UNIVERSIDAD DE LA FRONTERA

DOCTORADO EN CIENCIAS MENCIÓN MATEMÁTICA

BIRNBAUM-SAUNDERS QUANTILE REGRESSION MODELS

presented by

Luis Sánchez Melo

Doctoral thesis supervised by

Víctor Leiva

Temuco – Chile
2020

Luis Sánchez Melo

Birnbaum-Saunders quantile regression models

Doctoral thesis submitted to the
Doctorate Program in Sciences Mention Mathematics
Department of Mathematics and Statistics
as a requirement to obtain a
PhD in Sciences Mention Mathematics.

During the development of this work the author received
scholarships from Universidad de la Frontera and financing from
the grant project FONDECYT 1160868 of the Chilean government.

Temuco – Chile
2020

Resumen

La distribución Birnbaum-Saunders (BS) ha sido extensamente estudiada y aplicada. En un modelo de regresión, la respuesta media condicionada a valores de variables explicativas es descrita . Sin embargo, si la respuesta sigue una distribución asimétrica, la media no es una buena medida de centralidad para resumir los datos y la mediana debería ser usada. En ese caso, un modelo de regresión cuantil parece ser más apropiado para relacionar la respuesta de interés a las variables explicativas. La precisión de un estimador de la media (o mediana) puede ser mejorada si una componente espacial es considerada en el modelado. Proponemos modelos de regresión cuantil para datos independientes y también con dependencia espacial basados en la distribución BS. Estimamos sus parámetros usando el método de verosimilitud máxima y presentamos sus propiedades asintóticas para modelos de regresión cuantil BS para datos independientes. Asimismo, derivamos métodos de diagnóstico para evaluar la pertinencia de las suposiciones del modelo y detectar casos potencialmente influyentes. Ilustramos los resultados obtenidos con datos reales para mostrar sus aplicaciones potenciales y comparamos los modelos BS y gaussiano. Los resultados numéricos indican un buen desempeño de la regresión cuantil BS, demostrando que la distribución BS es una buena elección cuando se modelan datos que tienen soporte positivo y asimetría.

Palabras claves: análisis de datos; influencia local; método de verosimilitud máxima; modelos lineales generalizados; modelos espaciales; regresión de la mediana; software R; residuos.

Abstract

The Birnbaum-Saunders (BS) distribution has been largely studied and applied. In a regression model, the mean response conditional to values of explanatory variables is described. However, if the response follows a skew distribution, the mean is not a good centrality measure to summarize the data and the median should be used. In this case, a quantile regression model seems to be more suitable for relating the response of interest to explanatory variables. Accuracy of an estimator of the mean (or median) may be improved if a spatial component is considered in the modeling. We propose quantile regression models for independent data and also with spatial dependence based on the BS distribution. We estimate their parameters by using the maximum likelihood method and present their asymptotic properties for BS quantile regression models with independent data. Also, we derive diagnostic techniques to assess the suitability of the model assumptions and to detect potentially influential cases. We illustrate the obtained results with real data to show their potential applications and compare BS and Gaussian models. The numerical results report an adequate performance of the approach to quantile regression based on the BS distribution indicating that this distribution is a good modeling choice when dealing with data that have both positive support and asymmetry.

Keywords: data analysis; generalized linear models; local influence; maximum likelihood method; median regression; R software; residuals; spatial models.

Contents

1 Preliminaries 6
1.1 Summary 6
1.2 Introduction and bibliographical review 6
1.3 Background 8
1.4 Motivation of the thesis 10
1.5 Objectives of the thesis 10
1.6 Products of the thesis 11
1.7 Organization of the thesis 11

2 BS quantile regression models for independent data 12
2.1 Summary 12
2.2 Introduction 12
2.3 A BS distribution parametrized by its quantiles 13
2.4 A BS quantile regression model 14
2.5 Local influence analytics 18
2.6 Simulation study 21
2.7 Ilustrative example 25
2.8 Economic implications 32
2.9 Concluding remarks 33

3 BS spatial quantile regression models 34
3.1 Summary 34
3.2 Introduction 34
3.3 A parameterization of the multivariate BS distribution 35
3.4 Formulation and estimation of the spatial model 38
3.5 Model checking 41
3.6 Global influence diagnostics 42
3.7 Local influence diagnostics 43

3.8 Illustrative example 45
3.9 Concluding remarks 50

4 Discussion 51
4.1 Summary 51
4.2 Conclussions and limitations 51
4.3 Future research 52

References 61

List of Figures

2.1 Plots of the BS probability density function for $Q = 1.0$ (left) and $\alpha = 1.0$ (center), for $q = 0.25$ (a)-(b); for $q = 0.50$ (d)-(e); as well as for $q = 0.75$ (g)-(h); and of the BS variance against α for $Q = 2.0$ (right) with $q = 0.25$ (c), $q = 0.50$ (f) and $q = 0.75$ (i). 15
2.2 Plots of the indicated residual, link function and value of α with simulated data. 23
2.3 Plots of the indicated residual, link function and value of α with simulated data. 24
2.4 Histogram (a) and boxplots (b) for Chilean income data. 26
2.5 Scatters plots and correlations between each pair of variables for Chilean income data. 27
2.6 QQ plot with envelope of the residual $r^{(4)}$ for the BS median model (a) and of the standardized residual for the normal regression model (b); index plot of $r^{(4)}$ (c) and plot of $\widehat{z}_2$ against $\widehat{\eta}$ for the model fit with identity link (d) based on Chilean income data and the BS quantile regression model. 29
2.7 Index plots of C_i for $\boldsymbol{\beta}$ (a) and α (b) under case-weight perturbation; for $\boldsymbol{\beta}$ (c) and α (d) under response perturbation; for $\boldsymbol{\beta}$ (e) and α (f) under perturbation of the parameter α; for $\boldsymbol{\beta}$ (g) and α (h) under covariate perturbation X_2, using Chilean income data and the BS quantile regression model. 31

3.1 Reparameterized 2-variate BS density plots for (a) $\alpha_i = 0.5$, (b) $\alpha_i = 0.8$, (c) $\alpha_i = 1.5$, with $Q_i = 1.0$, and (d) $Q_i = 0.5$, (e) $Q_i = 0.8$, (f) $Q_i = 1.5$, with $\alpha_i = 1.0$, for $i = 1, 2$ and $\rho = 0.9$. 37
3.2 Histogram (a), boxplot (b) and semi-variogram (c) for the response variable with environmental data. 45
3.3 QQ plots for transformed residuals with environmental data. 47

3.4 plots of the CD with environmental data. 48
3.5 Perturbation in the response for $\boldsymbol{d}_{\max}$ (a), C_i (b) and B_i (c) and perturbation in the regressor for $\boldsymbol{d}_{\max}$ (d), C_i (e) and B_i (f) with environmental data. . . 49

Chapter 1

Preliminaries

1.1 Summary

In this chapter, we give an introduction and bibliographical review for our study. Next, the relation between the BS and standard normal distributions is described and the probability density and cumulative distribution functions for the univariate BS case are presented. Then, we establish a new parameterization of the BS distribution based on its quantiles. Thus, we introduce the multivariate BS distribution and propose a reparameterization based on the quantiles of the marginal distributions associated. Hence, we mention aspects related to motivations, aims and products of this thesis.

1.2 Introduction and bibliographical review

Life distributions are often positively skewed, unimodal and two-parameter models, in addition to having positive support; see Marshall and Olkin (2007) and Saunders (2007). A life distribution that has received a considerable attention in recent decades is the BS model. It was originated from a problem of material fatigue and has been largely applied to reliability and fatigue studies; see Birnbaum and Saunders (1969) and Leiva and Saunders (2015). The BS distribution relates the total time until the failure to some type of cumulative damage normally distributed. This attention for the BS distribution is due to its attractive properties and its relationship with the normal distribution. Extensive work has been done on the BS distribution with regard to its properties, inference and modeling. Its natural applications have been mainly focussed on engineering. However, today they range diverse fields including business, environment, industry and medicine; see Lio et al. (2010), Leiva et al. (2011, 2014a,c, 2015a,b); Leiva (2016); Leiva et al. (2016b, 2017, 2018), Lillo et al. (2018), Marchant et al. (2013a,b, 2016a,b, 2018), Saulo et al. (2013, 2015, 2019), Rojas et al. (2015), Wanke and Leiva (2015), Leão et al. (2017b, 2018b), Leiva and Saulo (2017), Desousa et al. (2018) and Leiva et al. (2018), for some of its more recent applications. For a comprehensive treatment on this model, its computational implementation, and details about new applications in fields beyond engineering, see the book by Leiva (2016) and the review article by Balakrishnan and Kundu (2019).

A random variable with BS distribution can be considered as a transformation of another random variable with standard normal distribution. Then, because all random variable following the BS distribution may be represented by another basis random variable, generalizations of this distribution might be obtained changing the distribution of this basis random variable. Diverse arguments can be used for supporting this change, which allows more general classes of models to be constructed. Several extensions and generalizations of the BS distribution have been conducted by an international and transdisciplinary group of researchers. The first extension of the BS distribution is attributed to Volodin and Dzhungurova (2000). Then, Díaz-García and Leiva (2005) introduced the generalized BS (GBS) distribution based on elliptical distributions. Owen (2006) proposed a three-parameter extension of the BS distribution. Vilca and Leiva (2006) derived a BS distribution based on skew-normal models. Gómez et al. (2009) extended the BS distribution from the slash-elliptic model. Guiraud et al. (2009) deducted a non-central version of the BS distribution. Leiva et al. (2009) provided a length-biased option of the BS distribution. Ahmed et al. (2010) truncated the BS distribution. Kotz et al. (2010) performed mixture models related to the BS distribution. Vilca et al. (2010) and Castillo et al. (2011) developed the epsilon-skew BS distribution. Balakrishnan et al. (2011) considered BS mixture distributions. Cordeiro and Lemonte (2011) defined the beta-BS distribution. Leiva et al. (2011) modeled wind energy flux using a shifted BS distribution. Athayde et al. (2012) viewed the BS distributions as part of the Johnson system, allowing location-scale BS distributions to be obtained. Santos-Neto et al. (2012, 2014, 2016) reparameterized the BS distribution obtaining interesting properties. Saulo et al. (2012) presented the Kumaraswamy-BS distribution. Fierro et al. (2013) generated the BS distribution from a non-homogeneous Poisson process. Lemonte (2013) studied the Marshall-Olkin-BS (MOBS) distribution. Bourguignon et al. (2014) derived the power-series BS class of distributions. Martinez et al. (2014) introduced an alpha-power extension of the BS distribution. Leiva et al. (2016b) proposed a zero-adjusted BS distribution. Bourguignon et al. (2017) derived the transmuted BS distribution.

Common regression models provide an estimate of the mean response given certain values of the explanatory variables. However, as mentioned, if the response variable follows a skew distribution, the mean is not a good central tendency measure to summarize the data. Quantile regression models were proposed by Koenker and Bassett (1978), who extended the median regression model attributed to Laplace (1818), generalizing the ordinary sample quantiles in the regression setting. The quantile regression aims at estimating either the conditional median or other quantiles of the response. Quantile regression models for the BS distribution do not exist in the literature; see Noufaily and Jones (2013) for a generalized gamma quantile regression, with the gamma distribution being a direct competitor of the BS distribution. In particular, we are interested in modeling the median of the BS distribution by regression. In the BS distribution, the median is one of its parameters, so that its modeling by quantile regression seems natural; see Leiva et al. (2014b) and Santos-Neto et al. (2016) for modeling using a mean-based reparameterization of the BS distribution, whereas Saulo et al. (2019) did a comparison between models based on the mean and median of the BS distribution.

Accuracy of an estimator of the mean (or median) may be improved if a spatial component is added in the modeling. A first idea of spatial quantile regression was proposed by Kostov (2009). Trzpiot (2013) performed a work about spatial quantile regression and proposed a general model based on the conditional quantile function. McMillen (2013) showed variants of the spatial quantile regression. Garcia-Papani et al. (2017, 2018a,b) introduced BS spatial models and their diagnostics for the conditional mean. Stochastic processes are needed when modeling data spatially, which have not been proposed for BS distributions, so that derivation of a BS process is also an open problem. For stochastic processes, we need to know he corresponding finite dimensional multivariate distributions, which have been proposed and studied for BS distributions by Caro-Lopera et al. (2012), Kundu et al. (2013), Vilca et al. (2014), Jamalizadeh and Kundu (2015), Khosravi et al. (2015), Kundu (2015a,b), Lemonte et al. (2015), Sánchez et al. (2015), Marchant et al. (2016a,b) and Garcia-Papani et al. (2017, 2018a,b).

1.3 Background

If $Z \sim \mathrm{N}(0,1)$, then the random variable T defined as

$$T = \frac{\beta}{4}\left(\alpha Z + \sqrt{\alpha^2 Z^2 + 4}\right)^2 \tag{1.1}$$

follows a BS distribution with parameters of shape $\alpha > 0$ and scale $\beta > 0$, which is denoted by $T \sim \mathrm{BS}(\alpha, \beta)$. The random variable T has positive support and the transformation given in (1.1) is one-to-one, which allows us to establish that

$$Z = \frac{1}{\alpha}\left(\sqrt{T/\beta} - \sqrt{\beta/T}\right) \sim \mathrm{N}(0,1).$$

The probability density and cumulative distribution functions of T are expressed respectively as

$$
\begin{aligned}
f_T(t) &= \frac{1}{2\alpha\beta\sqrt{2\pi}}\left[\sqrt{\beta/t}+\sqrt{(\beta/t)^3}\right]\exp\left[-\frac{1}{2\alpha^2}\left(\frac{t}{\beta}+\frac{\beta}{t}-2\right)\right],\ t>0,\\
F_T(t) &= \Phi\left[\frac{1}{\alpha}\left(\sqrt{t/\beta}-\sqrt{\beta/t}\right)\right],\ t>0,
\end{aligned}
$$

where Φ is the standard normal cumulative distribution function. Given $q \in (0,1)$ and based on (1.1), note that the $q\times 100$th quantile of the BS distribution is defined as

$$
Q = t_q = \frac{\beta}{4}\left(\alpha z_q + \sqrt{\alpha^2 z_q^2 + 4}\right)^2 = \frac{\beta}{4}\gamma_\alpha^2, \tag{1.2}
$$

where

$$
\gamma_\alpha = \alpha z_q + \sqrt{\alpha^2 z_q^2 + 4} \tag{1.3}
$$

and z_q is the $q\times 100$th quantile of the standard normal distribution. Note that $Q>0$.

Let the random vector $\boldsymbol{V} = (V_1,\ldots,V_n)^\top \in \mathbb{R}^n$ follow a multivariate normal distribution, denoted by $\boldsymbol{V}\sim \mathrm{N}_n(\boldsymbol{\mu},\boldsymbol{\Sigma})$, with mean vector $\boldsymbol{\mu}=(\mu_i)\in\mathbb{R}^n$ and variance-covariance matrix $\boldsymbol{\Sigma}=(\sigma_{jk})\in\mathbb{R}^{n\times n}$, with $\mathrm{rank}(\boldsymbol{\Sigma})=n$. The probability density function of $\boldsymbol{V}$ is given by

$$
f_{\boldsymbol{V}}(\boldsymbol{v};\boldsymbol{\mu},\boldsymbol{\Sigma}) = (2\pi)^{-n/2}\,|\boldsymbol{\Sigma}|^{-1/2}\exp\left(-\frac{1}{2}[\boldsymbol{v}-\boldsymbol{\mu}]^\top\boldsymbol{\Sigma}^{-1}[\boldsymbol{v}-\boldsymbol{\mu}]\right),\ \boldsymbol{v}=(v_1,\ldots,v_n)^\top\in\mathbb{R}^n \tag{1.4}
$$

and its cumulative distribution function is denoted by $F_{\boldsymbol{V}}(\boldsymbol{v};\boldsymbol{\mu},\boldsymbol{\Sigma})$. When the mean vector is zero, that is $\boldsymbol{\mu}=\mathbf{0}_{n\times 1}$, with $\mathbf{0}_{n\times 1}$ being an $n\times 1$ vector of zeros, we use the notation ϕ_n and Φ_n for the n-variate normal probability density and cumulative distribution functions, respectively.

The random vector $\boldsymbol{T}=(T_1,\ldots,T_n)^\top\in\mathbb{R}_+^n$ follows a multivariate BS distribution with parameters $\boldsymbol{\alpha}=(\alpha_1,\ldots,\alpha_n)^\top\in\mathbb{R}_+^n$, $\boldsymbol{\beta}=(\beta_1,\ldots,\beta_n)^\top\in\mathbb{R}_+^n$ and $\boldsymbol{\Gamma}\in\mathbb{R}^{n\times n}$ if $T_i = T(V_i;\alpha_i,\beta_i)$, for $i=1,\ldots,n$, where T is given in (1.1) and $\boldsymbol{V}=(V_1,\ldots,V_n)^\top\in\mathbb{R}^n\sim\mathrm{N}_n(\mathbf{0}_{n\times 1},\boldsymbol{\Gamma})$, with $\boldsymbol{\Gamma}\in\mathbb{R}^{n\times n}$ being the correlation matrix of $\boldsymbol{V}$. Hence, we denote the n-variate BS distribution by $\boldsymbol{T}\sim\mathrm{BS}_n(\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\Gamma})$. Thus, the cumulative distribution function and probability density function of $\boldsymbol{T}$ are, respectively, defined by

$$
\begin{aligned}
F_{\boldsymbol{T}}(\boldsymbol{t};\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\Gamma}) &= \Phi_n(\boldsymbol{A};\boldsymbol{\Gamma}), && (1.5)\\
f_{\boldsymbol{T}}(\boldsymbol{t};\boldsymbol{\alpha},\boldsymbol{\beta},\boldsymbol{\Gamma}) &= \phi_n(\boldsymbol{A};\boldsymbol{\Gamma})\,a(\boldsymbol{t};\boldsymbol{\alpha},\boldsymbol{\beta}),\ \ \boldsymbol{t}=(t_1,\ldots,t_n)\in\mathbb{R}_+^n, && (1.6)
\end{aligned}
$$

where $\boldsymbol{A}=\boldsymbol{A}(\boldsymbol{t};\boldsymbol{\alpha},\boldsymbol{\beta})=(A_1,\ldots,A_n)^\top$, with $A_j = A(t_j;\alpha_j,\beta_j)$,

$$
a(\boldsymbol{t};\boldsymbol{\alpha},\boldsymbol{\beta}) = \prod_{j=1}^{n} a(t_j;\alpha_j,\beta_j), \tag{1.7}
$$

and both $A(t_j;\alpha_j,\beta_j)$ and $a(t_j;\alpha_j,\beta_j)$ are given by

$$A(t_j;\alpha_j,\beta_j) = \frac{1}{\alpha_j}\left[\sqrt{t_j/\beta_j}-\sqrt{\beta_j/t_j}\right], \tag{1.8}$$

$$a(t_j;\alpha_j,\beta_j) = \frac{d}{dt_j}(A(t_j;\alpha_j,\beta_j)) = \frac{1}{2\alpha_j\beta_j}\left[\left\{\frac{\beta_j}{t_j}\right\}^{1/2}+\left\{\frac{\beta_j}{t_j}\right\}^{3/2}\right]. \tag{1.9}$$

1.4 Motivation of the thesis

According to the review of literature discussed in the Section 1.2 and the background presented in Section 1.3, we have mainly the next motivations to develop this thesis:

(1) Since there are not studies about quantile regression models for independent data based on the BS distribution, we propose to formulate models in this line and to realize a study on parameters estimation by the maximum likelihood (ML) method, asymptotic properties and simulation of residuals. In addition, analytics of local influence is derived.

(2) Because of the non-existence in the literature of BS spatial quantile regression models, we propose to formulate models in this line for correlated-spatially data. We are interested in carrying out estimation of parameters and diagnostics by using the global and local influence techniques.

1.5 Objectives of the thesis

Based on Section 1.4, the objectives of this work are:

(1) To formulate quantile regression models based on the univariate BS distribution.

(2) To develop BS spatial quantile regression models.

(3) To derive influence diagnostics for the formulated models.

(4) To evaluate the proposed results by simulations.

(5) To apply BS quantile regression models to real-world data.

(6) To implement the methodology obtained in the R software.

1.6 Products of the thesis

This thesis led to the following products:

(1) **Sánchez, L.**, Leiva, V. (2018). A quantile regression model for the Birnbaum-Saunders distribution. X Simposio Nororiental de Matematicas, Bucaramanga, Colombia.

(2) **Sánchez, L.**, Leiva, V. (2018) Diagnostics on a Birnbaum-Saunders quantile regression model. III International Workshop on Data Science, Vina del Mar, Chile.

(3) **Sánchez, L.**, Leiva, V., Galea, M., Saulo, H. (2020). Birnbaum-Saunders quantile regression and its diagnostics with application to economic data. Applied Stochastic Models in Business and Industry, pages in press.

(4) **Sánchez, L.**, Leiva, V., Galea, M., Saulo, H. (2020). Birnbaum-Saunders Quantile Regression Models with Application to Spatial Data. Under review.

(5) **Sánchez, L.**, Leiva, V., Galea, M., Saulo, H. (2020). Global and local diagnostic analytics for a geostatistical model based on Birnbaum-Saunders quantile regression with different distance measures. Under review.

1.7 Organization of the thesis

This thesis is organized as follows. In Chapter 2, we formulate BS quantile regression models for data whose observations are independent. Parameter estimation, residuals, local influence diagnostic analytics, and illustrations with real data are provided. In Chapter 3, we introduce BS spatial quantile regression models once again considering parameters estimation, diagnostics of global and local influence and illustrations with real data. In Chapter 4, we establish the conclusions and further research to be developed. Appendixes contain the corresponding Fisher information and perturbation matrices for inference and diagnostics in the BS spatial regression model.

Chapter 2

BS quantile regression models for independent data

2.1 Summary

In this chapter, we introduce a class of quantile regression models based on the BS distribution, which allows us to describe positive and asymmetric data when a quantile must be predicted using covariates. We use an approach based on a quantile parameterization to generate the model, permitting us to consider a similar framework to generalized linear models, providing wide flexibility. The methodology proposed includes a thorough study of theoretical properties and practical issues, such as parameter estimation by ML method and diagnostic analytics based on local influence and residuals. The performance of the residuals is evaluated by simulations, whereas an illustrative example of income data is conducted using the methodology to show its potential for applications. The economic implications of our investigation are discussed in the final section.

2.2 Introduction

The study of the BS distribution has received growing interest and a considerable amount of work is available; see the recent publications by Leiva (2016), Balakrishnan and Kundu (2019) and references therein, which summarize most of the works to the date. BS spatial models and their diagnostics for the conditional mean have been developed; see Garcia-Papani et al. (2017, 2018a,b). However, no quantile regression models based on the BS distribution have been derived.

Diagnostic analytics plays a relevant role in statistical modeling, which can be classified in global and local techniques. Residuals are well-known and often used as measures of global influence and for detecting the model adequacy (Krzanowski, 1998; Leiva et al., 2016), whereas the local influence technique is currently very popular. This technique allows us to evaluate the local effect of perturbations on the estimates of parameters and then to detect potentially influential cases in different models; see, for example, Santana et al. (2011) and Tapia et al. (2019).

The main objective of this chapter is to formulate quantile regression models based on the BS distribution and its diagnostics. We use a quantile parameterization to generate the new model, which allows us to consider a similar framework to generalized linear models, providing wide flexibility; see also the works by Mitnik and Baek (2013) and Noufaily and Jones (2013) for similar parameterizations but not identical. Note that it is not possible to make comparison between our model and that proposed by Noufaily and Jones (2013) because are models postulated in different contexts. In any case, future research about this issue is mentioned in the final section.

The remainder chapter is organized as follows. Section 2.3 presents the BS distribution in its original parameterization and a new parameterization of it which allows us to model a quantile. In Section 2.4, we formulate the regression model and provide estimation based on the ML method. In Section 2.5, we derive diagnostic analytics based on the local influence technique. Section 2.6 proposes four types of residuals for the BS quantile regression model and then we evaluate their performance by using Monte Carlo simulations. In Section 2.7, we also apply the obtained results to household income data, including formulation, estimation, inference and diagnostic analytics to illustrate the potential of the new model. In Section **??**, we discuss economic implications of our illustrative example. In Section 2.9, we present concluding remarks.

2.3 A BS distribution parametrized by its quantiles

Consider $T \sim \mathrm{BS}(\alpha,\beta)$. Let $q \in (0,1)$ be a fixed number and the transformation $(\alpha,\beta) \mapsto (\alpha,Q)$ be one-to-one, where Q is defined in (1.2). Then, we can define a parameterization of the BS model based on Q so that the associated cumulative distribution and probability density functions can be written, respectively, as

$$\begin{aligned} F(t) &= \Phi\left[\frac{1}{\alpha\gamma_\alpha}\sqrt{\frac{4Q}{t}}\left(\frac{t\gamma_\alpha^2}{4Q}-1\right)\right],\ t>0,\\ f_T(t) &= \frac{1}{\alpha\gamma_\alpha\sqrt{8\pi Q t}}\left(\frac{\gamma_\alpha^2}{2}+\frac{2Q}{t}\right)\exp\left[-\frac{2Q}{\alpha^2\gamma_\alpha^2 t}\left(\frac{t\gamma_\alpha^2}{4Q}-1\right)^2\right],\ t>0, \qquad (2.1)\end{aligned}$$

where γ_α is given by (1.3). Hence, under this parameterization, we write $T \sim \mathrm{BS}(\alpha, Q)$. The mean and variance of T are, respectively, given by

$$\mathrm{E}[T] = \frac{4Q}{\gamma_\alpha^2}\left(1 + \frac{\alpha^2}{2}\right), \quad \mathrm{Var}[T] = \frac{16Q^2\alpha^2}{\gamma_\alpha^4}\left(1 + \frac{5}{4}\alpha^2\right).$$

Figure 2.1 displays some shapes of the probability density function of $T \sim \mathrm{BS}(\alpha, Q)$ defined in (2.1). From Figures 2.1 (a), (d) and (g), observe that the parameter α modifies the skewness and kurtosis of the model, as expected since it is a shape parameter. From Figures 2.1 (b), (e) and (h), note that, as Q increases, the kurtosis decreases, also as expected because it is a quantile parameter so that, as it increases, less kurtosis is detected. Also there is more concentration around the quantile as α decreases and therefore the variability decreases. Furthermore, notice that, when α increases, the variance increases exponentially for the first and second quartiles –see Figures 2.1 (c) and (f)–, whereas it increases in a controlled way for the third quartile –see Figure 2.1 (i)–.

2.4 A BS quantile regression model

Let $T_1, \ldots, T_n$ be independent random variables, where $T_i \sim \mathrm{BS}(\alpha, Q_i)$, for $i = 1, \ldots, n$, and $\boldsymbol{t} = (t_1, \ldots, t_n)^\top$ be their associated observations. Then, we define a statistical model based on (2.1) by the systematic component

$$h(Q_i) = \eta_i = \boldsymbol{x}_i^\top \boldsymbol{\beta},\ i = 1, \ldots, n, \tag{2.2}$$

such that $Q_i = h^{-1}(\boldsymbol{x}_i^\top \boldsymbol{\beta})$, where $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_{p-1})^\top$, for $p < n$, is a vector of unknown regression parameters to be estimated, and $\boldsymbol{x}_i^\top = (1, x_{i1}, \ldots, x_{i(p-1)})$ represents the values of p covariates. In the model defined from (2.2), the link function h is invertible, has a positive support and at least twice differentiable. Examples of link functions are $h(u) = \log_k(u)$ and $h(u) = \sqrt[a]{u}$ with a, k being positive integer numbers. Also, can be considered $h(u) = u$, that is, the identity function, with $\mathbb{R}^+$ as domain of h.

