UNIVERSIDADE FEDERAL DE PERNAMBUCO
Centro de Ciências Exatas e da Natureza
Departamento de Estatística

Pós-graduação em Estatística

ESSAYS ON MULTIVARIATE GENERALIZED BIRNBAUM-SAUNDERS METHODS

by

CAROLINA MARCHANT

Doctoral thesis

Supervisor: **Prof. Dr. Francisco José de Azevêdo Cysneiros**
Co-supervisor: **Prof. Dr. Víctor Leiva**

Recife
2016

Carolina Marchant

Essays on multivariate generalized Birnbaum-Saunders methods

Doctoral thesis submitted to the Graduate Program in Statistics, Department of Statistics, Universidade Federal de Pernambuco as a requirement to obtain a PhD. in Statistics.

During the development of this work the author received scholarship from FACEPE and "Becas Chile" of Conicyt of the Chilean government

Recife
2016

Catalogação na fonte
Bibliotecário Jefferson Luiz Alves Nazareno CRB 4-1758

M215e Marchant Fuentes, Carolina Ivonne.
Essays on multivariate generalized Birnbaum-Saunders methods / Carolina Ivonne Marchant Fuentes. – 2016.
124f.: fig., tab.

Orientador: Francisco José de Azevêdo Cysneiros
Tese (Doutorado) – Universidade Federal de Pernambuco. CCEN. Estatística, Recife, 2016.
Inclui referências e apêndice.

1. Modelos de regressão multivariados. 2. Distribuição Birnbaum Saunders. 3. Métodos de influência,. I. Cysneiros, Francisco José de Azevêdo. (Orientador). II. Titulo.

310 CDD (22. ed.) UFPE-MEI 2016-155

CAROLINA IVONNE MARCHANT FUENTES

ESSAYS ON MULTIVARIATE GENERALIZED BIRNBAUM-SAUNDERS METHODS

Tese apresentada ao Programa de Pós-Graduação em Estatística da Universidade Federal de Pernambuco, como requisito parcial para a obtenção do título de Doutor em Estatística.

Aprovada em: 31 de outubro de 2016.

BANCA EXAMINADORA

Prof. Victor Eliseo Leiva Sánchez
Universidad Adolfo Ibáñez/Chile

Prof. Francisco Cribari Neto
UFPE

Prof. Abraão David Costa do Nascimento
UFPE

Prof. Raydonal Ospina Martínez
UFPE

Prof. Michelli Karinne Barros da Silva
UFCG

Acknowledgements

To God for his spiritual support throughout my life.

To my family, Flora, Víctor, Felipe and Lorena, who have given to me their support and unconditional affection always.

To my supervisors, professors Dr. Francisco José de Azevêdo Cysneiros and Dr. Víctor Leiva, for trusting on me, for their constant guidance, dedication, patience and good advises, as well as for their academic and personal support.

To the professors of the examining committee, for their constructive comments on an earlier version of this thesis, which resulted in this improved version, as well as for their suggestions of future research.

To the professors of the Department of Statistics of the Universidade Federal de Pernambuco. Especially, to who contributed in my PhD formation, Dr. Francisco Cribari-Neto, Dr. Klaus Vasconcellos and Dr. Francisco José de Azevêdo Cysneiros.

To my classmates of the PhD program in statistics of the Department of Statistics of the Universidade Federal de Pernambuco, who gave to me their friendship and professional support.

To the staff members of the Department of Statistics of the Universidade Federal de Pernambuco. Specially, to Valéria, who helped to me always with excellent disposition and efficiency.

Finally, to the organizations that provide to me scholarships in different periods of my PhD studies, without whose economical support I would not have been able to carry out these studies. Specifically, (i) to the Foundation for the Support of Science and Technology of Pernambuco (FACEPE) by its scholarship committee, for its support during the first year of my PhD studies; and (ii) to the National Commission of Science and Technology (Conicyt) of the government of Chile, by its scholarship committee "Becas Chile" for doctoral studies during the last 3 years of the program.

Thank you very much to all!

Resumo

Nas últimas décadas, o modelo Birnbaum-Saunders univariado recebeu considerável atenção na literatura. Esse modelo tem sido amplamente estudado e aplicado inicialmente à modelagem de fadiga de materiais. Com o passar dos anos surgiram trabalhos com aplicações em outras áreas do conhecimento. Em muitas das aplicações é necessário modelar diversas variáveis simultaneamente incorporando a correlação entre elas. Os modelos de regressão multivariados são uma ferramenta útil de análise multivariada, que leva em conta a correlação entre as variáveis de resposta. A análise de diagnóstico é um aspecto importante a ser considerado no modelo estatístico e verifica as suposições adotadas como também sua sensibilidade. Além disso, os gráficos de controle de qualidade multivariados são ferramentas visuais eficientes e simples para determinar se um processo multivariado está ou não fora de controle. Este gráfico mostra como diversas variáveis afetam conjuntamente um processo. Primeiro, propomos, derivamos e caracterizamos as distribuições Birnbaum-Saunders generalizadas logarítmicas multivariadas. Em seguida, propomos um modelo de regressão Birnbaum-Saunders generalizado multivariado. Métodos para estimação dos parâmetros do modelo, tal como o método de máxima verossimilhança baseado no algoritmo EM, foram desenvolvidos. Estudos de simulação de Monte Carlo foram realizados para avaliar o desempenho dos estimadores propostos. Segundo, realizamos uma análise de diagnóstico para modelos de regressão Birnbaum-Saunders generalizados multivariados. Consideramos a distância de Mahalanobis como medida de influência global de detecção de outliers multivariados utilizando-a para avaliar a adequacidade do modelo. Além disso, desenvolvemos medidas de diagnósticos baseadas em influência local sob alguns esquemas de perturbações. Implementamos a metodologia apresentada no software R, e ilustramos com dados reais multivariados de biomateriais. Terceiro, e finalmente, desenvolvemos uma metodologia robusta baseada em gráficos de controle de qualidade multivariados para a distribuição Birnbaum-Saunders generalizada usando a estatística de Hotelling. Baseado no método bootstrap paramétrico encontramos aproximações da distribuição desta estatística e obtivemos limites de controle para o gráfico proposto. Realizamos um estudo de simulação de Monte Carlo para avaliar a metodologia proposta indicando seu bom desempenho para fornecer alertas precoces de processos fora de controle. Uma ilustração com dados reais de qualidade do ar de Santiago-Chile é fornecida. Essa ilustração mostra que a metodologia proposta pode ser útil para alertar sobre episódios de poluição extrema do ar, evitando efeitos adversos na saúde humana.

Palavras-chave: algoritmo EM. bondade de ajuste. distribuições Birnbaum-Saunders generalizadas. gráficos de controle. estatística de Hotelling. influência local e global. método de máxima verossimilhança. simulação de Monte Carlo. software `R`. modelos de regressão.

Abstract

In the last decades, univariate Birnbaum-Saunders models have received considerable attention in the literature. These models have been widely studied and applied to fatigue, but they have also been applied to other areas of the knowledge. In such areas, it is often necessary to model several variables simultaneously. If these variables are correlated, individual analyses for each variable can lead to erroneous results. Multivariate regression models are a useful tool of the multivariate analysis, which takes into account the correlation between variables. In addition, diagnostic analysis is an important aspect to be considered in the statistical modeling. Furthermore, multivariate quality control charts are powerful and simple visual tools to determine whether a multivariate process is in control or out of control. A multivariate control chart shows how several variables jointly affect a process. First, we propose, derive and characterize multivariate generalized logarithmic Birnbaum-Saunders distributions. Also, we propose new multivariate generalized Birnbaum-Saunders regression models. We use the method of maximum likelihood estimation to estimate their parameters through the expectation-maximization algorithm. We carry out a simulation study to evaluate the performance of the corresponding estimators based on the Monte Carlo method. We validate the proposed models with a regression analysis of real-world multivariate fatigue data. Second, we conduct a diagnostic analysis for multivariate generalized Birnbaum-Saunders regression models. We consider the Mahalanobis distance as a global influence measure to detect multivariate outliers and use it for evaluating the adequacy of the distributional assumption. Moreover, we consider the local influence method and study how a perturbation may impact on the estimation of model parameters. We implement the obtained results in the `R` software, which are illustrated with real-world multivariate biomaterials data. Third and finally, we develop a robust methodology based on multivariate quality control charts for generalized Birnbaum-Saunders distributions with the Hotelling statistic. We use the parametric bootstrap method to obtain the distribution of this statistic. A Monte Carlo simulation study is conducted to evaluate the proposed methodology, which reports its performance to provide earlier alerts of out-of-control conditions. An illustration with air quality real-world data of Santiago-Chile is provided. This illustration shows that the proposed methodology can be useful for alerting episodes of extreme air pollution.

Keywords: EM algorithm. global and local influence. Hotelling statistic. goodness-of-fit. maximum likelihood method. Monte Carlo simulation. multivariate generalized Birnbaum-Saunders distributions. multivariate quality control charts. `R` software. multivariate regression models.

List of Figures

List of Tables

List of Algorithms

List of Acronyms

ARL	Average run length
BFGS	Broyden-Fletcher-Goldfarb-Shanno
BS	Birnbaum-Saunders
CDF	Cumulative distribution function
CL	Center line
CT	Computed tomography
CV	Coefficient of variation
DF	Degree of freedom
EC	Elliptically contoured
EM	Expectation-maximization
FAR	False alarm rate
GBS	Generalized Birnbaum-Saunders
GL	Generalized leverage
GOF	Goodness-of-fit
IID	Independent and identically distributed
KS	Kolmogorov-Smirnov
LCL	Lower control limit
LD	Likelihood displacement
log-BS	Logarithmic Birnbaum-Saunders
log-BS-t	Logarithmic Student-t Birnbaum-Saunders
log-GBS	Logarithmic generalized Birnbaum-Saunders
MC	Monte Carlo
MD	Mahalanobis distance
ML	Maximum likelihood
PDF	Probability density function
PP	Probability versus probability
RB	Relative bias
RMSE	Root of the mean squared error
RV	Random variable
SD	Standard deviation
SE	Standard error
SS	Standard symmetric
t	Student-t
UCL	Upper control limit
WH	Wilson-Hilferty

Contents

CHAPTER 1

Preliminaries

1.1 Resumo

Neste capítulo, apresentamos uma revisão bibliográfica sobre o tema da tese. Adicionalmente, oferecemos ferramentas teóricas que proporcionam ao leitor um melhor entendimento de nossa proposta. Aqui, são incluidos aspectos relacionados com as distribuições Birnbaum-Saunders generalizadas e Birnbaum-Saunders logarítmicas generalizadas univariadas e sua modelagem estatística. Aspectos das distribuições Birnbaum-Saunders generalizadas multivariadas são também mostradas. Além disso, motivações, objetivos e produtos desta tese são apresentados.

1.2 Introduction and bibliographical review

Univariate Birnbaum-Saunders (BS) distributions are a family of models originated from the cumulative damage law related to fatigue and strength of materials that has been widely studied. This origin allows the BS distribution to be interpreted as a life distribution, because it describes the time elapsed until that the extension of a crack exceeds a threshold conducting to a failure. This family is unimodal, positively skewed and useful for modeling data that take values greater than zero. The BS family has two parameters, which modify the shape and scale of the distribution; see Birnbaum and Saunders (1969), Leiva and Saunders (2015), Leiva (2016) and Leiva and Vivanco (2016).

The family of BS distributions has evolved over time since 1969 until the present. Three stages can be mentioned for these models. The first stage (1969-1999) contains few works and a slow growth for their development; see Birnbaum and Saunders (1969), Rieck and Nedelman (1991), Achcar (1993), Johnson et al. (1995), Dupuis and Mills (1998) and Owen and Padgett (1999). The second stage (2000-2010) is formed by works that include diverse estimation, modeling and diagnostic aspects, as well as generalizations, computational implementations and new applications which go beyond its genesis from science of materials, but these applications were still based on arguments from this genesis; see Owen and Padgett (2000), Volodin and Dzhungurova (2000), Tsionas (2001), Ng et al. (2003), Rieck (2003), Galea et al. (2004), Díaz and Leiva (2005), Owen (2006), Xie and Wei (2007), Lemonte et al. (2007, 2008), Leiva et al. (2007, 2008a,b, 2009), Cysneiros et al. (2008), Barros et al. (2008), Lemonte and Cordeiro (2009), Balakrishnan et al. (2009), and Vilca et al. (2010). The third stage (2011-present) is the openness to a new vision of BS distributions, which allows us to consider it in a general setting rather than to be restricted to failure time data, which extends their applicability to all areas of knowledge as: agriculture, air contamination, bio-engineering, business, economics, engineering, environment, finance, food, forest and textile industries, human and three mortality, informatics, insurance, inventory management, medicine, nutrition, pharmacology, psychology, neurology, quality control, queue theory, toxicology, water quality, and wind energy; see, for example, Bhatti (2010), Kotz et al. (2010), Balakrishnan et al. (2011), Leiva et al. (2010, 2011a,b, 2012, 2014a,b,c, 2016a, 2015b,c, 2016b, 2017), Vilca et al. (2010, 2011), Villegas et al. (2011), Azevedo et al. (2012), Ferreira et al. (2012), Paula et al. (2012), Santos-Neto et al. (2012, 2014, 2016), Vanegas et al. (2012), Marchant et al. (2013a,b, 2016a,b), Saulo et al. (2013, 2015), Barros et al. (2014), Rojas et al. (2015), Wanke and Leiva (2015), Garcia-Papani et al. (2016), and references therein.

Univariate BS distributions have been widely studied because of its good properties and its relation with the normal distribution. Specifically, every random variable following the BS distribution can be considered as a transformation of another random variable following a standard normal distribution; see Johnson et al. (1995, pp. 651-663) and Leiva (2016, p. 18). Then, because the cumulative damage is assumed to be normally distributed in the BS model, its parameter estimates obtained from the maximum likelihood (ML) method are sensitive to atypical observations. In order to attenuate this sensitivity, and using the relation between the normal and BS distributions, one can obtain a BS distribution based on the Student-t (called t hereafter) distribution. Thus, ML estimates of the parameters of the BS-t distribution attribute smaller weights to atypical observations than the BS distribution, producing robust parameter estimates; see Paula et al. (2012) and references therein. BS and

BS-t distributions are members of a wider family of distributions generated from elliptically contoured (EC) distributions, known as generalized BS (GBS) distributions; see, for example, Fang et al. (1990) for details of EC distributions and Díaz and Leiva (2005) for the seminal paper on GBS distributions.

Rieck and Nedelman (1991) introduced a logarithmic version of the BS distribution, in short log-BS, which is often useful to formulate BS log-linear regression models; see Tsionas (2001), Leiva et al. (2007), Xie and Wei (2007), Desmond et al. (2008), Lemonte et al. (2010), Xiao et al. (2010), Lemonte (2011) and Lemonte and Ferrari (2011a,b). Such models are helpful, for example, to predict fracture; see also Galea et al. (2004). Univariate log-linear models for generalized BS (GBS) distributions were studied by Barros et al. (2008), Li et al. (2012), and Paula et al. (2012).

Bivariate versions of the BS distribution were proposed by Kundu et al. (2010) and Vilca et al. (2014a). Multivariate versions of GBS distributions were derived by Kundu et al. (2013), whereas Caro-Lopera et al. (2012) and Sánchez et al. (2015) introduced matrix-variate GBS distributions. Kundu (2015a) and Vilca et al. (2014b) presented bivariate log-BS and log-GBS distributions. Moreover, Lemonte (2013) developed an independent multivariate BS log-linear regression model. Other works related to multivariate GBS distributions are attributed to Jamalizadeh and Kundu (2015), Khosravi et al. (2015, 2016), Kundu (2015b), Lemonte et al. (2015) and Garcia-Papani et al. (2016). However, no studies on multivariate log-GBS distributions nor GBS log-linear regression models have been published.

One of the main steps of a parametric statistical analysis is to validate the reference distribution on which this analysis relies. Specifically, one must be concerned about the adequacy of this distribution to the data under analysis. Goodness-of-fit (GOF) tests have been proposed to evaluate the hypotheses of distributional adequacy or discrepancy, with respect to a data set, where the vector of parameters that indexes the distribution can be known or unknown; see D'Agostino and Stephens (1986), Castro-Kuriss et al. (2009, 2010) and Barros et al. (2014).

Influence diagnostics is an important aspect to be considered in data modeling. This aspect is often carried out after the parameters are estimated allowing the stability of the estimation process to be assessed. Influence methods have been used in normal linear regression models and widely studied in the literature. Some classical books on the topic are Cook and Weisberg (1982) and Chatterjee and Hadi (1988). Cook (1986) introduced the local influence method to evaluate the effect of small perturbations in the model and/or data on the ML estimates. Such a method has played a significant role on regression diagnostics. After Cook (1986), who applied his method in the linear regression model under

normality, a number of researchers have studied influence diagnostics for the linear model with a number of variations or additional specific structures. For more details about diagnostics methods in more general models; see Paula (1993), Galea et al. (1997, 2004), Shi (1997), Haslett and Dillane (2004), Leiva et al. (2007), Osorio et al. (2007), Atkinson (2009), Barros et al. (2010), Santana et al. (2011), Villegas et al. (2011), Paula et al. (2012) and Leiva et al. (2014b,c,d). No studies on influence diagnostics in multivariate GBS log-linear regression models have been carried out until now.

Shewhart (1931) introduced univariate control charts to monitor the process quality through a statistical sample. However, many times there is a need to monitor several quality characteristics of a process simultaneously. If these characteristics are correlated, then using separate univariate control charts for individual monitoring may not be adequate, when detecting changes in the overall quality of the process. Thus, it is desirable to have tools that can jointly monitor all of these variables. Such tools are the multivariate control charts, which are the most common used for this joint monitoring; see Alt (1985). These charts can take into account the simultaneous nature of the control scheme and the correlation structure between the quality characteristics. A multivariate control chart has as objective to detect the presence of special causes of variation and can be used as a tool to detect multivariate outliers, mean shifts, and other distributional deviations from the in-control distribution. For studies on univariate quality control charts based on the BS distribution, see Lio and Park (2008) and Leiva et al. (2015b). However, to our knowledge, no studies on multivariate control charts based on GBS distributions are reported.

1.3 Background

A random variable (RV) T is said to follow a univariate BS distribution with parameters of shape $\alpha \in \mathbb{R}_+$ and scale $\lambda \in \mathbb{R}_+$, which is denoted by $T \sim \mathrm{BS}(\alpha, \lambda)$, if

$$T = \lambda \left[\frac{\alpha V}{2} + \left\{ \left(\frac{\alpha V}{2} \right)^2 + 1 \right\}^{\frac{1}{2}} \right]^2, \tag{1.1}$$

where $V \sim \mathrm{N}(0,1)$, with $V = [\{T/\lambda\}^{\frac{1}{2}} - \{\lambda/T\}^{\frac{1}{2}}]/\alpha$.

Let the assumption given in (1.1) to be relaxed by supposing V to follow any standard symmetric (SS) distribution in $\mathbb{R}$ (corresponding to EC distributions in the multivariate case), with kernel g for its probability density function (PDF). Then, we obtain the class of univariate GBS distributions, which is denoted by $T \sim \mathrm{GBS}(\alpha, \lambda, g)$, with cumulative distribution function (CDF) given by

$$F_T(t; \alpha, \lambda, g) = F(A(t; \alpha, \lambda), g), \quad t \in \mathbb{R}_+,$$

where $F(\cdot)$ is the CDF of an SS distribution and

$$A(t;\alpha,\lambda)=\frac{1}{\alpha}\left[\left\{\frac{t}{\lambda}\right\}^{\frac{1}{2}}-\left\{\frac{\lambda}{t}\right\}^{\frac{1}{2}}\right].$$

The PDF of T is

$$f_T(t;\alpha,\lambda,g)=f\left(A(t;\alpha,\lambda),g\right)a(t;\alpha,\lambda),\quad t\in\mathbb{R}_+, \tag{1.2}$$

where $f(\cdot,g)$ is the PDF of an SS distribution and $a(t;\alpha,\lambda)$ is the derivative of $A(t;\alpha,\lambda)$, so that

$$a(t;\alpha,\lambda)=\frac{1}{[2\alpha\lambda]}\left[\left\{\frac{\lambda}{t}\right\}^{\frac{1}{2}}+\left\{\frac{\lambda}{t}\right\}^{\frac{3}{2}}\right].$$

Some properties of GBS distributions are:

(A1) $kT\sim\text{GBS}(\alpha,\,k\,\lambda,\,g)$, with $k\in\mathbb{R}_+$.

(A2) $1/T\sim\text{GBS}(\alpha,1/\lambda,g)$.

(A3) $V^2=[T/\lambda+\lambda/T-2]/\alpha^2\sim\text{G}\chi^2(1,g)$, that is, V^2 follows a generalized χ^2 distribution with one degree of freedom (DF) and kernel g; see Fang et al. (1990) and Fang and Zhang (1990).

