30713	0.10713	0.46320	0.20529	0.00529	0.03719	0.20278	0.00278	0.00825
	$\beta_{(1)1} = 0.5$	0.54346	0.04346	0.46502	0.50423	0.00423	0.03999	0.50052	0.00052	0.00791
	$\beta_{(2)0} = 1$	1.09830	0.09830	0.64609	1.00368	0.00368	0.04167	0.99817	−0.00183	0.00807
	$\beta_{(2)1} = 2$	2.04050	0.04050	0.54142	2.00462	0.00462	0.04132	2.00093	0.00093	0.00815
2	α	2.01445	0.01445	0.06145	1.99856	−0.00144	0.00159	2.00002	2.34677e−5	9.92905e−5
	$\beta_{(1)0} = 0.2$	0.35064	0.15064	0.27488	0.20109	0.00109	0.00959	0.20030	0.00030	0.00010
	$\beta_{(1)1} = 0.5$	0.56994	0.06994	0.24613	0.50125	0.00125	0.01012	0.50057	5.77348e−5	9.76858e−5
	$\beta_{(2)0} = 1$	1.20834	0.20834	0.72936	1.00040	0.00040	0.00985	0.99979	−0.00021	9.96743e−5
	$\beta_{(2)1} = 2$	2.10056	0.10056	0.38308	2.00157	0.00157	0.01018	2.00010	0.00010	0.00010
4	α	5.74480	1.74480	31.44706	5.63512	1.63512	18.12119	4.21585	0.21585	4.01609
	$\beta_{(1)0} = 0.2$	0.84559	0.64559	3.61015	0.90587	0.70587	3.46159	0.26577	0.06577	0.36885
	$\beta_{(1)1} = 0.5$	0.77905	0.27905	0.90060	0.91386	0.41386	1.16450	0.53409	0.03409	0.10023
	$\beta_{(2)0} = 1$	3.18593	2.8593	40.9525	3.29328	2.29328	34.75177	1.22581	0.22581	4.40974
	$\beta_{(2)1} = 2$	3.09786	1.09786	10.54635	3.22245	1.22245	9.94170	2.10635	0.10635	0.97295

3.5 Application

3.5.1 Exploratory data analysis

We analyse a data set provided by Moulton and Halsey (1995) from a study of measles vaccines. Neutralization antibody levels were collected from 330 children at 12 months of age. The LDL was 0.1 international units (IU) or -2.306 in logarithm scale. Around 86 (26.1%) of the observations fell below the LDL and then recorded as 0.1. The following covariates were considered: X_1 indicates the type of vaccine used (0 if Schwartz and 1 if Edmonston-Zagreb); X_2 is the level of the dosage (0 if medium and 1 if high); and X_3 is the gender where 0 is male and 1 is female.

Table 3.2 shows the descriptive statistics of the observed neutralization antibody levels from the measles vaccines study. This table includes the median, mean, standard deviation (SD) and coefficients of variation (CV), skewness (CS) and kurtosis (CK) values. The CK and CS indicate the positively skewed nature and high kurtosis level of the data distribution. Figure 3.2 shows the scaled total time on test (TTT) plot, histogram and boxplots for the measles vaccine data. In particular, the TTT plot is a graphic tool to assess whether a particular distribution is suitable or not for a data set. This is done by characterizing the shape of a HR, namely, we can detect the type of HR that the data have and then choose a suitable distribution. Let $h_Y(y) = f_Y(y)/(1 - F_Y(y))$ be the HR of an RV Y, where f_Y and F_Y are the PDF and CDF of Y, respectively. Then, the TTT function is $W(u) = H^{-1}(u)/H^{-1}(1)$, for $0 \leq u \leq 1$, where $H^{-1}(u) = \int_0^{F_Y^{-1}(u)}(1 - F_Y(z))\mathrm{d}z$, with F_Y^{-1} denoting the inverse function of the CDF of Y. A plot of the points $(k/n, W_n(k/n))$ can approximate W, with $W_n(k/n) = (\sum_{i=1}^{k} y_{(i)} + (n-k)y_k)/\sum_{i=1}^{n} y_{(i)}$, for $k = 1, \ldots, n$, and $t_{(i)}$ denoting the ith observed order statistic; see, for example, Figure 1 in Azevedo et al. (2012) for some theoretical shapes of scaled TTT curves.