The log-likelihood function of the model given in (2.2) for $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \alpha)^\top$ is $\ell(\boldsymbol{\theta}; \boldsymbol{t}) = \ell(\boldsymbol{\theta}) = \sum_{i=1}^n \ell_i(Q_i, \alpha; t_i)$, where

$$\begin{aligned} \ell_i(Q_i, \alpha; t_i) = \ell_i(Q_i, \alpha) \quad &= \quad -\frac{1}{2}\log(8\pi t_i) - \log(\alpha\gamma_\alpha) - \frac{1}{2}\log(Q_i) \\ &\quad + \log\left(\frac{\gamma_\alpha^2}{2} + \frac{2Q_i}{t_i}\right) - \frac{2Q_i}{\alpha^2\gamma_\alpha^2 t_i}\left(\frac{t_i\gamma_\alpha^2}{4Q_i} - 1\right)^2. \end{aligned} \tag{2.3}$$

The score functions for β_j, with $j = 0, 1, \ldots, p-1$, and α are, respectively, expressed as

$$\dot{\ell}_{\beta_j} = \frac{\partial \ell(\boldsymbol{\theta})}{\partial \beta_j} = \sum_{i=1}^n \underbrace{\left(-\frac{1}{2Q_i} - \frac{2}{\alpha^2\gamma_\alpha^2 t_i} + \frac{\gamma_\alpha^2 t_i}{8\alpha^2 Q_i^2} + \frac{4}{t_i\gamma_\alpha^2 + 4Q_i}\right)}_{z_i} \underbrace{\frac{1}{h'(Q_i)}}_{a_i} x_{ij},$$

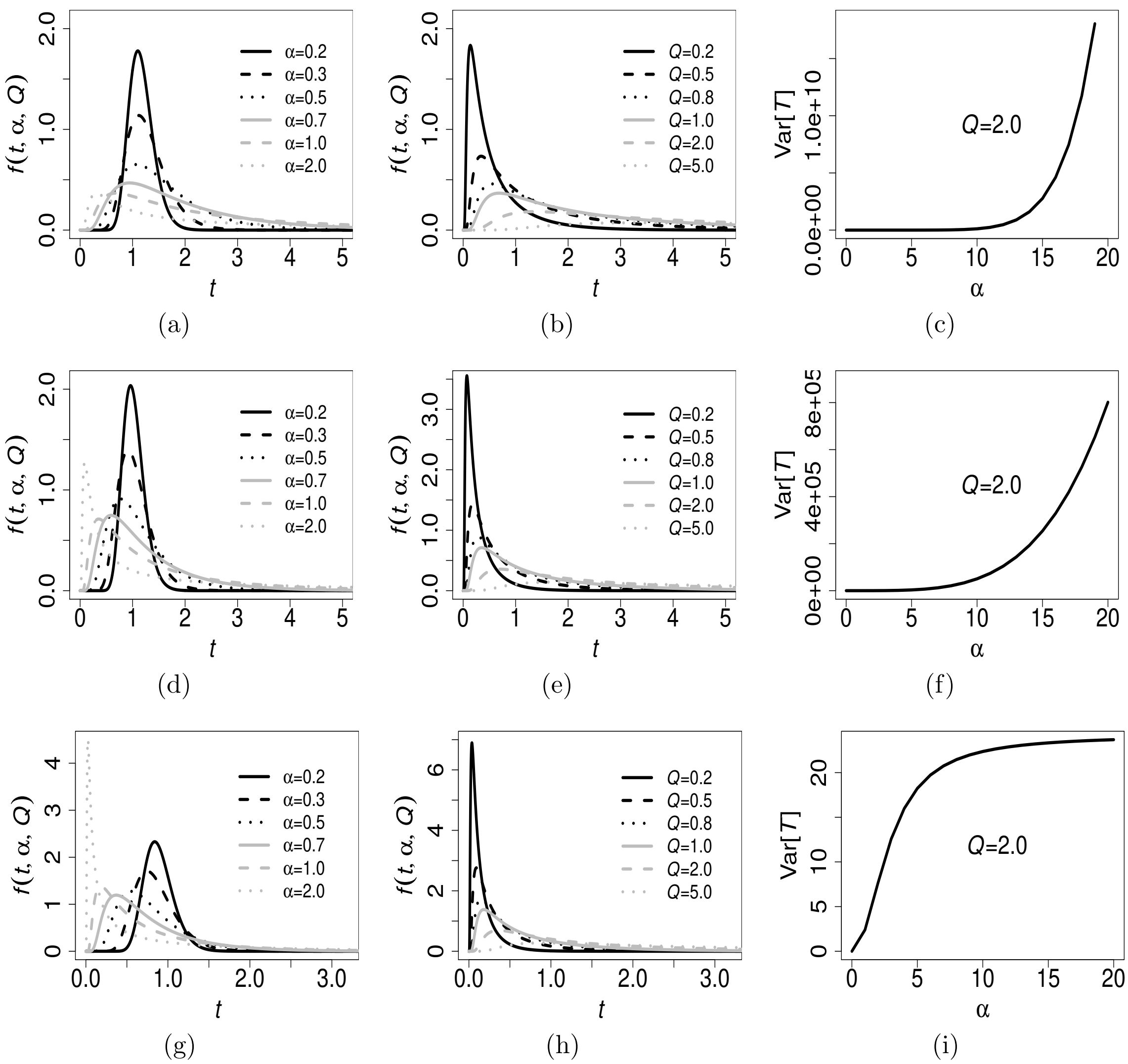


Figure 2.1: Plots of the BS probability density function for $Q = 1.0$ (left) and $\alpha = 1.0$ (center), for $q = 0.25$ (a)-(b); for $q = 0.50$ (d)-(e); as well as for $q = 0.75$ (g)-(h); and of the BS variance against α for $Q = 2.0$ (right) with $q = 0.25$ (c), $q = 0.50$ (f) and $q = 0.75$ (i).

$$\dot{\ell}_\alpha = \frac{\partial \ell(\boldsymbol{\theta})}{\partial \alpha} = \sum_{i=1}^{n} \underbrace{\left[-\frac{(\gamma_\alpha + \alpha\gamma'_\alpha)}{\alpha\gamma_\alpha} + \frac{2t_i\gamma_\alpha\gamma'_\alpha}{t_i\gamma_\alpha^2 + 4Q_i} - \frac{(\gamma_\alpha\gamma'_\alpha\alpha - \gamma_\alpha^2)t_i}{4Q_i\alpha^3} - \frac{2}{\alpha^3} + \frac{4Q_i(\gamma_\alpha + \alpha\gamma'_\alpha)}{(\alpha\gamma_\alpha)^3 t_i} \right]}_{b_i} \tag{2.4}$$

where h' is the derivative of h and γ'_α is the derivative of γ_α. Thus, we can write (2.4) in matrix form as

$$\dot{\boldsymbol{\ell}}_{\boldsymbol{\beta}} = (\dot{\ell}_{\beta_j}) = \boldsymbol{X}^\top \boldsymbol{D}(\boldsymbol{a})\boldsymbol{z}, \quad \dot{\ell}_\alpha = \mathrm{tr}(\boldsymbol{D}(\boldsymbol{b})),$$

where $\boldsymbol{X}^\top = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n)$, with $\boldsymbol{x}_i$ being defined in (2.2), for $i = 1, \ldots, n$, $\boldsymbol{z} = (z_1, \ldots, z_n)^\top$ and $\boldsymbol{D}$ denotes the diagonalization operator of a vector, such that $\boldsymbol{D}(\cdot) = \mathrm{diag}(\cdot)$, with $\boldsymbol{a} = (a_1, \ldots, a_n)^\top$ and $\boldsymbol{b} = (b_1, \ldots, b_n)^\top$. Then, the score vector is $\dot{\boldsymbol{\ell}}_{\boldsymbol{\theta}} = (\dot{\boldsymbol{\ell}}_{\boldsymbol{\beta}}, \ell_\alpha)^\top$.

The elements of the associated Hessian matrix are expressed as

$$\ddot{\ell}_{\beta_l\beta_j} = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\beta_l \partial\beta_j} = \sum_{i=1}^{n} \underbrace{\left[\frac{\partial^2 \ell_i(Q_i, \alpha)}{\partial Q_i^2} \left(\frac{dQ_i}{d\eta_i} \right)^2 + \frac{\partial \ell_i(Q_i, \alpha)}{\partial Q_i} \left(\frac{\partial}{\partial Q_i} \frac{dQ_i}{d\eta_i} \right) \frac{dQ_i}{d\eta_i} \right]}_{c_i} x_{ij} x_{il},$$

where

$$\frac{\partial \ell_i(Q_i, \alpha)}{\partial Q_i} = z_i, \quad \frac{\partial^2 \ell_i(Q_i, \alpha)}{\partial Q_i^2} = \frac{1}{2Q_i^2} - \frac{16}{(t_i\gamma_\alpha^2 + 4Q_i)^2} - \frac{\gamma_\alpha^2 t_i}{4\alpha^2 Q_i^3},$$
$$\frac{\mathrm{d}Q_i}{\mathrm{d}\eta_i} = a_i, \quad \frac{\partial}{\partial Q_i} \left(\frac{\mathrm{d}Q_i}{\mathrm{d}\eta_i} \right) = -\frac{h''(Q_i)}{(h'(Q_i))^2},$$

with h'' being the second derivative of h. Hence, we can group the expressions obtained in matrix form as $\ddot{\boldsymbol{\ell}}_{\boldsymbol{\beta\beta}} = \boldsymbol{X}^\top \boldsymbol{D}(\boldsymbol{c}) \boldsymbol{X}$, where $\boldsymbol{c} = (c_1, \ldots, c_n)^\top$. Furthermore, we have

$$\ddot{\ell}_{\beta_j\alpha} = \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\beta_j \partial\alpha} = \sum_{i=1}^{n} \underbrace{\left[-\frac{8t_i\gamma_\alpha\gamma'_\alpha}{(t_i\gamma_\alpha^2 + 4Q_i)^2} + \frac{(\gamma_\alpha\gamma'_\alpha\alpha - \gamma_\alpha^2)t_i}{4\alpha^3 Q_i^2} + \frac{4(\gamma_\alpha + \alpha\gamma'_\alpha)}{(\alpha\gamma_\alpha)^3 t_i} \right]}_{m_i} a_i x_{ij},$$

which can be represented in matrix form as $\ddot{\boldsymbol{\ell}}_{\boldsymbol{\beta}\alpha} = \boldsymbol{X}^\top \boldsymbol{D}(\boldsymbol{a})\, \boldsymbol{m}$, where $\boldsymbol{m} = (m_1, \ldots, m_n)$. In addition, we get

$$\begin{aligned}
\ddot{\ell}_{\alpha\alpha} &= \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\alpha^2} \\
&= \sum_{i=1}^{n} \underbrace{\left\{ -\mathcal{A} + \mathcal{B}_i \frac{t_i}{(t_i\gamma_\alpha^2 + 4Q_i)^2} + 2[\gamma_\alpha^3\gamma''_\alpha - \gamma_\alpha^2(\gamma'_\alpha)^2] \left(\frac{t_i}{t_i\gamma_\alpha^2 + 4Q_i} \right)^2 - \mathcal{C}_i t_i + \frac{6}{\alpha^4} + \mathcal{D}_i \frac{1}{t_i} \right\}}_{r_i},
\end{aligned}$$

where γ''_α is the second derivative of γ_α and

$$\begin{aligned}
\mathcal{A} &= \frac{(2\gamma'_\alpha + \alpha\gamma''_\alpha)(\alpha\gamma_\alpha) - (\gamma_\alpha + \alpha\gamma'_\alpha)^2}{(\alpha\gamma_\alpha)^2}, \quad \mathcal{B}_i = 8Q_i[(\gamma'_\alpha)^2 + \gamma_\alpha\gamma''_\alpha], \\
\mathcal{C}_i &= \frac{1}{4Q_i} \left[\frac{\alpha^2(\gamma'_\alpha)^2 + \alpha^2\gamma_\alpha\gamma''_\alpha - \alpha\gamma_\alpha\gamma'_\alpha - 3\gamma_\alpha\gamma'_\alpha\alpha + 3\gamma_\alpha^2}{\alpha^4} \right], \\
\mathcal{D}_i &= 4Q_i \left[\frac{(2\gamma'_\alpha + \alpha\gamma''_\alpha)(\alpha\gamma_\alpha) - 3(\gamma_\alpha + \alpha\gamma'_\alpha)^2}{(\alpha\gamma_\alpha)^4} \right].
\end{aligned}$$

In matrix notation, we can write $\ddot{\ell}_{\alpha\alpha} = \mathrm{tr}(\boldsymbol{D}(\boldsymbol{r}))$, where $\boldsymbol{r} = (r_1, \ldots, r_n)^\top$.

The associated expected Fisher information $\boldsymbol{K}_{\boldsymbol{\theta\theta}} = \mathrm{E}[-\ddot{\boldsymbol{\ell}}_{\boldsymbol{\theta\theta}}]$ can be expressed in form matrix as

$$\boldsymbol{K}_{\boldsymbol{\theta\theta}} = \begin{pmatrix} \boldsymbol{K}_{\boldsymbol{\beta\beta}} & \boldsymbol{K}_{\boldsymbol{\beta}\alpha} \\ \boldsymbol{K}_{\alpha\boldsymbol{\beta}} & K_{\alpha\alpha} \end{pmatrix},$$

where $\boldsymbol{K}_{\boldsymbol{\beta\beta}} = \boldsymbol{X}^\top \boldsymbol{D}(\boldsymbol{v})\, \boldsymbol{X}$, with $\boldsymbol{v} = (v_1, \ldots, v_n)^\top$, whose elements are

$$v_i = \left[16V_i(\boldsymbol{\theta}) - \frac{1}{2Q_i^2} + \frac{1}{\alpha^2 Q_i^2}\left(1+\frac{\alpha^2}{2}\right)\right]\frac{1}{[h'(Q_i)]^2} - \left[\frac{1}{2Q_i} + \frac{1}{2\alpha^2 Q_i}\left(1+\frac{\alpha^2}{2}\right) - 4W_i(\boldsymbol{\theta})\right]\cdot\frac{h''(Q_i)}{[h(Q_i)]^3}; \quad (2.5)$$

$\boldsymbol{K}_{\boldsymbol{\beta}\alpha} = \boldsymbol{K}_{\alpha\boldsymbol{\beta}}^\top = \boldsymbol{X}^\top \boldsymbol{D}(\boldsymbol{a})\, \boldsymbol{s}$, with $\boldsymbol{s} = (s_1, \ldots, s_n)$, whose elements are

$$s_i = 8\gamma_\alpha\gamma'_\alpha U_i(\boldsymbol{\theta}) - \frac{(\gamma_\alpha\gamma'_\alpha\alpha - \gamma_\alpha^2)}{\alpha^3\gamma_\alpha^2 Q_i^2}\left(1+\frac{\alpha^2}{2}\right) - \frac{(\gamma_\alpha + \alpha\gamma'_\alpha)}{\alpha^3\gamma_\alpha Q_i}\left(1+\frac{\alpha^2}{2}\right);$$

and $K_{\alpha\alpha} = \mathrm{tr}(\boldsymbol{D}(\boldsymbol{u}))$, for $\boldsymbol{u} = (u_1, \ldots, u_n)^\top$, whose elements are

$$u_i = \mathcal{A} - \mathcal{B}_i U_i(\boldsymbol{\theta}) - 2[\gamma_\alpha^3\gamma''_\alpha - \gamma_\alpha^2(\gamma'_\alpha)^2]S_i(\boldsymbol{\theta}) + \mathcal{C}_i\frac{4Q_i}{\gamma_\alpha^2}\left(1+\frac{\alpha^2}{2}\right) - \frac{6}{\alpha^4} - \mathcal{D}_i\frac{\gamma_\alpha^2}{4Q_i}\left(1+\frac{\alpha^2}{2}\right),$$

with

$$\begin{aligned}
S_i(\boldsymbol{\theta}) &= \int_0^\infty \left(\frac{t}{t\gamma_\alpha^2 + 4Q_i}\right)^2 f_{T_i}(t)\,\mathrm{d}t,\\
U_i(\boldsymbol{\theta}) &= \int_0^\infty \frac{t}{(t\gamma_\alpha^2 + 4Q_i)^2} f_{T_i}(t)\,\mathrm{d}t,\\
V_i(\boldsymbol{\theta}) &= \int_0^\infty \frac{1}{(t\gamma_\alpha^2 + 4Q_i)^2} f_{T_i}(t)\,\mathrm{d}t,\\
W_i(\boldsymbol{\theta}) &= \int_0^\infty \frac{1}{t\gamma_\alpha^2 + 4Q_i} f_{T_i}(t)\,\mathrm{d}t.
\end{aligned}$$

To estimate the model parameters by the ML method, we solve the equations $\dot{\boldsymbol{\ell}} = \boldsymbol{0}$. However, no closed-form expressions for the ML estimates are available. Following the definitions in Leiva et al. (2014b) and Santos-Neto et al. (2016), we can write the iterative algorithm

$$\boldsymbol{\theta}^{(m+1)} = (\tilde{\boldsymbol{X}}^\top \tilde{\boldsymbol{W}}^{(m)} \tilde{\boldsymbol{X}})^{-1} \tilde{\boldsymbol{X}}^\top \tilde{\boldsymbol{W}}^{(m)} \boldsymbol{z}^{*(m)}, \ m = 0, 1, 2, \ldots,$$

which uses the Fisher scoring method, where

$$\begin{aligned}
\tilde{\boldsymbol{X}} &= \begin{pmatrix} \boldsymbol{X} & 0 \\ 0 & 1 \end{pmatrix},\\
\tilde{\boldsymbol{W}} &= \begin{pmatrix} \boldsymbol{D}(\boldsymbol{v}) & \boldsymbol{D}(\boldsymbol{a})\boldsymbol{s} \\ \boldsymbol{s}^\top \boldsymbol{D}(\boldsymbol{a}) & \mathrm{tr}(\boldsymbol{D}(\boldsymbol{u})) \end{pmatrix},\\
\boldsymbol{z}^{*(m)} &= \tilde{\boldsymbol{X}}\boldsymbol{\theta}^{(m)} + (\tilde{\boldsymbol{W}}^{(m)})^{-1} \begin{pmatrix} \boldsymbol{D}(\boldsymbol{a})^{(m)} & 0 \\ 0 & \mathrm{tr}(\boldsymbol{D}(\boldsymbol{b}))^{(m)} \end{pmatrix} \begin{pmatrix} \boldsymbol{z}^{(m)} \\ 1 \end{pmatrix}.
\end{aligned}$$

Under usual regularity conditions, the asymptotic distribution of $\widehat{\boldsymbol{\theta}}$ is $\widehat{\boldsymbol{\theta}} \dot{\sim} \mathrm{N}_{p+1}(\boldsymbol{\theta}, \boldsymbol{\Sigma}_{\boldsymbol{\theta}})$, where $\dot{\sim}$ means 'approximately distributed' and $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$ is the asymptotic variance-covariance matrix of $\widehat{\boldsymbol{\theta}}$; see Cox and Hinkley (1974). This matrix can be obtained by using the inverse expected Fisher information matrix $\boldsymbol{K}_{\boldsymbol{\theta}\boldsymbol{\theta}}^{-1}$ and estimated by replacing $\widehat{\boldsymbol{\theta}}$ at $\boldsymbol{\theta}$. Thus, an approximate $100 \times (1-\xi)\%$, confidence region for $\boldsymbol{\theta}$ is defined as by $(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta})^{\top}\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}}^{-1}(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \leq \chi^2_{1-\xi}(p+1)$, for $\boldsymbol{\theta}$ in $\mathbb{R}^{p+1}$, where $\chi^2_{1-\xi}(p+1)$ is the $(1-\xi) \times 100$th quantile of the chi-squared distribution with $p+1$ degrees of freedom and $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\theta}}$ is a consistent estimator of $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$. Therefore, it is possible to construct asymptotic $100 \times (1-\xi)\%$ confidence bands for the linear predictor $Q(\boldsymbol{x}_{\mathrm{pred}}) = h^{-1}(\boldsymbol{x}_{\mathrm{pred}}^{\top}\boldsymbol{\beta})$, $\forall\, \boldsymbol{x}_{\mathrm{pred}} \in \mathbb{R}^p$, where $\boldsymbol{x}_{\mathrm{pred}}$ is an arbitrary $p \times 1$ vector. Note that the asymptotic distribution of $\widehat{\boldsymbol{\beta}}$ is given by $\widehat{\boldsymbol{\beta}} \dot{\sim} \mathrm{N}_p(\boldsymbol{\beta}, \boldsymbol{\Sigma}_{\boldsymbol{\beta}})$, where $\boldsymbol{\Sigma}_{\boldsymbol{\beta}}$ is the asymptotic variance-covariance matrix of $\widehat{\boldsymbol{\beta}}$, which can be obtained appropriately from $\boldsymbol{\Sigma}_{\boldsymbol{\theta}}$. Then, an approximate $100 \times (1-\xi)\%$ confidence region for $Q(\boldsymbol{x}_{\mathrm{pred}})$ is expressed as

$$\left\{h^{-1}\left[\boldsymbol{x}_{\mathrm{pred}}^{\top}\widehat{\boldsymbol{\beta}} - \sqrt{\chi^2_{1-\xi}(p)}\left(\boldsymbol{x}_{\mathrm{pred}}^{\top}\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\beta}}\,\boldsymbol{x}_{\mathrm{pred}}\right)^{1/2}\right], h^{-1}\left[\boldsymbol{x}_{\mathrm{pred}}^{\top}\widehat{\boldsymbol{\beta}} + \sqrt{\chi^2_{1-\xi}(p)}\left(\boldsymbol{x}_{\mathrm{pred}}^{\top}\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\beta}}\,\boldsymbol{x}_{\mathrm{pred}}\right)^{1/2}\right]\right\},$$

where $\widehat{\boldsymbol{\Sigma}}_{\boldsymbol{\beta}} = (\boldsymbol{X}^{\top}\widehat{\boldsymbol{V}}\boldsymbol{X})^{-1}$, with $\widehat{\boldsymbol{V}} = \boldsymbol{D}(\widehat{\boldsymbol{v}}) - \boldsymbol{D}(\widehat{\boldsymbol{a}})\widehat{\boldsymbol{s}}[\mathrm{tr}(\boldsymbol{D}(\widehat{\boldsymbol{u}}))]^{-1}\widehat{\boldsymbol{s}}^{\top}\boldsymbol{D}(\widehat{\boldsymbol{a}})$, $\boldsymbol{x}_{\mathrm{pred}} \in \mathbb{R}^p$ and $0 < \xi < 1$.

2.5 Local influence analytics

The influence local technique examines the effect of small perturbations in the data and/or the model assumptions on the estimated parameters. The likelihood distance (LD) is expressed by $\mathrm{LD}(\boldsymbol{\omega}) = 2[\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}})]$, where $\widehat{\boldsymbol{\theta}}_{\boldsymbol{\omega}}$ is the ML estimate of $\boldsymbol{\theta}$ for a perturbed model and $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)^{\top}$ is a perturbation vector. Cook (1986) studied the local behavior of $\mathrm{LD}(\boldsymbol{\omega})$ around the non-perturbed vector $\boldsymbol{\omega}_0$, such that $\mathrm{LD}(\boldsymbol{\omega}_0) = 0$. The normal curvature for $\widehat{\boldsymbol{\theta}}$ at the direction $\boldsymbol{d}$, with $\|\boldsymbol{d}\| = 1$, is defined as $C_{\boldsymbol{d}}(\widehat{\boldsymbol{\theta}}) = 2|\boldsymbol{d}^{\top}\boldsymbol{\Delta}^{\top}\ddot{\boldsymbol{\ell}}_{\widehat{\boldsymbol{\theta}}\widehat{\boldsymbol{\theta}}}^{-1}\boldsymbol{\Delta}\boldsymbol{d}|$, where $\ddot{\boldsymbol{\ell}}_{\widehat{\boldsymbol{\theta}}\widehat{\boldsymbol{\theta}}}$ is the Hessian matrix of $\boldsymbol{\ell}(\boldsymbol{\theta})$ evaluated at $\widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\Delta}$ is a $(p+1) \times n$ perturbation matrix also evaluated at $\widehat{\boldsymbol{\theta}}$ and at $\boldsymbol{\omega} = \boldsymbol{\omega}_0$, so that its elements are given by

$$\Delta_{ij} = \left.\frac{\partial^2 \ell_{\boldsymbol{\omega}}(\boldsymbol{\theta})}{\partial\theta_i \partial\omega_j}\right|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}},\, \boldsymbol{\omega}=\boldsymbol{\omega}_0}, \; i = 0, 1, \ldots, p, \; j = 1, \ldots, n, \tag{2.6}$$

and with $\ell_{\boldsymbol{\omega}}(\boldsymbol{\theta})$ being the log-likelihood function associated with the model perturbed by $\boldsymbol{\omega}$. For the model (2.2), the elements of $\ddot{\boldsymbol{\ell}}_{\boldsymbol{\theta}\boldsymbol{\theta}}$ are $\ddot{\boldsymbol{\ell}}_{\boldsymbol{\beta}\boldsymbol{\beta}} = \boldsymbol{X}^{\top}\boldsymbol{D}(\boldsymbol{c})\boldsymbol{X}$, $\ddot{\boldsymbol{\ell}}_{\boldsymbol{\beta}\alpha} = \ddot{\boldsymbol{\ell}}_{\alpha\boldsymbol{\beta}} = \boldsymbol{X}^{\top}\boldsymbol{D}(\boldsymbol{a})\boldsymbol{m}$ and $\ddot{\ell}_{\alpha\alpha} = \mathrm{tr}(\boldsymbol{D}(\boldsymbol{r}))$. We consider the direction $\boldsymbol{d}_{\max}$ as the eigenvector associated with the largest eigenvalue of the matrix

$$\boldsymbol{B} = -\boldsymbol{\Delta}^{\top}\ddot{\boldsymbol{\ell}}_{\widehat{\boldsymbol{\theta}}\widehat{\boldsymbol{\theta}}}^{-1}\boldsymbol{\Delta}. \tag{2.7}$$

The index plot of $\boldsymbol{d}_{\max}$ may be considered to detect cases that are potentially influential on $\widehat{\boldsymbol{\theta}}$. If our interest is only on the vector $\widehat{\boldsymbol{\beta}}$, then the normal curvature at the direction $\boldsymbol{d}$ is given by $C_{\boldsymbol{d}}(\widehat{\boldsymbol{\beta}}) = 2|\boldsymbol{d}^\top \boldsymbol{\Delta}^\top [\ddot{\boldsymbol{\ell}}^{-1}_{\widehat{\boldsymbol{\theta}}\widehat{\boldsymbol{\theta}}} - \ddot{\boldsymbol{\ell}}_1]\boldsymbol{\Delta}\, \boldsymbol{d}|$, where the $(p+1)\times(p+1)$ matrix $\ddot{\boldsymbol{\ell}}_1$ is expressed as

$$\ddot{\boldsymbol{\ell}}_1 = \begin{pmatrix} \mathbf{0} & 0 \\ \mathbf{0} & \ddot{\ell}^{-1}_{\widehat{\alpha}\widehat{\alpha}} \end{pmatrix}.$$

To study local influence on $\widehat{\alpha}$, the normal curvature in the direction of $\boldsymbol{d}$ is defined by $C_{\boldsymbol{d}}(\widehat{\alpha}) = 2|\boldsymbol{d}^\top \boldsymbol{\Delta}^\top (\ddot{\boldsymbol{\ell}}^{-1}_{\widehat{\boldsymbol{\theta}}\widehat{\boldsymbol{\theta}}} - \ddot{\boldsymbol{\ell}}_2)\boldsymbol{\Delta}\boldsymbol{d}|$, where the $(p+1)\times(p+1)$ matrix $\ddot{\boldsymbol{\ell}}_2$ is expressed as

$$\ddot{\boldsymbol{\ell}}_2 = \begin{pmatrix} \ddot{\boldsymbol{\ell}}^{-1}_{\widehat{\boldsymbol{\beta}}\widehat{\boldsymbol{\beta}}} & 0 \\ \mathbf{0} & 0 \end{pmatrix}.$$

The vector $\boldsymbol{d} = \boldsymbol{e}_{in}$, where $\boldsymbol{e}_{in}$ is an $n\times 1$ vector of zeros, with one at the ith position, is other relevant direction. In that case, the normal curvature, called total local influence of the case i, is calculated by $C_i = 2|\boldsymbol{e}_{in}\boldsymbol{B}\boldsymbol{e}_{in}| = 2|B_{ii}|$, where B_{ii} is the ith diagonal element of $\boldsymbol{B}$ defined in (2.7). Lesaffre and Verbeke (1998) proposed to pay attention to those cases with $C_i > 2\,\bar{C}$, where $\bar{C} = \sum_{i=1}^n C_i/n$.

Case-weight perturbation Let $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n)^\top$ be a weight vector. In this case, the perturbed log-likelihood function is defined by $\ell_{\boldsymbol{\omega}}(\boldsymbol{\theta}) = \sum_{i=1}^n \omega_i \ell_i(Q_i, \alpha)$, where $\ell_i(Q_i, \alpha)$ is given in (2.3), with $0 \leq \omega_i \leq 1$, for $i = 1, \ldots, n$. Hence, the perturbation matrix is expressed as $\boldsymbol{\Delta} = [\boldsymbol{\delta}_1, \ldots, \boldsymbol{\delta}_n]$, where for a_i, b_i and z_i are given previously, we have

$$\boldsymbol{\delta}_i = \begin{pmatrix} \boldsymbol{x}_i a_i z_i \\ b_i \end{pmatrix}, \ i = 1, \ldots, n.$$

Note that $\boldsymbol{\Delta}$ must be evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega} = \boldsymbol{\omega}_0 = (1, \ldots, 1)^\top$.