Consider the regression model proposed by Rieck and Nedelman (1991) given by

$$T_i=\exp(\boldsymbol{\beta}^\top\boldsymbol{x_i})\zeta_i,\quad i=1,\ldots,n, \tag{1.3}$$

where T_i is the response variable (called response hereafter), $\boldsymbol{x}_i=(x_{i1},\ldots,x_{ip})^\top$ is a $p\times 1$ vector containing the values of p regressor variables (called regressor hereafter); $\boldsymbol{\beta}^\top=(\beta_1,\ldots,\beta_p)$ is a $1\times p$ vector of unknown parameters to be estimated, with β_j corresponding to x_j, for $j=1,\ldots,p$ and $x_1=1$. In addition, $\zeta_i\sim\text{BS}(\alpha,1)$ is the model error. Thus, from this error distribution and property (A1), note that $T_i\sim\text{BS}(\alpha,\exp(\boldsymbol{\beta}^\top\boldsymbol{x_i}))$. Based on the model defined in (1.3), we have a BS log-linear regression model given by

$$Y_i=\boldsymbol{\beta}^\top\boldsymbol{x_i}+\varepsilon_i,\quad i=1,\ldots,n, \tag{1.4}$$

where $Y_i=\log(T_i)$ is the log-response, $\boldsymbol{x}_i$, $\boldsymbol{\beta}$ are as given in (1.3), and ε_i is the model error, with $\varepsilon_i=\log(\zeta_i)\sim\text{log-BS}(\alpha,0)$. Thus, Y follows a log-BS distribution with shape parameter $\alpha\in\mathbb{R}_+$, and mean $\mu=\text{E}[Y]=\log(\lambda)\in\mathbb{R}$, which is denoted by $Y\sim\text{log-BS}(\alpha,\mu)$, if

$$Y=\mu+2\,\text{arcsinh}\left(\frac{\alpha\,W}{2}\right), \tag{1.5}$$

where

$$W = \frac{2}{\alpha}\sinh\left(\frac{Y-\mu}{2}\right) \sim \mathrm{N}(0,1).$$

Similarly to the BS case, (1.5) can also be relaxed assuming W to follow any SS distribution in $\mathbb{R}$ with kernel g. Then, the notation $Y \sim \text{log-GBS}(\alpha, \mu, g)$ is used and its CDF is

$$F_Y(y;\alpha,\mu,g) = F(B(y;\alpha,\mu),g), \quad y \in \mathbb{R},$$

where $F(\cdot)$ is an SS CDF and

$$B(y;\alpha,\mu) = \frac{2}{\alpha}\sinh\left(\frac{y-\mu}{2}\right).$$

Note that $W = B(Y;\alpha,\mu)$. The corresponding PDF of Y is given by

$$f_Y(y;\alpha,\mu,g) = f(B(y;\alpha,\mu),g)\, b(y;\alpha,\mu), \quad y \in \mathbb{R}, \tag{1.6}$$

where $f(\cdot, g)$ is as given in (1.2) and $b(y;\alpha,\mu)$ is the derivative of $B(y;\alpha,\mu)$ expressed as

$$b(y;\alpha,\mu) = \frac{1}{\alpha}\cosh\left(\frac{y-\mu}{2}\right).$$

Some properties of $Y \sim \text{log-GBS}(\alpha, \mu, g)$ are:

(B1) $W = B(Y;\alpha,\mu)$ follows an SS distribution with location zero and kernel g.

(B2) $W^2 = B^2(Y;\alpha,\mu) \sim \mathrm{G}\chi^2(1, g)$.

The log-BS-t distribution is constructed replacing $W \sim \mathrm{N}(0,1)$ in (1.5) by the representation

$$S = U^{-\frac{1}{2}} Z \sim t(\nu), \tag{1.7}$$

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

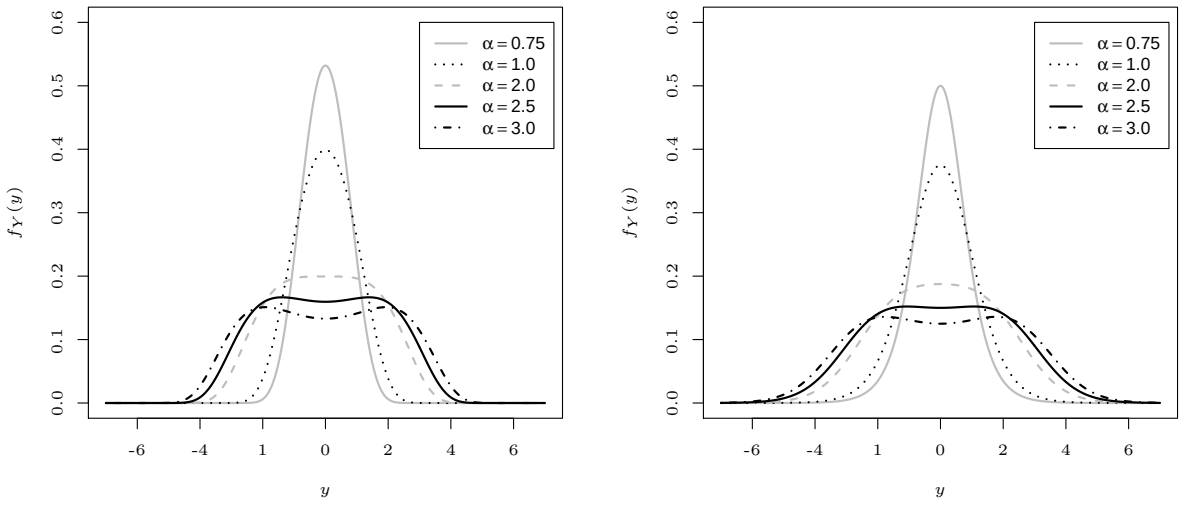
$$U \sim \text{Gamma}\left(\frac{\nu}{2}, \frac{\nu}{2}\right),$$

with U being independent of Z, that is, S has a t distribution with shape parameter $\nu \in \mathbb{R}_+$ (also known as DFs). Therefore, from (1.5) and (1.7), we have

$$Y = \mu + 2\,\text{arcsinh}\left(\frac{\alpha\, Z}{2\, U^{\frac{1}{2}}}\right) \sim \text{log-BS-}t(\alpha,\mu,\nu). \tag{1.8}$$

Figure 1.1 shows the PDFs for the univariate log-BS and log-BS-t distributions for $\alpha \in \{0.75, 1.0, 2.0, 2.5, 3.0\}$, $\mu = 0$ and $\nu = 4$. From this figure, note that the log-BS-t distribution has heavier tails than the log-BS distribution. Also, when $\alpha \leq 2$, the PDFs have just a mode,

Figure 1.1: PDF plots for the indicated log-GBS distributions.



(a) log-BS(α, 0) (b) log-BS-t(α, 0, ν=4)

Source: From the author.

whereas for $\alpha > 2$, we observe bimodality in log-GBS distributions. Then, the parameter α modifies the kurtosis and bimodality of the distribution while μ the location.

The usage of the t kernel allows us to obtain robust estimation to atypical data and implement the expectation-maximization (EM) algorithm; see Balakrishnan et al. (2009) and Paula et al. (2012). This algorithm is a powerful iterative method to find ML estimates of statistical parameters, when the model depends on (unobserved) latent variables. The EM algorithm is computationally simple, numerically stable and used in diverse applications; see details in McLachlan and Krishnan (1997) and Appendix 6.1.

Let $\boldsymbol{V} = (V_1, \ldots, V_m)^\top \in \mathbb{R}^m$ be a random vector with m-variate EC distribution, location vector equal to zero, scale matrix $\boldsymbol{\Sigma} = (\sigma_{kl}) \in \mathbb{R}^{m\times m}$ ($\text{rk}(\boldsymbol{\Sigma}) = m$), kernel density generator $g^{(m)}$, and all its moments exist. In this case, the notation $\boldsymbol{V} \sim \text{EC}_m(\mathbf{0}_{m\times 1}, \boldsymbol{\Sigma}, g^{(m)})$ is used, with $\mathbf{0}_{m\times 1}$ being a vector of zeros. The matrix $\boldsymbol{\Sigma}$ allows us to obtain the variance-covariance matrix of $\boldsymbol{V}$ by $\boldsymbol{\Sigma}_0 = c_0\,\boldsymbol{\Sigma}$, where $c_0 = \text{E}[V^2]$, with $V^2 \sim \text{G}\chi^2(m, g^{(m)})$. Thus, the correlation matrix $\boldsymbol{\Psi} = (\psi_{kl}) \in \mathbb{R}^{m\times m}$ (with $\text{rk}(\boldsymbol{\Psi}) = m$) of $\boldsymbol{V}$ is given by

$$\boldsymbol{\Psi} = \boldsymbol{D}(\boldsymbol{\Sigma}_0^{-1/2})\boldsymbol{\Sigma}_0\,\boldsymbol{D}(\boldsymbol{\Sigma}_0^{-1/2}) = \boldsymbol{D}(\boldsymbol{\Sigma}^{-1/2})\boldsymbol{\Sigma}\,\boldsymbol{D}(\boldsymbol{\Sigma}^{-1/2}), \tag{1.9}$$

where $\boldsymbol{\Sigma}_0 = (c_0\,\sigma_{kl}) \in \mathbb{R}^{m\times m}$, $\boldsymbol{D}(\boldsymbol{\Sigma}_0^{-1/2}) = \text{diag}([c_0\,\sigma_{11}]^{-1/2}, \ldots, [c_0\,\sigma_{mm}]^{-1/2})$ and $\boldsymbol{D}(\boldsymbol{\Sigma}^{-1/2}) =$

$\text{diag}(\sigma_{11}^{-1/2},\ldots,\sigma_{mm}^{-1/2})$. Therefore, we can express the variance-covariance matrix of $\boldsymbol{V}$ as

$$\boldsymbol{\Sigma}_0 = \boldsymbol{D}(\boldsymbol{\Sigma}_0^{1/2})\boldsymbol{\Psi}\,\boldsymbol{D}(\boldsymbol{\Sigma}_0^{1/2}) = c_0\begin{pmatrix} \sigma_{11} & \sqrt{\sigma_{11}\sigma_{22}}\psi_{12} & \cdots & \sqrt{\sigma_{11}\sigma_{mm}}\psi_{1m} \\ \sqrt{\sigma_{11}\sigma_{22}}\psi_{12} & \sigma_{22} & \cdots & \sqrt{\sigma_{22}\sigma_{mm}}\psi_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \sqrt{\sigma_{11}\sigma_{mm}}\psi_{1m} & \sqrt{\sigma_{22}\sigma_{mm}}\psi_{2m} & \cdots & \sigma_{mm} \end{pmatrix}. \tag{1.10}$$

The PDF of $\boldsymbol{V}$ is

$$f_{\text{EC}_m}(\boldsymbol{v};\boldsymbol{\Sigma},g^{(m)}) = c^{(m)}|\boldsymbol{\Sigma}|^{-1/2}g^{(m)}(\boldsymbol{v}^\top\boldsymbol{\Sigma}^{-1}\boldsymbol{v}), \quad \boldsymbol{v}=(v_1,\ldots,v_m)^\top\in\mathbb{R}^m, \tag{1.11}$$

with normalizing constant $c^{(m)}>0$, whereas its CDF is denoted by F_{EC_m}. For more details about EC distributions, see Fang et al. (1990), Fang and Zhang (1990), and Gupta et al. (2013).

Let $\boldsymbol{T}=(T_1,\ldots,T_m)^\top\in\mathbb{R}_+^m$ be a random vector with m-variate GBS distribution and parameters $\boldsymbol{\alpha}=(\alpha_1,\ldots,\alpha_m)^\top\in\mathbb{R}_+^m$, $\boldsymbol{\lambda}=(\lambda_1,\ldots,\lambda_m)^\top\in\mathbb{R}_+^m$, and EC kernel $g^{(m)}$, being its scale and correlation matrices $\boldsymbol{\Sigma}\in\mathbb{R}^{m\times m}$ and $\boldsymbol{\Psi}\in\mathbb{R}^{m\times m}$, respectively; see Kundu et al. (2013). Note that, for the GBS case, $\sigma_{kk}=1$, for all $k=1,\ldots,m$, and then from (1.9) and (1.10),

$$\boldsymbol{\Sigma} = \begin{pmatrix} 1 & \psi_{12} & \cdots & \psi_{1m} \\ \psi_{12} & 1 & \cdots & \psi_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{1m} & \psi_{2m} & \cdots & 1 \end{pmatrix} = \boldsymbol{\Psi}, \tag{1.12}$$

which we denote by $\boldsymbol{T}\sim\text{GBS}_m(\boldsymbol{\alpha},\boldsymbol{\lambda},\boldsymbol{\Psi},g^{(m)})$. Thus, the CDF and PDF of $\boldsymbol{T}$ are, respectively,

$$\begin{aligned} F_{\boldsymbol{T}}(\boldsymbol{t};\boldsymbol{\alpha},\boldsymbol{\lambda},\boldsymbol{\Psi},g^{(m)}) &= F_{\text{EC}_m}(\boldsymbol{A};\boldsymbol{\Psi},g^{(m)}), \\ f_{\boldsymbol{T}}(\boldsymbol{t};\boldsymbol{\alpha},\boldsymbol{\lambda},\boldsymbol{\Psi},g^{(m)}) &= f_{\text{EC}_m}(\boldsymbol{A};\boldsymbol{\Psi},g^{(m)})\,a(\boldsymbol{t};\boldsymbol{\alpha},\boldsymbol{\lambda}), \quad \boldsymbol{t}=(t_1,\ldots,t_m)^\top\in\mathbb{R}_+^m, \end{aligned}$$

where $\boldsymbol{A}=A(\boldsymbol{t};\boldsymbol{\alpha},\boldsymbol{\lambda})=(A_1,\ldots,A_m)^\top$, with $A_j=A(t_j;\alpha_j,\lambda_j)$ and

$$a(\boldsymbol{t};\boldsymbol{\alpha},\boldsymbol{\lambda}) = \prod_{j=1}^m a(t_j;\alpha_j,\lambda_j),$$

both $A(t_j;\alpha_j,\lambda_j)$, $a(t_j;\alpha_j,\lambda_j)$ given in (1.2). Some properties of $\boldsymbol{T}$ are:

(C1) $k\,\boldsymbol{T}\sim\text{GBS}_m(\boldsymbol{\alpha},k\,\boldsymbol{\lambda},\boldsymbol{\Psi},g^{(m)})$, with $k\in\mathbb{R}_+$.

(C2) $\boldsymbol{T}^*=(1/T_1,\ldots,1/T_m)^\top\sim\text{GBS}_m(\boldsymbol{\alpha},\boldsymbol{\lambda}^*,\boldsymbol{\Psi},g^{(m)})$, with $\boldsymbol{\lambda}^*=(1/\lambda_1,\ldots,1/\lambda_m)^\top$.

(C3) $A^\top(\boldsymbol{t};\boldsymbol{\alpha},\boldsymbol{\lambda})\boldsymbol{\Psi}^{-1}A(\boldsymbol{t};\boldsymbol{\alpha},\boldsymbol{\lambda})\sim\text{G}\chi^2(m,g^{(m)})$.

Table 1.1: Normalizing constant and kernel of the indicated distribution.

Distribution	c	$g(u)$	$c^{(m)}$	$g^{(m)}(u)$
Normal	$[2\pi]^{-1/2}$	$\exp(-u/2)$	$[2\pi]^{-\frac{m}{2}}$	$\exp(-u/2)$
t	$\dfrac{\Gamma([\nu+1]/2)}{[\nu\pi]^{1/2}\Gamma(\nu/2)}$	$\left[1+\dfrac{u}{\nu}\right]^{-\frac{[\nu+1]}{2}}$	$\dfrac{\Gamma([\nu+m]/2)}{[\nu\pi]^{\frac{m}{2}}\Gamma(\nu/2)}$	$\left[1+\dfrac{u}{\nu}\right]^{-\frac{[\nu+m]}{2}}$

where Γ is the gamma function.

Source: From the author.

Table 1.1 provides the constants and kernels of the univariate and m-variate normal and t distributions.

1.4 Motivation of the thesis

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

(1) Because no studies on the multivariate log-GBS distribution were known, we try to fill this gap.

(2) Since no studies on the multivariate GBS log-linear regression models were published to starting date of this thesis, we formulate these models and estimate their parameters. In addition, we implement their results in `R` language and apply them to real-world data.

(3) Due to that no studies on diagnostic analyses in multivariate GBS log-linear regression models were known, we carry out a diagnostic analysis for these models.

(4) No studies on multivariate GBS quality control charts were published to date. Then, we develop a methodology for multivariate quality control charts based on generalized Birnbaum-Saunders distributions with the Hotelling statistic.

1.5 Objectives of the thesis

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

(1) To derive, characterize and implement multivariate log-GBS distributions.

(2) To propose a methodology based on multivariate GBS log-linear regression models including their formulation, ML estimation of their parameters by means of the EM algorithm, their implementation, as well as their application to real-world data.

(3) To develop influence diagnostic methods for multivariate GBS log-linear regression models, implement then in `R` language and apply the developed methods to real-world data.

(4) To derive multivariate GBS quality control charts, implement them in the `R` language and apply these charts to real-world processes.

1.6 Products of the thesis

This thesis led to the following products:

(1) **Marchant, C.**, Leiva, V. and Cysneiros, F.J.A. (2016) A multivariate log-linear model for Birnbaum-Saunders distributions. IEEE Transactions on Reliability 65(2):816-827.
(2) **Marchant, C.**, Leiva, V., Cysneiros, F.J.A. and Vivanco, J.F. (2016) Diagnostics in multivariate Birnbaum-Saunders regression models. Journal of Applied Statistics 43:2829-2849.
(3) **Marchant, C.**, Leiva, V., Cysneiros, F.J.A. and Liu, S. Robust multivariate control charts based on Birnbaum-Saunders distributions. Under review.
(4) **Marchant, C.**, Leiva, V. and Cysneiros, F.J.A. (2015) Multivariate generalized Birnbaum-Saunders regression models for metal forming processes. 60th ISI World Statistics Congress, Rio de Janeiro, Brazil.
(5) **Marchant, C.**, Leiva, V., Cysneiros, F.J.A. and Vivanco J.F. (2015) A multivariate log-linear model for Birnbaum-Saunders distributions and its influence diagnostics. XLII Jornadas Nacionales de Estadística, Concepción, Chile.
(6) **Marchant, C.**, Leiva, V., Cysneiros, F.J.A. and Vivanco J.F. (2015) Influence diagnostics for multivariate generalized Birnbaum-Saunders regression models. VI SEEMI, Toledo, Brazil.
(7) **Marchant, C.**, Leiva, V., Cysneiros, F.J.A. and Vivanco J.F. (2016) A multivariate log-linear model for Birbaum-Saunders distributions: Influence diagnostic and applications. X Workshop on Statistics, Mathematics and Computation, Tomar, Portugal.
(8) **Marchant, C.**, Leiva, V., Cysneiros, F.J.A. and Vivanco J.F. (2016) Multivariate Birnbaum-Saunders regression models: Diagnostic analysis and applications. XXII SINAPE, Porto Alegre-RS, Brazil.

1.7 Organization of the thesis

This thesis contains five chapters taking into account this introduction chapter. In Chapter 2, we propose, derive and characterize multivariate logarithmic generalized Birnbaum-Saunders distributions. Also, we propose new multivariate generalized Birnbaum-Saunders regression models. In Chapter 3, we carry out a diagnostic analysis for this new model. In Chapter 4, we develop a methodology for multivariate quality control charts based on generalized Birnbaum-Saunders distributions with the Hotelling statistic. Finally, in Chapter 5, we present general conclusions for this thesis and propose some topics for future work. Appendices related to some mathematical derivations of this thesis, as well the data sets used and a general references list, are presented at the final pages.

CHAPTER 2

A multivariate log-linear model for Birnbaum-Saunders distributions

2.1 Resumo

Neste capítulo, derivamos e caracterizamos as distribuições Birnbaum-Saunders generalizadas logarítmicas multivariadas. Desenvolvemos uma nova metodologia baseada em modelos de regressão multivariados Birnbaum-Saunders generalizados. Usamos o método de máxima verossimilhança através do algoritmo EM para estimar os parâmetros do modelo. Também, discutimos o desempenho dos estimadores propostos, avaliando-os através de estudos de simulações de Monte Carlo (MC). Implementamos a metodologia na linguagem `R` e a ilustramos através de um conjunto de dados reais aplicado a fadiga multivariada. Os resultados obtidos podem permitir aos engenheiros de materiais agendar a troca de ferramentas em processos de conformação de metal e avaliar os seus custos.

2.2 Introduction

Fatigue is related to the failure of materials, mainly metals, which occurs after a long time of service, caused by stress. Fatigue of metals is understood as the propagation of cracks, originating from stress, causing their fracture. Thus, fatigue is a process composed

by the initiation of a crack and its propagation, until the material is ultimately fractured. Calculation of fatigue life is of great importance in determining the reliability of components or structures.