We note that the skewed nature reported in Table 3.2 is confirmed by the histogram of Figure 3.2(a), and the TTT plot displayed in Figure 3.2(b) suggests a decreasing HR. In addition, Figure 3.2(c) indicates that some outliers considered by the usual boxplot are not outliers when we consider the adjusted boxplot.

Table 3.2: *Descriptive summary for the measles vaccine data.*

n	Min	Max	Mean	Median	SD	CV	CS	CK
330	0.10	15.47	1.20	0.40	2.10	174.74%	3.46	14.37

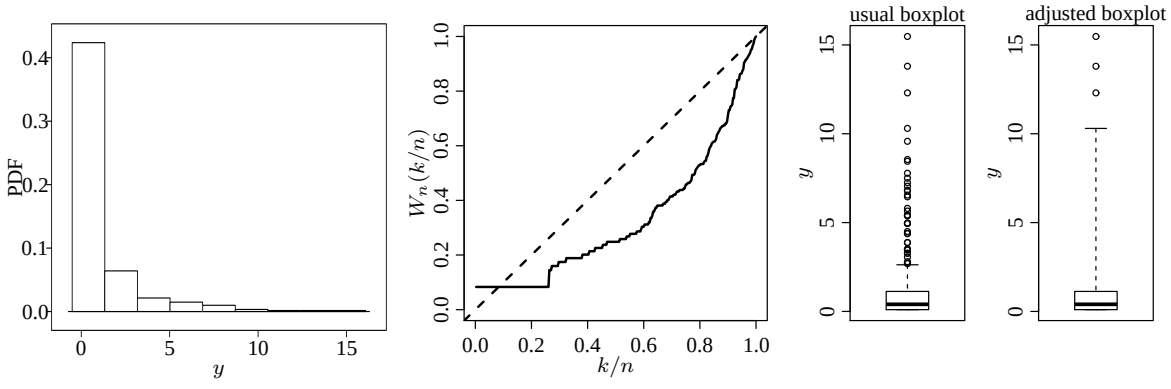

Figure 3.2: *Histogram (a), TTT plot (b), and boxplots (c) for the measles vaccine data.*

3.5.2 Results

We here present the estimation results for the Bernoulli/BS model along with those of the standard tobit, tobit-BS (Chapter 2) and Bernoulli/LPN (MARTÍNEZ-FLORES; BOLFARINE; GÓMEZ, 2013b) models. The Bernoulli/BS and Bernoulli/LPN have both a logit link. The covariates EZ and HI were used only in the logit component, and covariate FEM entered only in

the continuous component of the models. Table 3.3 shows the ML estimates, Akaike information criterion (AIC) values and standard errors for the considered models. A glance at the results indicates that, in the Bernoulli/BS model, the receiver of Edmonston-Zagreb strain does not contribute to the odds ratio of being above the detection limit, however, the receiver of a high dose impacts $\exp(1.499) = 4.472$ in the odds of being above the detection limit. Moreover, the Bernoulli/BS model suggests that girls have $\exp(-0.078) - 1 = -0.075$ less concentration of measles antibody concentration than boys. We observe that the Bernoulli/LPN and Bernoulli/BS models do not agree on the sign of the coefficient corresponding to the FEM variable, while the first model indicates that girls have a higher measles antibody concentration than boys, the other indicates the opposite. Also from Table 3.3, we note that the Bernoulli/BS model provides a better fit compared to the other models based on the AIC values. BIC was not computed because we did not have sufficient information to compute it for the Bernoulli/LPN model.