Response perturbation We consider now an additive perturbation on the response i by making $t_i(\omega_i) = t_i + \omega_i s_{T_i}$, where $\omega_i \in \mathbb{R}$ and s_{T_i} is a scale factor often represented by the sample SD of T, for $i = 1, \ldots, n$. Then, under the scheme of response perturbation, the log-likelihood function is given by $\ell_{\boldsymbol{\omega}}(\boldsymbol{\theta}) = \sum_{i=1}^n \ell_{\omega_i}(Q_i, \alpha)$, where

$$\begin{aligned} \ell_{\omega_i}(Q_i, \alpha) &= -\frac{1}{2}\log[8\pi t_i(\omega_i)] - \log(\alpha\gamma_\alpha) - \frac{1}{2}\log(Q_i) \\ &\quad + \log\left[\frac{\gamma_\alpha^2}{2} + \frac{2Q_i}{t_i(\omega_i)}\right] - \frac{2Q_i}{\alpha^2\gamma_\alpha^2 t_i(\omega_i)}\left[\frac{t_i(\omega_i)\gamma_\alpha^2}{4Q_i} - 1\right]^2. \end{aligned}$$

Hence, the column vectors of the matrix $\boldsymbol{\Delta}$ are expressed as

$$\boldsymbol{\delta}_i = \begin{pmatrix} \boldsymbol{x}_i a_i \psi_i \vartheta_i \\ \tau_i \vartheta_i \end{pmatrix}, \ i = 1, \ldots, n,$$

with $\boldsymbol{\Delta}$ being evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega} = \boldsymbol{\omega}_0 = (0, \ldots, 0)^\top$ so that

$$\begin{aligned}
\widehat{\psi}_i &= \frac{2}{\widehat{\alpha}^2\gamma_{\widehat{\alpha}}^2 t_i^2} + \frac{\gamma_{\widehat{\alpha}}^2}{8\widehat{\alpha}^2\widehat{Q}_i^2} - \frac{4\gamma_{\widehat{\alpha}}^2}{(t_i\gamma_{\widehat{\alpha}}^2 + 4\widehat{Q}_i)^2},\\
\widehat{\tau}_i &= 2\gamma_{\widehat{\alpha}}\gamma'_{\widehat{\alpha}}\left[\frac{t_i\gamma_{\widehat{\alpha}}^2 + 4\widehat{Q}_i - \gamma_{\widehat{\alpha}}^2 t_i}{(t_i\gamma_{\widehat{\alpha}}^2 + 4\widehat{Q}_i)^2}\right] - \frac{\gamma_{\widehat{\alpha}}\gamma'_{\widehat{\alpha}}\widehat{\alpha} - \gamma_{\widehat{\alpha}}^2}{4\widehat{Q}_i\widehat{\alpha}^3} - \frac{4\widehat{Q}_i(\gamma_{\widehat{\alpha}} + \widehat{\alpha}\gamma'_{\widehat{\alpha}})}{(\widehat{\alpha}\gamma_{\widehat{\alpha}})^3 t_i^2},\\
\widehat{\vartheta}_i &= s_{T_i},
\end{aligned}$$

where $\gamma_{\widehat{\alpha}}$ is γ_α evaluated at $\widehat{\alpha}$.

Perturbation of a continuos covariate Consider now an additive perturbation on a particular continuous covariate, namely x_t, for $t = 1, \ldots, p-1$, by making $x_{ti}(\omega_i) = x_{ti} + \omega_i s_{X_t}$, where s_{X_t} is a scale factor, which can be the sample SD of X_t, and $\omega_i \in \mathbb{R}$, for $i = 1, \ldots, n$. Then, under the scheme of covariate perturbation, the log-likelihood function is given by $\ell_{\boldsymbol{\omega}}(\boldsymbol{\theta}) = \sum_{i=1}^n \ell_{\omega_i}(Q_i, \alpha)$, where

$$\begin{aligned}
\ell_{\omega_i}(Q_i, \alpha) &= -\frac{1}{2}\log(8\pi t_i) - \log(\alpha\gamma_\alpha) - \frac{1}{2}\log[Q_i(\omega_i)]\\
&\quad + \log\left[\frac{\gamma_\alpha^2}{2} + \frac{2Q_i(\omega_i)}{t_i}\right] - \frac{2Q_i(\omega_i)}{\alpha^2\gamma_\alpha^2 t_i}\left[\frac{t_i\gamma_\alpha^2}{4Q_i(\omega_i)} - 1\right]^2,
\end{aligned}$$

with $Q_i(\omega_i) = h^{-1}[\boldsymbol{x}_i^\top(\omega_i)\boldsymbol{\beta}]$ and $\boldsymbol{x}_i^\top(\omega_i) = (1, x_{i1}, \ldots, x_{ti}(\omega_i), \ldots, x_{i(p-1)})^\top$. Hence, the perturbation matrix assumes the form

$$\boldsymbol{\Delta} = \begin{pmatrix} \boldsymbol{\Delta}_{\boldsymbol{\beta}} \\ \boldsymbol{\Delta}_\alpha \end{pmatrix},$$

where $\boldsymbol{\Delta} = (\Delta_{\beta_{ij}})$ is a $p \times n$ matrix with elements, when $j \neq t$, expressed as

$$\Delta_{\beta_{ij}} = s_X\beta_t a'_i x_{ij} q_i + s_X\beta_t x_{ij} a_i^2\left[\frac{1}{2Q_i^2} - \frac{16}{(t_i\gamma_\alpha^2 + 4Q_i)^2} - \frac{\gamma_\alpha^2 t_i}{4\alpha^2 Q_i^3}\right],$$

with $q_i = z_i$, whereas, when $j = t$, it is given by

$$\Delta_{\beta_{it}} = s_X a_i q_i + s_X\beta_t a'_i x_{it} q_i + s_X\beta_t x_{it} a_i^2\left[\frac{1}{2Q_i^2} - \frac{16}{(t_i\gamma_\alpha^2 + 4Q_i)^2} - \frac{\gamma_\alpha^2 t_i}{4\alpha^2 Q_i^3}\right],$$

where a'_i is the derivative of a_i defined in (2.4). In addition, $\boldsymbol{\Delta}_\alpha = (\zeta_1, \ldots, \zeta_n)$, where $\zeta_i = s_X\beta_t a_i m_i$. Recall that $\boldsymbol{\Delta}$ must be evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega} = \boldsymbol{\omega}_0 = (0, \ldots, 0)^\top$.

Perturbation of the parameter $\boldsymbol{\alpha}$ Here α is perturbed as $\alpha_i = \alpha/\omega_i$, with $\omega_i > 0$ and then the perturbed log-likelihood function is $\ell_{\boldsymbol{\omega}}(\boldsymbol{\theta}) = \sum_{i=1}^n \ell_{\omega_i}(Q_i, \alpha_i)$, where

$$\begin{aligned}
\ell_{\omega_i}(Q_i, \alpha_i) &= -\frac{1}{2}\log(8\pi t_i) - \log(\alpha_i\gamma_{\alpha_i}) - \frac{1}{2}\log(Q_i)\\
&\quad + \log\left(\frac{\gamma_{\alpha_i}^2}{2} + \frac{2Q_i}{t_i}\right) - \frac{2Q_i}{\alpha_i^2\gamma_{\alpha_i}^2 t_i}\left(\frac{t_i\gamma_{\alpha_i}^2}{4Q_i} - 1\right)^2,\ i = 1, \ldots, n.
\end{aligned}$$

Hence, the column vectors of $\boldsymbol{\Delta}$ are expressed as

$$\boldsymbol{\delta}_i = \begin{pmatrix} \boldsymbol{x}_i a_i \varpi_i \\ \varphi_i \end{pmatrix},$$

where $\boldsymbol{\varpi} = (\varpi_1, \ldots, \varpi_n)^\top$ and $\boldsymbol{\varphi} = (\varphi_1, \ldots, \varphi_n)^\top$, with $\varpi_i = -\alpha\, m_i$ and $\varphi_i = -r_i\,\alpha - b_i$, for $i = 1, \ldots, n$, with $\boldsymbol{\Delta}$ being evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ and $\boldsymbol{\omega} = \boldsymbol{\omega}_0 = (1, \ldots, 1)^\top$.

2.6 Simulation study

Note that an important aspect to be considered in all regression is the use of residuals to evaluate the model adequacy. In non-normal models, the identification of a suitable residual is not an easy task. We consider one of the possible simulation studies for our model based on the residuals. Comments on other possible simulation studies regarding our model are mentioned in the final section. Also, note that we are assessing the adequacy of these residuals to our BS quantile regression model and not in a general context. Therefore, this study is relevant and necessary to our model and not to others. Thus, in order to evaluate the fit of the our model to a data set, we consider the four following types of residuals proposed in the literature.

Pearson type residual First, we use a modification of the standardized Pearson residual given by

$$r_i^{(1)} = \frac{t_i - \widehat{Q}_i}{\sqrt{(1-h_i)\widehat{\mathrm{Var}}[T_i]}} = \frac{\widehat{\gamma}_\alpha^2 (t_i - \widehat{Q}_i)}{4\widehat{Q}_i \widehat{\alpha}\sqrt{(1-h_i)(1+5\widehat{\alpha}^2/4)}},\ i = 1, \ldots, n, \tag{2.8}$$

where α and Q_i are defined in (2.1) and (2.2), respectively, and h_i is the ith element of the matrix $\boldsymbol{H} = \boldsymbol{D}(\widehat{\boldsymbol{v}})^{1/2}\boldsymbol{X}(\boldsymbol{X}^\top \boldsymbol{D}(\widehat{\boldsymbol{v}})\boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{D}(\widehat{\boldsymbol{v}})^{1/2}$, which is equivalent to the hat matrix of regression but for generalized linear models.

Deviance type residual Second, we consider a deviance type residual, replacing the mean by the quantile, expressed as

$$r_i^{(2)} = \frac{\mathrm{sgn}(t_i - \widehat{Q}_i)\sqrt{D_i}}{\sqrt{1-h_i}},\ i = 1, \ldots, n, \tag{2.9}$$

where $D_i = 2[\ell(\tilde{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}})]$, with $\tilde{\boldsymbol{\theta}}$ being the ML estimate of $\boldsymbol{\theta}$ under the saturated model (with n parameters), $\widehat{\theta}$ is the ML estimate of $\boldsymbol{\theta}$ under the model of interest (with p parameters) and $\mathrm{sgn}(z)$ denotes the sign of z.

Likelihood residual Third, we derive a likelihood residual, which is a combination of the two previous residuals, given by

$$r_i^{(3)} = \mathrm{sgn}(t_i - \widehat{Q}_i)\left\{h_i[r_i^{(1)}]^2 + (1-h_i)[r_i^{(2)}]^2\right\}^{1/2},\ i = 1, \ldots, n, \tag{2.10}$$

with its elements previously defined.

Randomized quantile residual Fourth, we employ also the randomized quantile residual proposed by Dunn and Smyth (1996). In our case, this residual is given by

$$r_i^{(4)} = \Phi^{-1}(F(t_i; \widehat{\alpha}, \widehat{Q}_i)),\ i = 1, \ldots, n, \tag{2.11}$$

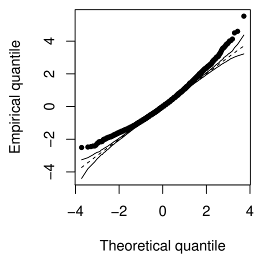
where F is the BS CDF defined in (2.1). This residual follows approximately a normal standard distribution.

A simulation is performed to evaluate the distributions of $r^{(1)}$, $r^{(2)}$, $r^{(3)}$ and $r^{(4)}$ formulated in (2.8), (2.9), (2.10) and (2.11), respectively. In this simulation study, we use the regression model

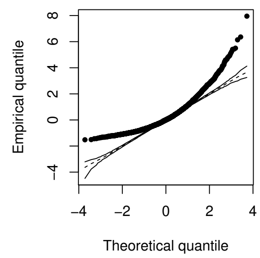
$$h(Q_i) = \beta_0 + \beta_1 x_i,\ i = 1, \ldots, 100, \tag{2.12}$$

considering the logarithm (log) and identity link functions, with $\beta_0 > 0$ and $\beta_1 > 0$ to guarantee $Q_i > 0$ in the case of the identity link. The true values of the parameters are taken as $\beta_0 = 0.3$ and $\beta_1 = 0.5$, whereas $\alpha \in \{0.2, 0.5, 1.0\}$, which indicates low, moderate and high asymmetry. Note that, using the identity link function, we have that $Q_i \in [0.30, 0.80]$, whereas for the log link function, we have that $Q_i \in [1.35, 2.23]$. Other ranges for Q_i need other values for β_0, β_1 and other link functions so that is a limitation of this simulation study. We assume that the values of the covariate X_i are generated from a uniform distribution in the interval $(0, 1)$. The number of Monte Carlo replications is 5000. Using the relation $Q_i = h^{-1}(\beta_0 + \beta_1 x_i)$, we calculate the values of Q_i. In each of the 5000 replications, we obtain the observations $\boldsymbol{t} = (t_1, \ldots, t_{100})^\top$ from the BS distribution with parameters α and Q_i, for $i = 1, \ldots, 100$. Then, the model given in (2.12) is fitted using the implemented functions in the `R` software.

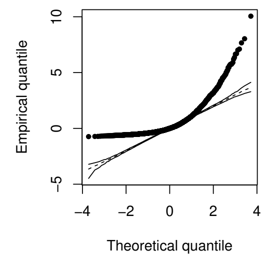
Statistical behavior of $r^{(1)}$, $r^{(2)}$, $r^{(3)}$ and $r^{(4)}$ can be graphically viewed when comparing the empirical distribution of the residuals and the standard normal distribution. We use a quantile against quantile (QQ) plot with simulated envelope to make this comparison; see Atkinson (1985). Figures 2.2 and 2.3 display QQ plots with these envelopes, one for each 5000 residuals $r^{(1)}$, $r^{(2)}$, $r^{(3)}$ and $r^{(4)}$, using log and identity links, with $\alpha \in \{0.2, 0.5, 1.0\}$. We observe that the QQ plot with simulated envelope of $r^{(1)}$ shows that it is further away from the diagonal line and outside of the envelope, which says us that this residual is not suitable for the proposed model. The QQ plots associated with $r^{(2)}$, $r^{(3)}$ and $r^{(4)}$ are adequately over the diagonal and inside of the envelope, indicating that these residuals follows approximately a standard normal distribution, at least when α is in $[0.2; 1.0]$, showing their adequacy for the BS quantile regression model. A study when the number of observations is greater than $n = 100$, namely $n = 200$ and $n = 500$, provides similar results than for $n = 100$. Other study for $q = 0.1$ and $q = 0.9$ with $n = 100$ reports that the residuals $r^{(1)}$, $r^{(2)}$ and $r^{(3)}$ do not follow a normal distribution, but $r^{(4)}$ does. Due to reasons of space, the obtained results are omitted here. Then, we recommend the use of $r^{(4)}$ for this model.



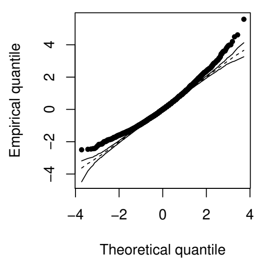
(a) $r^{(1)}$, $h = \log$, $\alpha = 0.2$



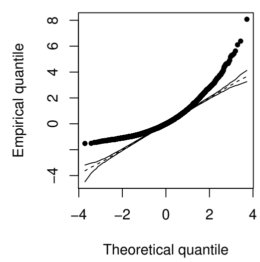
(b) $r^{(1)}$, $h = \log$, $\alpha = 0.5$



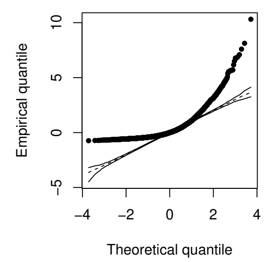
(c) $r^{(1)}$, $h = \log$, $\alpha = 1.0$



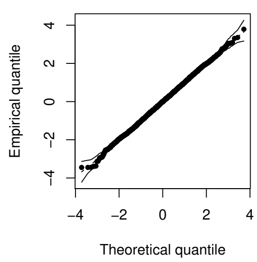
(d) $r^{(1)}$, $h = \mathrm{id}$, $\alpha = 0.2$



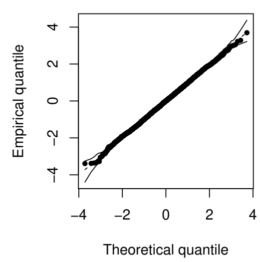
(e) $r^{(1)}$, $h = \mathrm{id}$, $\alpha = 0.5$



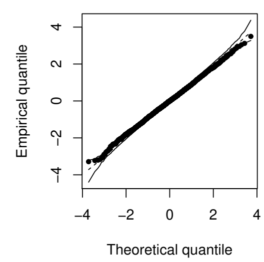
(f) $r^{(1)}$, $h = \mathrm{id}$, $\alpha = 1.0$



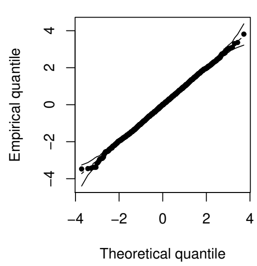
(g) $r^{(2)}$, $h = \log$, $\alpha = 0.2$



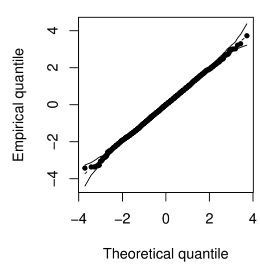
(h) $r^{(2)}$, $h = \log$, $\alpha = 0.5$



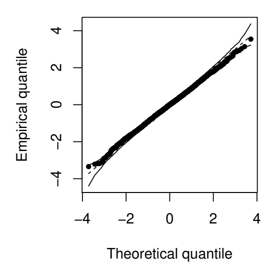
(i) $r^{(2)}$, $h = \log$, $\alpha = 1.0$



(j) $r^{(2)}$, $h = \mathrm{id}$, $\alpha = 0.2$

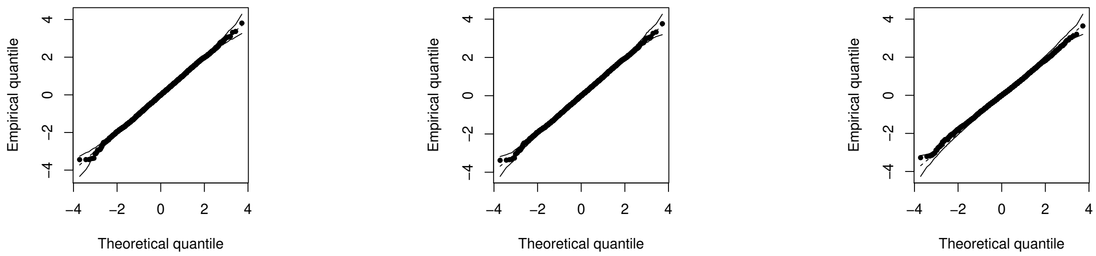


(k) $r^{(2)}$, $h = \mathrm{id}$, $\alpha = 0.5$

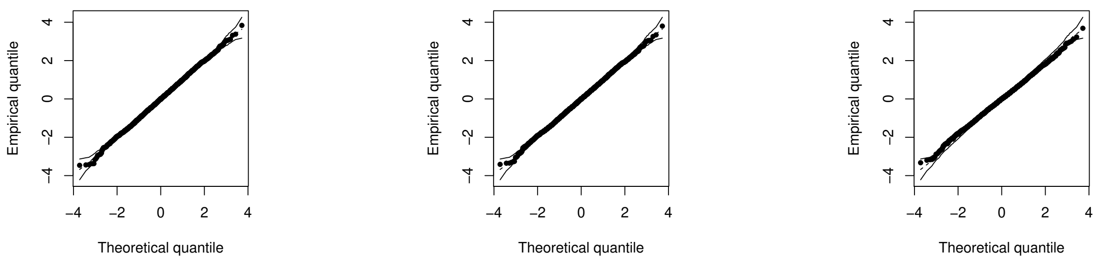


(l) $r^{(2)}$, $h = \mathrm{id}$, $\alpha = 1.0$

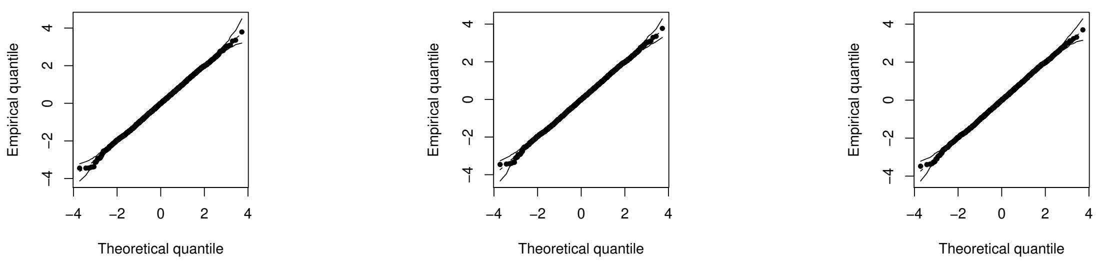
Figure 2.2: Plots of the indicated residual, link function and value of α with simulated data.



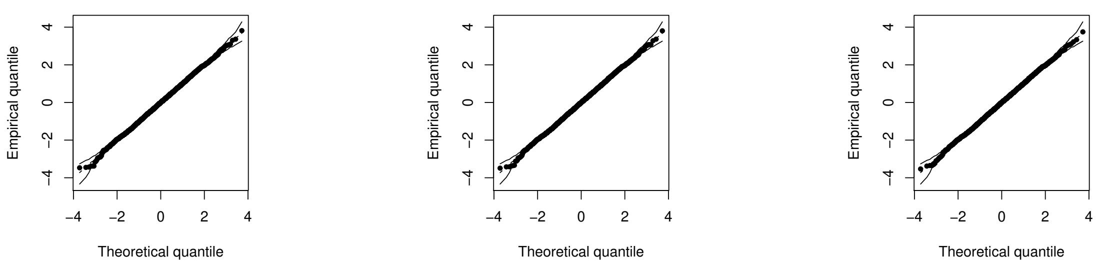
(a) $r^{(3)}$, $h = \log$, $\alpha = 0.2$ (b) $r^{(3)}$, $h = \log$, $\alpha = 0.5$ (c) $r^{(3)}$, $h = \log$, $\alpha = 1.0$



(d) $r^{(3)}$, $h = \mathrm{id}$, $\alpha = 0.2$ (e) $r^{(3)}$, $h = \mathrm{id}$, $\alpha = 0.5$ (f) $r^{(3)}$, $h = \mathrm{id}$, $\alpha = 1.0$



(g) $r^{(4)}$, $h = \log$, $\alpha = 0.2$ (h) $r^{(4)}$, $h = \log$, $\alpha = 0.5$ (i) $r^{(4)}$, $h = \log$, $\alpha = 1.0$



(j) $r^{(4)}$, $h = \mathrm{id}$, $\alpha = 0.2$ (k) $r^{(4)}$, $h = \mathrm{id}$, $\alpha = 0.5$ (l) $r^{(4)}$, $h = \mathrm{id}$, $\alpha = 1.0$

Figure 2.3: Plots of the indicated residual, link function and value of α with simulated data.

2.7 Ilustrative example

In economic scenarios, as the behavior of household income, the data can be follow a skew distribution and (as it is known) the mean is not a good central tendency measure to summarize the data, but, for example, the median is. Also, our interest can be study a quantile of the distribution of the data.

A motivation to consider a BS quantile regression model comes from a real data set corresponding to Chilean household income in the year 2016, collected by the National Institute of Statistics, Chile, which are available at:

`http://www.ine.cl/estadisticas/ingresos-y-gastos/esi/base-de-datos`.

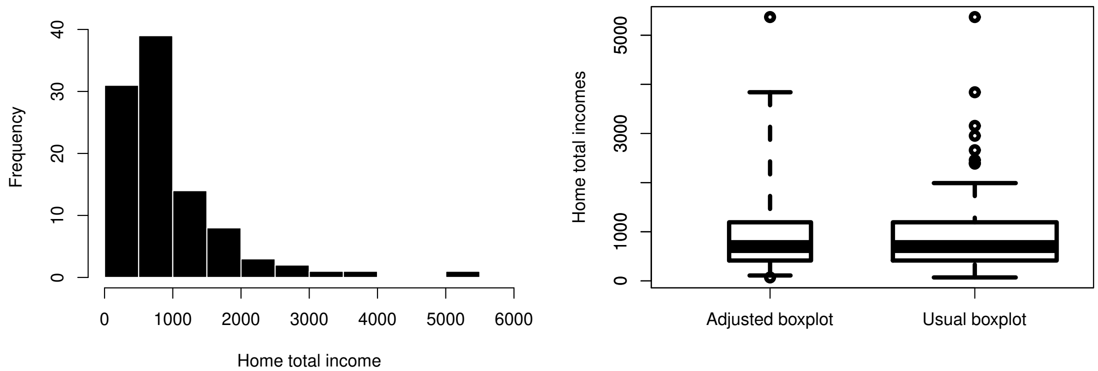
For an illustrative purpose in order to show potential applications of our model, we focus on a data subset which considers $n = 100$ cases randomly selected from the full data set. In this data subset, the response variable (T) is the household income, whereas the covariates to be considered in our analysis are: the total number of persons in the home work force (X_1), the total income due to salaries (X_2), the total income due to independent work (X_3), the total income due to retirements (X_4), the total income due to pensions (X_5), and the total income due to public subsidy (X_6). These covariates were selected from the full data set (which contains 107 variables including $T, X_1, \ldots, X_6$) based on economical and statistical criteria in relation to the response variable. All incomes are expressed in thousands of Chilean pesos; see `http://www.bancocentral.cl` for its equivalence in American dollars.

Table 2.1 provides a descriptive summary of the household income that includes sample median, mean, standard deviation (SD), coefficients of variation (CV), skewness (CS) and kurtosis (CK), as well as minimum ($t_{(1)}$) and maximum ($t_{(n)}$) values. Figure 3.2 shows the histogram as well as usual and adjusted boxplots of the household income; for details of the adjusted boxplot for asymmetric data, see Rousseeuw et al. (2016). Also, Figure 2.5 displays scatterplots of household income and each one of the covariates.

Table 2.1: Descriptive statistics for Chilean income data (in thousands of Chilean pesos).

median	mean	SD	CV	CS	CK	$t_{(1)}$	$t_{(n)}$	n
698.80	938.10	837.52	0.89	2.45	11.03	70	5369.90	100

Based on Table 2.1 and Figures 3.2-2.5, observe the following aspects. First, from the histogram displayed in Figure 3.2(a), note that the values of the household income have an empirical distribution which is unimodal and positively skewed, justifying the use of an asymmetric distribution for the response variable. Second, the boxplots presented in Figure 3.2(b) show some atypical cases for the household income which



(a) (b)

Figure 2.4: Histogram (a) and boxplots (b) for Chilean income data.

we analyze in next subsection. Third, from the scatterplots and correlations displayed in Figure 2.5, linear relationships between some variables are detected as well as some evidence of non-constant variance for T. Note that X_5 and X_6 have a low correlation with T, reason why we discard them in our illustrative data analysis. In addition, we detect a relatively high correlation between X_1 and X_2 so that we discard X_1 as well due to possible collinearity problems and also because it has less correlation with T than X_2. Therefore, we propose to use X_2, X_3 and X_4 for illustrating the BS quantile regression model, which has characteristics and properties suitable for describing the median of the data, the non-constant variance and the asymmetry detected in these data.