Manufacturing, quality and productivity improvements of metallic tools can be integrated into their process cycle. Thus, fracture of these tools can be predicted by considering simultaneously RVs related to this integration, such as deformation, die lifetime, manufacturing force, and stress. Because these RVs are often correlated, fracture should be predicted by multivariate models. Based on these models, engineers can make decisions after specifying priorities of these RVs with target values for each of them.

Multivariate regression models are useful tools of the multivariate analysis. The main advantage of a multivariate regression model over marginal regressions is that it takes into account the correlation between the response variables. Several responses could be correlated and then this correlation should be considered in the modeling through a multivariate regression. If no correlation exists, several marginal models, one for each response, can be considered. However, analyzing responses individually, if correlations between them exist, may yield wrong prediction.

The objectives of this chapter are (i) to propose multivariate log-GBS distributions, and (ii) to derive multivariate GBS regression models for describing fatigue data. In Section 2.3, we derive new multivariate log-GBS distributions. In Section 2.4, we propose multivariate GBS regression models, including their formulation and ML estimation by means of the EM algorithm. In Section 2.5, we evaluate the performance of the ML estimators with simulations based on the MC method. In Section 2.6, we validate the proposed models with a regression analysis of real-world multivariate fatigue data. Finally, in Section 2.7, we present some conclusions of this chapter.

2.3 Multivariate log-GBS distributions

Let $\boldsymbol{T} = (T_1, \ldots, T_m)^\top \sim \text{GBS}_m(\boldsymbol{\alpha}, \boldsymbol{\lambda}, \boldsymbol{\Psi}, g^{(m)})$ as given in Section 1.3. Then, $\boldsymbol{Y} = (\log(T_1), \ldots, \log(T_m))^\top$ follows an m-variate log-GBS distribution with shape vector $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m)^\top$, location vector

$$\boldsymbol{\mu} = \text{E}[\boldsymbol{Y}] = (\text{E}[Y_1], \ldots, \text{E}[Y_m])^\top = (\log(\lambda_1), \ldots, \log(\lambda_m))^\top \in \mathbb{R}^m,$$

$g^{(m)}$ is the EC kernel and correlation matrix $\boldsymbol{\Psi} \in \mathbb{R}^{m \times m}$ given in (1.12). This is denoted by $\boldsymbol{Y} \sim \text{log-GBS}_m(\boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\Psi}, g^{(m)})$. The CDF of $\boldsymbol{Y}$ is

$$F_{\boldsymbol{Y}}(\boldsymbol{y}; \boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\Psi}, g^{(m)}) = F_{\text{EC}_m}(\boldsymbol{B}; \boldsymbol{\Psi}, g^{(m)}), \quad \boldsymbol{y} = (y_1, \ldots, y_m)^\top \in \mathbb{R}^m, \tag{2.1}$$

where $\boldsymbol{B} = B(\boldsymbol{y};\boldsymbol{\alpha},\boldsymbol{\mu}) = (B_1,\ldots,B_m)^\top$, with $B_j = B(y_j;\alpha_j,\mu_j)$, for $j=1,\ldots,m$, as given in (1.6). The PDF of $\boldsymbol{Y}$ is given by

$$f_{\boldsymbol{Y}}(\boldsymbol{y};\boldsymbol{\alpha},\boldsymbol{\mu},\boldsymbol{\Psi},g^{(m)}) = f_{\mathrm{EC}_m}\left(\boldsymbol{B};\boldsymbol{\Psi},g^{(m)}\right) b(\boldsymbol{y};\boldsymbol{\alpha},\boldsymbol{\mu}),\ \ \boldsymbol{y}\in\mathbb{R}^m, \tag{2.2}$$

where f_{EC_m} is given in (1.11) and $b(\boldsymbol{y};\boldsymbol{\alpha},\boldsymbol{\mu}) = \prod_{j=1}^m b(y_i;\alpha_j,\mu_j)$, with $b(y_j;\alpha_j,\mu_j)$ as given in (1.6), for $j=1,\ldots,m$.

If $\boldsymbol{Y}\sim \text{log-GBS}_m(\boldsymbol{\alpha},\boldsymbol{\mu},\boldsymbol{\Psi},g^{(m)})$, from (B1) and (1.12), we have that:

(D1) $\boldsymbol{D}(\boldsymbol{\alpha})\,B(\boldsymbol{Y};\boldsymbol{\alpha},\boldsymbol{\mu}) \sim \mathrm{EC}_m(\boldsymbol{0},\boldsymbol{D}(\boldsymbol{\alpha})\boldsymbol{\Psi}\boldsymbol{D}(\boldsymbol{\alpha}),g^{(m)})$, where $\boldsymbol{D}(\boldsymbol{\alpha}) = \mathrm{diag}(\alpha_1,\ldots,\alpha_m)$ and

$$\boldsymbol{D}(\boldsymbol{\alpha})\boldsymbol{\Psi}\boldsymbol{D}(\boldsymbol{\alpha}) = \begin{pmatrix} \alpha_1^2 & \alpha_1\alpha_2\psi_{12} & \cdots & \alpha_1\alpha_m\psi_{1m} \\ \alpha_1\alpha_2\psi_{12} & \alpha_2^2 & \cdots & \alpha_2\alpha_m\psi_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1\alpha_m\psi_{1m} & \alpha_2\alpha_m\psi_{2m} & \cdots & \alpha_m^2 \end{pmatrix}. \tag{2.3}$$

(D2) From (C3), we have that: $B^\top(\boldsymbol{Y};\boldsymbol{\alpha},\boldsymbol{\mu})\boldsymbol{\Psi}^{-1}B(\boldsymbol{Y};\boldsymbol{\alpha},\boldsymbol{\mu}) \sim \mathrm{G}\chi^2(m,g^{(m)})$.

Next, Theorem 1 provides marginal and conditional distributions of multivariate log-GBS distributions, which are useful in the development of this work.

Theorem 1. *Let $\boldsymbol{Y}\sim$ log-GBS$_m(\boldsymbol{\alpha},\boldsymbol{\mu},\boldsymbol{\Psi},g^{(m)})$, and $\boldsymbol{Y}$, $\boldsymbol{\alpha},\boldsymbol{\mu},\boldsymbol{\Psi}$ be partitioned as*

$$\boldsymbol{Y} = \begin{pmatrix} \boldsymbol{Y}_1 \\ \boldsymbol{Y}_2 \end{pmatrix},\ \boldsymbol{\alpha} = \begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \end{pmatrix},\ \boldsymbol{\mu} = \begin{pmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{pmatrix},\ \boldsymbol{\Psi} = \begin{pmatrix} \boldsymbol{\Psi}_{11} & \boldsymbol{\Psi}_{12} \\ \boldsymbol{\Psi}_{21} & \boldsymbol{\Psi}_{22} \end{pmatrix}, \tag{2.4}$$

where $\boldsymbol{Y}_1$, $\boldsymbol{\alpha}_1$, $\boldsymbol{\mu}_1\in\mathbb{R}^q$, $\boldsymbol{\Psi}_{11}\in\mathbb{R}^{q\times q}$, and the remainder of elements defined suitably. Then,

(a) $\boldsymbol{Y}_1\sim$ log-GBS$_q(\boldsymbol{\alpha}_1,\boldsymbol{\mu}_1,\boldsymbol{\Psi}_{11},g^{(q)})$, with

$$g^{(q)}(s) = \left[\frac{\pi^{\frac{m-q}{2}}}{\Gamma\left(\frac{m-q}{2}\right)}\right]\int_0^\infty y^{\frac{m-q-2}{2}} g^{(m)}(s+y)\mathrm{d}y,$$

and $\boldsymbol{Y}_2\sim$ log-GBS$_{m-q}(\boldsymbol{\alpha}_2,\boldsymbol{\mu}_2,\boldsymbol{\Psi}_{22},g^{(m-q)})$, with $g^{(m-q)}$ being obtained such as $g^{(q)}$;

(b) The CDF of $\boldsymbol{Y}_1$ given $\boldsymbol{Y}_2=\boldsymbol{y}_2$ is

$$F_{\boldsymbol{Y}_1|\boldsymbol{Y}_2=\boldsymbol{y}_2}(\boldsymbol{y}_1) = F_{EC_q}(\boldsymbol{B}_1-\boldsymbol{\Psi}_{12}\boldsymbol{\Psi}_{22}^{-1}\boldsymbol{B}_2;\boldsymbol{\Psi}_{11.2},g^{(q)}_{\boldsymbol{B}_2^\top\boldsymbol{\Psi}_{22}^{-1}\boldsymbol{B}_2}),$$

where $\boldsymbol{B}_1 = (B_1,\ldots,B_q)^\top$, $\boldsymbol{B}_2 = (B_{q+1},\ldots,B_m)^\top$, $\boldsymbol{\Psi}_{11.2} = \boldsymbol{\Psi}_{11}-\boldsymbol{\Psi}_{12}\boldsymbol{\Psi}_{22}^{-1}\boldsymbol{\Psi}_{21}$, with $B_j = B(y_j;\alpha_j,\mu_j)$, for $j=1,\ldots,m$ as given in (1.6) and

$$g_w^{(q)}(s) = \frac{g^{(m)}(s+w)}{g^{(m-q)}(w)};$$

(c) The PDF of $\boldsymbol{Y}_1|\boldsymbol{Y}_2=\boldsymbol{y}_2$ *is*

$$f_{\boldsymbol{Y}_1|\boldsymbol{Y}_2=y_2}(y_1)=f_{EC_q}(\boldsymbol{B}_1-\boldsymbol{\Psi}_{12}\boldsymbol{\Psi}_{22}^{-1}\boldsymbol{B}_2,\boldsymbol{\Psi}_{11.2},g^{(q)}_{\boldsymbol{B}_2^\top\boldsymbol{\Psi}_{22}^{-1}\boldsymbol{B}_2})\prod_{j=1}^{q}b(y_j;\alpha_j,\mu_j).$$

See Appendix 1 for the proof of this theorem.

From (2.2), if $g^{(m)}$ is the m-variate normal or t kernel, then the m-variate log-BS or log-BS-t distribution is obtained, which is denoted by $\boldsymbol{Y}\sim\text{log-BS}_m(\boldsymbol{\alpha},\,\boldsymbol{\mu},\,\boldsymbol{\Psi})$ or $\boldsymbol{Y}\sim\text{log-BS-}t_m(\boldsymbol{\alpha},\,\boldsymbol{\mu},\,\boldsymbol{\Psi},\nu)$, respectively. Thus, from Table 1.1, the corresponding PDFs are

$$\begin{aligned}f_{\boldsymbol{Y}}(\boldsymbol{y};\boldsymbol{\alpha},\boldsymbol{\mu},\boldsymbol{\Psi}) &= \frac{1}{[2\pi]^{\frac{m}{2}}|\boldsymbol{\Psi}|^{\frac{1}{2}}}\exp\left(-\frac{1}{2}\boldsymbol{B}^\top\boldsymbol{\Psi}^{-1}\boldsymbol{B}\right)\prod_{j=1}^{m}\frac{1}{\alpha_j}\cosh\left(\frac{y_j-\mu_j}{2}\right),\ \boldsymbol{y}\in\mathbb{R}^m,\\ f_{\boldsymbol{Y}}(\boldsymbol{y};\boldsymbol{\alpha},\boldsymbol{\mu},\boldsymbol{\Psi},\nu) &= \frac{\Gamma(\frac{\nu+m}{2})}{\Gamma(\frac{\nu}{2})[\nu\pi]^{\frac{m}{2}}|\boldsymbol{\Psi}|^{\frac{1}{2}}}\left[1+\frac{\boldsymbol{B}^\top\boldsymbol{\Psi}^{-1}\boldsymbol{B}}{\nu}\right]^{-\frac{[\nu+m]}{2}}\prod_{j=1}^{m}\frac{1}{\alpha_j}\cosh\left(\frac{y_j-\mu_j}{2}\right),\ \boldsymbol{y}\in\mathbb{R}^m,\end{aligned}$$

respectively, where $\boldsymbol{B}$ is defined in (2.1).

From property (D2), for $\boldsymbol{\theta}=(\boldsymbol{\alpha}^\top,\boldsymbol{\mu}^\top,\text{svec}(\boldsymbol{\Psi})^\top)^\top$, with 'svec' denoting vectorization of a symmetric matrix, we obtain the Mahalanobis distance (MD) for the case i as

$$\text{MD}_i(\boldsymbol{\theta})=B^\top(\boldsymbol{Y}_i;\boldsymbol{\alpha},\boldsymbol{\mu})\boldsymbol{\Psi}^{-1}B(\boldsymbol{Y}_i;\boldsymbol{\alpha},\boldsymbol{\mu}),\quad i=1,\ldots,n. \tag{2.5}$$

Based on Lange et al. (1989) and Lange and Sinsheimer (1993), we have:
(i) $\text{MD}_i(\boldsymbol{\theta})\sim\chi^2(p)$, if $g^{(p)}$ is the multivariate normal kernel; and
(ii) $\text{MD}_i(\boldsymbol{\theta})/p\sim\mathcal{F}(p,\nu)$, if $g^{(p)}$ is the multivariate t kernel, where $\mathcal{F}(p,\nu)$ denotes the Fisher distribution with ν DFs in the numerator and p in the denominator. Note that MD_i for the multivariate t kernel is known as modified or generalized MD.
When evaluated at the ML estimate of $\boldsymbol{\theta}$, the MD for the case i defined in (2.5) is useful for assessing multivariate outliers and evaluating the goodness of fit in m-variate log-GBS distributions.

Let $\boldsymbol{S}\in\mathbb{R}^m$ be a random vector following an m-variate t distribution with $\nu\in\mathbb{R}_+$ DFs and scale matrix $\boldsymbol{\Sigma}$. Note that, similarly to (1.12), $\boldsymbol{\Sigma}$ is equal to the correlation matrix $\boldsymbol{\Psi}$ ($\text{rk}(\boldsymbol{\Sigma})=\text{rk}(\boldsymbol{\Psi})=m$). Thus, we use the notation $\boldsymbol{S}\sim t_m(\boldsymbol{0}_{m\times1},\boldsymbol{\Psi},\nu)$. Then, analogously to the univariate case given in (1.7), we have

$$\boldsymbol{S}=U^{-1/2}\boldsymbol{Z}\sim t_m(\boldsymbol{0},\boldsymbol{\Psi},\nu),\ \text{ where }\ \boldsymbol{Z}\sim\text{N}_m(\boldsymbol{0},\boldsymbol{\Psi}),\ \text{ and }\ U\sim\text{Gamma}\left(\frac{\nu}{2},\frac{\nu}{2}\right), \tag{2.6}$$

with U being independent of the elements of $\boldsymbol{Z}=(Z_1,\ldots,Z_m)^\top$. Based on (2.6), $\boldsymbol{S}|U=u\sim\text{N}_m(\boldsymbol{0},[1/u]\boldsymbol{\Psi})$; see McLachlan and Krishnan (1997, p. 73). Therefore, from (1.8) and (2.6), $\boldsymbol{Y}=(Y_1,\ldots,Y_m)^\top\sim\text{log-BS-}t_m(\boldsymbol{\alpha},\boldsymbol{\mu},\boldsymbol{\Psi},\nu)$, with

$$Y_j=\mu_j+2\,\text{arcsinh}\left(\frac{\alpha_j\,Z_j}{2\,U^{1/2}}\right),\ \text{for } j=1,\ldots,m.$$

Note that:

(E1) $\boldsymbol{Y}|U=u \sim \text{log-BS}_m(u^{-1/2}\boldsymbol{\alpha},\, \boldsymbol{\mu}, [\frac{1}{u}]\boldsymbol{\Psi})$;

(E2) $U|\boldsymbol{Y}=\boldsymbol{y} \sim \text{Gamma}\Big(\frac{\nu+m}{2},\, \frac{\nu+\boldsymbol{B}^\top\boldsymbol{\Psi}^{-1}\boldsymbol{B}}{2}\Big)$, with $\boldsymbol{B}$ as given in (2.1);

(E3) $\text{E}[U|\boldsymbol{Y}=\boldsymbol{y}] = \frac{[\nu+m]}{[\nu+\boldsymbol{B}^\top\boldsymbol{\Psi}^{-1}\boldsymbol{B}]}$;

(E4) $\text{E}[U^2|\boldsymbol{Y}=\boldsymbol{y}] = \frac{[\nu+m][\nu+m+2]}{[\nu+\boldsymbol{B}^\top\boldsymbol{\Psi}^{-1}\boldsymbol{B}]^2}$.

Note that (E4) is useful for implementing the EM algorithm in the ML estimation of the multivariate log-BS-t parameters; see Lange et al. (1989), McLachlan and Krishnan (1997), Balakrishnan et al. (2009), and Paula et al. (2012). Random vectors from m-variate log-BS and log-BS-t distributions can be generated using the Algorithms 1 and 2, respectively; see Leiva et al. (2008b) for generation of GBS and log-GBS numbers.

Algorithm 1 Generator of random vectors from m-variate log-BS distributions.

1: Make a Cholesky decomposition of $\boldsymbol{\Psi}$ as $\boldsymbol{\Psi}=\boldsymbol{L}\boldsymbol{L}^\top$, where $\boldsymbol{L}$ is a lower triangular matrix with real and positive diagonal entries.
2: Generate m independent standard normal random numbers, say, $\boldsymbol{W}=(W_1,\ldots,W_m)^\top$.
3: Compute $\boldsymbol{Z}=\boldsymbol{L}\boldsymbol{W}=(Z_1,\ldots,Z_m)^\top$.
4: Obtain $\boldsymbol{Y}$ with elements $Y_j=\mu_j+2\arcsin(\alpha_j\, Z_j/2)$ for $j=1,\ldots,m$.
5: Repeat Steps 1 to 4 until the required vector of data is generated.

Algorithm 2 Generator of random vectors from m-variate log-BS-t distributions.

1: Repeat Steps 1 to 3 of Algorithm 1.
2: Generate random numbers from $U \sim \text{Gamma}(\nu/2, \nu/2)$.
3: Obtain $\boldsymbol{Y}$ with elements $Y_j=\mu_j+2\arcsin(\alpha_j\, Z_j/(2\,U^{1/2}))$, for $j=1,\ldots,m$.
4: Repeat Steps 1 to 5 until the required vector of data is generated.

Figures 2.1, 2.2, 2.3, 2.4, 2.5 and 2.6 show contours of the PDFs for the bivariate log-BS and log-BS-t distributions, denoted log-BS_2 and log-BS-t_2, respectively, for $\psi=0.0$, $\boldsymbol{\alpha}=(\alpha_1,\alpha_2)^\top$, with $\alpha_1=0.75, 1.0, 2.0, 3.0$, $\alpha_2=0.75, 1.0, 2.0, 3.0$, $\boldsymbol{\mu}=(0,0)^\top$ and $\nu=4$, and Figures 2.7, 2.8, 2.9, 2.10, 2.11 and 2.12 for $\psi=0.5$ and the same values for $\boldsymbol{\alpha}$, $\boldsymbol{\mu}$ and ν. From these figures, note that the log-BS-t_2 distribution has heavier tails than the log-BS_2 and a more asymmetric behavior as α_l decreases, for $l=1,2,3,4$. Note that values for $\alpha_l>2$ provide bimodality in log-GBS_2 distributions as in the univariate case; see Figure 1.1, Leiva et al. (2007) and Barros et al. (2008). In addition, for $\alpha_l=4$ PDFs of

log-GBS distributions present four-modality. However, a number of studies have reported that this situation is unusual when fatigue data are analyzed, such is in the case of the data analysis reported in the part of application of this work provided in Section 2.6; see Birnbaum and Saunders (1969), Rieck and Nedelman (1991) and Lepadatu et al. (2005).