Table 3.3: *ML estimates (with SE in parentheses) and AIC values for the indicated models with the measles vaccine data*

		Logit component				Continuous component				
Model	AIC	INT	EZ	HI	FEM	α	INT	EZ	HI	FEM
tobit	1299.27					0.945***	0.597**	0.225	−0.228	0.271
						(0.047)	(0.288)	(0.297)	(0.295)	(0.296)
tobit-BS	1168.60					1.545***	−0.910***	0.188*	0.074	0.121
						(0.048)	(0.105)	(0.111)	(0.109)	(0.110)
Bernoulli/LPN	976.48	0.539***	0.757***	0.358		8.918**	−2.869***			0.222*
		(0.204)	(0.270)	(0.261)		(3.922)	(0.582)			(0.134)
Bernoulli/BS	760.64	7.568	4.219	1.499***		1.560***	0.123***			−0.078***
		(5.055)	(5.543)	(0.391)		(0.108)	(<0.001)			(<0.001)

Obs: Rejects H_0 at *10% of significance,** 5% of significance and ***1% of significance.

3.6 Concluding remarks

The paper presents a mixture model where the continuous part follows a Birnbaum-Saunders distribution. We performed estimation based on the maximum likelihood approach. The simulation study showed the good performance of the maximum likelihood estimators. An application to measles vaccine data showed that the Bernoulli/BS models fits the data better than the standard tobit, tobit-BS and Bernoulli/LPN models. Moreover, it was found that in the Bernoulli/BS model the variable FEM is significant, as in the Bernoulli/LPN model, but with opposite sign.

Chapter 4

Concluding remarks and future works

There is a lack of asymmetric regression models for censored data. In order to fill this gap, we developed two models that are adequate for this kind of data. We first developed a tobit model based on the Birnbaum-Saunders distribution. Then, a mixture model consisting of the Birnbaum-Saunders and Bernoulli distributions was introduced. Both models presented good performances when dealing with censored data, thus proving that the Birnbaum-Saunders distribution is a wise choice in the studied settings.

The process of expanding the standard tobit model to an asymmetric framework proved itself to be a fruitful area of research due to the huge amount of topics that remain unexploited. Here, we present some open problems:

i) The tobit-Birnbaum-Saunders models can only be used in asymmetric censored data that are strictly positive because the support of the Birnbaum-Saunders distribution does not include zero. It is of great interest to develop a model for asymmetric censored data that includes zero since, at least in economics, most of the censoring occurs at zero. To solve the problem one can develop a tobit model based on a re-parametrized Birnbaum-Saunders distribution inflated at zero, see Leiva et al. (2014), Leiva et al. (2016) and Santos-Neto et al. (2016);

ii) In the mixture model developed in Chapter 3 the logit link was used but others could also be used, for example, the probit link;

iii) Note that all models considered so far are univariate and being so they do not consider the possibility of correlated response variables. One ambitious, but factual, proposal is to develop a multivariate tobit model with the scale-mixture Birnbaum-Saunders family of distributions. Multivariate tobit models and the multivariate scale-mixture Birnbaum-Saunders family of distributions have been considered, respectively, by Amemiya (1974), Blundell and Meghir (1987), Chib (1992), Zhou and Tan (2009), Zhou and Liu (2009), Chen and Zhou (2011), Deng and Xue (2014), Castro et al. (2015) and Balakrishnan et al. (2009), Leiva et al. (2014), Kundu, Balakrishnan and Jamalizadeh (2013), Kundu (2015), Vilca, Balakrishnan and Zeller (2014), Vilca, Romeiro and Balakrishnan (2016).

Work on these problems is currently in progress and we hope to report these findings in future works.

Bibliography

AMEMIYA, T. Multivariate regression and simultaneous equation models when the dependent variables are truncated normal. **Econometrica: Journal of the Econometric Society**, JSTOR, p. 999–1012, 1974.