Based on previous subsection, we assume the response $T_i \sim \mathrm{BS}(\alpha, Q_i)$. We consider the logarithm, square root and identity link functions for the systematic component of the regression model on the median, which are expressed as: (L1) $\log(Q_i) = \boldsymbol{x}_i^\top \boldsymbol{\beta}$; (L2) $\sqrt{Q_i} = \boldsymbol{x}_i^\top \boldsymbol{\beta}$; (L3) $Q_i = \boldsymbol{x}_i^\top \boldsymbol{\beta}$, with $\boldsymbol{\beta} > 0$; for $i = 1, \ldots, 100$, where $\boldsymbol{\beta} = (\beta_0, \beta_2, \beta_3, \beta_4)^\top$ is the regression coefficient vector and $\boldsymbol{x}_i^\top = (1, x_{i2}, x_{i3}, x_{i4})$ is the observed value of $\boldsymbol{X}_i$. We emphasize that the median regression is used as it is a robust measure of centrality for right-skewed distributions; see Hao and Naiman (2007, p. 57) and Davino et al. (2014, p. 76). We fit the BS model by using the command `bsreg.fit()` that we have implemented in the `R` software. The values of the corrected Akaike information criterion (AIC) and of the log-likelihood function (in parenthesis) for the model with indicated link functions are: (L1) 1502.116 (-745.739); (L2) 1411.67 (-700.516); and (L3) 1396.313 (-692.837). With these values, we establish that the identity link function should be used in the modeling, as conjectured in Subsection **??**. The ML estimates for the model parameters with link function (L3), approximate estimated standard errors (SEs) and their significance at 5% are reported in Table 2.2, where X_2, X_3 and X_4 are identified as significant to explain T. Therefore, we consider

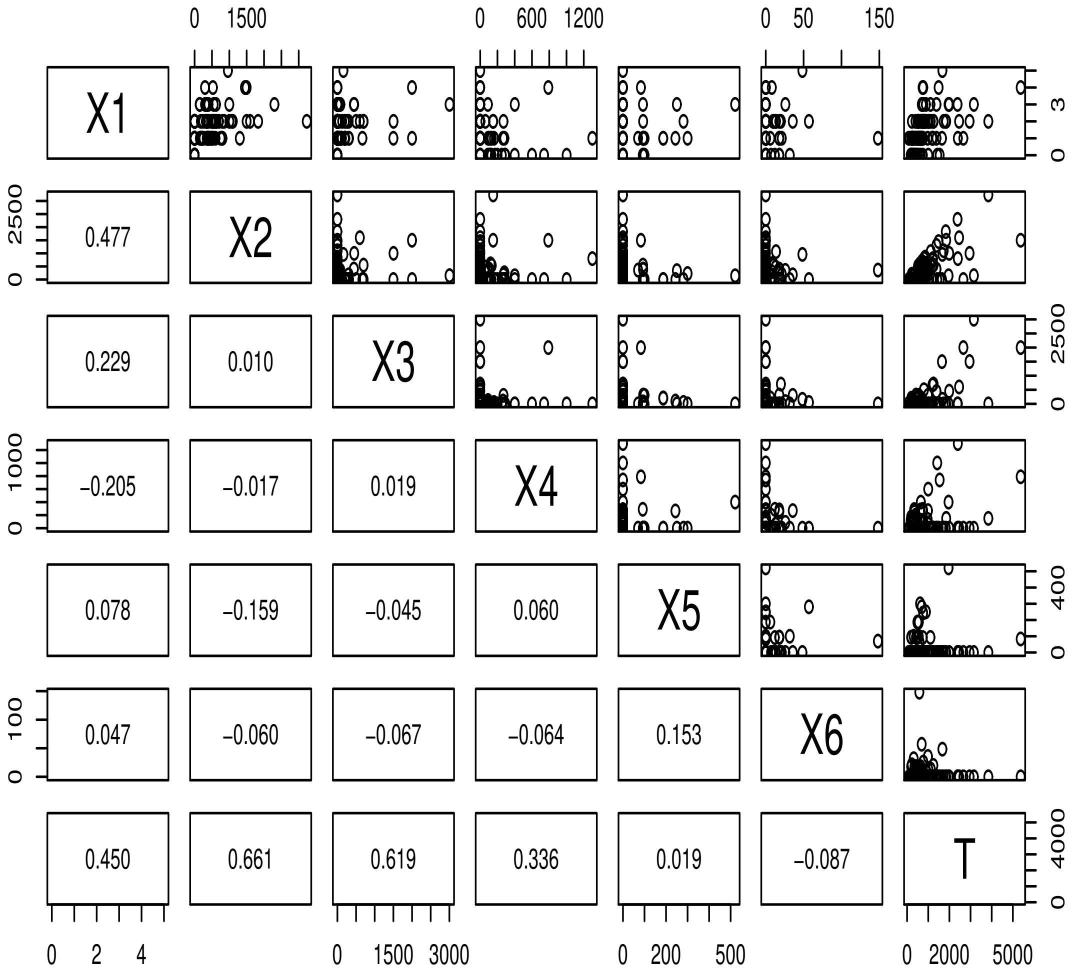


Figure 2.5: Scatters plots and correlations between each pair of variables for Chilean income data.

the model given by

$$Q_i = \beta_0 + \beta_1 x_{2i} + \beta_2 x_{3i} + \beta_3 x_{4i},\ i = 1, \ldots, 100. \tag{2.13}$$

Additionally, we estimate the model for $q = 0.1$, 0.25, 0.75 and 0.9 obtaining satisfactory fittings such as with the median so that, due to reasons of space, the obtained results are omitted here. Therefore, we decide to continue our illustration with the BS-median model.

Distributional assumption of the model given in (2.13) is verified by the QQ plot with envelope for the residual $r^{(4)}$ in Figure 2.6(a). This figure shows that the residuals follow approximately a standard normal distribution so that the assumption that the response variable follows a BS distribution does not seem to be unsuitable. In addition, no unusual features are detected by the plot of residuals presented in Figure 2.6(b), solving the problem of non-constant variance detected in subsection previous, but five outlying cases (#13, #27, #32, #80 and #87) are identified, which are analyzed in the diagnostic study presented next. When comparing our model with the normal regression model (where as known the mean is equal to the median), a better performance is detected in favor of the BS quantile regression model, based on the QQ plots with envelope for the residuals displayed in Figure 2.6(a) and (b) (in the case of the normal regression model, we use the usual standardized Pearson residual). Comparison of our model with other similar models based on, for instance, the gamma, lognormal or Weibull distributions (Noufaily and Jones, 2013) is not possible because this implies to derive models using such distributions with identical parameterizations to that used in our approach, which are not available in the literature.

Suitability of the identity link function used in the model given in (2.13) is verified by employing $\boldsymbol{z}_2 = \widehat{\boldsymbol{\eta}} + \widehat{\boldsymbol{v}}^* \odot \widehat{\boldsymbol{z}}$, such as in Leiva et al. (2014b), where $\widehat{\boldsymbol{v}}^* = (1/\widehat{v}_1, \ldots, 1/\widehat{v}_n)^\top$, with v_i being defined in (2.5) and $\odot$ being the Hadamard product. The plot of $\widehat{z}_{2i}$ against $\widehat{\eta}_i$ is utilized to verify the adequacy of the link function, where a linear tendency is requested. According to Figure 2.6(c), it is possible to note such a linear tendency, and therefore the identity function link is suitable for our model.

Table 2.2: Estimate, SE and significance at 5% of the indicated parameter for Chilean income data.

	$\widehat{\beta}_0$	$\widehat{\beta}_2$	$\widehat{\beta}_3$	$\widehat{\beta}_4$	$\widehat{\alpha}$
Estimate	198.0903	1.0440	1.1090	1.0865	0.3646
SE	22.3166	0.0871	0.1502	0.1759	0.0087
Significance	yes	yes	yes	yes	yes

Diagnostics based on the local influence technique for the BS quantile regression model given in (2.2) are displayed in Figure 2.7, which shows index plots of C_i. From there, cases #13, #27, #32, #80 and #87 are detected as potentially influential. Also, we analyze index plots of $\boldsymbol{l}_{\max}$, but the results are similar to those presented for the index plots of C_i, so that we omit these results here. Figure 2.7 presents plots of C_i against X_{2i}, which indicate that small values of X_2 have a moderate influence on the estimates; see for example case #87. Impact on the model inference is analyzed for three cases (#13, #80 and #87) identified as more potentially influential in the diagnostic analytics. Then, we remove the sets of cases {#13}, {#80}, {#87}, {#13, #80}, {#13,#87}, {#80,#87},{#13,#80,#87} and re-estimate the model parameters. Relative changes (RCs) in the parameter estimates

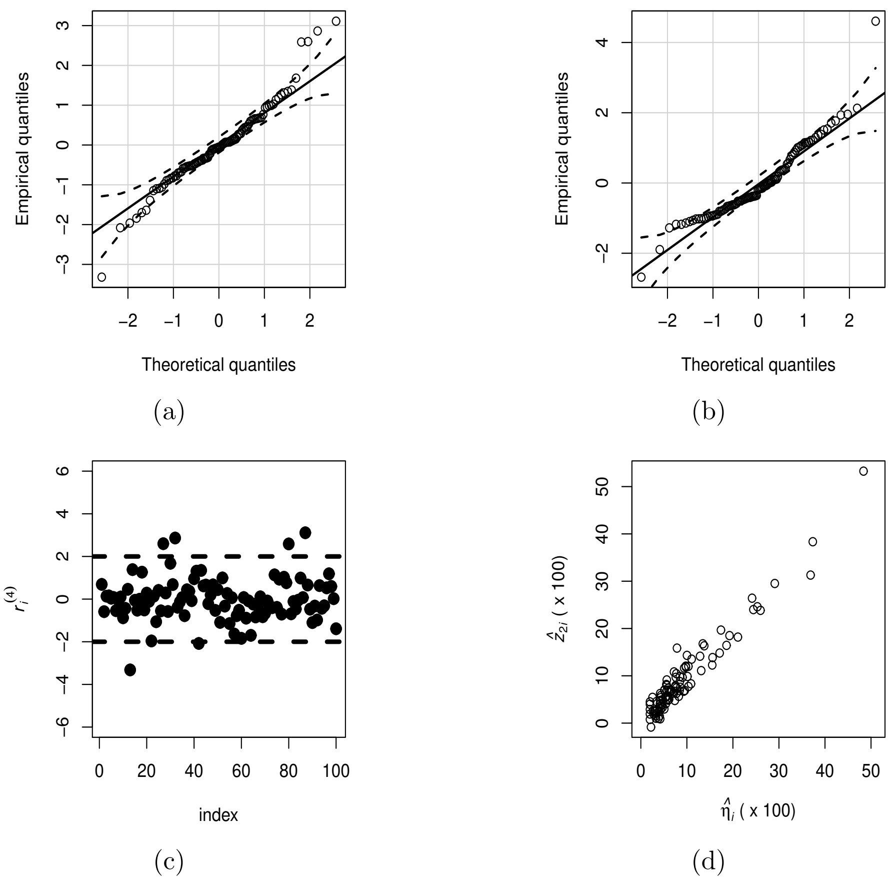


Figure 2.6: QQ plot with envelope of the residual $r^{(4)}$ for the BS median model (a) and of the standardized residual for the normal regression model (b); index plot of $r^{(4)}$ (c) and plot of $\widehat{z}_2$ against $\widehat{\eta}$ for the model fit with identity link (d) based on Chilean income data and the BS quantile regression model.

and in their associated estimated SEs, by using the income data, are calculated as

$$\mathrm{RC}_{\theta_{j(i)}} = \left|\frac{\widehat{\theta}_j - \widehat{\theta}_{j(i)}}{\widehat{\theta}_j}\right| \times 100\%, \quad \mathrm{RC}_{\mathrm{SE}(\widehat{\theta}_j)_{(i)}} = \left|\frac{\widehat{\mathrm{SE}}(\widehat{\theta}_j) - \widehat{\mathrm{SE}}(\widehat{\theta}_j)_{(i)}}{\widehat{\mathrm{SE}}(\widehat{\theta}_j)}\right| \times 100\%,$$

where $\widehat{\theta}_{j(i)}$ and $\widehat{\mathrm{SE}}(\widehat{\theta}_j)_{(i)}$ denote the ML estimates of θ_j and of the estimated SE of the associated estimator, respectively, obtained after removing case i, for $j = 1, \ldots, 5$ and $i = 1, \ldots, 100$, with $\theta_1 = \beta_0, \theta_2 = \beta_2, \theta_3 = \beta_3, \theta_4 = \beta_4, \theta_5 = \alpha$. Table 2.3 reports these RCs, from where the largest values of RCs are identifed when removing simultaneously the cases #80 and #87, which influences importantly on all parameters, with RCs until approximately 21%. However, no inferential changes are found. The results presented in this table show that the diagnostic measures derived in this study identify potentially influential cases, but these do not affect the inference of the model. We can conclude that the diagnostic analytics based on the local influence technique confirm that the BS quantile regression model presented in (2.13) is stable to the atypical cases detected and suitable for modeling the income data.

Table 2.3: RCs (in %) of ML estimates and of the associated estimated SEs for the indicated removed case(s), and respective p-values using Chilean income data and the BS quantile regression model.

Removed cases		β_0	β_1	β_2	β_3	α
None	RC($\widehat{\theta}$)	-	-	-	-	-
	RC($\widehat{\text{SE}}$)	-	-	-	-	-
	p-value	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
$\{13\}$	RC($\widehat{\theta}$)	9.46	2.98	2.41	4.74	5.63
	RC($\widehat{\text{SE}}$)	3.41	4.66	4.99	4.84	9.80
	p-value	< 0.01	< 0.01	< 0.01	0.01	< 0.01
$\{80\}$	RC($\widehat{\theta}$)	7.61	2.4	3.59	3.85	3.24
	RC($\widehat{\text{SE}}$)	3.58	2.7	2.04	2.77	5.48
	p-value	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
$\{87\}$	RC($\widehat{\theta}$)	10.19	3.22	4.84	5.17	5.00
	RC($\widehat{\text{SE}}$)	7.01	4.57	3.65	4.64	8.68
	p-value	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
$\{13, 80\}$	RC($\widehat{\theta}$)	1.51	0.53	0.98	0.88	8.37
	RC($\widehat{\text{SE}}$)	0.52	6.80	6.64	7.02	14.23
	p-value	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
$\{13, 87\}$	RC($\widehat{\theta}$)	0.50	0.08	1.9	0.14	10.04
	RC($\widehat{\text{SE}}$)	2.48	8.57	8.2	8.76	17.14
	p-value	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
$\{80, 87\}$	RC($\widehat{\theta}$)	20.92	6.70	10.31	10.87	9.84
	RC($\widehat{\text{SE}}$)	13.54	8.95	7.06	9.10	16.80
	p-value	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01
$\{13, 80, 87\}$	RC($\widehat{\theta}$)	11.27	3.48	7.00	5.67	14.09
	RC($\widehat{\text{SE}}$)	7.65	12.06	10.86	12.24	23.66
	p-value	< 0.01	< 0.01	< 0.01	< 0.01	< 0.01

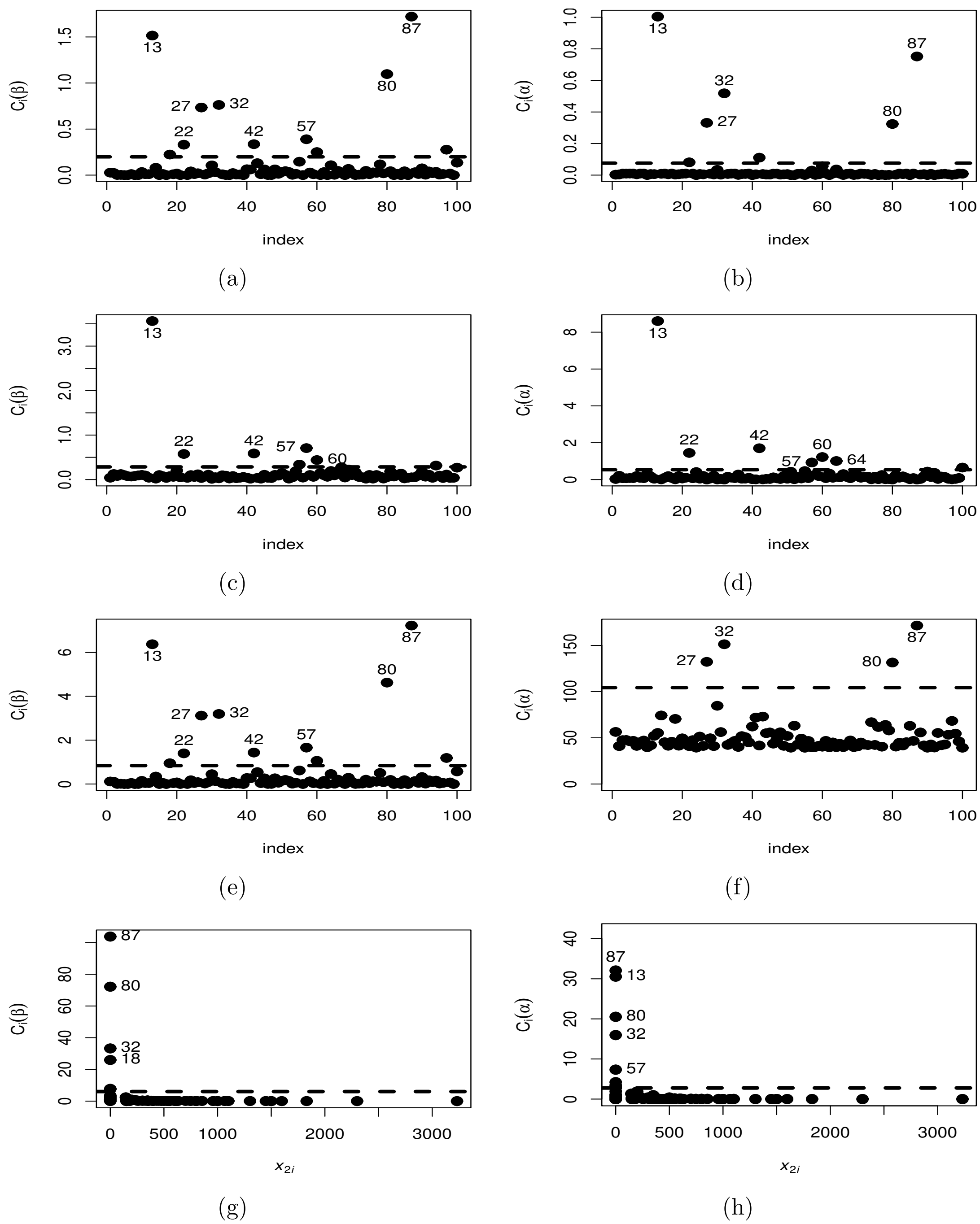


Figure 2.7: Index plots of C_i for $\boldsymbol{\beta}$ (a) and α (b) under case-weight perturbation; for $\boldsymbol{\beta}$ (c) and α (d) under response perturbation; for $\boldsymbol{\beta}$ (e) and α (f) under perturbation of the parameter α; for $\boldsymbol{\beta}$ (g) and α (h) under covariate perturbation X_2, using Chilean income data and the BS quantile regression model.

2.8 Economic implications

Quantile regression allows us to carry out a deeper analysis of the determinants of household income when compared to traditional ordinary least squares (OLS) regression, since such determinants may have different magnitude across income strata. Table 2.4 reports the estimated coefficients for various BS quantile regression models as well as for the OLS regression. In a quantile regression, the estimated coefficient is interpreted as the change in the $q \times 100$th percentile of the household income corresponding to a unit change in the covariate, whereas, in the OLS regression, the change is in the mean household income; see Hao and Naiman (2007, p. 57).

The results of Table 2.4 report that all the covariates affect positively the household income, as expected. Note that the effects of all the covariates increase with the household income (higher quantiles) in the BS quantile regression model. For example, an increase of one thousand Chilean pesos in salaries (X_2), increases the 10th percentile ($q = 0.10$) of the houlsehold income by an amount of \$657.1 Chilean pesos. In addition, it increases by \$1,659.1 Chilean pesos the 90th percentile ($q = 0.90$) of the houlsehold income. In other words, the effect of the salaries increases for individuals with higher household income (higher quantiles). As mentioned, see `http://www.bancocentral.cl` for the equivalence between Chilean pesos and American dollars.

Note also from Table 2.4 that the estimated coefficient for the total income due to independent work (X_3) in the BS quantile regression model for $q = 0.25$ is 0.8681, which is less than the estimated coefficient in the mean regression model (OLS), which is 0.9918. This suggests that while an increase of one thousand Chilean pesos of income due to pensions gives rise to an average increase of \$991.8 in household income. Observe that the increase would not be substantial for most of the population. Similarly, the estimated coefficient for the total income due to retirements (X_4) in the BS median regression model is 1.0865, which is greater than the corresponding estimated coefficient in the mean regression model.

In general, we can conclude that economic analyses are more informative using quantile regression. In the present example, the BS quantile regression model provided a thorough tool to analyses income data.

Table 2.4: OLS regression and BS quantile regression estimated coefficients (p-values in parentheses) at five different quantiles with household income data.

	BS quantile regression					OLS
	$q = 0.10$	$q = 0.25$	$q = 0.50$	$q = 0.75$	$q = 0.90$	
$\widehat{\beta}_0$	124.7264	155.0139	198.0903	251.4492	314.7382	126.7358
	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(<0.001)	(< 0.001)
$\widehat{\beta}_2$	0.6571	0.8170	1.0440	1.3369	1.6591	49.9244
	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)	(< 0.001)
$\widehat{\beta}_3$	0.6974	0.8681	1.1090	1.4218	1.7629	0.9918
	(<0.001)	(<0.001)	(<0.001)	(< 0.001)	(<0.001)	(< 0.001)
$\widehat{\beta}_4$	0.6840	0.8498	1.0865	1.3923	1.7259	1.0689
	(<0.001)	(<0.001)	(<0.001)	(< 0.001)	(<0.001)	(< 0.001)

2.9 Concluding remarks

In this chapter, we proposed an approach to quantile regression modeling based on the BS distribution. This approach used a quantile parameterization, which allowed us to consider a similar framework to generalized linear models, providing flexibility in the modeling. The ML method was considered for estimating the model parameters and for carrying out inference on these parameters. A diagnostic analytics based on the local influence technique under different perturbation schemes and a residual analysis were also conducted. The performance of four types of residuals was evaluated by simulations, with a randomized quantile residual being identified as the most suitable for our model. A data analysis of the new model was performed with Chilean household income data. Comparison of our model with the normal regression model reported a better performance in favor of the BS quantile regression model. This analysis showed an adequate performance of the approach, providing evidence that the BS distribution is a good modeling alternative for dealing with positive data and an asymmetric behavior, as the income data.

Chapter 3

BS spatial quantile regression models

3.1 Summary

In spatial regression, the mean of the response variable is modeled using explanatory variables. Typically, this modeling considers Gaussianity assuming the response follows a symmetric distribution. However, when this assumption is not satisfied, it is useful to suppose distributions with the same asymmetric behavior of the data, such as the BS distribution. In this chapter, we propose a geostatistical model based on BS quantile regression and its diagnostics by using the global and local influence. A new quantile parameterization is proposed here. The estimation of parameters and its local influence are conducted by the ML method. We consider global influence based on the Cook distance to compare with the local influence, in both cases to detect influential observations, whose detection and removal can modify the conclusions of a study. We illustrate the proposed methodology applying it to environmental data, which shows this situation. A comparison with Gaussian spatial regression is also conducted.

3.2 Introduction

Standard regression models describe the mean response given certain values of the explanatory variables. Nevertheless, if the response variable has a asymmetrical distribution, the mean is not a suitable centrality measure to summarize the data. Quantile regression was proposed by Koenker and Bassett (1978), extending the median regression model to the ordinary quantiles by means of the regression context. We focus and propose to model the median or other quantiles of the BS distribution by regression. Considering a spatial component in the modeling may improve the accuracy of an estimator of the mean (or median); see Diggle and Ribeiro (2007). A first idea of spatial quantile regression was suggested by Kostov (2009). Trzpiot (2013) derived a spatial regression model using the quantile function; see McMillen (2013) for some variants of spatial quantile regressions. Garcia-Papani et al. (2017, 2018a,b) introduced BS spatial models for the mean, which need multivariate BS distributions; see, for example, Kundu (2015b), Lemonte et al. (2015), Sánchez et al. (2015), Marchant et al. (2016a,b) and Garcia-Papani et al. (2017, 2018a,b). In the best of our knowledge, spatial BS quantile regression and its local influence diagnostics not have been formulated to the date.

As we have said, diagnostic analytics have a important part in statistical modeling. The Cook distance and residuals are well-known and often used as measures of global influence for detecting the model adequacy; see Krzanowski (1998) and Leiva et al. (2016). However, the local influence technique is very common to detect potentially influential cases in different models; see, for example, Díaz-García et al. (2003), Santana et al. (2011) and Garcia-Papani et al. (2017). Detection and removal of potentially influential cases can modify the conclusions of a study.

This chapter has as objective to propose a geostatistical model based on Birnbaum-Saunders quantile regression and its global and local influence diagnostics. We use a new quantile parameterization to generate the model, which permits us to consider a similar framework to generalized linear models, providing wide flexibility. A comparison with Gaussian spatial regression is performed, but with other natural competing models, as gamma, lognormal or Weibull, it is not possible because such models based on our new parameterization are not available in the literature.

In Section 3.3, we establish a new parameterization of it to model a quantile. Section 3.4 proposes the model and provides estimation of its parameters based on the ML method. In Section 3.5, tools for model checking are discussed. In Section 3.6, we derive global influence measures based on the Cook distance for detecting influential potentially observations. Section 3.7 introduces the local influence technique for the new model including two schemes of perturbation. Next, we illustrate the proposed methodology in Section 3.8 considering an example related to environmental data. Some conclusions are given in Section 3.9. All the numerical calculations were carried out with the aid of the `R` software; see R-Team (2018). Mathematical details of some results are provided in the appendix.

3.3 A parameterization of the multivariate BS distribution

Let $\boldsymbol{T} = (T_1, \ldots, T_n) \sim \mathrm{BS}_n(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Gamma})$ and $q \in (0, 1)$ is a fixed value. Then, we have a new parameterization of the n-variate BS distribution by the transformation expressed as $(\boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Gamma}) \mapsto (\boldsymbol{\alpha}, \boldsymbol{Q}, \boldsymbol{\Gamma})$, where Q_i and β_i are related by (1.2), for the marginal distribution of T_i, Q_i being the q-th quantile of the $\mathrm{BS}(\alpha_i, \beta_i)$ distribution, for all $i = 1, \ldots, n$. This new parameterization of the n-variate BS distribution is denoted by $\boldsymbol{T} \sim \mathrm{BS}_n(\boldsymbol{\alpha}, \boldsymbol{Q}, \boldsymbol{\Gamma})$, with distribution function and density given, respectively, by

$$\begin{aligned} F_{\boldsymbol{T}}(\boldsymbol{t}; \boldsymbol{\alpha}, \boldsymbol{Q}, \boldsymbol{\Gamma}) &= \Phi_n(\tilde{\boldsymbol{A}}; \boldsymbol{\Gamma}), \quad \boldsymbol{t} = (t_1, \ldots, t_n) \in \mathbb{R}_+^n, \\ f_{\boldsymbol{T}}(\boldsymbol{t}; \boldsymbol{\alpha}, \boldsymbol{Q}, \boldsymbol{\Gamma}) &= \phi_n(\tilde{\boldsymbol{A}}; \boldsymbol{\Gamma})\, \tilde{a}(\boldsymbol{t}; \boldsymbol{\alpha}, \boldsymbol{Q}), \quad \boldsymbol{t} = (t_1, \ldots, t_n) \in \mathbb{R}_+^n, \end{aligned} \tag{3.1}$$

where $\tilde{\boldsymbol{A}} = (\tilde{A}_1, \ldots, \tilde{A}_n)^\top$, for

$$\tilde{A}_j = \frac{1}{\alpha_j \gamma_{\alpha_j}} \sqrt{\frac{4Q_j}{t_j}} \left(\frac{t_j \gamma_{\alpha_j}^2}{4Q_j} - 1 \right), \quad \tilde{a}(\boldsymbol{t}; \boldsymbol{\alpha}, \boldsymbol{Q}) = \prod_{j=1}^{n} \frac{1}{\alpha_j \gamma_{\alpha_j} \sqrt{4Q_j t_j}} \left(\frac{\gamma_{\alpha_j}^2}{2} + \frac{2Q_j}{t_j} \right),$$

and γ_{α_j} being defined in (1.3).

Theorem Let $\boldsymbol{T} = (T_1, \ldots, T_n) \sim \mathrm{BS}_n(\boldsymbol{\alpha}, \boldsymbol{Q}, \boldsymbol{\Gamma})$, with $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_n)$ and $\boldsymbol{Q} = (Q_1, \ldots, Q_n)$. Then,

(i) $T_i \sim \mathrm{BS}(\alpha_i, Q_i)$, for $i = 1, \ldots, n$.

(ii) $(T_i, T_j) \sim \mathrm{BS}_2(\boldsymbol{\alpha}_{ij}, \boldsymbol{Q}_{ij}, \boldsymbol{\Gamma}_{ij})$, where $\boldsymbol{\alpha}_{ij} = (\alpha_i, \alpha_j)$, $\boldsymbol{Q}_{ij} = (Q_i, Q_j)$ and $\boldsymbol{\Gamma}_{ij}$ is the 2×2 matrix with diagonal of ones and each one of its other elements equal to element (i, j) of the matrix $\boldsymbol{\Gamma}$.

Proof Both results are deduced from Kundu et al. (2013) and using the new parametrization of the BS distribution. □

Theorem Let $\boldsymbol{T} = (T_1, T_2) \sim \mathrm{BS}_2(\boldsymbol{\alpha}, \boldsymbol{Q}, \boldsymbol{\Gamma})$, with $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$, $\boldsymbol{Q} = (Q_1, Q_2)$ and $\boldsymbol{\Gamma} = (\rho_{ij})$. Then,

(i) $\mathrm{E}(T_1 T_2) = \dfrac{4Q_1 Q_2}{\gamma_{\alpha_1}^2 \gamma_{\alpha_2}^2}(4 + 2(\alpha_1^2 + \alpha_2^2) + \alpha_1^2 \alpha_2^2 (1 + 2\rho_{12}^2)).$

(ii) $\mathrm{Cov}(T_1, T_2) = 8\rho_{12}^2 \, \dfrac{\alpha_1^2 \alpha_2^2 \, Q_1 Q_2}{\gamma_{\alpha_1}^2 \gamma_{\alpha_2}^2}.$

(iii) $\mathrm{Corr}(T_1, T_2) = 2\rho_{12}^2 \, \dfrac{\alpha_1 \alpha_2}{\sqrt{4 + 5\alpha_1^2} \sqrt{4 + 5\alpha_2^2}}.$

Proof These three results are obtained from Saulo et al. (2019) and using the new parametrization of the BS distribution.□

From Figure 3.1, note that some shapes of the density expressed in (3.1) with $n = 2$ are shown, varying the parameter α (a, b, c) and varying the parameter Q (plots d, e, f).