2.4 Multivariate GBS log-linear regression model

Consider a multivariate extension of the BS log-linear regression model defined in (1.4) as

$$\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{\beta} + \boldsymbol{E}, \tag{2.7}$$

where $\boldsymbol{Y} = (Y_{ij}) \in \mathbb{R}^{n\times m}$ is the log-response matrix, and $\boldsymbol{X} = (x_{is}) \in \mathbb{R}^{n\times p}$ the model matrix of rank p containing the values of p regressors. Here, $\boldsymbol{X}$ and $\boldsymbol{Y}$ are linked by a coefficient matrix $\boldsymbol{\beta} = (\beta_{sj}) = (\boldsymbol{\beta}_1, \ldots, \boldsymbol{\beta}_m) \in \mathbb{R}^{p\times m}$ to be estimated, and $\boldsymbol{E} = (\varepsilon_{ij}) \in \mathbb{R}^{n\times m}$ is the error matrix. Also, in the model given in (2.7), let $\boldsymbol{Y}_i^\top$, $\boldsymbol{x}_i^\top$ and $\boldsymbol{\varepsilon}_i^\top$ be the ith rows of $\boldsymbol{Y}$, $\boldsymbol{X}$ and $\boldsymbol{E}$, respectively. Thus, we can write

$$\boldsymbol{Y}_i = \boldsymbol{\mu}_i + \boldsymbol{\varepsilon}_i = \boldsymbol{\beta}^\top \boldsymbol{x}_i + \boldsymbol{\varepsilon}_i, \quad i = 1, \ldots, n, \tag{2.8}$$

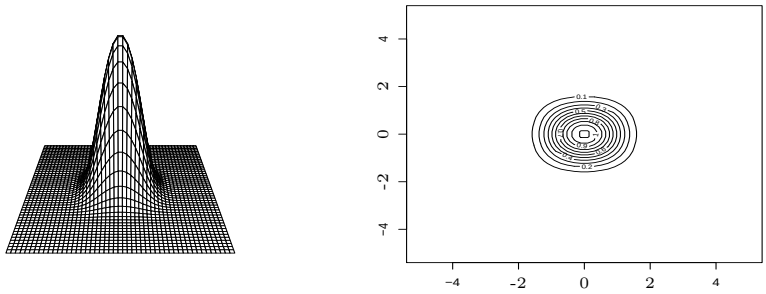
where $\boldsymbol{\varepsilon}_1, \ldots, \boldsymbol{\varepsilon}_n$ are independently and identically distributed (IID) log-GBS$_m(\alpha\mathbf{1}_{m\times 1}, \mathbf{0}_{m\times 1}, \boldsymbol{\Psi}, g^{(m)})$, with $\mathbf{1}_{m\times 1}$ being a vector of ones; see Liu and Rubin (1995) and Díaz-García et al. (2003).

From (2.8), note that the shape parameter (α) is assumed to be the same throughout the different log-responses and individuals, such as in Rieck and Nedelman's univariate model. This assumption could be restrictive, because the different responses might have different units of measurement. Then, two aspects about this assumption must be pointed out as follows:

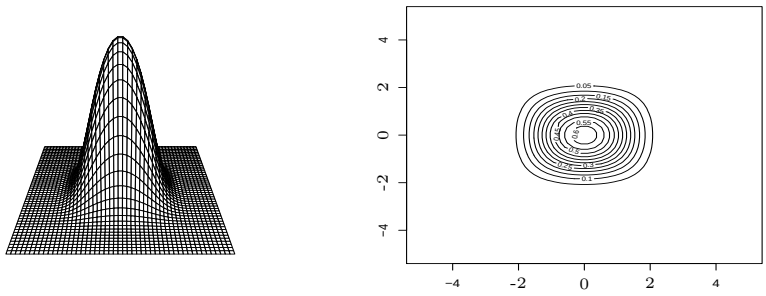
(1) Location and scale parameters have the same unit of measurement than the corresponding RV, facilitating their interpretation. However, meaning of shape parameters are distribution-dependent. In the case of the BS distribution, its shape parameter is dimensionless due to its genesis from material fatigue. This is because the BS shape parameter is expressed as a ratio between two parameters related to another RV used to generate the RV with BS distribution, causing it to be dimensionless; see Birnbaum and Saunders (1969). Thus, different units of measurement for the responses should not affect the shape of their marginal distributions.

(2) A distinct aspect to that pointed out in (1) is that the BS shape parameter changes with individuals, such as the median (in the original scale) or the mean (in the log-scale) to

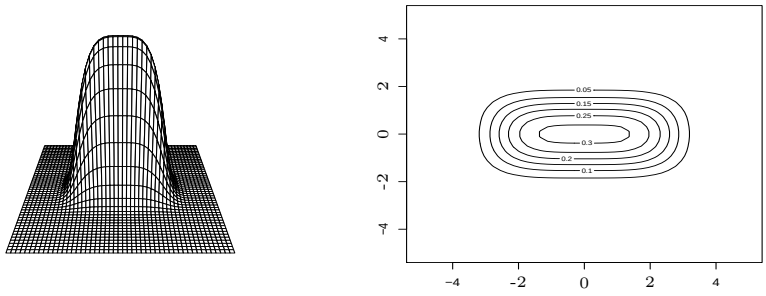
Figure 2.1: PDF and its contour plot for the log-BS_2 distribution with $\psi = 0.0$ and $\boldsymbol{\mu} = (0, 0)^\top$.



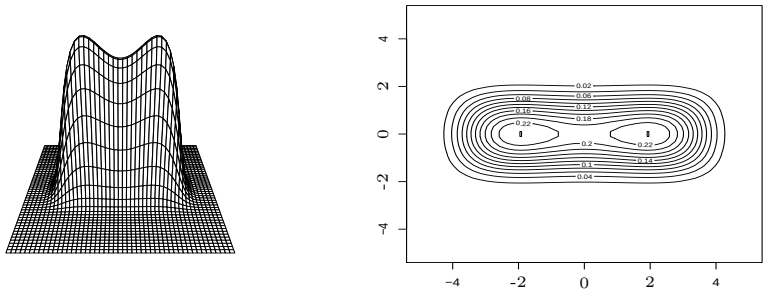
(a) $\boldsymbol{\alpha} = (0.75, 0.75)^\top$ (b) $\boldsymbol{\alpha} = (0.75, 0.75)^\top$



(c) $\boldsymbol{\alpha} = (1, 1)^\top$ (d) $\boldsymbol{\alpha} = (1, 1)^\top$



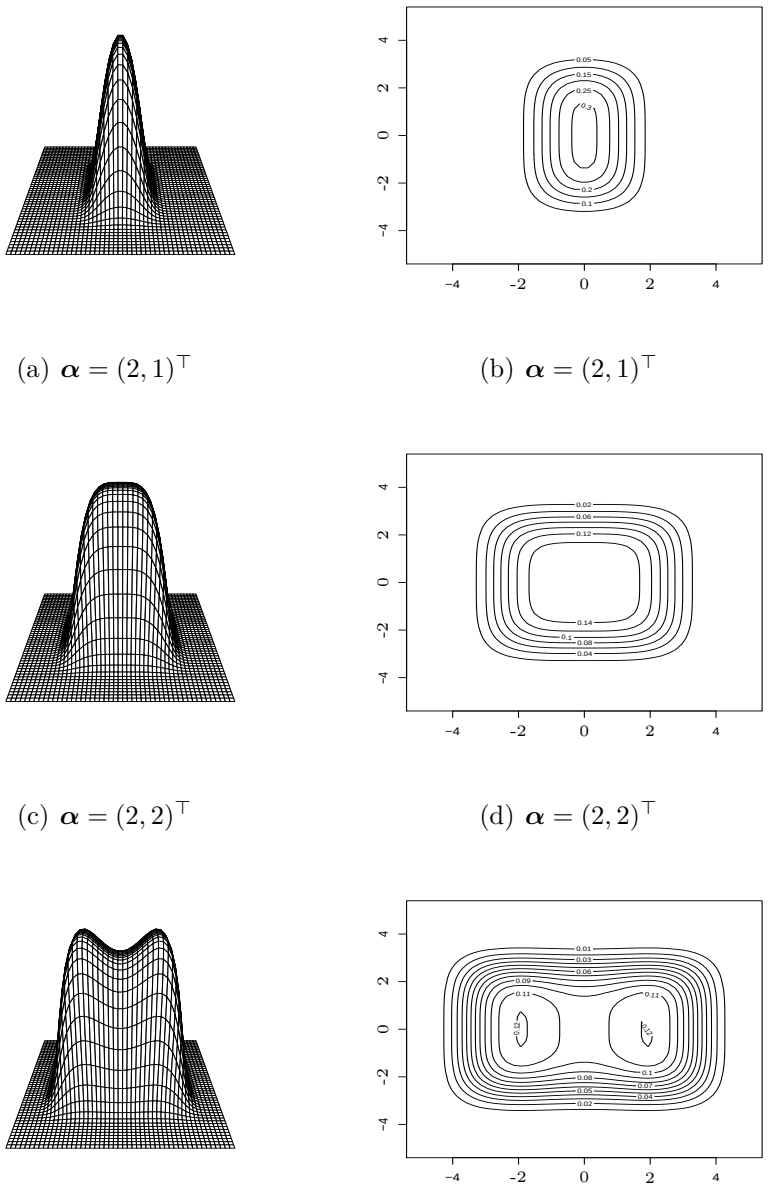
(e) $\boldsymbol{\alpha} = (1, 2)^\top$ (f) $\boldsymbol{\alpha} = (1, 2)^\top$



(g) $\boldsymbol{\alpha} = (1, 3)^\top$ (h) $\boldsymbol{\alpha} = (1, 3)^\top$

Source: From the author.

Figure 2.2: PDF and its contour plot for the log-BS_2 distribution with $\psi = 0.0$ and $\boldsymbol{\mu} = (0, 0)^\top$.



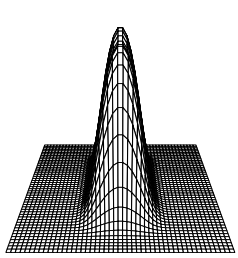
(a) $\boldsymbol{\alpha} = (2, 1)^\top$ (b) $\boldsymbol{\alpha} = (2, 1)^\top$

(c) $\boldsymbol{\alpha} = (2, 2)^\top$ (d) $\boldsymbol{\alpha} = (2, 2)^\top$

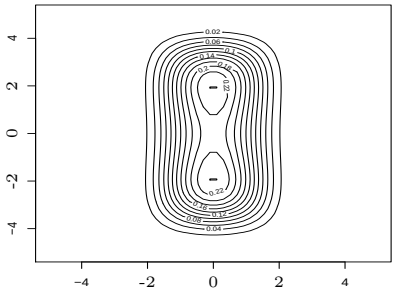
(e) $\boldsymbol{\alpha} = (2, 3)^\top$ (f) $\boldsymbol{\alpha} = (2, 3)^\top$

Source: From the author.

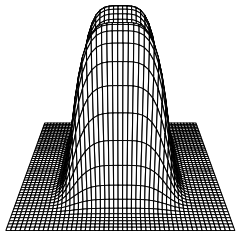
Figure 2.3: PDF and its contour plot for the log-BS_2 distribution with $\psi = 0.0$ and $\boldsymbol{\mu} = (0, 0)^\top$.



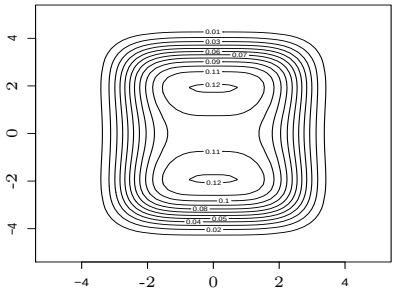
(a) $\boldsymbol{\alpha} = (3, 1)^\top$



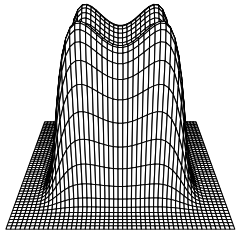
(b) $\boldsymbol{\alpha} = (3, 1)^\top$



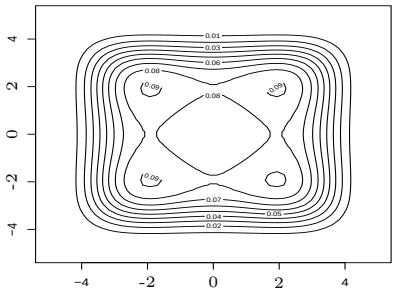
(c) $\boldsymbol{\alpha} = (3, 2)^\top$



(d) $\boldsymbol{\alpha} = (3, 2)^\top$



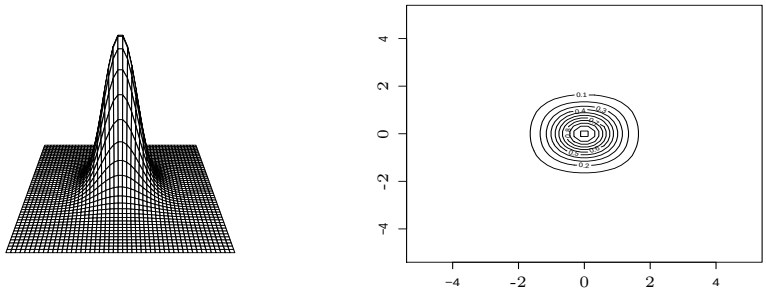
(e) $\boldsymbol{\alpha} = (3, 3)^\top$



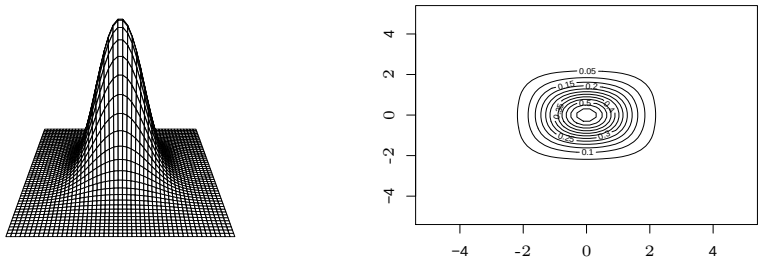
(f) $\boldsymbol{\alpha} = (3, 3)^\top$

Source: From the author.

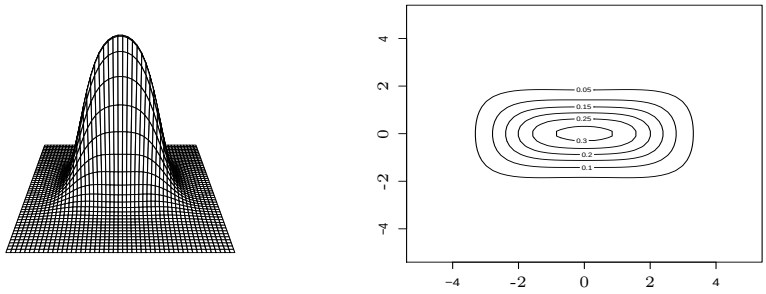
Figure 2.4: PDF and its contour plot for the log-BS-t_2 distribution with $\psi = 0.0$, $\boldsymbol{\mu} = (0, 0)^\top$ and $\nu = 4$.



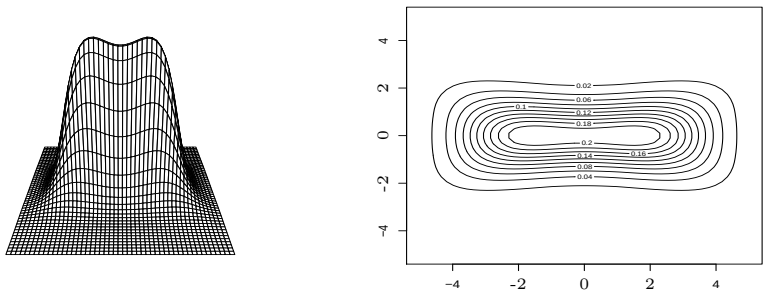
(a) $\boldsymbol{\alpha} = (0.75, 0.75)^\top$ (b) $\boldsymbol{\alpha} = (0.75, 0.75)^\top$



(c) $\boldsymbol{\alpha} = (1, 1)^\top$ (d) $\boldsymbol{\alpha} = (1, 1)^\top$



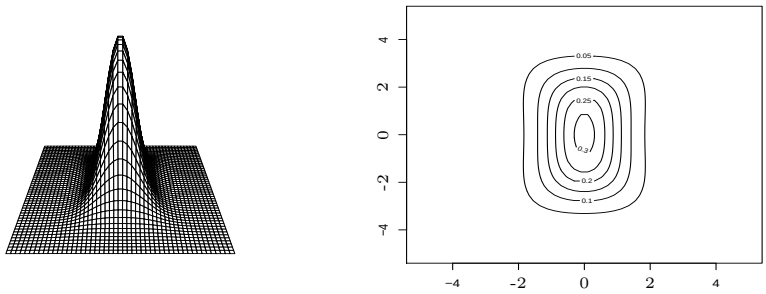
(e) $\boldsymbol{\alpha} = (1, 2)^\top$ (f) $\boldsymbol{\alpha} = (1, 2)^\top$



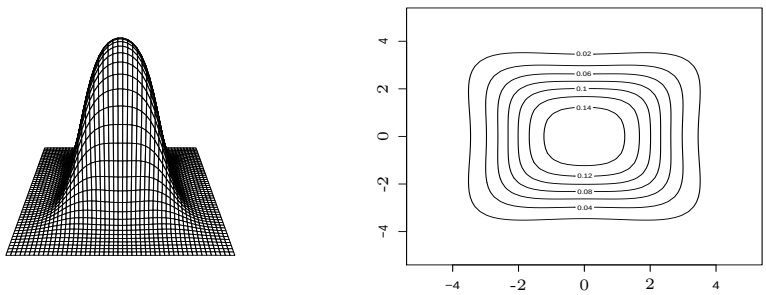
(g) $\boldsymbol{\alpha} = (1, 3)^\top$ (h) $\boldsymbol{\alpha} = (1, 3)^\top$

Source: From the author.

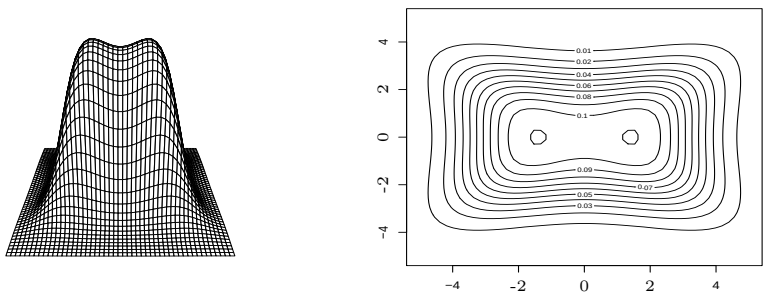
Figure 2.5: PDF and its contour plot for the log-BS-t_2 distribution with $\psi = 0.0$, $\boldsymbol{\mu} = (0, 0)^\top$ and $\nu = 4$.



(a) $\boldsymbol{\alpha} = (2, 1)^\top$ (b) $\boldsymbol{\alpha} = (2, 1)^\top$



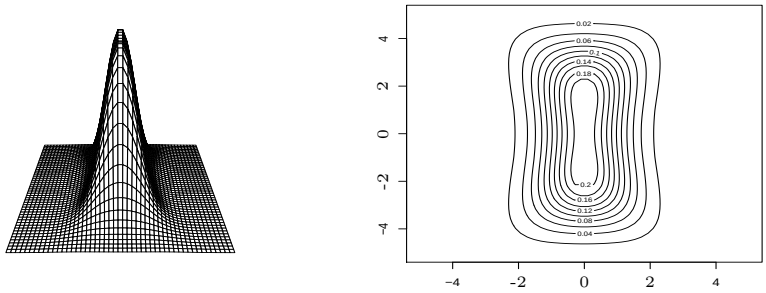
(c) $\boldsymbol{\alpha} = (2, 2)^\top$ (d) $\boldsymbol{\alpha} = (2, 2)^\top$



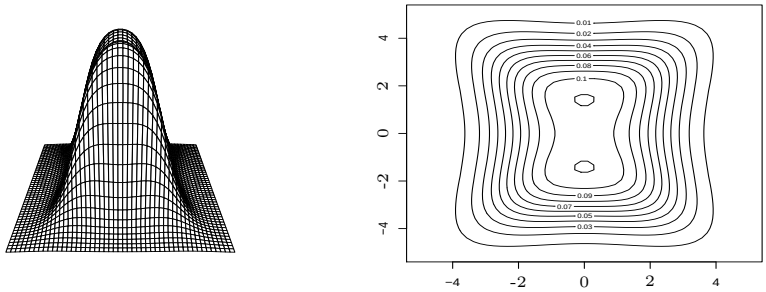
(e) $\boldsymbol{\alpha} = (2, 3)^\top$ (f) $\boldsymbol{\alpha} = (2, 3)^\top$

Source: From the author.

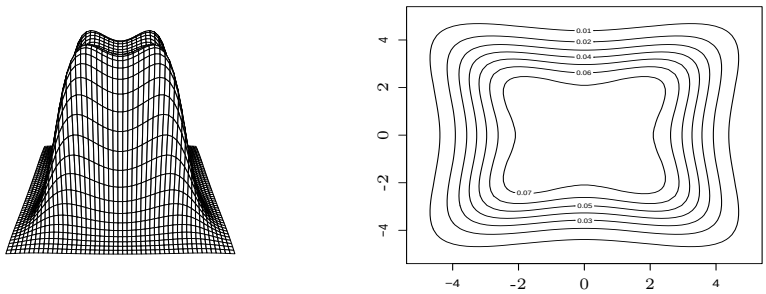
Figure 2.6: PDF and its contour plot for the log-BS-t_2 distribution with $\psi = 0.0$, $\boldsymbol{\mu} = (0, 0)^\top$ and $\nu = 4$.