AMEMIYA, T. Tobit models: A survey. **Journal of Econometrics**, v. 24, p. 3–61, 1984.

ARELLANO, R.; CASTRO, L. M.; GONZALEZ, G.; NOZ, K. A. M. Student-t censored regression model: Properties and inference. **Statistical Methods and Applications**, v. 21, p. 453–473, 2012.

AZEVEDO, C.; LEIVA, V.; ATHAYDE, E.; BALAKRISHNAN, N. Shape and change point analyses of the Birnbaum-Saunders-t hazard rate and associated estimation. **Computational Statistics and Data Analysis**, v. 56, p. 3887–3897, 2012.

BALAKRISHNAN, N.; LEIVA, V.; SANHUEZA, A.; VILCA, F. Estimation in the Birnbaum-Saunders distribution based on scale-mixture of normals and the EM-algorithm. **Statistics and Operations Research Transactions**, v. 33, p. 171–192, 2009.

BARROS, M.; GALEA, M.; GONZALEZ, M.; LEIVA, V. Influence diagnostics in the tobit censored response model. **Statistical Methods and Applications**, v. 19, p. 379–397, 2010.

BARROS, M.; GALEA, M.; LEIVA, V.; SANTOS-NETO, M. Generalized tobit models: Diagnostics and application in econometrics. **Journal of Applied Statistics**, p. Under review, 2017.

BARROS, M.; PAULA, G. A.; LEIVA, V. A new class of survival regression models with heavy-tailed errors: Robustness and diagnostics. **Lifetime Data Analysis**, v. 14, p. 316–332, 2008.

BHATTI, C. The Birnbaum-Saunders autoregressive conditional duration model. **Mathematics and Computers in Simulation**, v. 80, p. 2062–2078, 2010.

BIRNBAUM, Z. W.; SAUNDERS, S. C. A new family of life distributions. **Journal of Applied Probability**, v. 6, p. 319–327, 1969.

BLUNDELL, R.; MEGHIR, C. Bivariate alternatives to the tobit model. **Journal of Econometrics**, Elsevier, v. 34, n. 1-2, p. 179–200, 1987.

CASTRO, L. M.; COSTA, D. R.; PRATES, M. O.; LACHOS, V. H. Likelihood-based inference for tobit confirmatory factor analysis using the multivariate student-t distribution. **Statistics and Computing**, Springer, v. 25, n. 6, p. 1163–1183, 2015.

CHAI, H. S.; BAILEY, K. R. Use of log-skew-normal distribution in analysis of continuous data with a discrete component at zero. **Statistics in medicine**, Wiley Online Library, v. 27, n. 18, p. 3643–3655, 2008.

CHEN, S.; ZHOU, X. Semiparametric estimation of a bivariate tobit model. **Journal of econometrics**, Elsevier, v. 165, n. 2, p. 266–274, 2011.

CHIB, S. Bayes inference in the tobit censored regression model. **Journal of Econometrics**, Elsevier, v. 51, n. 1-2, p. 79–99, 1992.

COOK, R. D. Detection of influential observation in linear regression. **Technometrics**, v. 19, p. 15–18, 1977.

COX, D.; HINKLEY, D. **Theoretical Statistics**. London, UK: Chapman and Hall, 1974.

CRAGG, J. G. Some statistical models for limited dependent variables with application to the demand for durable goods. **Econometrica: Journal of the Econometric Society**, JSTOR, p. 829–844, 1971.

CRAMÉR, H. **Mathematical Methods of Statistics**. Princeton, US: Princeton University Press, 1946.

CROW, E. L.; SHIMIZU, K. **Lognormal Distributions: Theory and Applications**. New York, US: Dekker, 1988.

DENG, Q.; XUE, J. Multivariate tobit system estimation of education expenditure in urban china. **The Singapore Economic Review**, World Scientific, v. 59, n. 01, p. 1450005, 2014.