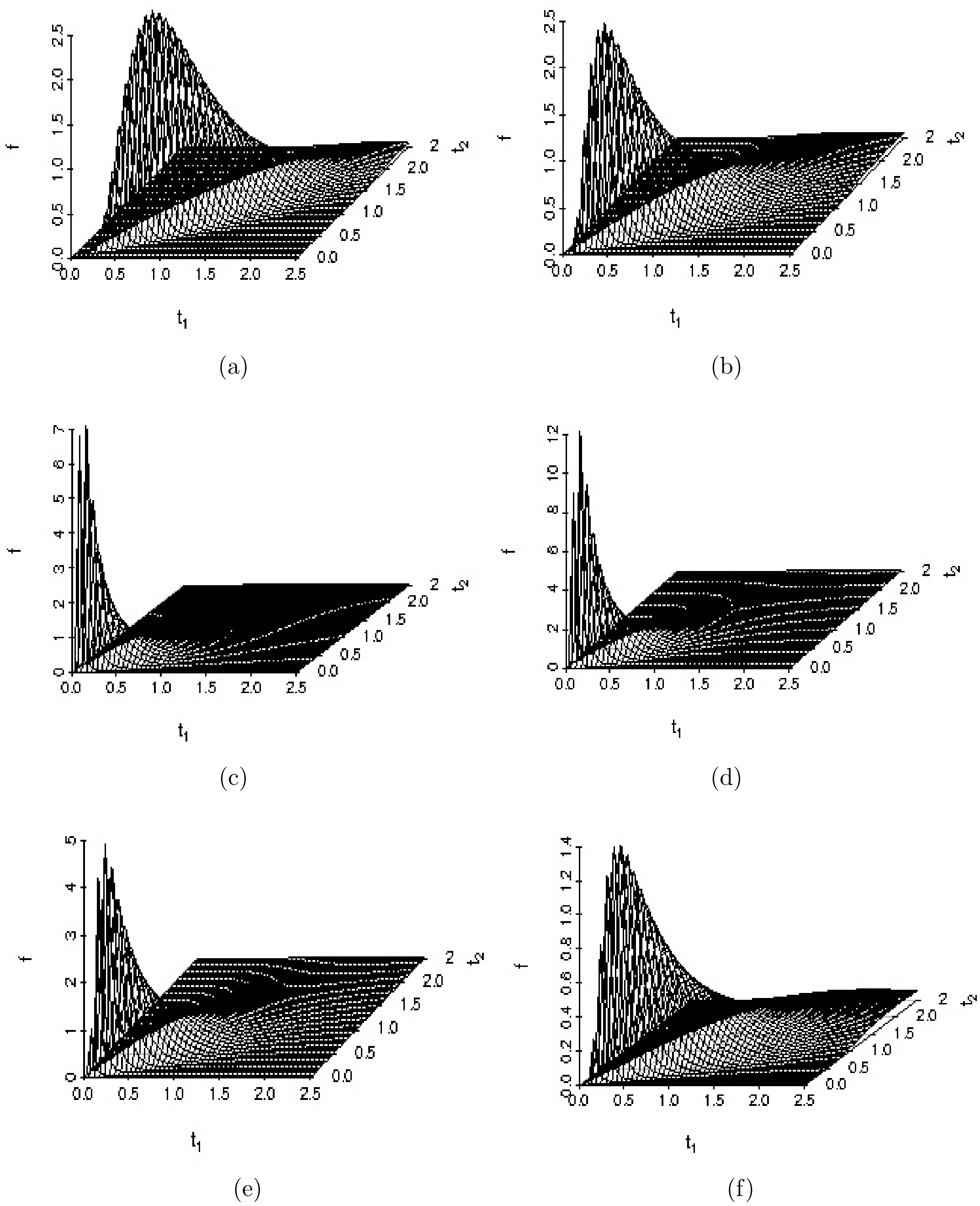


Figure 3.1: Reparameterized 2-variate BS density plots for (a) $\alpha_i = 0.5$, (b) $\alpha_i = 0.8$, (c) $\alpha_i = 1.5$, with $Q_i = 1.0$, and (d) $Q_i = 0.5$, (e) $Q_i = 0.8$, (f) $Q_i = 1.5$, with $\alpha_i = 1.0$, for $i = 1, 2$ and $\rho = 0.9$.

3.4 Formulation and estimation of the spatial model

To describe data dependent spatially, assume the stochastic process $\boldsymbol{T} = \{T(\boldsymbol{s}); \boldsymbol{s} \in \boldsymbol{D}\}$ defined on $\boldsymbol{D}$. We consider that $\boldsymbol{T}$ is stationary and isotropic, and that for spatial locations in $\boldsymbol{s}_i$, with $i = 1, \ldots, n$, the quantile function of the process may be formulated as

$$Q(T(\boldsymbol{s}_i)) = Q_i = h^{-1}(\eta_i) = h^{-1}(\boldsymbol{x}_i^\top \boldsymbol{\beta}), \ \ i = 1, 2, \ldots, n, \tag{3.2}$$

with h being a strictly monotone function of positive support and at least twice differentiable. Observe that $\boldsymbol{x}_i^\top = (1, x_{i2}, \ldots, x_{ip})$ are the values of p explanatory variables, with $x_{ij} = x_j(\boldsymbol{s}_i)$, for $j = 2, \ldots, p$, that is, x_{ij} is the value of the explanatory variable X_j at $\boldsymbol{s}_i$. Here, $\boldsymbol{\beta} = (\beta_1, \beta_2, \ldots, \beta_p)^\top$, for $p < n$, corresponds to a vector of regression coefficients to be estimated. In addition,

$$(T(\boldsymbol{s}_1), \ldots, T(\boldsymbol{s}_n)) = (T_1, \ldots, T_n) \sim \mathrm{BS}_n(\alpha \mathbf{1}_{n\times 1}, \boldsymbol{Q}(\boldsymbol{\beta}), \boldsymbol{\Gamma}), \quad \alpha > 0,$$

with $\mathbf{1}_{n\times 1}$ being an $n \times 1$ vector of ones and $\boldsymbol{\Gamma}$ being an $n \times n$ (non-singular) scale-dependence matrix.

Also, we suppose that the dependence over space is established by means of the $n \times n$ scale matrix already mentioned, which is symmetric, non-singular and positive definite. We observe $\boldsymbol{\Gamma}$ is proportional to $\mathrm{Cov}(T(\boldsymbol{s}_i), T(\boldsymbol{s}_j))$ and depends only on the Euclidean distance between $\boldsymbol{s}_i$ and $\boldsymbol{s}_j$, meaning

$$\mathrm{Cov}(T(\boldsymbol{s}_i), T(\boldsymbol{s}_j)) = C(\boldsymbol{s}_i, \boldsymbol{s}_j) = C(h_{ij}),\ \boldsymbol{s}_i, \boldsymbol{s}_j \in \boldsymbol{D}, \tag{3.3}$$

with $h_{ij} = \|\boldsymbol{s}_i - \boldsymbol{s}_j\|$. Further, we suppose the function C given in (3.3) is established by the spatial dependence parameter $\boldsymbol{\varphi} = (\varphi_1, \varphi_2, \varphi_3)^\top$ defined by $\boldsymbol{\Gamma} = \varphi_1 \boldsymbol{I}_n + \varphi_2 \boldsymbol{R}(\varphi_3)$, with $\boldsymbol{I}_n$ being the $n \times n$ identity matrix, $\varphi_1 \geq 0, \varphi_2 \geq 0$ being the nugget effect and partial sill, respectively, $\varphi_3 \geq 0$ is the spatial dependence radius, while $\boldsymbol{R}(\varphi_3) = r_{ij}$ is an $n \times n$ symmetric matrix, with main diagonal elements equal to one; see Mardia and Marshall (1984). Here, $\boldsymbol{R}(\varphi_3)$ depends on the covariance function used to model the dependence over space. If the family of Matérn model is used, we get

$$r_{ij} = \begin{cases} 1, \quad i = j, \\ \frac{1}{2^{\delta-1}\Gamma(\delta)} \left(\frac{h_{ij}}{\varphi_3}\right)^{\delta} K_\delta \left(\frac{h_{ij}}{\varphi_3}\right), \quad i \neq j, \end{cases} \tag{3.4}$$

with Γ being the usual gamma function, K_δ being the modified Bessel function of the third kind of order δ, and δ being a shape parameter; see Gradshteyn and Ryzhik (2000). If the family of power exponential models is used, for $i \neq j$, we have $r_{ij} = \exp(-(h_{ij}/\varphi_3)^p)$, with the shape parameter $0 < p \leq 2$. Table 3.1 provides particular cases of the family Matérn.

It is known that not all parameters in the covariance function (3.4) are consistently estimable under the fixed-domain asymptotic framework. In effect, the parameters φ_2 and φ_3 cannot be estimated consistently when the underlying process is observed in a bounded region of $\mathbb{R}^d$ for $d \leq 3$ Zhang (2004); Genton and Zhang (2012). However, Zhang (2004) also showed that if one fixes φ_3 at an arbitrary value, then the ML estimator for $\varphi = \varphi_2/\varphi_3^{2\delta}$ is consistent. The parameter φ is called microergodic parameter Stein (1999). Genton and Zhang (2012) showed that these same estimation problems remain for elliptically contoured random fields. In the present paper, to remove the identifiability problem, we impose restrictions on the parametric space. Specifically, we assume that the parameter φ_3 is fixed a priori Zhang (2004); Zhang and Wang (2010), and we use an alternative parameterization of the Matérn function suggested by Stein (1999) and given by

$$C(h_{ij}) = \varrho_1\delta_{ij} + \frac{\varrho_2}{2^{\delta-1}\Gamma(\delta)}(h_{ij}\varrho_3)^{\delta}K_\delta(h_{ij}/\varrho_3),\ \ h_{ij} \geq 0, i, j = 1, \ldots, n$$

with $\varrho_1 = \varphi_1$, $\varrho_2 = \varphi_2/\varphi_3^{2\delta}$ and $\varrho_3 = \varphi_3$. In this case, $\mathbf{\Gamma} = \varrho_1\boldsymbol{I}_n + \varrho_2\boldsymbol{R}$, where now the $n \times n$ matrix $\boldsymbol{R}$ has elements r_{ij} defined as

$$r_{ij} = \begin{cases} 1, & \text{if } i = j, \\ \dfrac{1}{2^{\delta-1}\Gamma(\delta)}(h_{ij}\varrho_3)^{\delta}K_\delta(h_{ij}/\varrho_3), & h_{ij} > 0, \end{cases}$$

for $i, j = 1, \ldots, n$.

Table 3.1: Special cases of the Matérn covariance function.

Smooth parameter	Covariance	Model
$\delta = 1/2$	$C(h) = \varphi_2 \exp(-h/\varphi_3)$	Exponential
$\delta = 1$	$C(h) = \varphi_2(h/\varphi_3)K_\delta(h/\varphi_3)$	Whittle
$\delta \to \infty$	$C(h) = \varphi_2 \exp(-(h/\varphi_3)^2)$	Gaussian

For the spatial model formulated in (3.2), with $\boldsymbol{\phi} = (\varrho_1, \varrho_2)$ and $\varrho_3 > 0$ being a fixed value, the corresponding parameter estimates can be estimated by the ML method using the log-likelihood function for $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\phi}^\top, \alpha)^\top$ based on the observations $\boldsymbol{t} = (t_1, \ldots, t_n)$ defined as

$$\ell(\boldsymbol{\theta}) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log(|\mathbf{\Gamma}|) - \frac{1}{2}\tilde{\boldsymbol{A}}^\top\mathbf{\Gamma}^{-1}\tilde{\boldsymbol{A}} + \log(\tilde{a}), \tag{3.5}$$

where $\boldsymbol{Q} = \boldsymbol{Q}(\boldsymbol{\beta})$, $\tilde{\boldsymbol{A}} = \tilde{\boldsymbol{A}}(\boldsymbol{t}; \alpha\mathbf{1}_{n\times1}, \boldsymbol{Q})$, $\tilde{a} = \tilde{a}(\boldsymbol{t}; \alpha\mathbf{1}_{n\times1}, \boldsymbol{Q})$ and $\mathbf{\Gamma} = \mathbf{\Gamma}(\boldsymbol{\phi}, \alpha, \boldsymbol{\beta})$. By the derivative of (3.5), with respect to $\boldsymbol{\theta}$ allows us to obtain the $(p+3) \times 1$ vector given by

$$\dot{\ell}(\boldsymbol{\theta}) = \left[\left(\frac{\partial\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}}\right)^\top, \left(\frac{\partial\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\phi}}\right)^\top, \left(\frac{\partial\ell(\boldsymbol{\theta})}{\partial\alpha}\right)^\top\right]^\top = \left(\dot{\ell}_{\beta_1}, \ldots, \dot{\ell}_{\beta_p}, \dot{\ell}_{\varrho_1}, \dot{\ell}_{\varrho_2}, \dot{\ell}_\alpha\right)^\top,$$

being

$$\begin{aligned}
\dot{\ell}_{\beta_j} &= -\tilde{\boldsymbol{A}}^\top \boldsymbol{\Gamma}^{-1} \frac{\partial \tilde{\boldsymbol{A}}}{\partial \beta_j} + \frac{\partial}{\partial \beta_j}[\log(\tilde{a})] \\
\dot{\ell}_{\varrho_j} &= -\frac{1}{2}\,\mathrm{tr}\left(\boldsymbol{\Gamma}^{-1}\frac{\partial \boldsymbol{\Gamma}}{\partial \varrho_j}\right) + \frac{1}{2}\tilde{\boldsymbol{A}}^\top \boldsymbol{\Gamma}^{-1}\frac{\partial \boldsymbol{\Gamma}}{\partial \varrho_j}\boldsymbol{\Gamma}^{-1}\tilde{\boldsymbol{A}} \\
\dot{\ell}_{\alpha} &= -\tilde{\boldsymbol{A}}^\top \boldsymbol{\Gamma}^{-1} \frac{\partial \tilde{\boldsymbol{A}}}{\partial \alpha} + \frac{\partial}{\partial \alpha}[\log(\tilde{a})],
\end{aligned}$$

where $\partial \tilde{\boldsymbol{A}}/\partial \beta_j = (\partial \tilde{A}_k/\partial \beta_j)$ and $\partial \tilde{\boldsymbol{A}}/\partial \alpha = (\partial \tilde{A}_k/\partial \alpha)$, with

$$\begin{aligned}
\frac{\partial \tilde{A}_k}{\partial \beta_j} &= -\frac{1}{\alpha\gamma_\alpha\sqrt{t_k Q_k}}\left(\frac{t_k\gamma_\alpha^2}{4Q_k}+1\right)\frac{1}{h'(Q_k)}x_{kj}, \\
\frac{\partial \tilde{A}_k}{\partial \alpha} &= \sqrt{\frac{4Q_k}{t_k}}\left\{-\frac{1}{(\alpha\gamma_\alpha)^2}(\gamma_\alpha+\alpha\gamma'_\alpha)\left(\frac{t_k\gamma_\alpha^2}{4Q_k}-1\right)+\frac{\gamma'_\alpha t_k}{2\alpha Q_k}\right\}, \\
\frac{\partial}{\partial \beta_j}[\log(\tilde{a})] &= \sum_{i=1}^{n}\left(-\frac{1}{2Q_i}+\frac{4}{t_i\gamma_\alpha^2+4Q_i}\right)\frac{1}{h'(Q_i)}x_{ij}, \\
\frac{\partial}{\partial \alpha}[\log(\tilde{a})] &= -\frac{n}{\alpha\gamma_\alpha}(\gamma_\alpha+\alpha\gamma'_\alpha)+\sum_{i=1}^{n}\frac{2t_i\gamma_\alpha\gamma'_\alpha}{t_i\gamma_\alpha^2+4Q_i},
\end{aligned}$$

and $\partial\boldsymbol{\Gamma}/\partial\varrho_1 = \boldsymbol{I}_n$, $\partial\boldsymbol{\Gamma}/\partial\varrho_2 = \boldsymbol{R}(\varrho_3)$. To estimate $\boldsymbol{\theta}$, $\dot{\ell}(\boldsymbol{\theta}) = \boldsymbol{0}_{(p+3)\times 1}$ must be solved. As this system does not have an analytical solution, $\widehat{\boldsymbol{\theta}}$ must be obtained with iterative procedures for non-linear systems; see Nocedal and Wright (1999) and Lange (2001).

Note that the Hessian matrix $\ddot{\ell}(\boldsymbol{\theta})$ for the BS spatial regression (3.2) is a $(p+3)\times(p+3)$ diagonal block matrix. The Hessian matrix is obtained by taking the second derivative of (3.5), with respect to the corresponding parameters, and is given by

$$\ddot{\ell}(\boldsymbol{\theta}) = \begin{pmatrix}
\dfrac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\beta}^\top} & \dfrac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}\partial\boldsymbol{\phi}^\top} & \dfrac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\beta}\partial\alpha} \\
\dfrac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\phi}\partial\boldsymbol{\beta}^\top} & \dfrac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\phi}\partial\boldsymbol{\phi}^\top} & \dfrac{\partial^2\ell(\boldsymbol{\theta})}{\partial\boldsymbol{\phi}\partial\alpha} \\
\dfrac{\partial^2\ell(\boldsymbol{\theta})}{\partial\alpha\partial\boldsymbol{\beta}^\top} & \dfrac{\partial^2\ell(\boldsymbol{\theta})}{\partial\alpha\partial\boldsymbol{\phi}^\top} & \dfrac{\partial^2\ell(\boldsymbol{\theta})}{\partial\alpha^2}
\end{pmatrix} = \begin{pmatrix}
\ddot{\ell}_{\boldsymbol{\beta\beta}} & \ddot{\ell}_{\boldsymbol{\beta\phi}} & \ddot{\ell}_{\boldsymbol{\beta}\alpha} \\
\ddot{\ell}_{\boldsymbol{\phi\beta}} & \ddot{\ell}_{\boldsymbol{\phi\phi}} & \ddot{\ell}_{\boldsymbol{\phi}\alpha} \\
\ddot{\ell}_{\alpha\boldsymbol{\beta}} & \ddot{\ell}_{\alpha\boldsymbol{\phi}} & \ddot{\ell}_{\alpha\alpha}
\end{pmatrix}, \tag{3.6}$$

where the $p\times p$, $p\times 2$ and 2×2 sub-matrices $\ddot{\ell}_{\boldsymbol{\beta}}$, $\ddot{\ell}_{\boldsymbol{\beta\phi}} = [\ddot{\ell}_{\boldsymbol{\phi\beta}}]^\top$ and $\ddot{\ell}_{\boldsymbol{\phi}}$, respectively, have elements detailed in Appendix A. Therefore, for the BS spatial regression model, the $(p+2)\times(p+2)$ expected Fisher information matrix, obtained from (3.6), is expressed as

$$\boldsymbol{K}(\boldsymbol{\theta}) = \mathrm{E}(-\ddot{\ell}_{\boldsymbol{\theta}}) = \begin{pmatrix}
\boldsymbol{K}_{\boldsymbol{\beta\beta}} & \boldsymbol{K}_{\boldsymbol{\beta\phi}} & \boldsymbol{K}_{\boldsymbol{\beta}\alpha} \\
\boldsymbol{K}_{\boldsymbol{\phi\beta}} & \boldsymbol{K}_{\boldsymbol{\phi\phi}} & \boldsymbol{K}_{\boldsymbol{\phi}\alpha} \\
\boldsymbol{K}_{\alpha\boldsymbol{\beta}} & \boldsymbol{K}_{\alpha\boldsymbol{\phi}} & K_{\alpha\alpha}
\end{pmatrix},$$

where $\boldsymbol{K}_{\boldsymbol{\beta\beta}} = \mathrm{E}(-\ddot{\ell}_{\boldsymbol{\beta\beta}})$, $\boldsymbol{K}_{\boldsymbol{\phi\phi}} = \mathrm{E}(-\ddot{\ell}_{\boldsymbol{\phi\phi}})$ and $\boldsymbol{K}_{\boldsymbol{\beta\phi}} = (\boldsymbol{K}_{\boldsymbol{\phi\beta}})^\top = \mathrm{E}(-\ddot{\ell}_{\boldsymbol{\beta\phi}})$ are $p \times p$, 2×2 and $p\times 2$ sub-matrices, whereas $\boldsymbol{K}_{\boldsymbol{\beta}\alpha} = (\boldsymbol{K}_{\alpha\boldsymbol{\beta}})^\top = \mathrm{E}(-\ddot{\ell}_{\boldsymbol{\beta}\alpha})$ and $\boldsymbol{K}_{\boldsymbol{\phi}\alpha} = (\boldsymbol{K}_{\alpha\boldsymbol{\phi}})^\top = \mathrm{E}(-\ddot{\ell}_{\boldsymbol{\phi}\alpha})$ are $p \times 1$ and 2×1 vectors, respectively.

3.5 Model checking

In order to evaluate the fit of the spatial model, we consider a property of the BS_n distribution related to the Mahalanobis distance, which may be used to validate the model in practice. Let

$$u_i = \tilde{\boldsymbol{A}}_{(i)}^\top \boldsymbol{\Gamma}^{-1} \tilde{\boldsymbol{A}}_{(i)}, i = 1, \ldots n, \tag{3.7}$$

where $\tilde{\boldsymbol{A}}_{(i)} = (\tilde{A}_{1(i)}, \ldots, \tilde{A}_{n(i)})^\top$, with

$$\tilde{A}_{j(i)} = \frac{1}{\alpha_j \gamma_{\alpha_j}} \sqrt{\frac{4h^{-1}(\boldsymbol{x}_j^\top \widehat{\boldsymbol{\beta}}_{(i)})}{t_j}} \left(\frac{t_j \gamma_{\alpha_j}^2}{4h^{-1}(\boldsymbol{x}_j^\top \widehat{\boldsymbol{\beta}}_{(i)})} - 1 \right), j = 1, \ldots n,$$

and $\widehat{\boldsymbol{\beta}}_{(i)}$ being the ML estimate of $\boldsymbol{\beta}$ obtained using the data set without the observation i. A Newton-Raphson one-step approximation to $\widehat{\boldsymbol{\theta}}_{(i)}$ can be obtained by

$$\widehat{\boldsymbol{\theta}}_{(i)} = \widehat{\boldsymbol{\theta}} + (-\boldsymbol{H}_{(i)}(\widehat{\boldsymbol{\theta}}))^{-1} \boldsymbol{U}_{(i)}(\widehat{\boldsymbol{\theta}}), i = 1, \ldots n,$$

where $\boldsymbol{H}_{(i)}(\boldsymbol{\theta})$ and $\boldsymbol{U}_{(i)}(\boldsymbol{\theta})$ are the Hessian matrix and score vector of the BS spatial model with its parameters estimated by the ML method without the observation i. Then, under the assumption

$$\boldsymbol{T} \sim \mathrm{BS}_n(\alpha \mathbf{1}_{n\times 1}, \boldsymbol{Q}(\boldsymbol{\beta}); \boldsymbol{\Gamma}),$$

we know that u_i defined in (3.7) is an observation from the χ^2 distribution with $n - 1$ degrees of freedom, for $i = 1, \ldots, n$. Thus, by using the Wilson-Hilferty approximation **?**, for $i = 1, \ldots, n$, we have that

$$z_i = \frac{\left(\frac{u_i}{n-1}\right)^{1/3} - \left(1 - \frac{2}{9(n-1)}\right)}{\left(\frac{2}{9(n-1)}\right)^{1/2}} \tag{3.8}$$

is an observation from a standard normal distribution. Hence, a QQ plot for z_i given in (3.8) can be used to evaluate the model fit. Besides the approximation of Wilson-Hilferty, the randomized quantile residual defined by Dunn and Smyth (1996) may be employed to evaluate the fit of the BS spatial log-linear model. In the case of this model, such a residual is given by

$$r_i = \Phi^{-1}\left(F(u_i)\right), i = 1, \ldots, n, \tag{3.9}$$

where Φ^{-1} is the inverse of the $\mathrm{N}(0, 1)$ CDF and F is the $\chi^2(n - 1)$ CDF. Because the randomized quantile residual has a $\mathrm{N}(0, 1)$ distribution, a QQ plot of r_i defined in (3.9) may also be employed for evaluating the model fit.

3.6 Global influence diagnostics

A global influence technique of case-deletion is based on the likelihood distance (LD) and established as

$$\mathrm{LD}_i(\boldsymbol{\theta}) = 2(\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}}_{(i)})),\ i = 1, \ldots, n, \tag{3.10}$$

where ℓ is the log-likelihood function, and $\widehat{\boldsymbol{\theta}}, \widehat{\boldsymbol{\theta}}_{(i)}$ are, respectively, the ML estimates of $\boldsymbol{\theta}$ considering the full data set and the data set without the case i; see Cook et al. (1988). The expression (3.10) measures the change in the LD with estimated parameters when the case i is deleted and may be employed as global influence technique to assess the potential influence of this case.

The Cook distance (CD) is other global influence technique based on case deletion and an alternative to the measure defined in (3.10). This has been generalized to several non-normal models; see Desousa et al. (2018). The usual expression for the CD is given by

$$\mathrm{CD}_i(\boldsymbol{\theta}) = (\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_{(i)})^\top \boldsymbol{M}(\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_{(i)}),\ i = 1, \ldots, n, \tag{3.11}$$

where $\boldsymbol{M}$ is an appropriately chosen positive definite matrix, which can be, for example, the inverse of the asymptotic covariance matrix. Thus, a measure based on the CD defined in (3.11) is stated as

$$\mathrm{CD}_i^{(1)}(\boldsymbol{\theta}) = (\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_{(i)})^\top [-\ddot{\ell}_{(i)}(\widehat{\boldsymbol{\theta}})](\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_{(i)}),\ i = 1, \ldots, n, \tag{3.12}$$

where

$$\ddot{\ell}_{(i)}(\boldsymbol{\theta}) = \frac{\partial \ell^2_{(i)}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta} \partial \boldsymbol{\theta}^\top},$$

with $\ell_{(i)}$ being the log-likelihood function obtained after deleting the case i. Note that

$$\boldsymbol{M} = [-\ddot{\ell}_{(i)}(\widehat{\boldsymbol{\theta}})],$$

defined in (3.11) is the inverse of $[-\ddot{\ell}_{(i)}(\widehat{\boldsymbol{\theta}})]^{-1}$ that is an estimate of the asymptotic covariance matrix. If n is too large, the computation of $\ddot{\ell}_{(i)}(\widehat{\boldsymbol{\theta}})$ may became difficult and, in this case, $\ddot{\ell}(\widehat{\boldsymbol{\theta}})$ can be used instead of $\ddot{\ell}_{(i)}(\widehat{\boldsymbol{\theta}})$; see De Bastiani et al. (2018). Then, an alternative measure of global influence based on the CD defined in (3.12) is given by

$$\mathrm{CD}_i^{(2)}(\boldsymbol{\theta}) = (\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_{(i)})^\top [-\ddot{\ell}(\widehat{\boldsymbol{\theta}})](\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_{(i)}),\ i = 1, \ldots, n. \tag{3.13}$$

Other measure based on the CD defined in (3.13) uses the first order approximation $\widehat{\boldsymbol{\theta}} - \widehat{\boldsymbol{\theta}}_{(i)} \approx \ddot{\ell}^{-1}_{(i)}(\widehat{\boldsymbol{\theta}})\dot{\ell}_{(i)}(\widehat{\boldsymbol{\theta}})$, which considers a Taylor expansion around $\widehat{\boldsymbol{\theta}}$, until the second order term, and the one-step-late Newton-Raphson estimate. This third measure based on (3.13) is expressed as

$$\mathrm{CD}_i^{(3)}(\boldsymbol{\theta}) = \left(\dot{\ell}_{(i)}(\widehat{\boldsymbol{\theta}})\right)^\top \left(\ddot{\ell}_{(i)}(\widehat{\boldsymbol{\theta}})\right)^{-1} \left(\dot{\ell}_{(i)}(\widehat{\boldsymbol{\theta}})\right),\ i = 1, \ldots, n, \tag{3.14}$$

where

$$\dot{\ell}_{(i)}(\widehat{\boldsymbol{\theta}}) = \frac{\partial \ell_{(i)}(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}}.$$

Other alternative measures based on the CD and similar to (3.14) can be seen in Garcia-Papani et al. (2018b). In many cases $\mathrm{CD}_i(\boldsymbol{\theta})$ is preferred to $\mathrm{LD}(\widehat{\boldsymbol{\theta}}_{(i)})$, because of its heavier computational burden. A large value of $\mathrm{CD}_{(i)}(\boldsymbol{\theta})$ means that the case i is potentially influential. A definition of what is large has been an unresolved aspect, but Cook et al. (1982) established this depends on the problem.

3.7 Local influence diagnostics

Again, we use the elements defined in the Section 2.5, which are, in fact, applicable to our spatial model. Specifically, we consider the direction $\boldsymbol{d}_{\max}$ and the normal curvature associate to the case i, that is, C_i, for detecting cases that are potentially influential on $\widehat{\boldsymbol{\theta}}$.