(a) $\boldsymbol{\alpha} = (3, 1)^\top$ (b) $\boldsymbol{\alpha} = (3, 1)^\top$



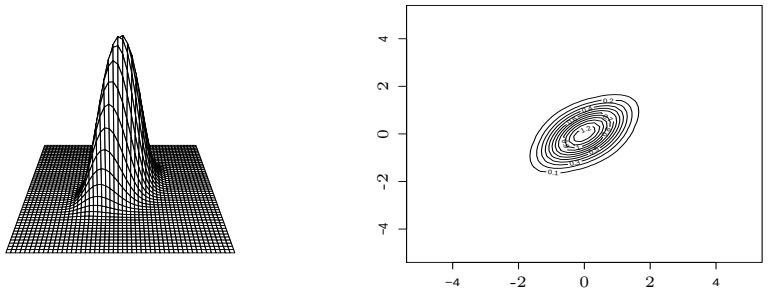
(c) $\boldsymbol{\alpha} = (3, 2)^\top$ (d) $\boldsymbol{\alpha} = (3, 2)^\top$



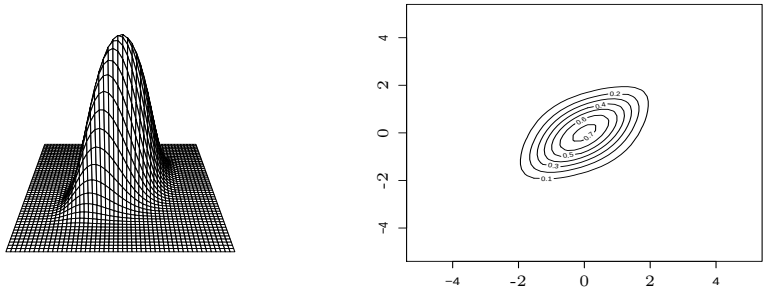
(e) $\boldsymbol{\alpha} = (3, 3)^\top$ (f) $\boldsymbol{\alpha} = (3, 3)^\top$

Source: From the author.

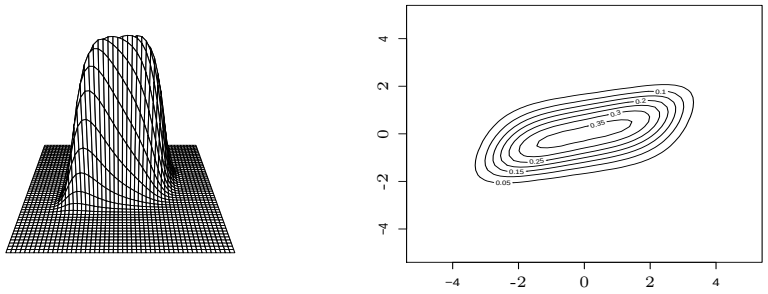
Figure 2.7: PDF and its contour plot for the log-BS_2 distribution with $\psi = 0.5$ and $\boldsymbol{\mu} = (0, 0)^\top$.



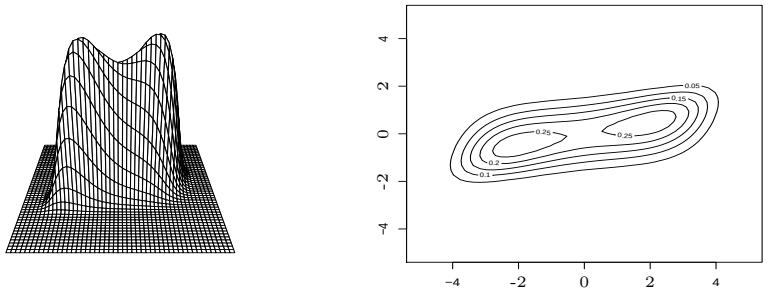
(a) $\boldsymbol{\alpha} = (0.75, 0.75)^\top$ (b) $\boldsymbol{\alpha} = (0.75, 0.75)^\top$



(c) $\boldsymbol{\alpha} = (1, 1)^\top$ (d) $\boldsymbol{\alpha} = (1, 1)^\top$



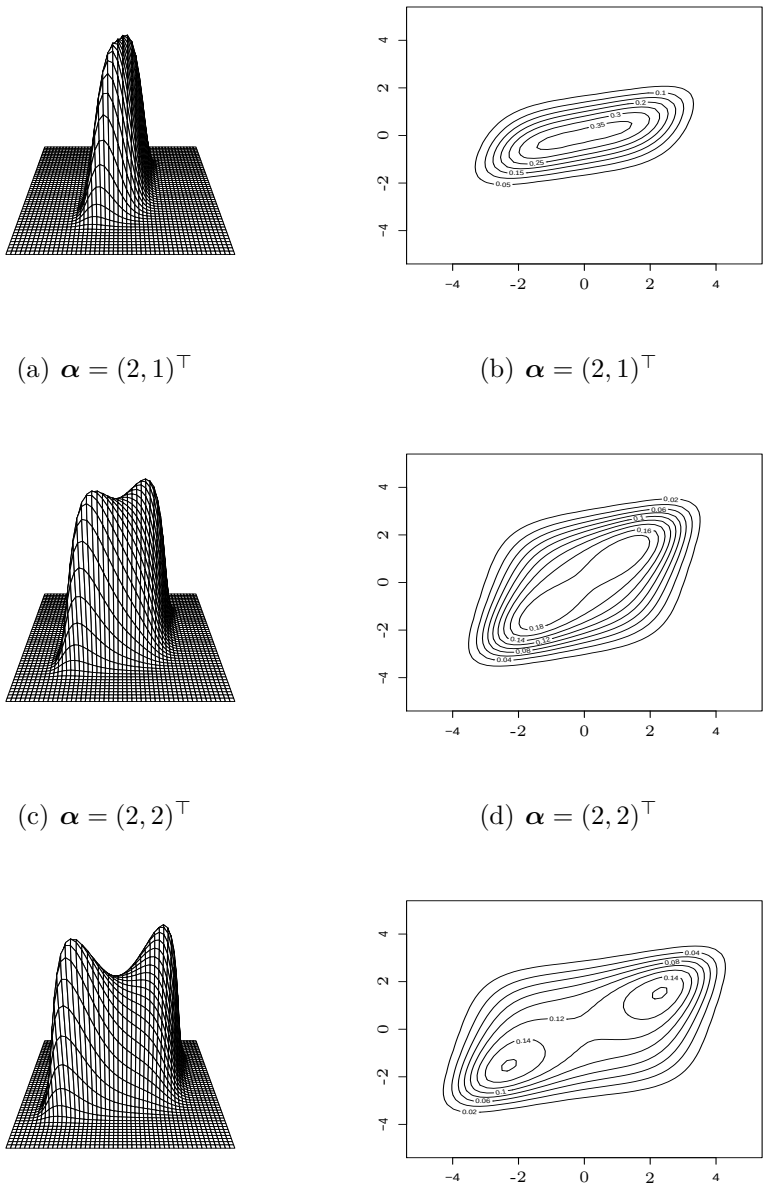
(e) $\boldsymbol{\alpha} = (1, 2)^\top$ (f) $\boldsymbol{\alpha} = (1, 2)^\top$



(g) $\boldsymbol{\alpha} = (1, 3)^\top$ (h) $\boldsymbol{\alpha} = (1, 3)^\top$

Source: From the author.

Figure 2.8: PDF and its contour plot for the log-BS_2 distribution with $\psi = 0.5$ and $\boldsymbol{\mu} = (0, 0)^\top$.



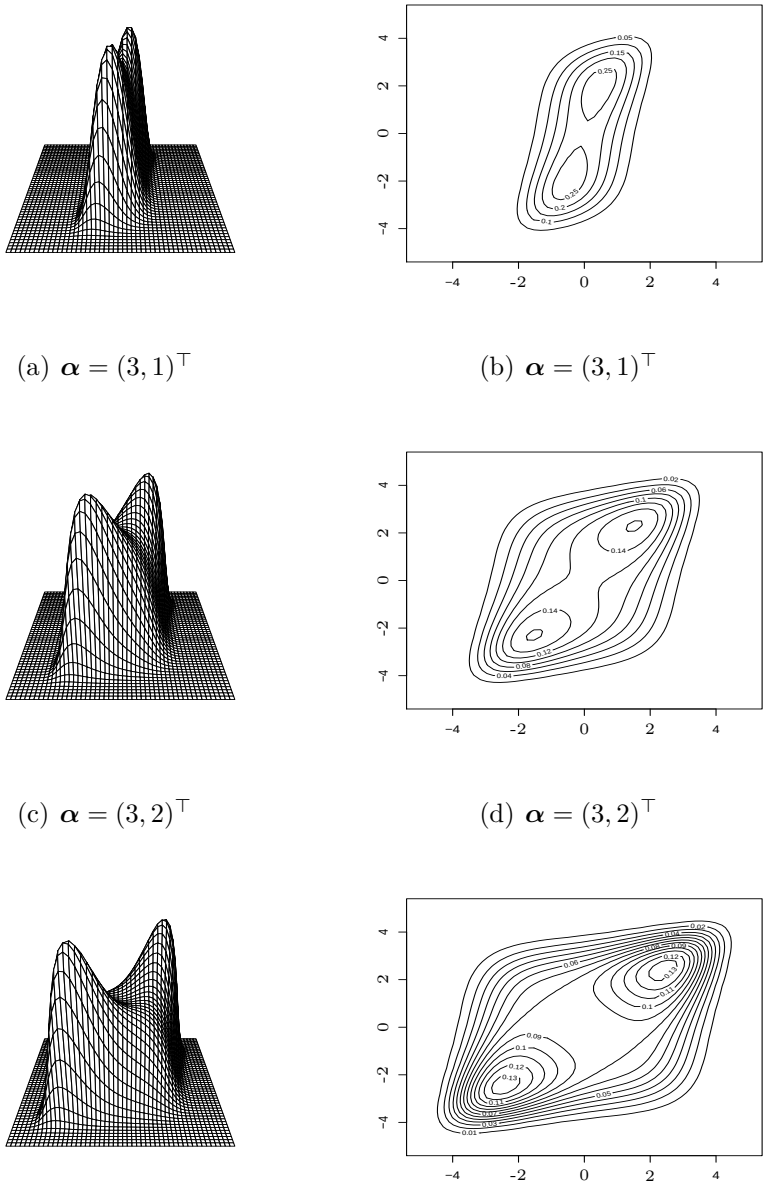
(a) $\boldsymbol{\alpha} = (2, 1)^\top$ (b) $\boldsymbol{\alpha} = (2, 1)^\top$

(c) $\boldsymbol{\alpha} = (2, 2)^\top$ (d) $\boldsymbol{\alpha} = (2, 2)^\top$

(e) $\boldsymbol{\alpha} = (2, 3)^\top$ (f) $\boldsymbol{\alpha} = (2, 3)^\top$

Source: From the author.

Figure 2.9: PDF and its contour plot for the log-BS_2 distribution with $\psi = 0.5$ and $\boldsymbol{\mu} = (0, 0)^\top$.



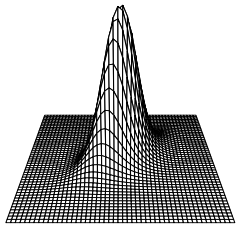
(a) $\boldsymbol{\alpha} = (3, 1)^\top$ (b) $\boldsymbol{\alpha} = (3, 1)^\top$

(c) $\boldsymbol{\alpha} = (3, 2)^\top$ (d) $\boldsymbol{\alpha} = (3, 2)^\top$

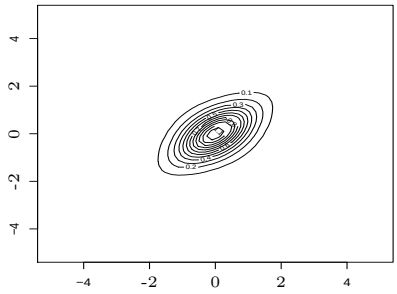
(e) $\boldsymbol{\alpha} = (3, 3)^\top$ (f) $\boldsymbol{\alpha} = (3, 3)^\top$

Source: From the author.

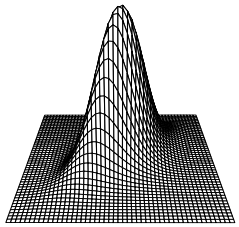
Figure 2.10: PDF and its contour plot for the log-BS-t_2 distribution with $\psi = 0.5$, $\boldsymbol{\mu} = (0, 0)^\top$ and $\nu = 4$.



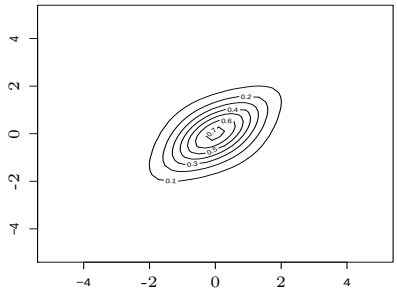
(a) $\boldsymbol{\alpha} = (0.75, 0.75)^\top$



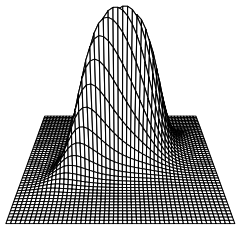
(b) $\boldsymbol{\alpha} = (0.75, 0.75)^\top$



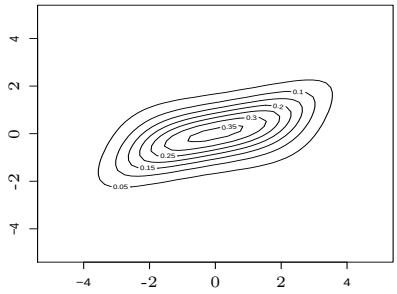
(c) $\boldsymbol{\alpha} = (1, 1)^\top$



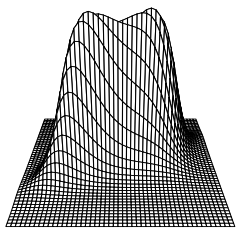
(d) $\boldsymbol{\alpha} = (1, 1)^\top$



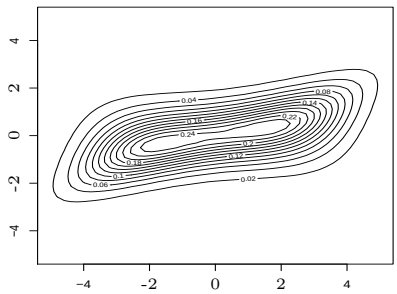
(e) $\boldsymbol{\alpha} = (1, 2)^\top$



(f) $\boldsymbol{\alpha} = (1, 2)^\top$



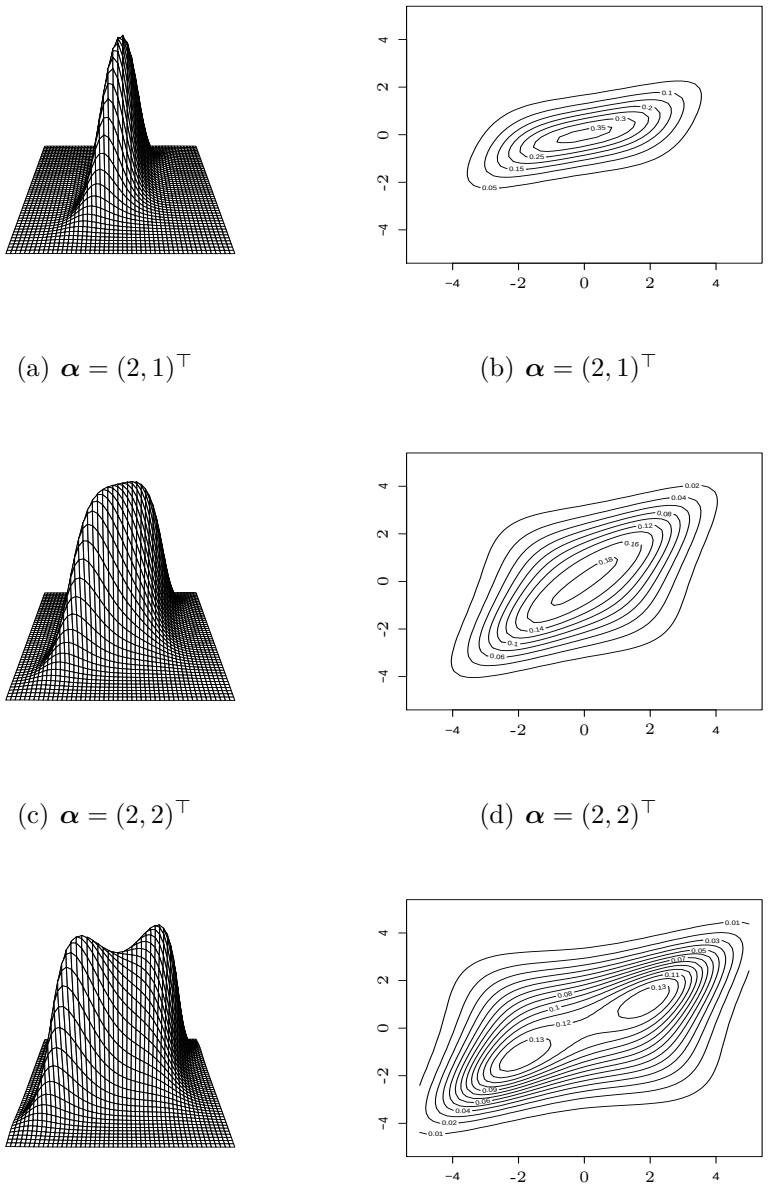
(g) $\boldsymbol{\alpha} = (1, 3)^\top$



(h) $\boldsymbol{\alpha} = (1, 3)^\top$

Source: From the author.

Figure 2.11: PDF and its contour plot for the log-BS-t_2 distribution with $\psi = 0.5$, $\boldsymbol{\mu} = (0, 0)^\top$ and $\nu = 4$.



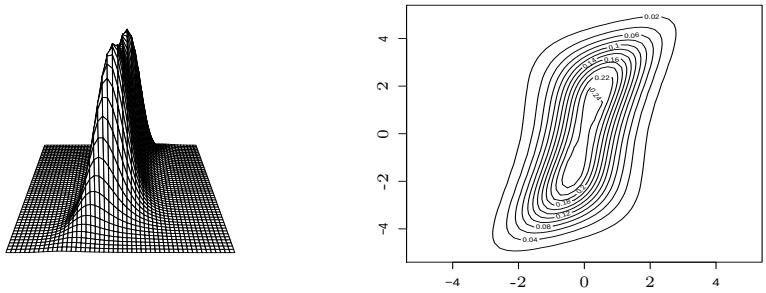
(a) $\boldsymbol{\alpha} = (2, 1)^\top$ (b) $\boldsymbol{\alpha} = (2, 1)^\top$

(c) $\boldsymbol{\alpha} = (2, 2)^\top$ (d) $\boldsymbol{\alpha} = (2, 2)^\top$

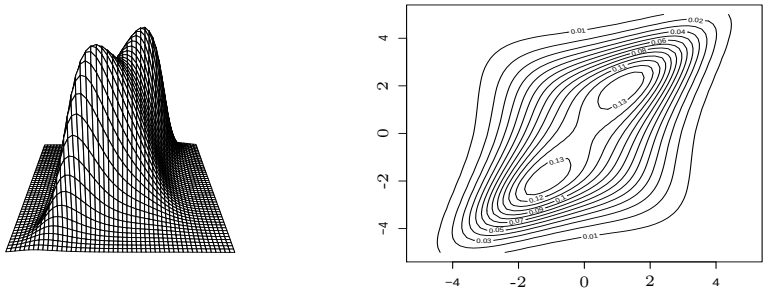
(e) $\boldsymbol{\alpha} = (2, 3)^\top$ (f) $\boldsymbol{\alpha} = (2, 3)^\top$

Source: From the author.

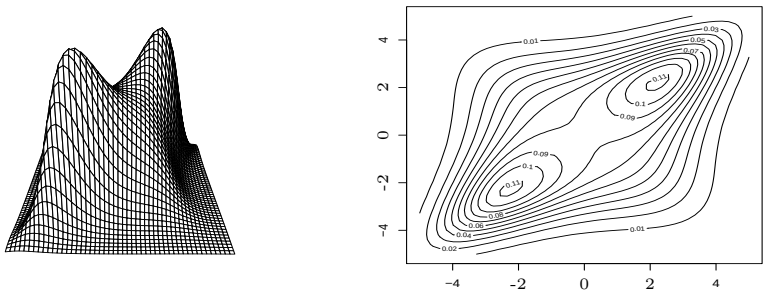
Figure 2.12: PDF and its contour plot for the log-BS-t_2 distribution with $\psi = 0.5$, $\boldsymbol{\mu} = (0, 0)^\top$ and $\nu = 4$.



(a) $\boldsymbol{\alpha} = (3, 1)^\top$ (b) $\boldsymbol{\alpha} = (3, 1)^\top$



(c) $\boldsymbol{\alpha} = (3, 2)^\top$ (d) $\boldsymbol{\alpha} = (3, 2)^\top$



(e) $\boldsymbol{\alpha} = (3, 3)^\top$ (f) $\boldsymbol{\alpha} = (3, 3)^\top$

Source: From the author.

be modeled in this chapter. Empirical evidence has demonstrated that often the BS shape parameter does not change; see the data analysis reported in Section 2.6, Birnbaum and Saunders (1969), and Leiva et al. (2014b). A test for homogeneity of the BS shape parameter in univariate BS regression models was derived by Xie and Wei (2007). However, empirical evidence with real-world fatigue data used in that work proved that the BS shape parameter does not change with the specimens, which supports our assumption. To derive a test for homogeneity of shape parameters in multivariate BS regression models can be the subject of another study, as it is beyond the scope of the present work. We recall the multivariate regression model here proposed allows us to describe changes in the location, such as in Rieck and Nedelman's univariate model.