DESMOND, A. Stochastic models of failure in random environments. **Canadian Journal of Statistics**, v. 13, p. 171–183, 1985.

EFRON, B.; HINKLEY, D. Assessing the accuracy of the maximum likelihood estimator: Observed vs. expected Fisher information. **Biometrika**, v. 65, p. 457–487, 1978.

FERREIRA, M.; GOMES, M. I.; LEIVA, V. On an extreme value version of the Birnbaum-Saunders distribution. **REVSTAT Statistical Journal**, v. 10, p. 181–210, 2012.

GARAY, A.; BOLFARINE, H.; LACHOS, V.; CABRAL, C. Bayesian analysis of censored linear regression models with scale mixtures of normal distributions. **Journal of Applied Statistics**, v. 42, p. 2694–2714, 2015.

GARCIA-PAPANI, F.; URIBE-OPAZO, M. A.; LEIVA, V.; AYKROYD, R. G. Birnbaum-Saunders spatial modelling and diagnostics applied to agricultural engineering data. **Stochastic Environmental Research and Risk Assessment**, p. in press available at <http://dx.doi.org/10.1007/s00477--015--1204--4>, 2016.

GOLDBERGER, A. **Econometric Theory**. New York, US: Wiley, 1964.

HELSEL, D. R. **Statistics for Censored Environmental Data Using Minitab and R**. New York, US: Wiley, 2011.

JOHNSON, N.; KOTZ, S.; BALAKRISHNAN, N. **Continuous Univariate Distributions**. New York, US: Wiley, 1995. v. 2.

JONES, M. C. On reciprocal symmetry. **Journal of Statistical Planning and Inference**, v. 138, p. 3039–3043, 2008.

KLEIBER, C.; ZEILEIS, A. **Applied Econometrics with R**. New York, US: Springer, 2008.

KLEIBER, C.; ZEILEIS, A. **R package ARE: Applied Econometrics with R**. 2015. <http://CRAN.R-project.org/package=AER>.

KOTZ, S.; LEIVA, V.; SANHUEZA, A. Two new mixture models related to the inverse Gaussian distribution. **Methodology and Computing in Applied Probability**, v. 12, p. 199–212, 2010.

KUNDU, D. Bivariate log birnbaum–saunders distribution. **Statistics**, Taylor & Francis, v. 49, n. 4, p. 900–917, 2015.

KUNDU, D.; BALAKRISHNAN, N.; JAMALIZADEH, A. Generalized multivariate birnbaum–saunders distributions and related inferential issues. **Journal of Multivariate Analysis**, Elsevier, v. 116, p. 230–244, 2013.

KUNDU, D.; KANNAN, N.; BALAKRISHNAN, N. On the hazard function of Birnbaum-Saunders distribution and associated inference. **Computational Statistics and Data Analysis**, v. 52, p. 2692–2702, 2008.

LEAO, J.; LEIVA, V.; SAULO, H.; TOMAZELLA, V. Birnbaum-Saunders frailty regression models: Diagnostics and application to medical data. **Biometrical Journal**, p. in press available at <http://dx.doi.org/10.1002/bimj.200800143>, 2017.

LEIVA, V. **The Birnbaum-Saunders Distribution**. New York, US: Academic Press, 2016.

LEIVA, V.; BARROS, M.; PAULA, G. A.; GALEA, M. Influence diagnostics in log-Birnbaum-Saunders regression models with censored data. **Computational Statistics and Data Analysis**, v. 51, p. 5694–5707, 2007.

LEIVA, V.; FERREIRA, M.; GOMES, M. I.; LILLO, C. Extreme value Birnbaum-Saunders regression models applied to environmental data. **Stochastic Environmental Research and Risk Assessment**, v. 30, p. 1045–1058, 2016.