In addition to the normal curvature of Cook (1987), other measures of local influence have been studied and employed. Poon and Poon (1999) defined the conformal curvature as

$$B_i = \frac{C_i}{\operatorname{tr} \boldsymbol{B}},\ i = 1, \ldots, n, \tag{3.15}$$

which demands similar work of computation than for C_i. The measure indicated in (3.15) is a standardized measure because is invariant under conformal reparameterization. Hence, it is not difficult to establish a cut-off point for it. According to Poon and Poon (1999), if for the case i we obtain

$$B_i > 2\sum_{i=1}^{n} \frac{B_i}{n} = 2\bar{B},\ i = 1, \ldots, n,$$

where $\bar{B}$ is the arithmetic mean of the basic conformal curvatures, that is, of $B_1, \ldots, B_n$, then the case i is potentially influential. Another cut-off point implies consider the case i as potentially influential if

$$B_i > \bar{B} + 2\mathrm{SD}(B),\ i = 1, \ldots, n,$$

where $\mathrm{SD}(B)$ is the standard deviation (SD) of $B_1, \ldots, B_n$.

Perturbation scheme in the response We assume the perturbation

$$\boldsymbol{T}_{\boldsymbol{\omega}}(\boldsymbol{s}) = \boldsymbol{T}(\boldsymbol{s}) + \boldsymbol{A}\boldsymbol{w},$$

where $\boldsymbol{A}$ is a symmetric, non-singular matrix and $\boldsymbol{\omega} = (\omega_1, \ldots, \omega_n) \in \mathbb{R}^n$ is a perturbation vector. It is clear that $\boldsymbol{\omega}_0 = \mathbf{0}_{n\times 1}$ is the non-perturbation vector. In this case perturbation scheme, the perturbed log-likelihood function is given by

$$\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log(|\boldsymbol{\Gamma}|) - \frac{1}{2}\tilde{\boldsymbol{A}}_{\boldsymbol{\omega}}^{\top}\boldsymbol{\Gamma}^{-1}\tilde{\boldsymbol{A}}_{\boldsymbol{\omega}} + \log(\tilde{a}_{\boldsymbol{\omega}}), \tag{3.16}$$

where $\tilde{\boldsymbol{A}}_{\boldsymbol{\omega}} = (\tilde{A}_1(\boldsymbol{\omega}), \ldots, \tilde{A}_n(\boldsymbol{\omega}))^{\top}$ and $\tilde{a}_{\boldsymbol{\omega}} = \tilde{a}(\boldsymbol{t}(\boldsymbol{\omega}); \alpha, \boldsymbol{Q})$, with

$$\tilde{A}_k(\boldsymbol{\omega}) = A(t_k(\boldsymbol{\omega}); \alpha, Q_k),\ k = 1, \ldots, n.$$

Zhu et al. (2007) established that the perturbation $\boldsymbol{\omega}$ is appropriate if and only if $\boldsymbol{G}(\boldsymbol{\theta}|\boldsymbol{\omega}_0) = c\boldsymbol{I}_n$, where $c > 0$ and

$$\boldsymbol{G}(\boldsymbol{\theta}|\boldsymbol{\omega}) = \mathrm{E}[\dot{\ell}(\boldsymbol{\theta}|\boldsymbol{\omega})\dot{\ell}^{\top}(\boldsymbol{\theta}|\boldsymbol{\omega})],$$

with $\dot{\ell}(\boldsymbol{\theta}|\boldsymbol{\omega}) = \partial\ell(\boldsymbol{\theta}|\boldsymbol{\omega})/\partial\boldsymbol{\omega}$. Obtaining the matrix $\boldsymbol{G}(\boldsymbol{\theta}|\boldsymbol{\omega}_0)$ can be a very difficult. In this paper, we assume that the form of $\boldsymbol{A}$ to obtain an appropriate perturbation $\boldsymbol{\omega}$ is the same obtained in Garcia-Papani et al. (2018b), that is,

$$\boldsymbol{A} = \left(\frac{\alpha}{4}\boldsymbol{\Gamma}^{1/2} - \frac{1}{\alpha}\boldsymbol{\Gamma}^{-1/2}\right)^{-1}, \tag{3.17}$$

where $\boldsymbol{\Gamma}^{1/2}$ is the square root matrix of $\boldsymbol{\Gamma}$, that is, $\boldsymbol{\Gamma}^{1/2}\boldsymbol{\Gamma}^{1/2} = \boldsymbol{\Gamma}$. For details of computations for this square root matrix, see De Bastiani et al. (2015). Therefore, we assume that an appropriate perturbation scheme for the response is given by

$$\boldsymbol{T}_{\boldsymbol{\omega}}(\boldsymbol{s}) = \boldsymbol{T}(\boldsymbol{s}) + \left(\frac{\alpha}{4}\boldsymbol{\Gamma}^{1/2} - \frac{1}{\alpha}\boldsymbol{\Gamma}^{-1/2}\right)^{-1}\boldsymbol{\omega}.$$

Perturbation in a continuous explanatory variable Now, we consider a perturbation scheme in a single continuous explanatory variable, labelled as X_l namely, and the other explanatory variables are not perturbed. Thus, we have

$$\boldsymbol{x}_{l,\boldsymbol{w}}(\boldsymbol{s}) = \boldsymbol{x}_l(\boldsymbol{s}) + \boldsymbol{A}\boldsymbol{w}, \quad \boldsymbol{x}_{j,\boldsymbol{w}}(\boldsymbol{s}) = \boldsymbol{x}_j(\boldsymbol{s}), \quad j \neq l, j = 1, \ldots, p,$$

where $\boldsymbol{w} \in \mathbb{R}^n$ and $\boldsymbol{w}_0 = \boldsymbol{0}_{n\times 1}$. Therefore, in this case perturbation scheme, the perturbed log-likelihood function is given by

$$\ell(\boldsymbol{\theta}|\boldsymbol{\omega}) = -\frac{n}{2}\log(2\pi) - \frac{1}{2}\log(|\boldsymbol{\Gamma}|) - \frac{1}{2}\tilde{\boldsymbol{A}}_{\boldsymbol{\omega}}^{\top}\boldsymbol{\Gamma}^{-1}\tilde{\boldsymbol{A}}_{\boldsymbol{\omega}} + \log(\tilde{a}_{\boldsymbol{\omega}}), \tag{3.18}$$

where $\tilde{\boldsymbol{A}}_{\boldsymbol{\omega}} = (\tilde{A}_1(\boldsymbol{\omega}), \ldots, \tilde{A}_n(\boldsymbol{\omega}))^{\top}$ and $\tilde{a}_{\boldsymbol{\omega}} = \tilde{a}(\boldsymbol{t}; \alpha, \boldsymbol{Q}(\boldsymbol{\omega}))$, with

$$\tilde{A}_k(\boldsymbol{\omega}) = A(t_k; \alpha, Q_k(\boldsymbol{\omega})),\ k = 1, \ldots, n.$$

Once again, in order to obtain the matrix $\boldsymbol{A}$ for an appropriate perturbation in X_l can be a hard work. As in the case of the response perturbation with $\boldsymbol{A}$ given in (3.17), we assume that the most appropriate explanatory variable perturbation scheme is expressed as

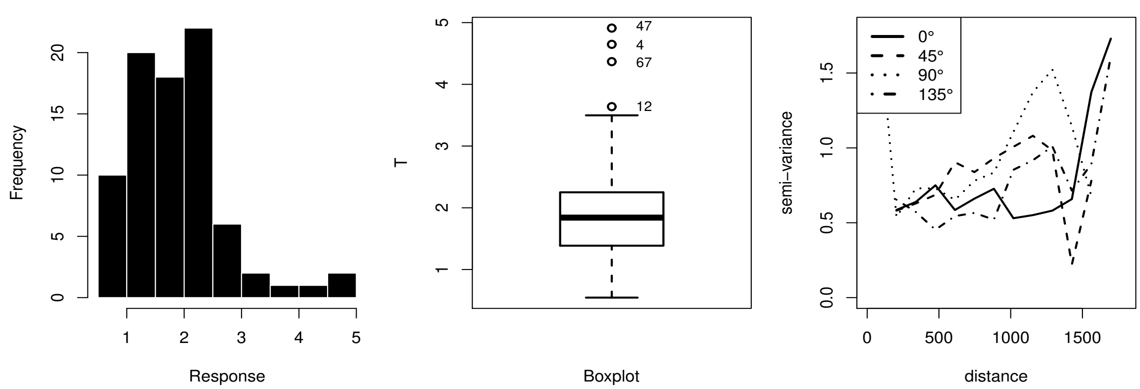
$$\boldsymbol{x}_{t,\boldsymbol{\omega}}(\boldsymbol{s}) = \boldsymbol{x}_t(\boldsymbol{s}) + \left(\frac{\alpha}{4}\boldsymbol{\Gamma}^{1/2} - \frac{1}{\alpha}\boldsymbol{\Gamma}^{-1/2}\right)^{-1}\boldsymbol{\omega}.$$

Details of the $\boldsymbol{\Delta}$ matrix given by (2.6) for the two perturbations considered are shown in Appendix B.

3.8 Illustrative example

The methodology presented in this chapter is illustrated considering an environmental data set related to key nutrients in the soil. The data set belongs to 82 locations of an area in Brazil, which contain levels of magnesium (Mg) affecting the development of the root system and calcium (Ca) and competing with Mg for absorption of nutrients. With the matrix of coordinates corresponding to these data, we construct the distance matrix indicated in (3.3). Considering to the model of Matérn family, with $\varrho_1 = 0.3482$, $\varrho_2 = 2.71 \times 10^{-7}$, $\varrho_3 = 1086.751$ (which are estimated using the weighted square linear method) and $\delta = 1.0$, we generate a scale matrix and, with it, an observation from the multivariate BS model of dimension 82, considering the values $\beta_0 = 0.3691$ and $\beta_1 = 0.1782$, with a square root link function, and $\alpha = 0.2673$. For the explanatory variables, we consider the vector of Ca values in the environmental data.

Descriptive statistics for the response with environmental dat are: median = 1.84054; mean = 1.9015; SD = 0.8567; coefficient of variation = 0.4505; skewness = 1.1822; kurtosis = 5.1592; minimum = 0.5448; maximum = 4.911; and n = 82. This summary shows the asymmetric behavior of the response variable, which is also observed in the histogram of Figure 3.2 (a), whereas the boxplot (b) of the values of the response T allows us to observe four outliers, which are #4, #12, #47 and #67. The directional variogram of Figure 3.2(c) indicates that there is no preferred direction, meaning an omni-directional semi-variogram is suitable. Hence, we can consider the associated stochastic process as isotropic.



(a) (b) (c)

Figure 3.2: Histogram (a), boxplot (b) and semi-variogram (c) for the response variable with environmental data.

We estimate the spatial dependence parameters assuming a variogram model in the Matérn family with $\delta = 1.0$. We suppose that

$$(T(\boldsymbol{s}_1), \ldots, T(\boldsymbol{s}_n)) = (T_1, \ldots, T_n) \sim \mathrm{BS}_n(\alpha \mathbf{1}_{n\times 1}, \boldsymbol{Q}(\boldsymbol{\beta}), \boldsymbol{\Gamma}),$$

considering three cases for the link function h defined in (3.2), that is, logarithm, square root and identity functions, which are expressed as

$$\begin{aligned}\log(Q_i) &= \boldsymbol{x}_i^\top \boldsymbol{\beta},\\ \sqrt{Q_i} &= \boldsymbol{x}_i^\top \boldsymbol{\beta},\\ Q_i &= \boldsymbol{x}_i^\top \boldsymbol{\beta}, \quad i = 1, \ldots, 82,\end{aligned}$$

with $\boldsymbol{\beta} = (\beta_0, \beta_1)^\top$ being the regression coefficient vector and $\boldsymbol{x}_i^\top = (1, x_{i1})$ being the value of $\boldsymbol{X}_i$.

In order to compare spatial regression models, we employ the Schwarz Bayesian information criterion (BIC) and corrected Akaike information criterion (CAIC) defined as

$$\text{BIC} = d\log(n) - 2\ell(\widehat{\boldsymbol{\theta}}), \quad \text{CAIC} = 2d - 2\ell(\widehat{\boldsymbol{\theta}}) + \frac{2d^2 + 2d}{n - d - 1},$$

where d is the number of model parameters, n the dimension of the data set, and $\ell(\widehat{\boldsymbol{\theta}})$ corresponds to the log-likelihood function for $\boldsymbol{\theta}$ for the model evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$. BIC and CAIC use the log-likelihood function and penalize a model with more parameters. When small information is obtained from a model in relation a specific data set, then a higher value is obtained for this model being it less adequate. This, a better model is that with a less value for BIC or CAIC; see Ferreira et al. (2012). Table 3.2 reports the values of the log-likelihood function, CAIC and BIC for the model with link functions defined in (3.19). Also, we compare the models given in (3.19) with the Gaussian spatial regression model applied to the data set, which considers the description of the mean (or equivalently the median) with identity link function; see last row of the Table 3.2. Only when the model Gaussian is compared with the BS case with identity link function, its values are better. From this table, we conclude that the BS spatial quantile regression model with square root link function should be used. The ML estimates of the selected model parameters and their corresponding estimated asymptotic standard errors (in parentheses) are: $\widehat{\beta}_0 = 0.2527(0.5720)$, $\widehat{\beta}_1 = 0.1576(0.3483)$, $\widehat{\varrho}_1 = 0.2893(0.7920)$, $\widehat{\varrho}_2 = 9.40 \times 10^{-8}(3.06 \times 10^{-12})$ and $\widehat{\alpha} = 0.2758(0.3027)$. With this information, the estimated model is $\widehat{Q}_i = (0.2527 + 0.1576x_{i1})^2$, for $i = 1, \ldots, 82$, while the scale-dependence matrix is estimated as

$$\widehat{\boldsymbol{\Gamma}} = 0.2893\boldsymbol{I}_{82} + 9.40 \times 10^{-8}\boldsymbol{R}(86875.55),$$

with $\boldsymbol{R}(\varrho_3)$ defined in (3.4) for $\delta = 1.0$ and evaluated at $\widehat{\varrho}_3 = 86875.55$.

Having estimated the spatial parameters, we can calculate the relative nugget effect (RNE) as

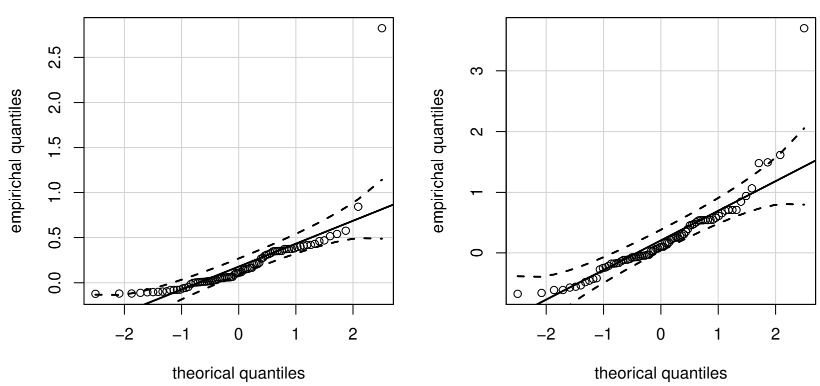
$$\text{RNE} = \frac{\varrho_1}{\varrho_1 + \varrho_2\varrho_3^{2\delta}}.$$

Table 3.2: Values of log-likelihood, CAIC and BIC for indicated models with environmental data.

Model	$\ell(\widehat{\boldsymbol{\theta}})$	CAIC	BIC
BS - logarithm link	-17.0303	44.8500	56.0942
BS - square root link	-12.9637	36.7169	47.9610
BS - identity link	-37.3900	85.5694	96.8135
Gaussian	-36.7378	81.9950	91.1024

This indicates the degree of spatial dependence (see Cambardella et al., 1994) following this RNE: if RNE < 0.25, the data present a strong spatial dependence; if $0.25 <$ RNE < 0.75, the data indicate an average spatial dependence; and if RNE ≥ 0.75, the data show a weak spatial dependence. For our environmental data, we have $\widehat{\text{RNE}} = 0.0004$, which means that an strong spatial dependence is presented. Therefore, this supports the use of the spatial model suggested in our research.

The quantile versus quantile (QQ) plot of the residuals transformed by the Wilson-Hilferty approximation (see Marchant et al., 2016b) is shown in Figure 3.4(a). An alternative to evaluate to fit of the model could be employ the randomized quantile residual defined by Dunn and Smyth (1996). Observe that most of the residuals are inside of the bands (at 1%). When the observations #11 and #25 are removed, a better fit is detected and almost every point is inside the envelope; see Figure 3.4(b). Thus, the BS spatial quantile regression model seems to be appropriate to describe the environmental data. However, if we use a heavy-tailed asymmetric distribution, such as the BS-Student-t model, we could obtain a better fitting. This implies a new research line for this work.



(a) (b)

Figure 3.3: QQ plots for transformed residuals with environmental data.

Figure 3.4(c) presents the potentially influential cases in the ML estimates of the parameter vector $\boldsymbol{\theta}$ considering the CD as criterion of global influence. It is possible to see that cases #4 and #13 are potencially influential for the estimate of $\boldsymbol{\theta}$ because their values of CD are outside of the bands.

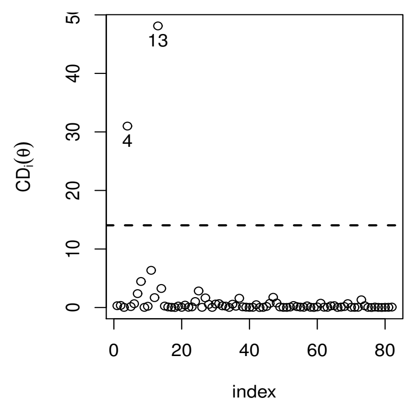


Figure 3.4: plots of the CD with environmental data.

For the local influence study, we assume two types of scheme: (i) perturbation in the response; and (ii) perturbation in the explanatory variable X. We consider three measures of local influence: (i) the absolute value of the components of $\boldsymbol{d}_{\max}$; (ii) normal curvature in the direction of basis vectors (C_i); and (iii) conformal curvature in the same direction (B_i). Figure 3.5 displays the local influence graphs corresponding to perturbations in the response and explanatory variable X. Note that cases #4 and #13 detected in the global influence plots are not locally influential by the plots associated with $\boldsymbol{d}_{\max}$, C_i and B_i when either the response or explanatory variable are perturbed. For perturbation in the response, we can observe that four cases (#7, #11, #28 and #81) are detected as potentially influential points in two plots. The plots associated with explanatory variable perturbation again detect cases #7, #11 and #81 as potentially influential; see plots (f) and (g). Observe that only the outlier #4 is detected as potentially influential in plots of diagnostics, that is, in spatial statistics, an influential point is not necessarily an outlier and viceversa.

We study the RC when the cases detected as potentially influential are removed, that is, cases #4, #7, #11, #13 and #81, which are the points detected for the most of the plots in Figures 3.4(c) and 3.5. We consider removing individual cases and combinations of them. The impact of the influential cases on the parameter estimates is evaluated by computing

$$\mathrm{RC}_{\theta_{j(I_k)}} = \left|(\widehat{\theta}_j - \widehat{\theta}_{j(I_k)})/\widehat{\theta}_j\right| \times 100\%,$$

where $\widehat{\theta}_{j(I_k)}$ is the ML estimate of θ_j after removing the set I_k, for $j = 1, \ldots, 5$ and

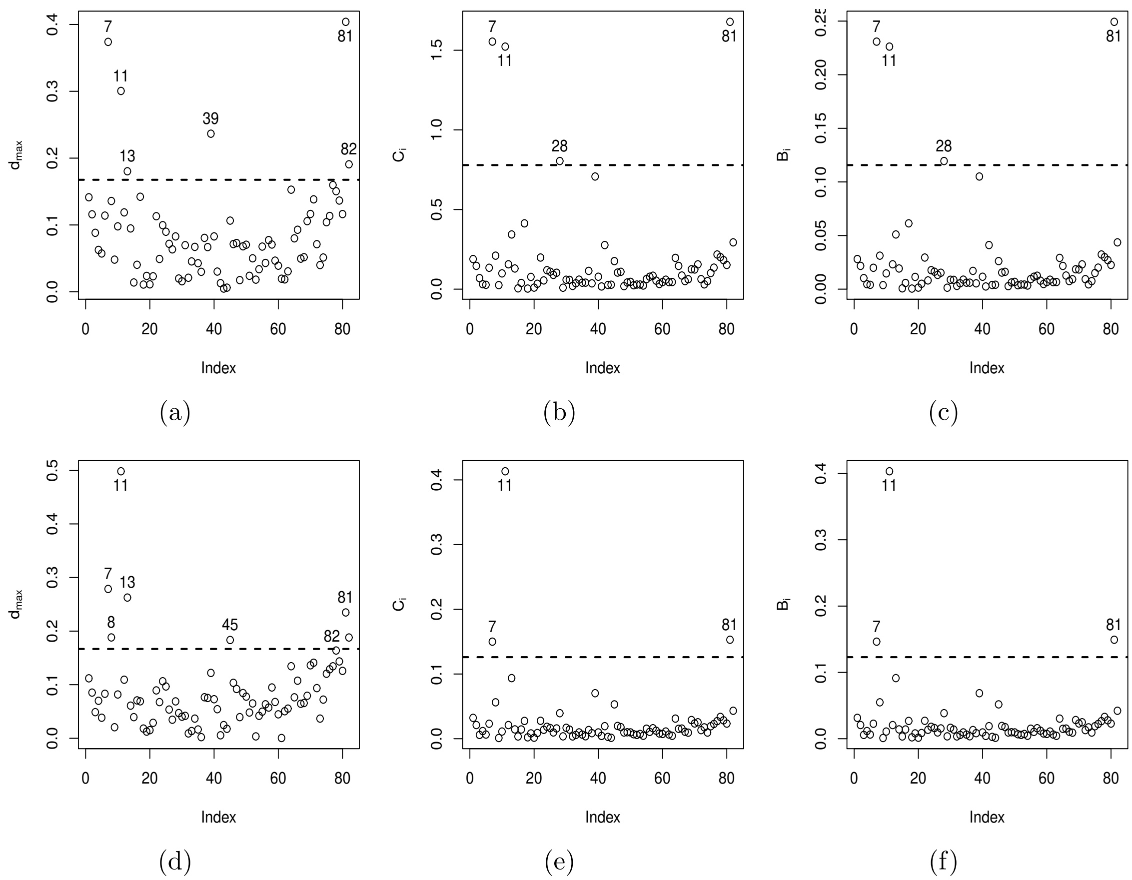


Figure 3.5: Perturbation in the response for $\boldsymbol{d}_{\max}$ (a), C_i (b) and B_i (c) and perturbation in the regressor for $\boldsymbol{d}_{\max}$ (d), C_i (e) and B_i (f) with environmental data.

$k = 1, \ldots, 31$, with $\theta_1 = \beta_0$, $\theta_2 = \beta_1$, $\theta_3 = \phi_1$, $\theta_4 = \phi_2$, $\theta_5 = \alpha$. The RCs in the parameter estimates obtained by considering the data with removed cases are presented in Table 3.3. In all cases, the RCs are larger for the parameters ϕ_1 and ϕ_2, with a more pronounced change for ϕ_2. Also, observe that, in all cases, the α parameter varies more than the parameters β_0 and β_1. From this table, we conclude that removing the potentially influential cases changes the spatial dependence of the data, because the variance, that is $\phi_1 + \phi_2$, involves the values of ϕ_2 and ϕ_1. Therefore, removal of these influential cases modify the spatial dependence and then our predictive model can be affected altering the conclusions of the study.

Table 3.3: RC of ML estimates for the indicated parameter and removed cases.

Removed case(s)	β_0	β_1	ϕ_1	ϕ_2	α
#4	12.6620	25.3177	71.7022	8.70×10^9	48.1703
#7	13.2343	25.6437	80.9952	5.84×10^9	76.9562
#11	3.3467	30.0718	81.6095	5.65×10^9	75.2459
#13	10.0901	25.9289	79.7180	6.24×10^9	70.5713
#81	8.4974	26.7649	75.4347	7.55×10^9	59.9401
#4, #7	15.9962	24.6032	71.8327	8.66×10^9	45.3186
#4, #11	0.8139	29.1265	76.6311	7.19×10^9	55.5973
#4, #13	12.8796	24.8727	74.8085	7.75×10^9	52.9534
#4, #81	11.5588	25.6413	69.8460	9.27×10^9	44.4082
#7, #11	0.0524	29.3297	78.4446	6.63×10^9	58.4609
#7, #13	13.3166	25.2482	75.0938	7.66×10^9	50.5123
#7, #81	12.4580	25.8690	74.8004	7.75×10^9	54.6182
#11, #13	3.2636	29.6689	79.8486	6.20×10^9	62.4288
#11, #81	7.8047	31.3709	80.2035	6.09×10^9	69.5418
#13, #81	8.5076	26.3895	80.6544	5.95×10^9	75.6726
#4, #7, #11	2.6465	28.3643	75.4187	7.55×10^9	48.4652
#4, #7, #13	16.1583	24.1749	76.0150	7.38×10^9	53.2131
#4, #7, #81	15.6057	24.7173	69.5971	9.35×10^9	40.7535
#4, #11, #13	0.6465	28.6897	82.3491	5.43×10^9	73.5712
#4, #11, #81	4.9677	30.3450	72.3805	8.49×10^9	43.7327
#4, #13, #81	11.6564	25.2312	71.4906	8.77×10^9	44.6538
#7, #11, #13	0.0746	28.9446	80.8501	5.89×10^9	62.9117
#7, #11, #81	3.6430	30.4025	81.7136	5.62×10^9	72.8290
#7, #13, #81	12.4002	25.5139	80.9331	5.86×10^9	73.0851
#11, #13, #81	7.8703	31.0100	81.4499	5.70×10^9	69.8762
#4, #7, #11, #13	2.7511	27.9460	83.0843	5.20×10^9	73.2831
#4, #7, #11, #81	0.7074	29.3445	72.8199	8.36×10^9	41.8980
#4, #7, #13, #81	15.6331	24.3282	69.8930	7.16×10^9	37.6196
#4, #11, #13, #81	4.9412	29.9484	76.7024	5.77×10^9	51.6835
#7, #11, #13, #81	3.7878	30.0646	81.2223	9.26×10^9	65.2296
#4, #7, #11, #13, #81	0.7615	28.9716	80.9607	5.85×10^9	64.1229

3.9 Concluding remarks

In this chapter, we have formulated spatial regression to model a quantile of a response variable that follows the Birnbaum-Saunders distribution. We have proposed a new parameterization of the multivariate Birnbaum-Saunders distribution to formulate the new model. Its fit has been evaluated using the Wilson-Hilferty approximation based on residuals of this new spatial quantile regression model. In addition, we have derived diagnostics techniques to detect observations that can be global or locally influential in a potential manner. In the last case, we have considered two perturbation cases, that is, schemes for the response and continuos explanatory variables. Furthermore, we have applied the derived methodology to an environmental data set, showing the suitability of the Birnbaum-Saunders spatial quantile regression models when consider strictly positive data with a distribution which is asymmetric to the right. An relevant aspect of the proposed methodology in this research is that detection of influential cases and their removal can modify the spatial dependence and then the predictive model may be affected altering the conclusions of the study.

Chapter 4

Discussion

4.1 Summary

In this brief chapter, we present conclusions and discuss aspects not treated in this thesis entitled "Birnbaum-Saunders quantile regression models". This is based on the two main achievements of this work, that is, to propose Birnbaum-Saunders quantile regression models based on the Birnbaum-Saunders distribution and to develop a methodology associated with them for independent data and for correlated-spatially data. At the end, product of this study, we delineate potential ideas for future research.

4.2 Conclussions and limitations

In this thesis, first, we have proposed quantile regression based on the Birnbaum-Saunders distribution with independent observations. We have established a new parameterization of the univariate Birnbaum-Saunders distribution based on its quantiles. For the new model, we have developed (i) parameter estimation using the maximum likelihood method, (ii) asymptotic inference for the estimators, (iii) diagnostic analytics based on the local influence technique, (iv) residuals, and (v) conducted a simulation study to evaluate performance of these residuals. We have applied the methodology to a income data set. A limitation of our proposal is that covariates can affect simultaneously the quantiles and shape parameter. Another limitation is that a comparison of our model with other similar models based on, for instance, the gamma, lognormal or Weibull distributions (Noufaily and Jones, 2013) is not possible because this implies to derive models using such distributions with identical parameterizations to that used in our approach, which are not available in the literature.

In second place, we have formulated spatial quantile regression models considering a new parameterization of the multivariate Birnbaum-Saunders distribution. We have developed (i) estimation of parameters with the maximum likelihood method, (ii) global and local influence techniques, and (iii) an illustrative example through an environmental data set, showing the potential applications of these models. An aspect that was not developed, and necessary for conduct inference, is the study of the asymptotic behavior and performance of maximum likelihood estimators. It is known the difficulty of the asymptotic frameworks for spatial data, because there being at least two relevant frameworks. The behavior of these frameworks can be very distinct when estimating the spatial dependence parameters; see Zhang and Zimmerman (2005). In addition of this difficulty, there are not works in the literature for the case of the Birnbaum-Saunders distribution.