Let $\boldsymbol{Y} = (\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n)^\top$ be a sample from a multivariate log-GBS distribution with $\mathrm{E}[\boldsymbol{Y}_i] = \boldsymbol{\beta}^\top \boldsymbol{x}_i$ (multivariate GBS log-linear regression structure), and $\boldsymbol{y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)^\top$ their observations. Then, the log-likelihood function for $\boldsymbol{\theta} = (\alpha, \mathrm{vec}(\boldsymbol{\beta})^\top, \mathrm{svec}(\boldsymbol{\Psi})^\top)^\top$, with 'vec' denoting the vectorization of a matrix, is

$$\ell(\boldsymbol{\theta}; \boldsymbol{y}) = \sum_{i=1}^{n} \log(f_{\mathrm{EC}_m}(\boldsymbol{\phi}_i; \boldsymbol{\Psi}, g^{(m)})) + \sum_{i=1}^{n}\sum_{j=1}^{m} \log(\xi_{ij}), \tag{2.9}$$

where $\boldsymbol{\phi}_i = (\phi_{i1}, \ldots, \phi_{im})^\top$, with

$$\phi_{ij} = B(y_{ij}; \alpha, \mu_{ij}) = \left[\frac{2}{\alpha}\right] \sinh\left(\frac{y_{ij} - \mu_{ij}}{2}\right) \quad \text{and} \quad \xi_{ij} = 2\, b(y_{ij}; \alpha, \mu_{ij}) = \left[\frac{2}{\alpha}\right] \cosh\left(\frac{y_{ij} - \mu_{ij}}{2}\right),$$

with $\mu_{ij} = \boldsymbol{\beta}_j^\top \boldsymbol{x}_i$, for $i = 1, \ldots, n, j = 1, \ldots, m$. From (2.9), with $g^{(m)}$ being the m-variate t kernel, the log-likelihood function for $\boldsymbol{\theta}$ is

$$\ell(\boldsymbol{\theta}; \boldsymbol{y}) = c_1 - \frac{n}{2}\log(|\boldsymbol{\Psi}|) - \left[\frac{\nu + m}{2}\right] \sum_{i=1}^{n} \log(\nu + \boldsymbol{\phi}_i^\top \boldsymbol{\Psi}^{-1} \boldsymbol{\phi}_i) + \sum_{i=1}^{n}\sum_{j=1}^{m} \log(\xi_{ij}), \tag{2.10}$$

where c_1 is a constant independent of $\boldsymbol{\theta}$, and $\xi_{ij}, \boldsymbol{\phi}_i$ are given in (2.9). From (2.10), note that it is not possible to find a closed form solution for the ML estimate of $\boldsymbol{\theta}$. Note that the parameter $\boldsymbol{\theta}$ does not involve the DFs ν of the multivariate log-BS-t distribution, which must be fixed. Thus, we work with a profile type log-likelihood function. As pointed by Lucas (1997), the influence function based on the t distribution is bounded only when ν is fixed, producing robust parameter estimates. However, it is unbounded if ν is estimated by the ML method, indicating non-robustness. This result is also valid for the estimation of the corresponding standard error (SE); see Paula et al. (2012) for the case of the univariate BS-t log-linear regression model.

For the multivariate BS log-linear regression model, the log-likelihood function for $\boldsymbol{\theta}$ can be analogously constructed as in (2.10). Despite its mathematically complex form,

the log-likelihood function given in (2.10) could be directly maximized with a numerical procedure; see MacDonald (2014). The author pointed out that, having excellent numerical optimizers now available freely, one does have reasons to avoid, when it corresponds, the direct maximization of a likelihood function. However, we decide to use, instead of this direct maximization, the EM algorithm due to the following. This algorithm comprises the expectation (E) and maximization (M) steps in each iteration. It has several advantages associated with the proposed multivariate regression model, related to properties of the t distribution given in (2.6) and (E1)-(E4), which are shared by the log-BS-t distribution, and transferred to the mentioned regression model. These advantages are:
(i) The EM algorithm can be easily implemented from a hierarchical structure based on the representation given in (2.6), allowing us to find an estimation method for both log-BS-t_m and log-BS$_m$ cases simultaneously;
(ii) Only by its E-step (and not from the direct maximization), it is possible to show the inherent robustness of the estimation method proposed for our models; and
(iii) Under the multivariate BS-t model, its M-step is equivalent to the direct maximization of the likelihood function under the multivariate BS model (normal kernel). These three points are described in detail below.

Again let $\boldsymbol{Y} = (\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n)^\top$ be a sample from the multivariate log-BS-t distribution and $\boldsymbol{y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)^\top$ their observations. In addition, let $\boldsymbol{U} = (U_1, \ldots, U_n)^\top$ be the latent vector and $\boldsymbol{u} = (u_1, \ldots, u_n)^\top$ their fixed values, such that $\boldsymbol{Y}^{(c)} = (\boldsymbol{Y} \vdots \boldsymbol{U})$ is the complete sample, with $\boldsymbol{y}^{(c)} = (\boldsymbol{y}, \boldsymbol{u})$ being the complete-data vector. Then, the complete-data log-likelihood function for $\boldsymbol{\theta}$ is given by

$$\ell^{(c)}(\boldsymbol{\theta}) = \ell^{(c)}(\boldsymbol{\theta}; \boldsymbol{y}^{(c)}) = c_2 - \frac{n}{2}\log(|\boldsymbol{\Psi}|) - \sum_{i=1}^{n} \frac{u_i^2}{2} \boldsymbol{\phi}_i^\top \boldsymbol{\Psi}^{-1} \boldsymbol{\phi}_i + \sum_{i=1}^{n}\sum_{j=1}^{m} \log(\xi_{ij}), \tag{2.11}$$

where c_2 is a constant independent of the parameter vector $\boldsymbol{\theta}$, and $\xi_{ij}, \boldsymbol{\phi}_i$ are as given in (2.9), for $i = 1, \ldots, n, j = 1, \ldots, m$. Now, from property (E1), we have

$$\begin{aligned} \boldsymbol{Y}_i | U_i = u_i &\overset{\text{ind}}{\sim} \text{log-BS}_m\left(\frac{1}{u_i^{1/2}}\alpha \mathbf{1}_{m\times 1}, \boldsymbol{\beta}^\top \boldsymbol{x}_i, \left[\frac{1}{u_i}\right]\boldsymbol{\Psi}\right), \\ U_i &\overset{\text{iid}}{\sim} \text{Gamma}\left(\frac{\nu}{2}, \frac{\nu}{2}\right), \quad i = 1, \ldots, n, \end{aligned} \tag{2.12}$$

where $\overset{\text{ind}}{\sim}$ denotes that the components of the random vector are independent, whereas $\overset{\text{iid}}{\sim}$ denotes IID RVs. Thus, based on (2.12) and properties (E2)-(E4),

$$\text{E}[U_i^2 | \boldsymbol{Y} = \boldsymbol{y}] = \frac{[\nu + m][\nu + m + 2]}{[\nu + \boldsymbol{\phi}_i^\top \boldsymbol{\Psi}^{-1} \boldsymbol{\phi}_i]^2}, \quad i = 1, \ldots, n. \tag{2.13}$$

Therefore, the expected value of $\ell^{(c)}(\boldsymbol{\theta}; \boldsymbol{Y}^{(c)})$ given from (2.11), conditional on $\boldsymbol{Y}_i = \boldsymbol{y}_i$, may be defined as a function of $\boldsymbol{\theta}$ by

$$Q(\boldsymbol{\theta}) = \mathrm{E}[\ell^{(c)}(\boldsymbol{\theta}; \boldsymbol{Y}^{(c)})|\boldsymbol{Y} = \boldsymbol{y}] = c_2 - \frac{n}{2}\log(|\boldsymbol{\Psi}|) - \frac{1}{2}\sum_{i=1}^{n} w_i \boldsymbol{\phi}_i^\top \boldsymbol{\Psi}^{-1}\boldsymbol{\phi}_i + \sum_{i=1}^{n}\sum_{j=1}^{m}\log(\xi_{ij}), \quad (2.14)$$

where $w_i = \bar{u}_i^2 = \mathrm{E}[U_i^2|\boldsymbol{Y}_i = \boldsymbol{y}_i]$ given in (2.13) acts as a weight function, such as in the univariate case, allowing this estimation procedure to be robust; see Paula et al. (2012).

Algorithm 3 describes an EM approach to estimate the parameters of the model given in (2.7) with t kernel, maximizing the expected value given in (2.14) with respect to $\boldsymbol{\theta}$. Because this expected value is function of $\boldsymbol{\theta}$, Algorithm 3 needs a starting value $\widehat{\boldsymbol{\theta}}^{(0)}$ to initiate the maximization. However, it reduces to finding a starting value for w_i given in (2.14) as

$$\widehat{w}_i^{(0)} = \mathrm{E}[U_i^2|\boldsymbol{Y}_i = \boldsymbol{y}_i; \widehat{\boldsymbol{\theta}}^{(0)}] = \frac{[\nu_{\mathrm{k}} + m][\nu_{\mathrm{k}} + m + 2]}{\left[\nu_{\mathrm{k}} + \widehat{\boldsymbol{\phi}}_i^{(0)\top}\{\widehat{\boldsymbol{\Psi}}^{(0)}\}^{-1}\widehat{\boldsymbol{\phi}}_i^{(0)}\right]^2}, \quad (2.15)$$

where ν_{k} is a value of ν to be fixed and $\widehat{\boldsymbol{\phi}}_i^{(0)}$ has elements

$$\widehat{\phi}_{ij}^{(0)} = \left[\frac{2}{\widehat{\alpha}^{(0)}}\right]\sinh\left(\frac{y_{ij} - \widehat{\mu}_{ij}^{(0)}}{2}\right),$$

which must be evaluated at initial values for α and μ. Then, an initial value for μ can be obtained as $\widehat{\mu}_{ij}^{(0)} = \widehat{\boldsymbol{\beta}}_j^{(0)\top}\boldsymbol{x}_i$, with $\widehat{\boldsymbol{\beta}}_j^{(0)} = \widetilde{\boldsymbol{\beta}}_j$ being computed from the corresponding vector of the ordinary least square estimate $\widetilde{\boldsymbol{\beta}} = (\boldsymbol{X}^\top\boldsymbol{X})^{-1}\boldsymbol{X}^\top\boldsymbol{Y}$. In addition, from the ML method for the univariate BS log-linear model, an initial value for α may be obtained as

$$\widehat{\alpha}^{(0)} = \frac{1}{m}\sum_{j=1}^{m}\widehat{\alpha}_j^{(0)} = \frac{1}{m}\sum_{j=1}^{m}\left[\frac{4}{n}\sum_{i=1}^{n}\sinh^2\left(\frac{y_{ij} - \widehat{\mu}_{ij}^{(0)}}{2}\right)\right]^{1/2}. \quad (2.16)$$

Furthermore, to compute the weight function given in (2.15), we need the matrix of starting values $\widehat{\boldsymbol{\Psi}}^{(0)} = (\widehat{\psi}_{kl}^{(0)})$. Based on Property (D1), expressions in (2.3) and (2.16), and analogously to Kundu et al. (2013), we have

$$\widehat{\psi}_{kl}^{(0)} = \left[\frac{4}{n}\right]\left[\widehat{\alpha}_k^{(0)}\,\widehat{\alpha}_l^{(0)}\right]^{-1}\sum_{i=1}^{n}\sinh\left(\frac{y_{ik} - \widehat{\mu}_{ik}^{(0)}}{2}\right)\sinh\left(\frac{y_{il} - \widehat{\mu}_{il}^{(0)}}{2}\right).$$

Note that if $U = 1$ in Algorithm 3, that is, the RV U used in the representation given in (2.12) is degenerate, then this algorithm also is useful for the multivariate BS regression model.

Since the EM algorithm is an iterative procedure, the function $Q(\boldsymbol{\theta})$ given in (2.14) to be maximized must be evaluated at a previous value to the rth iteration of $\boldsymbol{\theta}$, inducting the notation $Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(r-1)})$. Algorithm 3 must be iterated until reaching convergence, for example, when

$|\ell^{(c)}(\widehat{\boldsymbol{\theta}}^{\,(r)})-\ell^{(c)}(\widehat{\boldsymbol{\theta}}^{\,(r-1)})| < 10^{-5}$, where $\widehat{\boldsymbol{\theta}}^{\,(r)}$ is the current ML estimate of $\boldsymbol{\theta}$ and $\widehat{\boldsymbol{\theta}}^{\,(r-1)}$ its previous estimate, with $\ell^{(c)}$ being given in (2.11); see McLachlan and Krishnan (1997, pp. 21-23). Note that, in some cases, the EM algorithm does not admit an analytical solution in its E-step or M-step. Then, it becomes necessary to use iterative methods for the computation of the expectation or for the maximization. For variants of the EM algorithm based on approximations of its E-step or M-step, which preserve its convergence properties, see Meng and Rubin (1993). In our case, in the M-step of Algorithm 3, an iterative method for solving nonlinear optimization problems is needed. For instance, the Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-Newton method can be used; see Nocedal and Wright (1999), Lange (2000), Mittelhammer et al. (2000), Paula et al. (2012) and Leiva et al. (2014d). The BFGS method is implemented in the `R` software by the functions `optim` and `optimx`; see `www.R-project.org` and R-Team (2016).

Algorithm 3 EM approach for estimating the multivariate BS-t regression model parameters

E-step. Given $\widehat{\boldsymbol{\theta}}^{(r-1)}$, compute $Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(r-1)})$, for $r = 1, 2, \ldots$, which reduces to compute $\widehat{w}_i^{(r)}$ from (2.14);

M-step. Find $\widehat{\boldsymbol{\theta}}^{(r)} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(r-1)})$, for $r = 1, 2, \ldots$, where $Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(r-1)})$ is given in (2.14), leading to the iterative procedure:

$$
\begin{aligned}
\widehat{\alpha}^{(r)} &= \Big[\frac{4}{nm}\sum_{i=1}^{n}\widehat{w}_i^{(r-1)}\widehat{\boldsymbol{\varrho}}_i^{(r-1)\top}\big\{\widehat{\boldsymbol{\Psi}}^{(r-1)}\big\}^{-1}\widehat{\boldsymbol{\varrho}}_i^{(r-1)}\Big]^{1/2},\\
\widehat{\boldsymbol{\Upsilon}}^{(r)} &= \sum_{i=1}^{n}\widehat{w}_i^{(r-1)}\widehat{\boldsymbol{\phi}}_i^{(r-1)}\widehat{\boldsymbol{\phi}}_i^{(r-1)\top},\\
\widehat{\boldsymbol{\Psi}}^{(r)} &= \mathrm{diag}(1/\widehat{\alpha}^{(r)},\ldots,1/\widehat{\alpha}^{(r)})\,\widehat{\boldsymbol{\Upsilon}}^{(r)}\,\mathrm{diag}(1/\widehat{\alpha}^{(r)},\ldots,1/\widehat{\alpha}^{(r)}),\\
\mathbf{0} &= \frac{1}{2}\sum_{i=1}^{n}\boldsymbol{D}(\boldsymbol{X})\Big[\widehat{w}_i^{(r-1)}\boldsymbol{D}(\widehat{\boldsymbol{\xi}}_i^{(r-1)})\big\{\widehat{\boldsymbol{\Psi}}^{\,(r-1)}\big\}^{-1}\widehat{\boldsymbol{\phi}}_i^{\,(r-1)}-\boldsymbol{D}(\{\widehat{\boldsymbol{\xi}}_i^{\,(r-1)}\}^{-1})\widehat{\boldsymbol{\phi}}_i^{\,(r-1)}\Big],
\end{aligned}
$$

where $\widehat{\boldsymbol{\varrho}}_i^{(r-1)} = [\widehat{\alpha}^{(r-1)}/2]\widehat{\boldsymbol{\phi}}_i^{(r-1)}$, with $\widehat{\boldsymbol{\phi}}_i^{(r-1)} = (\widehat{\phi}_{i1}^{(r-1)},\ldots,\widehat{\phi}_{im}^{(r-1)})^\top$, whose elements are $\widehat{\phi}_{ij}^{(r-1)} = [2/\widehat{\alpha}^{(r-1)}]\sinh([y_{ij}-\widehat{\mu}_{ij}^{(r-1)}]/2)$, $\boldsymbol{D}(\widehat{\boldsymbol{\xi}}_i^{(r-1)}) = \mathrm{diag}(\widehat{\xi}_{i1}^{(r-1)},\ldots,\widehat{\xi}_{im}^{(r-1)})$, with $\widehat{\xi}_{ij}^{(r-1)} = [2/\widehat{\alpha}^{(r-1)}]\cosh([y_{ij}-\widehat{\mu}_{ij}^{(r-1)}]/2)$, $\boldsymbol{D}(\{\widehat{\boldsymbol{\xi}}_i^{\,(r-1)}\}^{-1}) = \mathrm{diag}(1/\widehat{\xi}_{i1}^{(r-1)},\ldots,1/\widehat{\xi}_{im}^{(r-1)})$, for $i = 1,\ldots,n, j = 1,\ldots,m$, and $\boldsymbol{D}(\boldsymbol{X})$ is a block diagonal matrix with elements $\boldsymbol{x}_i^\top$.

2.5 Simulation study

We evaluate the performance of the ML estimators with simulations based on the MC method. For this simulation study, we consider the multivariate log-linear regression model

$$\begin{aligned} \boldsymbol{Y}_i &= \boldsymbol{\mu}_i + \boldsymbol{\varepsilon}_i \\ &= \boldsymbol{\beta}^\top \boldsymbol{x}_i + \boldsymbol{\varepsilon}_i \\ &= \begin{pmatrix} \beta_{11} & \beta_{12} & \beta_{13} \\ \beta_{21} & \beta_{22} & \beta_{23} \end{pmatrix}^\top \begin{pmatrix} 1 \\ x_{i2} \end{pmatrix} + \boldsymbol{\varepsilon}_i, \quad i = 1, \ldots, n, \end{aligned} \tag{2.17}$$

where $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \varepsilon_{i2}, \varepsilon_{i3}) \sim \text{log-GBS}_3(\alpha \mathbf{1}_{3\times 1}, \mathbf{0}_{3\times 1}, \boldsymbol{\Psi}_{3\times 3}, g^{(3)})$, with $g^{(3)}$ being the 3-variate normal or t kernel. Note that $\boldsymbol{\mu}_i = (\mu_{ij})$, where $\mu_{ij} = \beta_{1j} + \beta_{2j} x_{i2}$.

The true parameter values are

$$\alpha = 0.5, \quad \boldsymbol{\beta} = \begin{pmatrix} 5.0 & 1.0 & 1.0 \\ 4.0 & 2.0 & 0.5 \end{pmatrix}, \quad \boldsymbol{\Psi} = \begin{pmatrix} 1.0 & 0.8 & 0.3 \\ 0.8 & 1.0 & 0.5 \\ 0.3 & 0.5 & 1.0 \end{pmatrix},$$

whereas the sample size $n \in \{20, 50, 100\}$. The number of MC replications is 5,000. In each of these replications, we generate the matrix of observations $\boldsymbol{y} = (y_{ij})$, for $i = 1, \ldots, n$ and $j = 1, 2, 3$. Specifically, we use (2.17), algorithms 1 and 2 and the representations:
(i) $y_{ij} = \beta_{1j} + \beta_{2j} x_{i2} + 2 \arcsin(\alpha w_{ij}/2)$ given from (1.5) for the normal case, with w_{ij} being a value generated from $\boldsymbol{W} = (W_1, W_2, W_3) \sim \text{N}_3(\mathbf{0}, \boldsymbol{\Psi})$; or
(ii) $y_{ij} = \beta_{1j} + \beta_{2j} x_{i2} + 2 \arcsin(\alpha\, z_{ij}/[2\, u_i^{1/2}])$ given from (1.8) for the t case, with z_{ij} being a value generated from $\boldsymbol{Z} = (Z_1, Z_2, Z_3)^\top \sim \text{N}_3(\mathbf{0}, \boldsymbol{\Psi})$ and u_i a value generated from $U \sim \text{Gamma}(\nu/2, \nu/2)$, with $\nu = 4$.
In addition, the values of the regressor x_{i2} are obtained from a uniform distribution in the interval $(0, 1)$. Then, we fit the model given in (2.17) using Algorithm 3, which we have implemented in `R` code. For each parameter and sample size, we report the empirical mean, relative bias (RB) in absolute value, and root of the mean squared error (RMSE) of the ML estimators in Tables 2.1-2.2, for BS_3 and BS-t_3 log-linear regression models. In this table, observe that the ML estimators of the parameters α, $\boldsymbol{\beta}$, and $\boldsymbol{\Psi}$ present good statistical properties as n increases, that is, the RB and RMSE values decrease as the sample size increases, as expected, for the true values assumed for these parameters. In addition, note that the estimators of α and $\boldsymbol{\beta}$ are more efficient under the BS-t_3 model for all sample sizes considered, but the estimator of $\boldsymbol{\Psi}$ is slightly more efficient under the BS_3 model.