LEIVA, V.; MARCHANT, C.; RUGGERI, F.; SAULO, H. A criterion for environmental assessment using Birnbaum-Saunders attribute control charts. **Environmetrics**, v. 26, p. 463–476, 2015.

LEIVA, V. et al. Capability indices for Birnbaum-Saunders processes applied to electronic and food industries. **Journal of Applied Statistics**, v. 41, p. 1881–1902, 2014.

LEIVA, V.; ROJAS, E.; GALEA, M.; SANHUEZA, A. Diagnostics in Birnbaum-Saunders accelerated life models with an application to fatigue data. **Applied Stochastic Models in Business and Industry**, v. 30, p. 115–131, 2014.

LEIVA, V.; RUGGERI, F.; SAULO, H.; VIVANCO, J. F. A methodology based on the Birnbaum-Saunders distribution for reliability analysis applied to nano-materials. **Reliability Engineering and System Safety**, v. 157, p. 192–201, 2017.

LEIVA, V.; SANTOS-NETO, M.; CYSNEIROS, F. J. A.; BARROS, M. Birnbaum-Saunders statistical modelling: A new approach. **Statistical Modelling**, v. 14, p. 21–48, 2014.

LEIVA, V.; SANTOS-NETO, M.; CYSNEIROS, F. J. A.; BARROS, M. A methodology for stochastic inventory models based on a zero-adjusted Birnbaum-Saunders distribution. **Applied Stochastic Models in Business and Industry**, v. 32, p. 74–89, 2016.

LEIVA, V.; SAULO, H.; LEÃO, J.; MARCHANT, C. A family of autoregressive conditional duration models applied to financial data. **Computational Statistics and Data Analysis**, v. 79, p. 175–191, 2014.

LEIVA, V. et al. Modeling neural activity with cumulative damage distributions. **Biological Cybernetics**, v. 109, p. 421–433, 2015.

LESAFFRE, E.; VERBEKE, G. Local influence in linear mixed models. **Biometrics**, v. 54, p. 570–582, 1998.

LI, A. P.; CHEN, Z. X.; XIE, F. C. Diagnostic analysis for heterogeneous log-Birnbaum-Saunders regression models. **Statistical and Probability Letters**, v. 89, p. 1690–1698, 2012.

MARCHANT, C.; LEIVA, V.; CYSNEIROS, F. J. A. A multivariate log-linear model for Birnbaum-Saunders distributions. **IEEE Transactions on Reliability**, v. 65, p. 816–827, 2016.

MARCHANT, C.; LEIVA, V.; CYSNEIROS, F. J. A.; VIVANCO, J. F. Diagnostics in multivariate generalized Birnbaum-Saunders regression models. **Journal of Applied Statistics**, v. 43, p. 2829–2849, 2016.

MARSHALL, A.; OLKIN, I. **Life Distributions**. New York, US: Springer, 2007.

MARTÍNEZ-FLORES, G.; BOLFARINE, H.; GÓMEZ, H. W. The alpha-power tobit model. **Communications in Statistics: Theory and Methods**, v. 42, p. 633–643, 2013.

MARTÍNEZ-FLORES, G.; BOLFARINE, H.; GÓMEZ, H. W. Asymmetric regression models with limited responses with an application to antibody response to vaccine. **Biometrical Journal**, v. 55, p. 156–172, 2013.

MASSUIA, M. B.; CABRAL, C. R. B.; MATOS, L. A.; LACHOS, V. H. Influence diagnostics for student-t censored linear regression models. **Statistics**, v. 49, p. 1074–1094, 2015.

MOULTON, L. H.; HALSEY, N. A. A mixture model with detection limits for regression analyses of antibody response to vaccine. **Biometrics**, v. 51, p. 1570–1578, 1995.

QU, H.; XIE, F. Diagnostics analysis for log-Birnbaum-Saunders regression models with censored data. **Statistica Neerlandica**, v. 65, p. 1–21, 2011.