4.3 Future research

We are considering to study some new aspects related to this thesis in future works. For instance,

(1) As mentioned, covariates can affect simultaneously the quantiles and shape parameter of the Birnbaum-Saunders distribution. Then, this topic will studied in a future investigation following the line of the recent work of Ventura et al. (2019) about joint modeling of two parameters

(2) How our models perform and the statistical evaluation of the estimation process by means of Monte Carlo simulations.

(3) A comparison of our model with other similar models based on, for instance, the gamma, lognormal or Weibull distributions.

(4) An exploration of the novel quantile regression approach proposed in this work considering other distributions.

(5) Cobb-Douglas and tobit type models can be considered in the context of the present investigation; see Desousa et al. (2018), Cysneiros et al. (2019) and De la Fuente-Mella et al. (2020). The use of censored data can be also of interest to be analyzed; see Villegas et al. (2011) and Leão et al. (2018b).

(6) A study of the asymptotic behavior and performance of maximum likelihood estimators for the Birnbaum-Saunders spatial quantile regression model.

(7) Studying Birnbaum-Saunders-Student-t spatial quantile regression models is a relevant work because the parameter estimation in spatial quantile regression models can be affected by atypical cases. That is due to the Birnbaum-Saunders dis-

tribution is based on the normal distribution. Therefore, considering Birnbaum-Saunders-Student-t models can decrease their effects; see Athayde et al. (2019).

(8) Adding random effects by mixed models may produce a more refined Birnbaum-Saunders spatial quantile regression model and also closer to reality; see Villegas et al. (2011).

Appendix A: Fisher information matrix for the BS spatial model

To obtain the Fisher information matrix, $-\ddot{\ell}(\boldsymbol{\theta})$ must be evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$. For the BS spatial regression model presented in (3.2), the elements of Hessian matrix can be expressed as

$$\ddot{\ell}_{\beta_j\beta_l} = -\left[\left(\frac{\partial\tilde{\boldsymbol{A}}}{\partial\beta_l}\right)^\top \boldsymbol{\Gamma}^{-1}\frac{\partial\tilde{\boldsymbol{A}}}{\partial\beta_j} + \tilde{\boldsymbol{A}}^\top\boldsymbol{\Gamma}^{-1}\frac{\partial^2\tilde{\boldsymbol{A}}}{\partial\beta_j\partial\beta_l}\right] + \frac{\partial}{\partial\beta_l}\left(\frac{\partial\log(\tilde{a})}{\partial\beta_j}\right) \tag{4.1}$$

$$\ddot{\ell}_{\beta_j\varrho_l} = \tilde{\boldsymbol{A}}^\top\left(\boldsymbol{\Gamma}^{-1}\frac{\partial\boldsymbol{\Gamma}}{\partial\varrho_l}\boldsymbol{\Gamma}^{-1}\right)\frac{\partial\tilde{\boldsymbol{A}}}{\partial\beta_j} \tag{4.2}$$

$$\ddot{\ell}_{\varrho_j\varrho_l} = -\frac{1}{2}\frac{\partial}{\partial\varrho_l}\left[\mathrm{tr}\left(\boldsymbol{\Gamma}^{-1}\frac{\partial\boldsymbol{\Gamma}}{\partial\varrho_j}\right)\right] + \tag{4.3}$$

$$\frac{1}{2}\tilde{\boldsymbol{A}}^\top\left\{\left(-\boldsymbol{\Gamma}^{-1}\frac{\partial\boldsymbol{\Gamma}}{\partial\varrho_l}\boldsymbol{\Gamma}^{-1}\frac{\partial\boldsymbol{\Gamma}}{\partial\varrho_j} + \boldsymbol{\Gamma}^{-1}\frac{\partial^2\boldsymbol{\Gamma}}{\partial\varrho_l\partial\varrho_j}\right)\boldsymbol{\Gamma}^{-1} - \boldsymbol{\Gamma}^{-1}\frac{\partial\boldsymbol{\Gamma}}{\partial\varrho_j}\boldsymbol{\Gamma}^{-1}\frac{\partial\boldsymbol{\Gamma}}{\partial\varrho_l}\boldsymbol{\Gamma}^{-1}\right\}\tilde{\boldsymbol{A}} \tag{4.4}$$

where

$$\frac{\partial^2\tilde{A}_k}{\partial\beta_j\partial\beta_l} = \frac{1}{\alpha\gamma_\alpha\sqrt{t_kQ_k}}\left[\left(\frac{3t_k\gamma_\alpha^2}{8Q_k^2} + \frac{1}{2Q_k}\right)\frac{1}{h'(Q_k)}\right. \tag{4.5}$$

$$\left. + \left(\frac{t_k\gamma_\alpha^2}{4Q_k} + 1\right)\frac{h''(Q_k)}{[h'(Q_k)]^2}\right]\frac{1}{h'(Q_k)}x_{kj}x_{kl}\,(*) \tag{4.6}$$

$$\frac{\partial}{\partial\beta_l}\left(\frac{\partial\log(\tilde{a})}{\partial\beta_j}\right) = \sum_{i=1}^{n}\left\{\left[\frac{1}{2Q_i^2} - \frac{16}{(t_i\gamma_\alpha^2 + 4Q_i)^2}\right]\frac{1}{[h'(Q_i)]}\right. \tag{4.7}$$

$$\left. + \left[\frac{1}{2Q_i} - \frac{4}{t_i\gamma_\alpha^2 + 4Q_i}\right]\frac{h''(Q_i)}{[h'(Q_i)]^2}\right\}\frac{1}{h'(Q_k)}x_{ij}x_{il} \tag{4.8}$$

and

$$\frac{\partial^2 \mathbf{\Gamma}}{\partial \varrho_1 \varrho_2} = \frac{\partial^2 \mathbf{\Gamma}}{\partial \varrho_2 \partial \varrho_1} = \frac{\partial^2 \mathbf{\Gamma}}{\partial \varrho_1^2} = \frac{\partial^2 \mathbf{\Gamma}}{\partial \varrho_2^2} = 0, \tag{4.9}$$

In addition, the $p \times 1$ and 3×1 vectors $\ddot{\ell}_{\boldsymbol{\beta}\alpha} = [\ddot{\ell}_{\alpha\boldsymbol{\beta}}]^\top$ and $\ddot{\ell}_{\boldsymbol{\phi}\alpha} = [\ddot{\ell}_{\alpha\boldsymbol{\phi}}]^\top$, respectively, have elements given by

$$\begin{aligned}
\ddot{\ell}_{\alpha\beta_j} &= -\left[\left(\frac{\partial \tilde{\boldsymbol{A}}}{\partial \beta_j}\right)^\top \mathbf{\Gamma}^{-1}\frac{\partial \tilde{\boldsymbol{A}}}{\partial \alpha} + \tilde{\boldsymbol{A}}^\top \mathbf{\Gamma}^{-1}\frac{\partial^2 \tilde{\boldsymbol{A}}}{\partial \alpha \partial \beta_j}\right] + \frac{\partial}{\partial \alpha}\left(\frac{\partial \log(\tilde{a})}{\partial \beta_j}\right) && (4.10)\\
\ddot{\ell}_{\alpha\varrho_j} &= \tilde{\boldsymbol{A}}^\top\left(\mathbf{\Gamma}^{-1}\frac{\partial \mathbf{\Gamma}}{\partial \varrho_j}\mathbf{\Gamma}^{-1}\right)\frac{\partial \tilde{\boldsymbol{A}}}{\partial \alpha}, && (4.11)
\end{aligned}$$

where $\partial^2 \tilde{\boldsymbol{A}}/\partial\alpha\partial\beta_j = (\partial^2 \tilde{A}_k/\partial\alpha\partial\beta_j)$, with

$$\begin{aligned}
\frac{\partial^2 \tilde{A}_k}{\partial \alpha \partial \beta_j} &= \left[\frac{1}{(\alpha\gamma_\alpha)^2}(\gamma_\alpha + \alpha\gamma_\alpha')\left(\frac{t_k\gamma_\alpha^2}{4Q_k}+1\right) - \frac{1}{\alpha\gamma_\alpha}\left(\frac{t_k\gamma_\alpha\gamma_\alpha'}{2Q_k}\right)\right] && (4.12)\\
&\quad \times \frac{1}{\sqrt{t_k Q_k}}\frac{1}{h'(Q_k)}x_{kj}\ (*) && (4.13)\\
\frac{\partial}{\partial \alpha}\left(\frac{\partial \log(\tilde{a})}{\partial \beta_j}\right) &= -\sum_{i=1}^n \left(\frac{8t_i\gamma_\alpha\gamma_\alpha'}{(t_i\gamma_\alpha^2 + 4Q_i)^2}\right)\frac{1}{h'(Q_i)}x_{ij} && (4.14)
\end{aligned}$$

Furthermore, we have

$$\ddot{\ell}_{\alpha\alpha} = -\left[\left(\frac{\partial \tilde{\boldsymbol{A}}}{\partial \alpha}\right)^\top \mathbf{\Gamma}^{-1}\frac{\partial \tilde{\boldsymbol{A}}}{\partial \alpha} + \tilde{\boldsymbol{A}}^\top \mathbf{\Gamma}^{-1}\frac{\partial^2 \tilde{\boldsymbol{A}}}{\partial \alpha^2}\right] + \frac{\partial^2 \log(\tilde{a})}{\partial \alpha^2}, \tag{4.15}$$

where $\partial^2 \tilde{\boldsymbol{A}}/\partial\alpha^2 = (\partial^2 \tilde{A}_k/\partial\alpha^2)$, with

$$\begin{aligned}
\frac{\partial^2 \tilde{A}_k}{\partial \alpha^2} &= \sqrt{\frac{4Q_k}{t_k}}\left\{\left(\frac{t_k\gamma_\alpha^2}{4Q_k}-1\right)\left[\frac{2}{(\alpha\gamma_\alpha)^3}(\gamma_\alpha+\alpha\gamma_\alpha')^2 - \frac{2\gamma_\alpha' + \alpha\gamma_\alpha''}{(\alpha\gamma_\alpha)^2}\right]\right. && (4.16)\\
&\quad \left. -\frac{(\gamma_\alpha+\alpha\gamma_\alpha')t_k\gamma_\alpha\gamma_\alpha'}{2Q_k(\alpha\gamma_\alpha)^2} + \frac{t_k}{2Q_k}\left(\frac{\gamma_\alpha''\alpha - \gamma_\alpha'}{\alpha^2}\right)\right\} && (4.17)\\
\frac{\partial^2 \log(\tilde{a})}{\partial \alpha^2} &= -n\frac{(2\gamma_\alpha' + \alpha\gamma_\alpha'')(\alpha\gamma_\alpha) - (\gamma_\alpha+\alpha\gamma_\alpha')^2}{(\alpha\gamma_\alpha)^2} && (4.18)\\
&\quad + \sum_{i=1}^n 2t_i\frac{([\gamma_\alpha']^2 + \gamma_\alpha\gamma_\alpha'')(t_i\gamma_\alpha^2+4Q_i) - 2t_i\gamma_\alpha^2(\gamma_\alpha')^2}{(t_i\gamma_\alpha^2+4Q_i)^2} && (4.19)
\end{aligned}$$

Appendix B: Elements of perturbation matrices for the BS spatial model

For the model defined by (3.2), we have

$$\frac{\partial \ell(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \omega_j} = -\tilde{\boldsymbol{A}}_{\boldsymbol{\omega}}^{\top} \boldsymbol{\Gamma}^{-1} \frac{\partial \tilde{\boldsymbol{A}}_{\boldsymbol{\omega}}}{\partial \omega_j} + \frac{\partial}{\partial \omega_j} \left[\log(\tilde{a}_{\boldsymbol{\omega}})\right]. \tag{4.20}$$

The corresponding $(p+3)\times n$ perturbation matrix is given by $\boldsymbol{\Delta} = \left(\frac{\partial \ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \theta_r \omega_j}\right)$, where $r = 1, \cdots, p+3$ and $j = 1, \ldots, n$, with $\theta_1 = \beta_1, \ldots, \theta_p = \beta_p$, $\theta_{p+1} = \phi_1$, $\theta_{p+2} = \phi_2$ and $\theta_{p+3} = \alpha$. The elements of this matrix are given by

$$\begin{aligned}
\frac{\partial \ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \beta_r \partial w_j} &= -\left[\left(\frac{\partial \tilde{\boldsymbol{A}}_{\boldsymbol{\omega}}}{\partial \beta_r}\right)^{\top} \boldsymbol{\Gamma}^{-1} \frac{\partial \tilde{\boldsymbol{A}}_{\boldsymbol{\omega}}}{\partial \omega_j} + \tilde{\boldsymbol{A}}_{\boldsymbol{\omega}}^{\top} \boldsymbol{\Gamma}^{-1} \frac{\partial^2 \tilde{\boldsymbol{A}}_{\boldsymbol{\omega}}}{\partial \omega_j \partial \beta_r}\right] + \frac{\partial}{\partial \beta_r}\left(\frac{\partial \log(\tilde{a}_{\boldsymbol{\omega}})}{\partial \omega_j}\right), \\
\frac{\partial \ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \phi_r \partial w_j} &= -\left[\left(\frac{\partial \tilde{\boldsymbol{A}}_{\boldsymbol{\omega}}}{\partial \phi_r}\right)^{\top} \boldsymbol{\Gamma}^{-1} \frac{\partial \tilde{\boldsymbol{A}}_{\boldsymbol{\omega}}}{\partial \omega_j} + \tilde{\boldsymbol{A}}_{\boldsymbol{\omega}}^{\top} \frac{\partial \boldsymbol{\Gamma}^{-1}}{\partial \phi_r} \frac{\partial \tilde{\boldsymbol{A}}_{\boldsymbol{\omega}}}{\partial \omega_j} + \tilde{\boldsymbol{A}}_{\boldsymbol{\omega}}^{\top} \boldsymbol{\Gamma}^{-1} \frac{\partial^2 \tilde{\boldsymbol{A}}_{\boldsymbol{\omega}}}{\partial \omega_j \partial \phi_r}\right] \qquad (4.21) \\
&\quad + \frac{\partial}{\partial \beta_r}\left(\frac{\partial \log(\tilde{a}_{\boldsymbol{\omega}})}{\partial \omega_j}\right), \\
\frac{\partial \ell^2(\boldsymbol{\theta}|\boldsymbol{\omega})}{\partial \alpha \partial w_j} &= -\left[\left(\frac{\partial \tilde{\boldsymbol{A}}_{\boldsymbol{\omega}}}{\partial \alpha}\right)^{\top} \boldsymbol{\Gamma}^{-1} \frac{\partial \tilde{\boldsymbol{A}}_{\boldsymbol{\omega}}}{\partial \omega_j} + \tilde{\boldsymbol{A}}_{\boldsymbol{\omega}}^{\top} \boldsymbol{\Gamma}^{-1} \frac{\partial^2 \tilde{\boldsymbol{A}}_{\boldsymbol{\omega}}}{\partial \omega_j \partial \alpha}\right] + \frac{\partial}{\partial \alpha}\left(\frac{\partial \log(\tilde{a}_{\boldsymbol{\omega}})}{\partial \omega_j}\right) \qquad (4.22)
\end{aligned}$$

Perturbation in the response:

In the case of perturbation in the response, we have

$$\begin{aligned}
\frac{\partial \tilde{A}_k(\boldsymbol{\omega})}{\partial \omega_j} &= a(t_k(\boldsymbol{\omega}); \alpha, Q_k) \cdot A_{kj} \\
\frac{\partial \tilde{A}_k(\boldsymbol{\omega})}{\partial \beta_r} &= -\frac{1}{\alpha\gamma_\alpha\sqrt{t_k(\boldsymbol{\omega})Q_k}}\left(\frac{t_k(\boldsymbol{\omega})\gamma_\alpha^2}{4Q_k}+1\right)\frac{1}{h'(Q_k)}x_{kr} \\
\frac{\partial^2 \tilde{A}_k(\boldsymbol{\omega})}{\partial \omega_j \partial \beta_r} &= \left(-\frac{\gamma_\alpha^2}{4Q_k}+\frac{1}{t_k(\boldsymbol{\omega})}\right)\frac{1}{2\alpha\gamma_\alpha\sqrt{Q_k t_k(\boldsymbol{\omega})}}\frac{1}{h'(Q_k)}x_{kr}A_{kj}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial \beta_r}\left(\frac{\partial \log(\tilde{a}\boldsymbol{\omega})}{\partial \omega_j}\right) &= -4\sum_{k=1}^{n}\left(\frac{t_k(\boldsymbol{\omega})\gamma_\alpha}{t_k^2(\boldsymbol{\omega})\gamma_\alpha^2+4Q_k t_k(\boldsymbol{\omega})}\right)^2\frac{1}{h'(Q_k)}x_{kr}A_{kj} \\
\frac{\partial \tilde{A}_k(\boldsymbol{\omega})}{\partial \phi_r} &= a(t_k(\boldsymbol{\omega}); \alpha, Q_k)\frac{\partial t_k(\boldsymbol{\omega})}{\partial \phi_r} \\
\frac{\partial t_k(\boldsymbol{\omega})}{\partial \phi_r} &= \{k\text{-th file of } \frac{\partial \boldsymbol{A}}{\partial \phi_r}\} \cdot \boldsymbol{\omega} \\
\frac{\partial \boldsymbol{A}}{\partial \phi_r} &= \boldsymbol{A}\left(\frac{1}{\alpha}\boldsymbol{\Gamma}^{-1/2}\frac{\partial \boldsymbol{\Gamma}^{1/2}}{\partial \phi_r}\boldsymbol{\Gamma}^{-1/2}+\frac{\alpha}{4}\frac{\partial \boldsymbol{\Gamma}^{1/2}}{\partial \phi_r}\right)\boldsymbol{A} \\
\frac{\partial^2 \tilde{A}_k(\boldsymbol{\omega})}{\partial \omega_j \partial \phi_r} &= \frac{1}{\alpha\gamma_\alpha\sqrt{4Q_k}}\left[-\frac{1}{2t_k^{3/2}(\boldsymbol{\omega})}\left(\frac{\gamma_\alpha^2}{2}+\frac{2Q_k}{t_k(\boldsymbol{\omega})}\right)-\frac{2Q_k}{t_k^{3/2}(\boldsymbol{\omega})}\right]\cdot \\
&\quad \frac{\partial t_k(\boldsymbol{\omega})}{\partial \phi_r}A_{kj}+a(t_k(\boldsymbol{\omega}); \alpha, Q_k)\cdot\frac{\partial A_{kj}}{\partial \phi_r} \\
\frac{\partial}{\partial \phi_r}\left(\frac{\partial \log(\tilde{a}\boldsymbol{\omega})}{\partial \omega_j}\right) &= \sum_{k=1}^{n}\left[\frac{1}{2}\frac{1}{t_k^2(\boldsymbol{\omega})}\frac{\partial t_k(\boldsymbol{\omega})}{\partial \phi_r}+4\frac{Q_k}{(t_k^2(\boldsymbol{\omega})\gamma_\alpha^2+4Q_k t_k(\boldsymbol{\omega}))^2}\cdot\right. \\
&\quad \left.\cdot\left(2t_k(\boldsymbol{\omega})\frac{\partial t_k(\boldsymbol{\omega})}{\partial \phi_r}\gamma_\alpha^2+2t_k^2(\boldsymbol{\omega})\gamma_\alpha\gamma_\alpha'+4Q_k\frac{\partial t_k(\boldsymbol{\omega})}{\partial \phi_r}\right)\right]\cdot A_{kj} \\
&\quad +\sum_{k=1}^{n}\left[-\frac{1}{2}\frac{1}{t_k(\boldsymbol{\omega})}-4\frac{Q_k}{t_k^2(\boldsymbol{\omega})\gamma_\alpha^2+4Q_k t_k(\boldsymbol{\omega})}\right]\cdot\frac{\partial A_{kj}}{\partial \phi_r} \\
\frac{\partial \tilde{A}_k(\boldsymbol{\omega})}{\partial \alpha} &= \frac{1}{\sqrt{4Q_k}}\left[-\frac{\gamma_\alpha+\alpha\gamma_\alpha'}{\alpha^2\gamma_\alpha^2}\left(\sqrt{t_k(\boldsymbol{\omega})}\gamma_\alpha^2-4Q_k t_k^{-1/2}(\boldsymbol{\omega})\right)\right. \\
&\quad \left.+\frac{1}{\alpha\gamma_\alpha}\left(\frac{1}{2\sqrt{t_k(\boldsymbol{\omega})}}\frac{\partial t_k(\boldsymbol{\omega})}{\partial \alpha}\gamma_\alpha^2+2\sqrt{t_k(\boldsymbol{\omega})}\gamma_\alpha\gamma_\alpha'+2Q_k t_k^{-3/2}(\boldsymbol{\omega})\frac{\partial t_k(\boldsymbol{\omega})}{\partial \alpha}\right)\right] \\
\frac{\partial t_k(\boldsymbol{\omega})}{\partial \alpha} &= \{k\text{-th file of } \boldsymbol{A}\left(\frac{1}{\alpha^2}\boldsymbol{\Gamma}^{-1/2}+\frac{1}{4}\boldsymbol{\Gamma}^{1/2}\right)\boldsymbol{A}\}\cdot\boldsymbol{\omega}
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial^2 \tilde{A}_k(\boldsymbol{\omega})}{\partial\omega_j\partial\alpha} &= \frac{1}{\sqrt{4Q_k}}\left[\left(-\frac{\gamma_\alpha+\alpha\gamma'_\alpha}{\alpha\gamma_\alpha}\frac{1}{\sqrt{t_k(\boldsymbol{\omega})}}-\frac{1}{2\alpha\gamma_\alpha t_k^{3/2}(\boldsymbol{\omega})}\frac{\partial t_k(\boldsymbol{\omega})}{\partial\alpha}\right)\cdot\right.\\
&\quad\left.\cdot\left(\frac{\gamma_\alpha^2}{2}+\frac{2Q_k}{t_k(\boldsymbol{\omega})}\right)+\frac{1}{\alpha\gamma_\alpha}\left(\gamma_\alpha\gamma'_\alpha-\frac{2Q_k}{t_k^2(\boldsymbol{\omega})}\frac{\partial t_k(\boldsymbol{\omega})}{\partial\alpha}\right)\right]\cdot A_{kj}\\
&\quad+a(t_k\boldsymbol{\omega};\alpha,Q_k)\cdot\{\text{element } kj \text{ of } \frac{\partial\boldsymbol{A}}{\partial\alpha}\},\\
\frac{\partial}{\partial\alpha}\left(\frac{\partial\log\tilde{a}\boldsymbol{\omega}}{\partial\omega_j}\right) &= \sum_{k=1}^{n}\left[\frac{1}{2}\frac{1}{t_k^2(\boldsymbol{\omega})}\frac{\partial t_k(\boldsymbol{\omega})}{\partial\alpha}+4\frac{Q_k}{(t_k^2(\boldsymbol{\omega})\gamma_\alpha^2+4Q_kt_k(\boldsymbol{\omega}))^2}\cdot\right.\\
&\quad\left.\cdot\left(2t_k(\boldsymbol{\omega})\frac{\partial t_k(\boldsymbol{\omega})}{\partial\alpha}\gamma_\alpha^2+2\gamma_\alpha\gamma'_\alpha t_k^2(\boldsymbol{\omega})+4Q_k\frac{\partial t_k(\boldsymbol{\omega})}{\partial\alpha}\right)\right]\cdot A_{kj}\\
&\quad+\sum_{k=1}^{n}\left[-\frac{1}{2}\frac{1}{t_k(\boldsymbol{\omega})}-4\frac{Q_k}{t_k^2(\boldsymbol{\omega})\gamma_\alpha^2+4Q_kt_k(\boldsymbol{\omega})}\right]\cdot\frac{\partial A_{kj}}{\partial\alpha},
\end{aligned}
$$

where $\dfrac{\partial\boldsymbol{\Gamma}^{-1}}{\partial\phi_r}=-\boldsymbol{\Gamma}^{-1}\dfrac{\partial\boldsymbol{\Gamma}}{\partial\phi_r}\boldsymbol{\Gamma}^{-1}$, $\dfrac{\partial\boldsymbol{A}}{\partial\alpha}=\boldsymbol{A}\left(\dfrac{1}{\alpha^2}\boldsymbol{\Gamma}^{-1/2}+\dfrac{1}{4}\boldsymbol{\Gamma}^{1/2}\right)\boldsymbol{A}$ and to calculate $\dfrac{\partial\boldsymbol{\Gamma}^{1/2}}{\partial\phi_r}$, see De Bastiani et al. (2015).

Perturbation in the covariate:

The elements defined in (4.22) has as components the next expressions

$$
\begin{aligned}
\frac{\partial\tilde{A}_k(\boldsymbol{\omega})}{\partial\omega_j} &= -\frac{1}{\alpha\gamma_\alpha}\left(\frac{t_k}{Q_k(\boldsymbol{\omega})}\right)^{1/2}\left(\frac{\gamma_\alpha^2}{4Q_k(\boldsymbol{\omega})}+1\right)\frac{1}{h'(Q_k(\boldsymbol{\omega}))}\beta_lA_{kj},\\
\frac{\partial\tilde{A}_k(\boldsymbol{\omega})}{\partial\beta_r} &= -\frac{1}{\alpha\gamma_\alpha}\left(\frac{t_k}{Q_k(\boldsymbol{\omega})}\right)^{1/2}\left(\frac{\gamma_\alpha^2}{4Q_k(\boldsymbol{\omega})}+1\right)\frac{1}{h'(Q_k(\boldsymbol{\omega}))}Z_{kr},
\end{aligned}
$$

$$
\begin{aligned}
\frac{\partial^2\tilde{A}_k(\boldsymbol{\omega})}{\partial\omega_j\partial\beta_r} &= \frac{1}{\alpha\gamma_\alpha}\left(\frac{t_k}{Q_k(\boldsymbol{\omega})}\right)^{1/2}\left[\left(\frac{3\gamma_\alpha^2}{8Q_k^2(\boldsymbol{\omega})}+\frac{1}{2Q_k(\boldsymbol{\omega})}\right)\right.\\
&\quad\left.+\left(\frac{\gamma_\alpha^2}{4Q_k(\boldsymbol{\omega})}+1\right)\frac{h''(Q_k(\boldsymbol{\omega}))}{(h'(Q_k(\boldsymbol{\omega})))^2}\right]\frac{1}{h'(Q_k(\boldsymbol{\omega}))}\beta_lA_{kj}Z_{kr}\\
\frac{\partial}{\partial\beta_r}\left(\frac{\partial\log(\tilde{a}\boldsymbol{\omega})}{\partial\omega_j}\right) &= \sum_{k=1}^{n}\left[\left(\frac{1}{2Q_k^2(\boldsymbol{\omega})}-\frac{16}{(t_k\gamma_\alpha^2+4Q_k(\boldsymbol{\omega}))^2}\right)\frac{1}{h'(Q_k(\boldsymbol{\omega}))}\right.\\
&\quad\left.+\left(\frac{1}{2Q_k(\boldsymbol{\omega})}-\frac{4}{t_k\gamma_\alpha^2+4Q_k(\boldsymbol{\omega})}\right)\frac{h''(Q_k(\boldsymbol{\omega}))}{(h'(Q_k(\boldsymbol{\omega})))^2}\right]\cdot\frac{1}{h'(Q_k(\boldsymbol{\omega}))}\beta_lA_{kj}Z_{kr}\\
&\quad+\sum_{k=1}^{n}\left[-\frac{1}{2Q_k(\boldsymbol{\omega})}+\frac{4}{t_k\gamma_\alpha^2+4Q_k(\boldsymbol{\omega})}\right]\cdot\frac{1}{h'(Q_k(\boldsymbol{\omega}))}A_{kj}\rho_{rl}
\end{aligned}
$$