Table 2.1: Empirical mean, RB, and RMSE for the indicated estimator, n, and model with simulated data.

	n					
	20	50	100	20	50	100
	log-BS_3 model					
	$\widehat{\alpha}$			$\widehat{\beta}_{01}$		
True value	0.5000	0.5000	0.5000	5.0000	5.0000	5.0000
Mean	0.4544	0.4869	0.4976	4.9795	4.9916	4.9964
RB	0.0912	0.0262	0.0005	0.0041	0.0017	0.0007
RMSE	0.0878	0.0692	0.0599	0.2958	0.1734	0.0893
	$\widehat{\beta}_{02}$			$\widehat{\beta}_{03}$		
True value	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Mean	0.9834	0.9926	0.9952	0.9823	0.9875	0.9926
RB	0.0166	0.0074	0.0048	0.0177	0.0125	0.0074
RMSE	0.2906	0.1738	0.0886	0.2962	0.1716	0.0908
	$\widehat{\beta}_{11}$			$\widehat{\beta}_{12}$		
True value	4.0000	4.0000	4.0000	2.0000	2.0000	2.0000
Mean	4.0598	4.0400	4.0346	2.0562	2.0420	2.0392
RB	0.0149	0.0100	0.0087	0.0281	0.0210	0.0196
RMSE	0.5090	0.2991	0.1607	0.5012	0.3010	0.2110
	$\widehat{\beta}_{13}$			$\widehat{\psi}_{12}$		
True value	0.5000	0.5000	0.5000	0.8000	0.8000	0.8000
Mean	0.5563	0.5463	0.5425	0.7873	0.7925	0.7960
RB	0.1126	0.0927	0.0851	0.0158	0.0094	0.0051
RMSE	0.5155	0.2947	0.2114	0.0936	0.0562	0.0287
	$\widehat{\psi}_{13}$			$\widehat{\psi}_{23}$		
True value	0.3000	0.3000	0.3000	0.5000	0.5000	0.5000
Mean	0.2970	0.3036	0.3060	0.4978	0.5054	0.5080
RB	0.0100	0.0119	0.0201	0.0044	0.0108	0.0161
RMSE	0.2100	0.1293	0.0928	0.1771	0.1070	0.0761

Source: From the author.

Table 2.2: Empirical mean, RB, and RMSE for the indicated estimator, n, and model with simulated data.

	n					
	20	50	100	20	50	100
	log-BS-t_3 model					
	$\widehat{\alpha}$			$\widehat{\beta}_{01}$		
True value	0.5000	0.5000	0.5000	5.0000	5.0000	5.0000
Mean	0.4521	0.4910	0.4925	5.0073	5.0062	5.0068
RB	0.0958	0.0179	0.0150	0.0015	0.0012	0.0014
RMSE	0.0794	0.0522	0.0382	0.0429	0.0284	0.0233
	$\widehat{\beta}_{02}$			$\widehat{\beta}_{03}$		
True value	1.0000	1.0000	1.0000	1.0000	1.0000	1.0000
Mean	1.0072	1.0060	1.0010	1.0056	1.0040	1.0050
RB	0.0072	0.0060	0.0010	0.0056	0.0040	0.0050
RMSE	0.0427	0.0286	0.0229	0.0426	0.0277	0.0235
	$\widehat{\beta}_{11}$			$\widehat{\beta}_{12}$		
True value	4.0000	4.0000	4.0000	2.0000	2.0000	2.0000
Mean	3.9843	3.9893	3.9946	1.9840	1.9876	1.9898
RB	0.0039	0.0027	0.0013	0.0080	0.0062	0.0051
RMSE	0.0753	0.0571	0.0380	0.0751	0.0406	0.0400
	$\widehat{\beta}_{13}$			$\widehat{\psi}_{12}$		
True value	0.5000	0.5000	0.5000	0.8000	0.8000	0.8000
Mean	0.4917	0.4988	0.4919	0.7777	0.7956	0.7959
RB	0.0166	0.0023	0.0162	0.0279	0.0056	0.0051
RMSE	0.0761	0.0547	0.0398	0.1008	0.0583	0.0406
	$\widehat{\psi}_{13}$			$\widehat{\psi}_{23}$		
True value	0.3000	0.3000	0.3000	0.5000	0.5000	0.5000
Mean	0.2723	0.2905	0.2945	0.4823	0.4972	0.4979
RB	0.0924	0.0315	0.0183	0.0354	0.0057	0.0041
RMSE	0.2184	0.1385	0.1016	0.1831	0.1120	0.0842

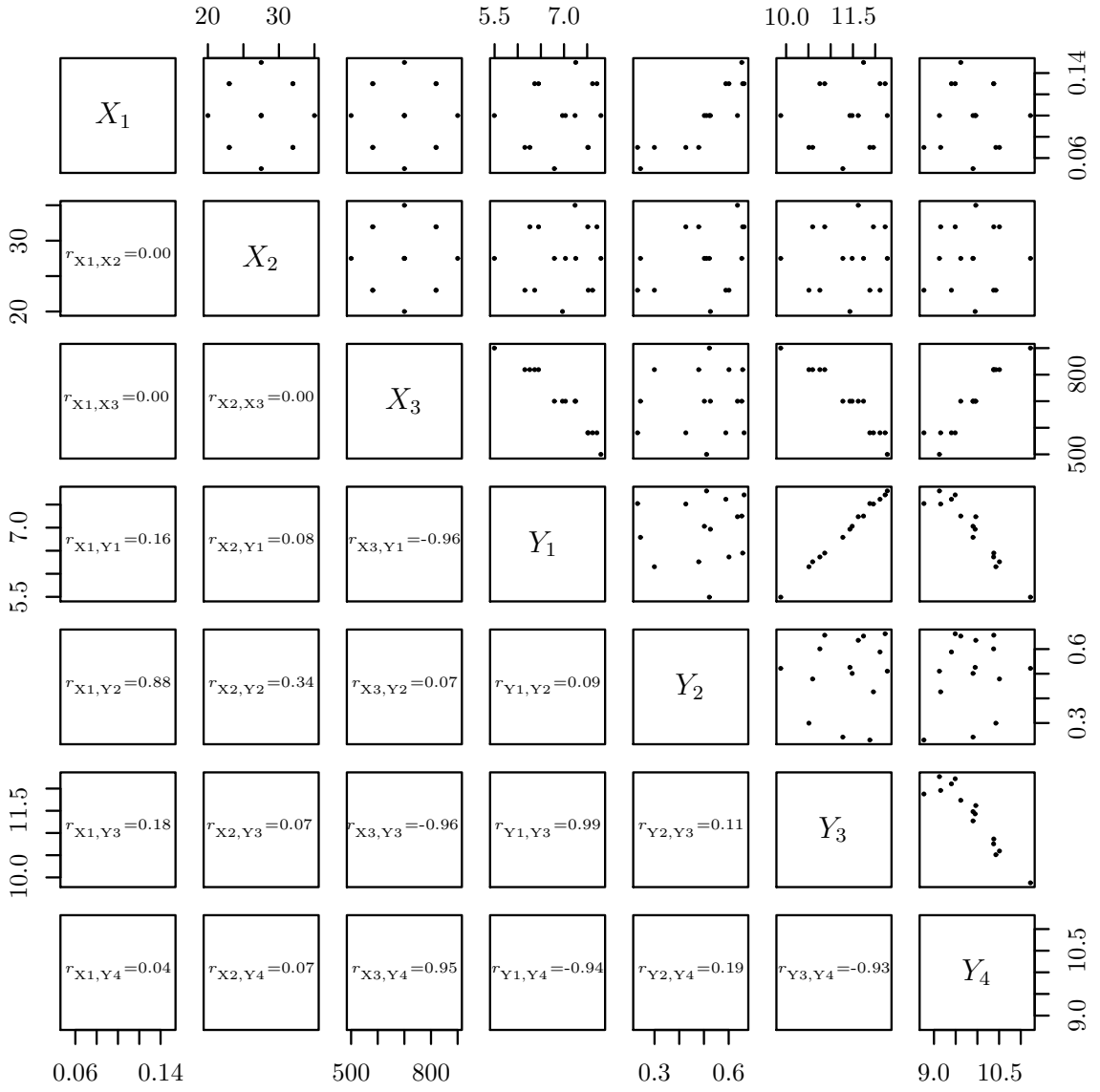
Source: From the author.

2.6 Data analysis

Die fracture is a typical metal fatigue caused by cyclic stress in the course of the service life cycle of dies (die lifetime). Although this fatigue could be mainly determined by die lifetime, other RVs can also be considered as responses to this fatigue. The objective of the present data analysis is to model fatigue in a metal forming process. We consider as responses to:
(i) Von Mises stress (T_1, in Newton/millimeter^2–N/mm^2–),
(ii) Maximum deformation (T_2, dimensionless),
(iii) Manufacturing force (T_3, in Newton –N–), and
(iv) Die lifetime (T_4, in number of cycles).
The regressors that could affect these responses are:
(v) Friction coefficient (X_1, dimensionless),
(vi) Angle of die (X_2, in $^{\circ}$), and
(vii) Work temperature (X_3, in $^{\circ}$C).
We illustrate our multivariate models with real-world fatigue data associated with these variables taken from Lepadatu et al. (2005); see data in Table 6.1 of Appendix 6.3. The main advantage of a multivariate regression model over marginal regressions is that it takes into account the correlation between the responses. As mentioned, several responses related to metal fatigue could be interacting in the process cycle. Then, this interaction must be studied by means of a correlation statistical analysis. It must provide information about if correlation must be considered in the modeling through a multivariate model. Otherwise, several marginal models, one for each response, can be considered. However, analyzing these variables individually, if statistical correlations between them exist, may yield wrong prediction. For more details about this issue, see Lepadatu et al. (2005).

First, we explore the data computing (linear) correlations between each pair from the set $\{T_1, T_2, T_3, T_4, X_1, X_2, X_3\}$ to justify the use of multivariate models, to discard any collinearity among X_1, X_2, X_3, and to propose log-linear regression models. Note that some log-linear relationships are detected between some responses T_1, T_2, T_3, T_4 and regressors X_1, X_2, X_3. Therefore, linear relationships can be considered between some log-responses $Y_j = \log(T_j)$ and X_1, X_2, X_3, for $j = 1, 2, 3, 4$. Figure 2.13 displays the scatter-plots for all the log-responses and regressors. From this figure we detect that:
(i) No correlations exist between each pair from the set $\{X_1, X_2, X_3\}$, discarding any possible collinearity problem in our model;
(ii) Large and reasonable correlations between each pair from the set $\{Y_1, Y_2, Y_3, Y_4\}$, justifying the use of multivariate distributions;

Figure 2.13: Scatter-plots with their corresponding correlations for the indicated variables.



Source: From the author.

(iii) Large, reasonable and small correlations between (X_1, Y_2), $\{(X_1, Y_1), (X_1, Y_3)\}$, and (X_1, Y_4), respectively;
(iv) Reasonable and small correlations between (X_2, Y_2) and $\{(X_2, Y_1), (X_2, Y_3), (X_2, Y_4)\}$, respectively;

(v) Large and small correlations between $\{(X_3, Y_1), (X_3, Y_3), (X_3, Y_4)\}$ and (X_3, Y_2), respectively; all which support the use of a log-linear regression model. This should be confirmed by the inferential analysis.

Second, based on this exploratory data analysis, arguing the fatigue principles of the BS distribution, and considering the robust estimation in BS-t models, to describe $\{T_1, T_2, T_3, T_4\}$ in function of $\{X_1, X_3\}$ (because X_2 is discarded due to its small correlations with the responses), we propose the multivariate log-linear regression model

$$\boldsymbol{Y}_i = \boldsymbol{\beta}^\top \boldsymbol{x}_i + \boldsymbol{\varepsilon}_i, \quad i = 1, \ldots, n,$$

where $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \varepsilon_{i2}, \varepsilon_{i3}, \varepsilon_{i4})^\top \sim \text{log-GBS}_4(\alpha \mathbf{1}_{4\times 1}, \mathbf{0}_{4\times 1}, \boldsymbol{\Psi}_{4\times 4}, g^{(4)})$, with $g^{(4)}$ being the 4-variate normal or t kernel. We estimate the parameters of the multivariate BS and BS-t log-linear regression models via the EM approach described in Algorithm 3. Starting values, $\widehat{\boldsymbol{\theta}}^{(0)}$ say, used in the maximization procedure are $\widehat{\alpha}^{(0)} = 0.13996$,

$$\widetilde{\boldsymbol{\beta}} = \begin{pmatrix} 10.4663 & 0.0075 & 15.0101 & 6.1768 \\ 3.5917 & 4.3166 & 4.3798 & 0.7771 \\ -0.0055 & 0.0001 & -0.0058 & 0.0052 \end{pmatrix},$$

$$\widehat{\boldsymbol{\Psi}}^{(0)} = \begin{pmatrix} 1.0000 & 0.2416 & 0.9628 & -0.4843 \\ 0.2416 & 1.0000 & 0.2427 & 0.5425 \\ 0.9628 & 0.2427 & 1.0000 & -0.3760 \\ -0.4843 & 0.5425 & -0.3760 & 1.0000 \end{pmatrix}$$

whereas $\nu_{\text{k}} = 4$; the interested reader is referred to Barros et al. (2008) and Paula et al. (2012) and references therein for a justification about this value. We have verified that $\nu_{\text{k}} = 4$ corresponds to the value that maximizes the log-likelihood function within a range of values for ν.

As mentioned, in the iterations of M-step, we maximize $Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(r-1)})$ with the BFGS procedure, because a non-linear optimization problem is presented. Note that, when comparing the multivariate BS and BS-t log-linear regression models, the maximized log-likelihood value is equivalent to using an information criteria (such as Akaike or Schwartz), because the value of ν is assumed to be known and, therefore, both models (BS and BS-t) have the same number of parameters. For the data under analysis, the maximum values of the log-likelihood function are 120.297 and 115.761 for the BS_4 and BS-t_4 models, respectively, indicating, according this criterion, that the BS_4 log-linear regression model is more appropriate to describe these data.

GBS_m log-linear regression model checking can be conducted using the MD, given by $\text{MD}_i(\boldsymbol{\theta}) = \boldsymbol{\phi}_i^\top \boldsymbol{\Psi}^{-1} \boldsymbol{\phi}_i$, for $i = 1, \ldots, n$, where $\boldsymbol{\phi}_i$ is defined in (2.9). Based on property (D2)

and (2.5), note that this distance follows the $\chi^2(m=4)$ or $4\mathcal{F}(m=4,\nu=4)$ distribution if $g^{(4)}$ is the 4-variate normal or t kernel, respectively. We plug the ML estimator $\widehat{\boldsymbol{\theta}}$ in $\mathrm{MD}_i(\boldsymbol{\theta})$, obtaining $\mathrm{MD}_i(\widehat{\boldsymbol{\theta}})$, which has asymptotically the same distribution of $\mathrm{MD}_i(\boldsymbol{\theta})$; see Lange et al. (1989). We use the Wilson-Hilferty (WH) approximation for transforming to normality this distance; see Ibacache-Pulgar et al. (2014) and references therein. The WH approximation, based on $\mathrm{MD}_i(\hat{\boldsymbol{\theta}})$, for normal kernel is given by

$$\vartheta_i^N = \frac{\left[\left(\frac{\mathrm{MD}_{M_i}(\hat{\boldsymbol{\theta}})}{m}\right)^{1/3} - \left(1-\frac{2}{9m}\right)\right]}{(\frac{2}{9})^{1/2}} \dot{\sim} \mathrm{N}(0,1), \quad i=1,\ldots,n, \tag{2.18}$$

where $\dot{\sim}$ means "asymptotically distributed". For t kernel the WH approximation is given by

$$\vartheta_i^t = \frac{\left[\left(1-\frac{2}{9m}\right)\left(\frac{\mathrm{MD}_i(\hat{\boldsymbol{\theta}})}{m}\right)^{1/3} - \left(1-\frac{2}{9m}\right)\right]}{\left[\left(\frac{2}{9m}\right)\left(\frac{\mathrm{MD}_i(\hat{\boldsymbol{\theta}})}{m}\right)^{2/3} + \left(\frac{2}{9m}\right)\right]^{1/2}} \dot{\sim} \mathrm{N}(0,1), \quad i=1,\ldots,n. \tag{2.19}$$

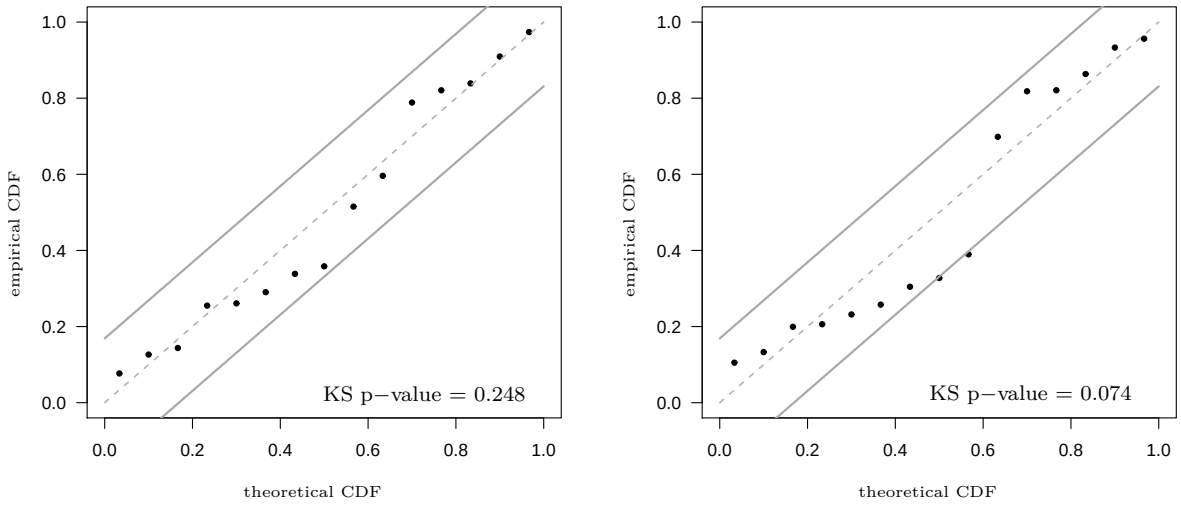
Then, we check normality of the transformed distances with the WH approximation using goodness-of-fit (GOF) techniques. Algorithm 4 details how to construct probability-probability (PP) plots and compute their acceptance bands associated with the Kolmogorov-Smirnov (KS) statistic for testing normality; for more details, see Barros et al. (2014) and Castro-Kuriss et al. (2014).

Algorithm 4 PP-plots with acceptance bands for testing normality.

1: Consider the data $y_1,\ldots,y_n$ and order them as $y_{1:n},\ldots,y_{n:n}$.
2: Estimate the mean and standard deviation (SD) of the normal distribution by using $\bar{y}=\sum_{j=1}^n y_j/n$ and $s_y=[\sum_{j=1}^n\{\widehat{y}_j-\bar{y}\}^2/\{n-1\}]^{1/2}$, respectively.
3: Compute $o_{j:n}=\Phi([y_{j:n}-\bar{y}]/s_y)$, for $j=1,\ldots,n$, where Φ is the N(0, 1) CDF.
4: Draw the PP-plot with points $q_{j:n}=[2j-1]/[2n]$ versus $o_{j:n}$, for $j=1,\ldots,n$.
5: Specify a significance level $1-\eta$.
6: Construct acceptance bands according to $(\max\{q-d_\eta+1/[2n],0\},\min\{q+d_\eta-1/[2n],1\})$, where d_η is the $100\times\eta$th percentile of the KS distribution and q is a continuous version of $q_{j:n}$.
7: Determine the p-value of the KS statistic and reject the null hypothesis of a normal distribution for the specified significance level based on this p-value.
8: Corroborate coherence between steps 6 and 7.

Figure 2.14 shows PP-plots with acceptance bands of a 10% significance level for MD transformed to normality with the WH approximation by using the fatigue data under analysis. From this figure, we detect that the BS_4 model provides a better fit than the BS-t_4 model, which is corroborated by the p-values 0.248 and 0.074 for both models, respectively, of the KS test associated with the PP-plot. Note that some points are outside the bands in the plot for the BS-t_4 model. Therefore, we can conclude that the BS_4 log-linear regression model provided a better fit to our data.

Figure 2.14: PP-plot with KS acceptance bands at 10% for transformed MDs obtained from metal fatigue data under BS_4 and BS-t_4 models.