R-TEAM. **R: A Language and Environment for Statistical Computing**. Vienna, Austria, 2016. Disponível em: <www.R-project.org>.

RIECK, J. R.; NEDELMAN, J. R. A log-linear model for the Birnbaum-Saunders distribution. **Technometrics**, v. 3, p. 51–60, 1991.

ROCHA, G. H. M. A.; ARELLANO, R.; LOSCHI, R. H. Maximum likelihood methods in a robust censored errors-in-variables model. **Test**, v. 24, p. 857–877, 2015.

ROUSSEEUW, P. et al. **Robustbase: Basic robust statistics**. [S.l.], 2015. R package version 0.92-5. Disponível em: <http://CRAN.R-project.org/package=robustbase>.

SANTOS-NETO, M. **tobitdiag: local influence for tobit models**. [S.l.], 2016. R package version 0.0.1.

SANTOS-NETO, M.; CYSNEIROS, F. J. A.; LEIVA, V.; BARROS, M. Reparameterized Birnbaum-Saunders regression models with varying precision. **Electronic Journal of Statistics**, v. 10, p. 2825–2855, 2016.

SAULO, H.; LEIVA, V.; ZIEGELMANN, F. A.; MARCHANT, C. A nonparametric method for estimating asymmetric densities based on skewed Birnbaum-Saunders distributions applied to environmental data. **Stochastic Environmental Research and Risk Assessment**, v. 27, p. 1479–1491, 2013.

SCOTT, J. **Regression Models of Categorical and Limited Dependent Variables**. California, US: Sage, 1997.

THORARINSDOTTIR, T. L.; GNEITING, T. Probabilistic forecasts of wind speed: Ensemble model output statistics by using heteroscedastic censored regression. **Journal of the Royal Statistical Society A**, v. 173, p. 371–388, 2010.

TOBIN, J. Estimation of relationships for limited dependent variables. **Econometrica**, v. 26, p. 24–36, 1958.

VANEGAS, L. H.; PAULA, G. A. An extension of log-symmetric regression models: R codes and applications. **Journal of Statistical Simulation and Computation**, v. 86, p. 1709–1735, 2016.

VANEGAS, L. H.; PAULA, G. A. Log-symmetric distributions: statistical properties and parame-

ter estimation. **Brazilian Journal of Probability and Statistics**, v. 30, p. 196–220, 2016.

VILCA, F.; BALAKRISHNAN, N.; ZELLER, C. B. A robust extension of the bivariate birnbaum–saunders distribution and associated inference. **Journal of Multivariate Analysis**, v. 124, p. 418–435, 2014.

VILCA, F.; ROMEIRO, R. G.; BALAKRISHNAN, N. A bivariate birnbaum–saunders regression model. **Computational Statistics & Data Analysis**, Elsevier, v. 97, p. 169–183, 2016.

VILLEGAS, C.; PAULA, G. A.; LEIVA, V. Birnbaum-Saunders mixed models for censored reliability data analysis. **IEEE Transactions on Reliability**, v. 60, p. 748–758, 2011.

WANKE, P.; LEIVA, V. Exploring the potential use of the Birnbaum-Saunders distribution in inventory management. **Mathematical Problems in Engineering**, Article ID 827246, p. 1–9, 2015.

ZHOU, X.; LIU, X. The monte carlo em method for estimating multivariate tobit latent variable models. **Journal of Statistical Computation and Simulation**, Taylor & Francis, v. 79, n. 9, p. 1095–1107, 2009.

ZHOU, X.; TAN, C. Maximum likelihood estimation of tobit factor analysis for multivariate t-distribution. **Communications in Statistics-Simulation and Computation**, Taylor & Francis, v. 39, n. 1, p. 1–16, 2009.

ZHU, H.; ZHANG, H. A diagnostic procedure based on local influence. **Biometrika**, v. 91, p. 579–589, 2004.