$$\begin{aligned}
\frac{\partial \tilde{A}_k(\boldsymbol{\omega})}{\partial \phi_r} &= -\frac{1}{\alpha\gamma_\alpha}\left(\frac{t_k}{Q_k(\boldsymbol{\omega})}\right)^{1/2}\left(\frac{\gamma_\alpha^2}{4Q_k(\boldsymbol{\omega})}+1\right)\frac{1}{h'(Q_k(\boldsymbol{\omega}))}\frac{\partial A_k}{\partial \phi_r}\boldsymbol{\omega} \\
\frac{\partial^2 \tilde{A}_k(\boldsymbol{\omega})}{\partial \omega_j \partial \phi_r} &= \beta_l\left\{\frac{1}{\alpha\gamma_\alpha}\left(\frac{t_k}{Q_k(\boldsymbol{\omega})}\right)^{1/2}\left[\left(\frac{3\gamma_\alpha^2}{8Q_k^2(\boldsymbol{\omega})}+\frac{1}{2Q_k(\boldsymbol{\omega})}\right)\frac{1}{h'(Q_k(\boldsymbol{\omega}))}+\right.\right. \\
&\quad \left. +\left(\frac{\gamma_\alpha^2}{4Q_k(\boldsymbol{\omega})}+1\right)\frac{h''(Q_k(\boldsymbol{\omega}))}{(h'(Q_k(\boldsymbol{\omega})))^2}\right]\frac{1}{h'(Q_k(\boldsymbol{\omega}))}\frac{\partial Q_k(\boldsymbol{\omega})}{\partial \phi_r}\frac{\partial A_k}{\partial \phi_r}\boldsymbol{\omega}A_{kj} \\
&\quad \left. -\frac{1}{\alpha\gamma_\alpha}\left(\frac{t_k}{Q_k(\boldsymbol{\omega})}\right)^{1/2}\left(\frac{\gamma_\alpha^2}{4Q_k(\boldsymbol{\omega})}+1\right)\frac{1}{h'(Q_k(\boldsymbol{\omega}))}\frac{\partial A_{kj}}{\partial \phi_r}\right\}
\end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial \phi_r}\left(\frac{\partial \log \tilde{a}\boldsymbol{\omega}}{\partial \omega_j}\right) &= \beta_l\sum_{k=1}^{n}\left\{\left[\left(\frac{1}{2Q_k^2(\boldsymbol{\omega})}-\frac{16}{(t_k\gamma_\alpha^2+4Q_k(\boldsymbol{\omega}))^2}\right)\frac{1}{h'(Q_k(\boldsymbol{\omega}))}\right.\right. \\
&\quad \left. +\left(\frac{1}{2Q_k(\boldsymbol{\omega})}-\frac{4}{t_k\gamma_\alpha^2+4Q_k(\boldsymbol{\omega})}\right)\frac{h''(Q_k(\boldsymbol{\omega}))}{(h'(Q_k(\boldsymbol{\omega})))^2}\right]\frac{1}{h'(Q_k(\boldsymbol{\omega}))}\frac{\partial Q_k(\boldsymbol{\omega})}{\partial \phi_r}A_{kj} \\
&\quad \left. +\left[-\frac{1}{2Q_k(\boldsymbol{\omega})}+\frac{4}{t_k\gamma_\alpha^2+4Q_k(\boldsymbol{\omega})}\right]\frac{1}{h'(Q_k(\boldsymbol{\omega}))}\frac{\partial A_{kj}}{\partial \phi_r}\right\},
\end{aligned}$$

$$\begin{aligned}
\frac{\partial \tilde{A}_k(\boldsymbol{\omega})}{\partial \alpha} = \frac{1}{\sqrt{4t_k}}&\left[-\frac{\gamma_\alpha+\alpha\gamma_\alpha'}{\alpha^2\gamma_\alpha^2}\left(\frac{t_k\gamma_\alpha^2-4Q_k(\boldsymbol{\omega})}{\sqrt{Q_k(\boldsymbol{\omega})}}\right)+\frac{1}{\alpha\gamma_\alpha Q_k(\boldsymbol{\omega})}\cdot\right. \\
&\left.\left(\left(2t_k\gamma_\alpha\gamma_\alpha'-4\frac{\partial Q_k(\boldsymbol{\omega})}{\partial \alpha}\right)\sqrt{Q_k(\boldsymbol{\omega})}-\frac{1}{2\sqrt{Q_k(\boldsymbol{\omega})}}\frac{\partial Q_k(\boldsymbol{\omega})}{\partial \alpha}(t_k\gamma_\alpha^2-4Q_k(\boldsymbol{\omega}))\right)\right],
\end{aligned}$$

$$\begin{aligned}
\frac{\partial Q_k(\boldsymbol{\omega})}{\partial \alpha} &= \frac{1}{h'(Q_k(\boldsymbol{\omega}))}\beta_l\frac{\partial A_k}{\partial \alpha}\boldsymbol{\omega}, \\
\frac{\partial^2 \tilde{A}_k(\boldsymbol{\omega})}{\partial \omega_j \partial \alpha} &= -\beta_l\left\{-\frac{\gamma_\alpha+\alpha\gamma_\alpha'}{\alpha^2\gamma_\alpha^2}\left(\frac{t_k}{Q_k(\boldsymbol{\omega})}\right)^{1/2}\left(\frac{\gamma_\alpha^2}{4Q_k(\boldsymbol{\omega})}+1\right)\frac{1}{h'(Q_k(\boldsymbol{\omega}))}A_{kj}\right. \\
&\quad +\frac{1}{\alpha\gamma_\alpha}\left[-\frac{1}{2}\left(\frac{Q_k(\boldsymbol{\omega})}{t_k}\right)^{-3/2}\frac{1}{t_k}\frac{\partial Q_k(\boldsymbol{\omega})}{\partial \alpha}\right]\left(\frac{\gamma_\alpha^2}{4Q_k(\boldsymbol{\omega})}+1\right)\frac{1}{h'(Q_k(\boldsymbol{\omega}))}A_{kj} \\
&\quad +\frac{1}{\alpha\gamma_\alpha}\left(\frac{t_k}{Q_k(\boldsymbol{\omega})}\right)^{1/2}\cdot\left[\frac{8\gamma_\alpha\gamma_\alpha' Q_k(\boldsymbol{\omega})-4\frac{\partial Q_k(\boldsymbol{\omega})}{\partial \alpha}\gamma_\alpha^2}{16Q_k^2(\boldsymbol{\omega})}\right]\frac{1}{h'(Q_k(\boldsymbol{\omega}))}A_{kj} \\
&\quad +\frac{1}{\alpha\gamma_\alpha}\left(\frac{t_k}{Q_k(\boldsymbol{\omega})}\right)^{1/2}\left(\frac{\gamma_\alpha^2}{4Q_k(\boldsymbol{\omega})}+1\right)\left(-\frac{1}{(h'(Q_k(\boldsymbol{\omega})))^2}\right)h''(Q_k(\boldsymbol{\omega}))\frac{\partial Q_k(\boldsymbol{\omega})}{\partial \alpha}A_{kj} \\
&\quad \left. +\frac{1}{\alpha\gamma_\alpha}\left(\frac{t_k}{Q_k(\boldsymbol{\omega})}\right)^{1/2}\left(\frac{\gamma_\alpha^2}{4Q_k(\boldsymbol{\omega})}+1\right)\frac{1}{h'(Q_k(\boldsymbol{\omega}))}\frac{\partial A_{kj}}{\partial \alpha}\right\},
\end{aligned}$$

$$
\begin{aligned}
\frac{\partial}{\partial\alpha}\left(\frac{\partial \log \tilde{a}\boldsymbol{\omega}}{\partial\omega_j}\right) \quad = \quad & \beta_l \sum_{k=1}^{n} \left\{ \left[\frac{1}{2Q_k^2(\boldsymbol{\omega})} \frac{\partial Q_k(\boldsymbol{\omega})}{\partial\alpha} - 4 \frac{2t_k\gamma_\alpha\gamma_\alpha' + 4\frac{\partial Q_k(\boldsymbol{\omega})}{\partial\alpha}}{(t_k\gamma_\alpha^2 + 4Q_k(\boldsymbol{\omega}))^2} \right] \frac{1}{h'(Q_k(\boldsymbol{\omega}))} A_{kj} \right. \\
& + \left[-\frac{1}{2Q_k(\boldsymbol{\omega})} + \frac{4}{t_k\gamma_\alpha^2 + 4Q_k(\boldsymbol{\omega})} \right] \cdot \left(-\frac{h''(Q_k(\boldsymbol{\omega}))}{(h'(Q_k(\boldsymbol{\omega})))^2} \right) \frac{\partial Q_k(\boldsymbol{\omega})}{\partial\alpha} A_{kj} \\
& \left. + \left[-\frac{1}{2Q_k(\boldsymbol{\omega})} + \frac{4}{t_k\gamma_\alpha^2 + 4Q_k(\boldsymbol{\omega})} \right] \frac{1}{h'(Q_k(\boldsymbol{\omega}))} \frac{\partial A_{kj}}{\partial\alpha} \right\}
\end{aligned}
$$

where $\rho_{rl} = 1$, if $r = l$ and $\rho_{rl} = 0$, if $r \neq l$, $Z_{kr} = 1$ for $r = 1$, $Z_{kr} = X_{kl}(\boldsymbol{\omega})$, if $r = l$, and $Z_{kr} = X_{kr}$, for $r \neq 1, l$

References

Ahmed, S., Castro-Kuriss, C., Leiva, V., Flores, E., Sanhueza, A. (2010). Truncated version of the Birnbaum-Saunders distribution with an application in financial risk. *Pakistan Journal of Statistics*, **26**:293–311.

Athayde, E., Azevedo, A., Barros, M., Leiva, V. (2019). Failure rate of Birnbaum-Saunders distributions: shape, change-point, estimation and robustness. *Brazilian Journal of Probability and Statistics*, **33**: 301–328.

Athayde, E., Azevedo, C., Leiva, V., Sanhueza, A. (2012). About Birnbaum-Saunders distributions based on the Johnson system. *Communications in Statistics: Theory and Methods*, **41**:2061–2079.

Atkinson, A.C. (1985). *Plots, Transformations and Regression.* Oxford University, Oxford, UK.

Balakrishnan, N., Gupta, R., Kundu, D., Leiva, V., Sanhueza, A. (2011). On some mixture models based on the Birnbaum-Saunders distribution and associated inference. *Journal of Statistical Planning and Inference*, **141**:2175–2190.

Balakrishnan, N., Kundu, D. (2019). Birnbaum-Saunders distribution: A review of models, analysis, and application. *Applied Stochastic Models in Business and Industry*, **35**:4–49.

Birnbaum, Z.W., Saunders, S.C. (1969). A new family of life distributions. *Journal of Applied Probability*, **6**:319–327.

Bourguignon, M., Leão, J., Leiva, V., Santos-Neto, M. (2017). The transmuted Birnbaum-Saunders distribution. *REVSTAT Statistical Journal*, **15**:601–628.

Bourguignon, M., Silva, R., Cordeiro, G. (2014). A new class of fatigue life distributions. *Journal of Statistical Computation and Simulation*, **84**:2619–2635.

Cambardella, C., Moorman, T., Novak, J., Parkin, T., Karlen, D., Turco, R., Konopka, A. (1994). Field-scale variability of soil properties in central Iowa soils. *Soil Science Society of America Journal*, **58**: 1501–1511.

Caro-Lopera, F., Leiva, V., Balakrishnan, N. (2012). Connection between the Hadamard and matrix products with an application to matrix-variate Birnbaum-Saunders distributions. *Journal of Multivariate Analysis*, **104**:126–139.

Castillo, N., Gómez, H., Bolfarine, H. (2011). Epsilon Birnbaum-Saunders distribution family: Properties and inference. *Statistical Papers*, **52**:871–883.

Cook, R.D. (1986). Assessment of local influence. *Journal of the Royal Statistical Society B*, **48**:133–169.

Cook, R.D. (1987). Influence assessment. *Journal of Applied Statistics*, **14**, 117-131.

Cook, R.D., Peña, D., Weisberg, S. (1988). The likelihood displacement: a unifying principle for influence measures. *Communications in Statistics - Theory and Methods*, **17**: 623-640.

Cook, R.D., Weisberg, S. (1982). *Residuals and Influence in Regression*. Chapman and Hall, London, UK.

Cordeiro, M., Lemonte, J. (2011). The β-Birnbaum-Saunders distribution: An improved distribution for fatigue life modeling. *Computational Statistics and Data Analysis*, **55**:1445–1461.

Cox, D.R., Hinkley, D.V. (1974). *Theoretical Statistics*. Chapman and Hall, London, UK.

Cysneiros, F.J.A., Leiva, V., Liu, S., Marchant, C., Scalco, P. (2019). A Cobb-Douglas type model with stochastic restrictions: formulation, local influence diagnostics and data analytics in economics. *Quality and Quantity*, **53**:1693–1719

Davino, C., Furno, M., Vistocco, D. (2014). *Quantile Regression*. Wiley, Chichester, UK.

De Bastiani, F., Cysneiros, A.H.M.A., Uribe-Opazo, M.A., Galea, M. (2015). Influence diagnostics in elliptical spatial linear models. *TEST*, **24**: 322-340.

De Bastiani F., Uribe-Opazo, M.A., Galea, M., Cysneiros, A.H.M.A. (2018). Case-deletion diagnostics for spatial linear mixed models. *Spatial Statistics*, **28**: 284–303.

De la Fuente-Mella, H., Rojas-Fuentes, J.L., Leiva, V (2020). Econometric modeling of productivity and technical efficiency in the Chilean manufacturing industry. *Computers and Industrial Engineering*, **139**:105793.

Desousa, M.F., Saulo, H., Leiva, V., Scalco, P. (2018). On a tobit-Birnbaum-Saunders model with an application to antibody response to vaccine. *Journal of Applied Statistics*, **45**: 932–955.

Díaz-García, J.A., Galea, M., Leiva, V. (2003). Influence diagnostics for elliptical multivariate linear regression models. *Communications in Statistics: Theory and Methods*, **32**:625–641.

Díaz-García, J., Leiva, V. (2005). A new family of life distributions based on elliptically contoured distributions. *Journal of Statistical Planning and Inference*, **128**:445–457.

Diggle, P., Ribeiro, P. (2007). *Model-Based Geoestatistics*. Springer, New York, US.

Dunn, P., Smyth, G. (1996). Randomized quantile residuals. *Journal of Computational and Graphical Statistics*, **5**:236–244.

Ferreira, M., Gomes, M.I., Leiva, V. (2012). On an extreme value version of the Birnbaum- Saunders distribution. *REVSTAT - Statistical Journal*, **10**:181–210.

Fierro, R., Leiva, V., Ruggeri, F., Sanhueza, A. (2013). On a Birnbaum-Saunders distribution arising from a non-homogeneous Poisson process. *Statistical and Probability Letters*, **83**:1233–1239.

Garcia-Papani, F., Leiva, V., Ruggeri, F., Uribe-Opazo, M.A. (2018a). Kriging with external drift in a Birnbaum-Saunders geostatistical model. *Stochastic Environmental Research and Risk Assessment*, **32**:1517-1530.

Garcia-Papani, F., Leiva, V., Uribe-Opazo, M.A., Aykroyd, R.G. (2018b). Birnbaum-Saunders spatial regression models: Diagnostics and application to chemical data. *Chemometrics and Intelligent Laboratory Systems*, **177**:114–128.

Garcia-Papani, F., Uribe-Opazo, M.A., Leiva, V., Aykroyd, R.G. (2017). Birnbaum-Saunders spatial modelling and diagnostics applied to agricultural engineering data. *Stochastic Environmental Research and Risk Assessment*, **31**:105–124.

Genton, MG., Zhang, H. (2012). Identifiability problems in some non-Gaussian spatial random fields. *Chilean Journal of Statistics*, **3**:171–179.

Gómez, H., Olivares-Pacheco, J.F., Bolfarine, H. (2009). An extension of the generalized Birnbaum-Saunders distribution. *Statistical and Probability Letters*, **79**:331–338.

Gradshteyn. I., Ryzhik, I. (2000). *Tables of Integrals, Series and Products*. Academic Press, New York.

Guiraud, P., Leiva, V., Fierro, R. (2009). A non-central version of the Birnbaum-Saunders distribution for reliability analysis. *IEEE Transactions on Reliability*, **58**:152–160.

Hao, L., Naiman, D. Q. (2007). *Quantile Regression*. Sage Publications, US.

Jamalizadeh, A., Kundu, D. (2015). Multivariate Birnbaum-Saunders distribution based on multivariate skew-normal distribution. *Journal of the Japan Statistical Society*, **45**:1–20.

Khosravi, M., Kundu, D., Jamalizadeh, A. (2015). On bivariate and mixture of bivariate Birnbaum-Saunders distributions. *Statistical Methodology*, **23**:1–17.

Koenker, R., Bassett, G. (1978). Regression quantiles. *Econometrica*, **46**:33–50.

Kostov, P. (2009). A spatial quantile regression hedonic model of agricultural land prices. *Spatial Economic Analysis*, **4**:53–72.

Kotz, S., Leiva, V., Sanhueza, A. (2010). Two new mixture models related to the inverse Gaussian distribution. *Methodology and Computing in Applied Probability*, **12**:199–212.

Krzanowski, W. (1998). *An Introduction to Statistical Modelling*. Arnold, London, UK.

Kundu, D., Balakrishnan, N., Jamalizadeh, A. (2013). Generalized multivariate Birnbaum-Saunders distributions and related inferential issues. *Journal of Multivariate Analysis*, **116**:230–244.

Kundu, D. (2015a). Bivariate log-Birnbaum-Saunders distribution. *Statistics*, **49**:900–917.

Kundu, D. (2015b). Bivariate sinh-normal distribution and a related model. *Brazilian Journal of Probability and Statistics*, **20**:590–607.

Lange, K. (2001). *Numerical analysis for statisticians*. Springer, New York

Laplace, P. (1818). *Th'eorie Analytique des Probabilit'es*. Editions Jacques Gabayr, Paris, France.

Leão, J., Leiva, V., Saulo, H., Tomazella, V. (2017b). Birnbaum-Saunders frailty regression models: Diagnostics and application to medical data. *Biometrical Journal*, **59**:291–314.

Leão, J., Leiva, V., Saulo, H., Tomazella, V. (2018b). A survival model with Birnbaum-Saunders frailty for uncensored and censored cancer data. *Brazilian Journal of Probability and Statistics*, **32**:707–729.

Leiva, V. (2016). *The Birnbaum-Saunders Distribution.* Academic Press, New York, US.

Leiva, V., Athayde, E., Azevedo, C., Marchant, C. (2011). Modeling wind energy flux by a Birnbaum-Saunders distribution with unknown shift parameter. *Journal of Applied Statistics*, **38**:2819–2838.

Leiva, V., Ferreira, M., Gomes, M.I., and Lillo, C. (2016). Extreme value Birnbaum-Saunders regression models applied to environmental data. *Stochastic Environmental Research and Risk Assessment*, **30**:1045–1058.

Leiva, V., Lillo, C., Morrás, R. (2018). On a business confidence index and its data analytics: A Chilean case. In Oliveira, T., Kitsos, C., Oliveira, A., and Grilo, L.M., editors, *Recent Studies on Risk Analysis and Statistical Modeling*, pages 61–78. Springer, Switzerland.

Leiva, V., Marchant, C., Ruggeri, F., Saulo, H. (2015a). A criterion for environmental assessment using Birnbaum-Saunders attribute control charts. *Environmetrics*, **26**:463–476.

Leiva, V., Marchant, C., Saulo, H., Aslam, M., Rojas, F. (2014a). Capability indices for Birnbaum-Saunders processes applied to electronic and food industries. *Journal of Applied Statistics*, **41**:1881–1902.

Leiva, V., Ruggeri, F., Saulo, H., Vivanco, J.F. (2017). A methodology based on the Birnbaum-Saunders distribution for reliability analysis applied to nano-materials. *Reliability Engineering and System Safety*, **157**:192–201.

Leiva, V., Sanhueza, A., Angulo, J.M. (2009). A length-biased version of the Birnbaum-Saunders distribution with application in water quality. *Stochastic Environmental Research and Risk Assessment*, **23**:299–307.

Leiva, V., Santos-Neto, M., Cysneiros, F. J.A., Barros, M. (2014b). Birnbaum-Saunders statistical modelling: A new approach. *Statistical Modelling*, **14**:21–48.

Leiva, V., Santos-Neto, M., Cysneiros, F. J.A., Barros, M. (2016b). A methodology for stochastic inventory models based on a zero-adjusted Birnbaum-Saunders distribution. *Applied Stochastic Models in Business and Industry*, **32**:74–89.

Leiva, V., Saulo, H. (2017). Environmental applications based on birnbaum-saunders models. In Adhikari, M.R., Chaubey, Y.P., and Adhikari, A., editors, *Mathematical and Statistical Applications in Life Sciences and Engineering*, pages 283–304. Springer, Singapore.

Leiva, V., Saulo, H., Leão, J., Marchant, C. (2014c). A family of autoregressive conditional duration models applied to financial data. *Computational Statistics and Data Analysis*, **79**:175–191.

Leiva, V., Saunders, S.C. (2015). Cumulative damage models. *Wiley StatsRef: Statistics Reference Online*, pages 1–10.

Leiva, V., Tejo, M., Guiraud, P., Schmachtenberg, O., Orio, P., Marmolejo, F. (2015b). Modeling neural activity with cumulative damage distributions. *Biological Cybernetics*, **109**:421–433.

Lemonte, A. (2013). A new extension of the Birnbaum Saunders distribution. *Brazilian Journal of Probability and Statistics*, **27**:133–149.

Lemonte, A., Martínez-Flores, G., Moreno-Arenas, G. (2015). Multivariate Birnbaum-Saunders distribution: Properties and associated inference. *Journal of Statistical Computation and Simulation*, **85**:374–392.

Lesaffre, E., Verbeke, G. (1998). Local influence in linear mixed models. *Biometrics*, **54**:570–582.

Lillo, C., Leiva, V., Nicolis, O., Aykroyd, R.G. (2018). L-moments of the Birnbaum-Saunders distribution and its extreme value version: Estimation, goodness of fit and application to earthquake data. *Journal of Applied Statistics*, **45**:187–209.

Lio, Y.L., Tsai, T.R., Wu, S.J. (2010). Acceptance sampling plans from truncated life tests based on the Birnbaum-Saunders distribution for percentiles. *Communications in Statistics: Simulation and Computation*, **39**:119–136.

Marchant, C., Bertin, K., Leiva, V., Saulo, H. (2013a). Generalized Birnbaum-Saunders kernel density estimators and an analysis of financial data. *Computational Statistics and Data Analysis*, **63**:1–15.

Marchant, C., Leiva, V., Cavieres, M., Sanhueza, A. (2013b). Air contaminant statistical distributions with application to PM10 in Santiago, Chile. *Reviews of Environmental Contamination and Toxicology*, **223**:1–31.

Marchant, C., Leiva, V., Cysneiros, F. J.A. (2016a). A multivariate log-linear model for Birnbaum-Saunders distributions. *IEEE Transactions on Reliability*, **65**:816–827.

Marchant, C., Leiva, V., Cysneiros, F. J.A., Liu, S. (2018). Robust multivariate control charts based on Birnbaum-Saunders distributions. *Journal of Statistical Computation and Simulation*, **88**:182–202.

Marchant, C., Leiva, V., Cysneiros, F. J.A., Vivanco, J.F. (2016b). Diagnostics in multivariate generalized Birnbaum-Saunders regression models. *Journal of Applied Statistics*, **43**:2829–2849.

Mardia, K., Marshall, R. (1984). Maximum likelihood estimation of models for residual covariance in spatial regression. *Biometrika*, **71**:135-146.

Marshall, A., Olkin, I. (2007). *Life Distributions*. Springer, New York, US.

Martinez, G., Bolfarine, H., Gómez, H. (2014). An alpha-power extension for the Birnbaum-Saunders distribution. *Statistics*, **48**:896–912.

Martinez, S., Giraldo, R., Leiva, V. (2019). Birnbaum-Saunders functional regression models for spatial data. *Stochastic Environmental Research and Risk Assessment*, **33**:1765–1780.

Mitnik, P.A. and Baek, S. (2019). The Kumaraswamy distribution: median-dispersion re-parameterizations for regression modeling and simulation-based estimation. *Statistical Papers*, **54**:177–192.

McMillen, D. (2013). *Quantile Regression for Spatial Data*. Springer, New York, US.

Nocedal, J., Wright, S. (1999). *Numerical optimization*. Springer, New York.

Noufaily, A., Jones, M. (2013). Parametric quantile regression based on the generalized gamma distribution. *Journal of the Royal Statistical Society Series C Applied Statistics*, **62**:723–740.

Owen, W. (2006). A new three-parameter extension to the Birnbaum-Saunders distribution. *IEEE Transactions on Reliability*, **55**:475–479.

Poon, W.Y., Poon, Y.S. (1999). Conformal normal curvature and assessment of local influence. *Journal of the Royal Statistical Society: Series B*, **61**:51-61.

R-Team (2018). *R: a language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna.

Rojas, F., Leiva, V., Wanke, P., Marchant, C. (2015). Optimization of contribution margins in food services by modeling independent component demand. *Revista Colombiana de Estadística*, **38**:1–30.

Rousseeuw, P., Croux, C., Todorov, V., Ruckstuhl, A., Salibian-Barrera, M., Verbeke, T., Koller, M., Maechler, M. (2016). robustbase: Basic Robust Statistics. R package version 0.92-6. Available at `http://CRAN.Rproject.org/package=robustbase`.

Sánchez, L., Leiva, V., Caro-Lopera, F.J., Cysneiros, F. J.A. (2015). On matrix-variate Birnbaum-Saunders distributions and their estimation and application. *Brazilian Journal of Probability and Statistics*, **29**:790–812.

Santana, L., Vilca, F., Leiva, V. (2011). Influence analysis in skew-Birnbaum-Saunders regression models and applications. *Journal of Applied Statistics*, **38**:1633–1649.

Santos-Neto, M., Cysneiros, F. J.A., Leiva, V., Ahmed, S. (2012). On new parameterizations of the Birnbaum-Saunders distribution. *Pakistan Journal of Statistics*, **28**:1–26.

Santos-Neto, M., Cysneiros, F. J.A., Leiva, V., Barros, M. (2014). On new parameterizations of the Birnbaum-Saunders distribution and its moments, estimation and application. *REVSTAT Statistical Journal*, **12**:247–272.

Santos-Neto, M., Cysneiros, F. J.A., Leiva, V., Barros, M. (2016). Reparameterized Birnbaum-Saunders regression models with varying precision. *Electronic Journal of Statistics*, **10**:2825–2855.

Saulo, H., Leão, V., Bourguignon, M. (2012). The Kumaraswamy Birnbaum-Saunders distribution. *Journal of Statistical Theory and Practice*, **1**:1–13.

Saulo, H., Leão, J., Leiva, V., Aykroyd, R.G. (2019). Birnbaum-Saunders autoregressive conditional duration models applied to high-frequency financial data. *Statistical Papers*, **60**:1605–1629.

Saulo, H., Leão, J., Vila, R., Leiva, V., Tomazella, V.(2019). On mean-based bivariate Birnbaum-Saunders distributions: Properties,inference and application. *Communications in Statistics-Theory and Methods*, DOI:10.1080/03610926.2019.1626425.

Saulo, H., Leiva, V., Ruggeri, F. (2015). Monitoring environmental risk by a methodology based on control charts. In Kitsos, C., Oliveira, T., Rigas, A., and Gulati, S., editors, *Theory and Practice of Risk Assessment*, pages 177–197. Springer, Switzerland.

Saulo, H., Leiva, V., Ziegelmann, F.A., Marchant, C. (2013). A nonparametric method for estimating asymmetric densities based on skewed Birnbaum-Saunders distributions applied to environmental data. *Stochastic Environmental Research and Risk Assessment*, **27**:1479–1491.

Saunders, S.C. (2007). *Reliability, Life Testing and Prediction of Services Lives.* Springer, New York, US.

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

Tapia, H., Leiva, V., Diaz, M.P., and Giampaoli, V. (2019). Influence diagnostics in mixed effects logistic regression models. *TEST*, 28:920–942.

Trzpiot, G. (2013). Spatial quantile regression. *Comparative Economic Research*, **15**:265–279.

Ventura, M., Saulo, H., Leiva, V., Monsueto, S. (2019). Log-symmetric regression models: information criteria, application to movie business and industry data with economic implications. *Applied Stochastic Models in Business and Industry*, **35**:963–977.

Vilca, F., Balakrishnan, N., Zeller, C. (2014). A robust extension of the bivariate Birnbaum-Saunders distribution and associated inference. *Journal of Multivariate Analysis*, **124**:418–435.

Vilca, F., Leiva, V. (2006). A new fatigue life model based on the family of skew-elliptical distributions. *Communications in Statistics: Theory and Methods*, **35**:229–244.

Vilca, F., Sanhueza, A., Leiva, V., Christakos, G. (2010). An extended Birnbaum-Saunders model and its application in the study of environmental quality in Santiago, Chile. *Stochastic Environmental Research and Risk Assessment*, **24**:771–782.

Villegas, C., Paula, G.A., Leiva, V. (2011). Birnbaum-Saunders mixed models for censored reliability data analysis. *IEEE Transactions on Reliability*, **60**:748–758.

Volodin, I., Dzhungurova, O. (2000). On limit distribution emerging in the generalized Birnbaum-Saunders model. *Journal of Mathematical Science*, **99**:1348–1366.

Wanke, P., Leiva, V. (2015). Exploring the potential use of the Birnbaum-Saunders distribution in inventory management. *Mathematical Problems in Engineering*, Article ID 827246:1–9.

Zhang, H. (2004). Inconsistent estimation and asymptotically equivalent interpolations in model-based geostatistics. *Journal of the American Statistical Association*, **99**:250–261.

Zhang, H., Wang, Y. (2010). Kriging and cross-validation for massive spatial data. *Environmetrics*, **21**:290–304.

Zhang, H., Zimmerman, D. L. (2005). Towards reconciling two asymptotic frameworks in spatial statistics. *Biometrika*, **92**:921–936.

Zhu, H., Ibrahim, J.G., Lee, S., Zhang, H. (2007). Perturbation selection and influence measures in local influence analysis. *Annals of Statistics*, **35**:2565–2588.