(a) BS_4 (b) BS-t_4

Source: From the author.

Table 2.3 contains the model parameter estimates, the estimated SEs of the corresponding ML estimators for both models, and p-values of each t-test. As usual, we use the root of the diagonal elements of the observed Fisher information inverse matrix to approximate the corresponding estimated SEs; see Efron and Hinkley (1978) and Paula et al. (2012). From this table, and the at 1% significance level, we arrive at the following conclusions:
(i) Most of the estimated correlations from the BS_4 and BS-t_4 models are statistically significant, thus corroborating our conjectures from the exploratory analysis.
(ii) The coefficients β_1 and β_2 must be discarded in the prediction of T_2, but both of them must be considered for T_1, T_3, and T_4. In addition, β_3 must also be discarded in the predic-

tion of T_4, but it must be considered for T_1, T_2, and T_3. Thus, we obtain a practical model to predict metal forming processes simultaneously describing T_1, T_2, T_3, T_4 as a function of X_1, X_3.

(iii) We can use the estimated regression coefficients to find the ML estimates of the scale parameters (medians) λ_{ij} of the BS distribution, which along with the ML estimate of its shape parameter α can be useful to determine the reliability of the metal forming process or to evaluate its failure rate.

Table 2.3: ML estimate of the indicated parameter and model with its estimated SE and p-value for metal fatigue data.

	BS_4 log-linear model			BS-t_4 log-linear model		
Parameter	Estimate	SE	p-value	Estimate	SE	p-value
ψ_{12}	0.21256	0.29848	0.47637	0.29631	0.39670	0.45511
ψ_{13}	0.94767	0.06221	< 0.0001	0.96941	0.13193	< 0.0001
ψ_{14}	-0.28525	0.09196	0.00192	-0.24120	0.08522	0.00465
ψ_{23}	0.25966	0.11558	0.02467	0.39877	0.11283	0.00041
ψ_{24}	0.47697	0.13588	0.00045	0.21981	0.33568	0.51259
ψ_{34}	-0.39672	0.13703	0.00379	-0.31354	0.12867	0.00391
β_{11}	10.59099	0.08575	< 0.0001	10.43469	0.08388	< 0.0001
β_{12}	0.02878	0.11467	0.80184	0.03759	0.06513	0.56388
β_{13}	15.12231	0.09023	< 0.0001	15.03395	0.13166	< 0.0001
β_{14}	6.09211	0.27178	< 0.0001	6.14291	0.15007	< 0.0001
β_{21}	-0.00572	0.00001	< 0.0001	-0.00550	0.00005	< 0.0001
β_{22}	0.00006	0.00012	0.58437	0.00005	0.00005	0.31731
β_{23}	-0.00598	0.00003	< 0.0001	-0.00586	0.00005	< 0.0001
β_{24}	0.00529	0.00027	< 0.0001	0.00522	0.00016	< 0.0001
β_{31}	3.59670	0.89320	< 0.0001	3.58761	0.77892	< 0.0001
β_{32}	4.29974	0.49856	< 0.0001	4.30641	0.58762	< 0.0001
β_{33}	4.37982	0.90850	< 0.0001	4.38079	0.71917	< 0.0001
β_{34}	0.77876	1.20388	0.51771	0.77169	1.08114	0.47537
α	0.14207	0.01006	–	0.14000	0.03914	–

Source: From the author.

2.7 Concluding remarks

In this chapter, we have proposed new multivariate generalized Birnbaum-Saunders regression models to describe metal fatigue data. Specifically, we have derived a new multivariate logarithmic generalized Birnbaum-Saunders distribution. Then, based on it, we have formulated multivariate generalized Birnbaum-Saunders regression models, which can be useful for predicting metal forming processes in practice considering, for example, die lifetime, manufacturing force, deformation, and stress, as a function of the friction, angle die, and work temperature. In addition, our models can be used for determining the reliability of the metal forming process, or its failure rate. We have provided a parameter estimation procedure based on the maximum likelihood method with an efficient computation expectation-maximization algorithm. These models were applied to real-world multivariate metal fatigue data. Goodness-of-fit tests concluded that they are appropriate to describe these data.

CHAPTER 3

Diagnostics in multivariate Birnbaum-Saunders regression models

3.1 Resumo

Neste capítulo, propomos uma metodologia baseada em técnicas de diagnóstico para modelos de regressão multivariados Birnbaum-Saunders generalizados. Consideramos a distância de Mahalanobis como medida de influência global de detecção de outliers multivariados e para avaliar a adequacidade do modelo. Desenvolvemos medidas de diagnóstico baseado em influência local sob alguns esquemas de perturbação, a saber: ponderação de casos, perturbação da matriz de correlação, das variáveis explicativas e da resposta. Além disso, obtemos o método de alavanca generalizado. Desenvolvemos rotinas computacionais para implementar a metodologia no software `R`. Aplicações com dados de fadiga foram apresentadas para ilustrar a metodologia proposta. Especificamente, analisamos dados de fadiga de biomateriais.

3.2 Introduction

Rieck and Nedelman (1991) were the pioneers in proposing BS regression models and Galea et al. (2004) in deriving diagnostic analyses for these models. BS regression models

are often based on the log-BS distribution; see Rieck and Nedelman (1991), Santana et al. (2011) and Leiva (2016). Marchant et al. (2016a) developed multivariate log-GBS distributions, which are generated from EC kernels, such as normal and t, being the last one more general than the normal kernel and widely flexible to deal with data following heavy-tailed distributions. Then, the ML estimation procedure based on the multivariate log-BS-t distribution is non-sensitive to atypical observations. In addition, Marchant et al. (2016a) derived multivariate GBS log-linear regression models for describing fatigue data.

Diagnostics should be addressed in all statistical modeling to assess its suitability and stability. Diagnostics can be conducted by GOF techniques and global/local influence methods. Goodness of fit allows us to assess the adequacy of a model to a data set; see Barros et al. (2014). Global influence eliminates cases and evaluates their effect on the fitted model; see Cook and Weisberg (1982). Local influence allows us to detect the effect of perturbations on the estimates of parameters; see Cook (1987). For the use of GOF and diagnostic methods in non-normal models, see Billor and Loynes (1999), Díaz-García et al. (2003), Osorio et al. (2007), Atkinson (2009), Barros et al. (2010), Villegas et al. (2011), Uribe-Opazo et al. (2012), Assumpção et al. (2014), Leiva et al. (2014b,c,d, 2015a), and Liu et al. (2016).

The main goal of this chapter is to carry out diagnostics in multivariate GBS log-linear regression models by (i) GOF techniques; (ii) global and local influence methods; and (iii) applying them to real-world data. To meet these objectives, we estimate the parameters of multivariate BS and BS-t models with the ML method. Then, we derive fitting and diagnostic tools to evaluate the adequacy and stability of these multivariate models from three perspectives. First, we use the MD to test goodness of fit in the distributional assumption employed for the proposed multivariate models. Second, we consider the MD as a global influence method to identify multivariate outliers. Third, we derive the total local influence method assuming perturbations of (i) case-weight, (ii) correlation matrix, (iii) response, and (iv) a continuous explanatory variable. In addition, we obtain the generalized leverage (GL) for detecting effect of the observed response on its own estimated value. We implement the obtained results by a computational routine in the `R` statistical software; see `www.R-project.org` and R-Team (2016). We use this routine for carrying out an illustration with real-world multivariate data collected by the authors. Specifically, we consider multivariate data useful for regression models based on computed tomography (CT) to study the bone quality as discussed in the application section. Note that the BS distribution is appropriate for modeling physical properties of bone, such as its densities extracted from CT data; see Vivanco et al. (2014). These properties are related to mechanical properties, as strength, which plays a role as human bone ages affecting its fatigue properties that can

be described by the BS distribution.

This chapter is organized as follows. In Section 3.3, we mention the multivariate GBS log-linear regression models proposed by Marchant et al. (2016a) and derive the corresponding MD, the score vector, robustness aspects, the Hessian matrix, the observed information matrix and inferential aspects. In Section 3.4, we derive diagnostics for multivariate GBS log-linear regression models considering local influence, as well as global influence by the MD. In Section 3.5, we apply these tools to real-world multivariate biomaterials data. Finally, in Section 3.6, we discuss some conclusions and future studies related to the topic of this chapter.

3.3 Modelling

In this section, we mention the regression model, the associated MD, the score vector, robustness aspects, the information matrix, and asymptotic inference.

3.3.1 Formulation

Consider the multivariate BS log-linear regression model proposed by Marchant et al. (2016a) and given in (2.7). From (2.9) and considering $\ell(\boldsymbol{\theta};\boldsymbol{y}) = \sum_{i=1}^{n} \ell_i(\boldsymbol{\theta})$, if $g^{(m)}$ is the multivariate normal or t kernel, then the log-likelihood functions for $\boldsymbol{\theta}$ are, respectively, given by

$$\begin{aligned}
\ell_i(\boldsymbol{\theta}) &= -m\log(2) - \frac{m}{2}\log(2\pi) - \frac{1}{2}\log(|\boldsymbol{\Psi}|) - \frac{1}{2}\boldsymbol{\phi}_i^{\top}\boldsymbol{\Psi}^{-1}\boldsymbol{\phi}_i + \sum_{j=1}^{m}\log(\xi_{ij}), \qquad (3.1)\\
\ell_i(\boldsymbol{\theta}) &= -m\log(2) - \log\left(\Gamma\left(\frac{\nu}{2}\right)\right) + \log\left(\Gamma\left(\frac{\nu+m}{2}\right)\right) - \frac{m}{2}\log(\nu\pi) + \left[\frac{\nu+m}{2}\right]\log(\nu)\\
&\quad -\frac{1}{2}\log(|\boldsymbol{\Psi}|) - \left[\frac{\nu+m}{2}\right]\log(\nu + \boldsymbol{\phi}_i^{\top}\boldsymbol{\Psi}^{-1}\boldsymbol{\phi}_i) + \sum_{j=1}^{m}\log(\xi_{ij}),
\end{aligned}$$

where ξ_{ij} and $\boldsymbol{\phi}_i$ were defined in (2.9).

3.3.2 Mahalanobis distance

As mentioned in Section 2.6, from property (D2) and (2.5), we can obtain the MD for multivariate BS log-linear regression model from the following quadratic form

$$\mathrm{MD}_i = \boldsymbol{\phi}_i^{\top}\boldsymbol{\Psi}^{-1}\boldsymbol{\phi}_i, \quad i = 1,\ldots,n, \qquad (3.2)$$

which is useful, as mentioned, for detecting outliers in multivariate regression models and testing goodness of fit in these models.

Multivariate log-GBS distributions are generated from EC kernels, say $g^{(m)}$. In this context, a result of interest is

$$\zeta(u) = \frac{g^{(m)'}(u)}{g^{(m)}(u)}, \quad u > 0, \tag{3.3}$$

where $g^{(m)'}(u)$ is the derivative of $g^{(m)}(u)$ with respect to u. If $g^{(m)}$ is continuous and decreasing, its maximum, say u_g, exists and is finite and positive. Moreover, if $g^{(m)}$ is continuous and differentiable, then u_g is the solution to $\zeta(u)+m/(2u) = 0$, where $\zeta(u)$ is given in (3.3). In several of the EC distributions, the kernel $g^{(m)}$ depends on an additional shape parameter, which we denote by ν and allows us to control the kurtosis of the distribution. It is known that $u_g = m$ for both normal and t kernels. Therefore, for the normal and t kernels, we have, respectively,

$$\zeta(u) = -\frac{1}{2}, \zeta'(u) = 0 \quad \text{(normal case)}, \quad \zeta(u) = -\frac{\nu+m}{2[\nu+u]}, \zeta'(u) = \frac{\nu+m}{2[\nu+u]^2} \quad (t \text{ case}).$$

3.3.3 Score vector and robustness

Consider the log-likelihood function for $\boldsymbol{\theta}$ given in (3.1) and $\boldsymbol{\Psi} = \boldsymbol{\Psi}(\boldsymbol{\rho})$, where $\boldsymbol{\rho} = \text{svec}(\boldsymbol{\Psi})^\top = (\rho_1, \ldots, \rho_l)^\top$, with $l = m[m-1]/2$. By taking the derivative of $\ell(\boldsymbol{\theta}; \boldsymbol{y})$ with respect to $\alpha, \boldsymbol{\beta}, \boldsymbol{\rho}$, we obtain the score vector for $\boldsymbol{\theta}$ given by $\dot{\boldsymbol{\ell}} = (\dot{\ell}_\alpha, \dot{\boldsymbol{\ell}}_{\boldsymbol{\beta}}^\top, \dot{\boldsymbol{\ell}}_{\boldsymbol{\rho}}^\top)^\top$, where

$$\begin{aligned}
\dot{\ell}_\alpha &= -\frac{2}{\alpha}\sum_{i=1}^{n} \zeta_i \,\text{MD}_i - \frac{m}{\alpha}, \\
\dot{\boldsymbol{\ell}}_{\boldsymbol{\beta}} &= -\sum_{i=1}^{n} \boldsymbol{D}(\boldsymbol{X})\Big[\zeta_i \boldsymbol{D}(\boldsymbol{\xi}_i)\boldsymbol{\Psi}^{-1}(\boldsymbol{\rho}) + \frac{1}{2}\boldsymbol{D}(\boldsymbol{\xi}_i^{-1})\Big]\boldsymbol{\phi}_i, \\
\dot{\boldsymbol{\ell}}_{\boldsymbol{\rho}} &= (\ell_{\rho_1}, \ldots, \dot{\ell}_{\rho_l})^\top,
\end{aligned} \tag{3.4}$$

with $\zeta_i = \zeta(\text{MD}_i)$, ζ given in (3.3),

$$\dot{\ell}_{\rho_k} = -\frac{n}{2}\text{tr}\left(\boldsymbol{\Psi}^{-1}(\boldsymbol{\rho})\frac{\partial \boldsymbol{\Psi}(\boldsymbol{\rho})}{\partial \rho_k}\right) - \sum_{i=1}^{n} \zeta_i \boldsymbol{\phi}_i^\top \boldsymbol{\Psi}^{-1}(\boldsymbol{\rho})\frac{\partial \boldsymbol{\Psi}(\boldsymbol{\rho})}{\partial \rho_k}\boldsymbol{\Psi}^{-1}(\boldsymbol{\rho})\boldsymbol{\phi}_i, \quad k = 1, \ldots, l,$$

$\boldsymbol{\phi}_i$ as given in (2.9), $\boldsymbol{D}(\boldsymbol{\xi}_i) = \text{diag}(\xi_{i1}, \ldots, \xi_{im})$, ξ_{ij} given in (2.9), $\boldsymbol{D}(\boldsymbol{\xi}_i^{-1}) = \text{diag}(\xi_{i1}^{-1}, \ldots, \xi_{im}^{-1})$ and $\boldsymbol{D}(\boldsymbol{X})$ is a block diagonal matrix with elements $\boldsymbol{x}_i^\top$.

In order to obtain the ML estimates of the model parameters formulated in (2.8), we must set the elements of the score vector given in (3.4) equal to zero. To maximize the log-likelihood function given in (3.1), the corresponding likelihood equations must be solved by

an iterative procedure for non-linear optimization problems. Initial values for the iterative procedure can be obtained from:
(i) The ordinary least square estimate $\widehat{\boldsymbol{\beta}}^{(0)} = (\boldsymbol{X}^\top \boldsymbol{X})^{-1}\boldsymbol{X}^\top \boldsymbol{Y}$;
(ii)

$$\widehat{\alpha}^{(0)} = \frac{1}{m}\sum_{j=1}^{m}\left[\frac{4}{n}\sum_{i=1}^{n}\operatorname{sinh}^2\left(\frac{y_{ij}-\widehat{\mu}_{ij}^{(0)}}{2}\right)\right]^{1/2},$$

where $\widehat{\mu}_{ij}^{(0)} = \widehat{\boldsymbol{\beta}}_j^{(0)\top}\boldsymbol{x}_i$, with $\widehat{\boldsymbol{\beta}}_j^{(0)}$ being computed from (i); and
(iii) $\widehat{\boldsymbol{\Psi}}^{(0)} = \boldsymbol{D}(\widehat{\boldsymbol{\Sigma}}^{(0)})^{-1/2}\widehat{\boldsymbol{\Sigma}}^{(0)}\boldsymbol{D}(\widehat{\boldsymbol{\Sigma}}^{(0)})^{-1/2}$, where $\widehat{\boldsymbol{\Sigma}}^{(0)} = \frac{1}{n}\sum_{i=1}^{n}\boldsymbol{\phi}_i^{(0)}(\boldsymbol{\phi}_i^{(0)})^\top$, with $\widehat{\boldsymbol{\phi}}_i^{(0)}$ having elements

$$\widehat{\phi}_{ij}^{(0)} = \left[\frac{2}{\widehat{\alpha}^{(0)}}\right]\operatorname{sinh}\left(\frac{y_{ij}-\widehat{\mu}_{ij}^{(0)}}{2}\right), \quad i = 1,\ldots,n, \quad j = 1,\ldots,m.$$

Note that, for the log-BS-t distribution, as ν approaches ∞, one has $-2\zeta_i$ approaching to 1, for all $i = 1,\ldots,n$. Therefore, in this case the scores given in (3.4) reduce to the BS case, as expected; see Paula et al. (2012) for the univariate case. Thus, the quantity ζ_i may be interpreted as a kind of weight in the multivariate BS-t log-linear regression model. Hence, since this weight is inversely proportional to MD_i, if the case i has a large value for its MD, then it should have a small weight in the ML estimation procedure. Thus, this procedure assigns less weight to outlying observations in the sense of the MD given in (3.2). In addition, note that the parameter ν of the multivariate log-BS-t distribution can be assumed to be fixed or known from the data due to a statistical robustness aspect; see Lucas (1997) for the t distribution, and Paula et al. (2012) for the univariate BS-t case.

3.3.4 Information matrix and asymptotic inference

The observed information matrix is defined by $\boldsymbol{I}(\boldsymbol{\theta}) = -\ddot{\boldsymbol{\ell}}$, where $\ddot{\boldsymbol{\ell}}$ is the Hessian matrix given by

$$\ddot{\boldsymbol{\ell}} = \frac{\partial^2 \ell}{\partial\boldsymbol{\theta}\partial\boldsymbol{\theta}^\top} = \begin{pmatrix} \ddot{\ell}_{\alpha\alpha} & \ddot{\boldsymbol{\ell}}_{\alpha\boldsymbol{\beta}} & \ddot{\boldsymbol{\ell}}_{\alpha\boldsymbol{\rho}} \\ & \ddot{\boldsymbol{\ell}}_{\boldsymbol{\beta}\boldsymbol{\beta}} & \ddot{\boldsymbol{\ell}}_{\boldsymbol{\beta}\boldsymbol{\rho}} \\ & & \ddot{\boldsymbol{\ell}}_{\boldsymbol{\rho}\boldsymbol{\rho}} \end{pmatrix}. \tag{3.5}$$

The Hessian matrix is obtained from the log-likelihood function given in (3.1) with elements

$$\begin{aligned}
\ddot{\ell}_{\alpha\alpha} &= \frac{1}{\alpha^2}\sum_{i=1}^{n}\left[\{6\zeta_i - 2\zeta_i'\mathrm{MD}_i\}\mathrm{MD}_i + m\right], \\
\ddot{\boldsymbol{\ell}}_{\alpha\boldsymbol{\beta}} &= \frac{2}{\alpha}\sum_{i=1}^{n}[\zeta_i + \zeta_i'\mathrm{MD}_i]\boldsymbol{D}(\boldsymbol{X})\boldsymbol{D}(\boldsymbol{\xi}_i)\boldsymbol{\Psi}^{-1}(\boldsymbol{\rho})\boldsymbol{\phi}_i, \\
\ddot{\boldsymbol{\ell}}_{\alpha\boldsymbol{\rho}} &= (\ddot{\ell}_{\alpha\rho_1}, \ldots, \ddot{\ell}_{\alpha\rho_l})^\top, \\
\ddot{\boldsymbol{\ell}}_{\boldsymbol{\beta}\boldsymbol{\rho}} &= (\ddot{\boldsymbol{\ell}}_{\boldsymbol{\beta}\rho_1}, \ldots, \ddot{\boldsymbol{\ell}}_{\boldsymbol{\beta}\rho_l})^\top, \\
\ddot{\boldsymbol{\ell}}_{\boldsymbol{\rho}\boldsymbol{\rho}} &= (\ddot{\ell}_{\rho_k\rho_s}) \in \mathbb{R}^{l\times l}, \\
\ddot{\boldsymbol{\ell}}_{\boldsymbol{\beta}\boldsymbol{\beta}} &= \frac{1}{2}\sum_{i=1}^{n}\boldsymbol{D}(\boldsymbol{X})\boldsymbol{D}(\boldsymbol{\xi}_i)\boldsymbol{\Psi}^{-1}(\boldsymbol{\rho})[\zeta_i'\boldsymbol{\phi}_i\boldsymbol{\phi}_i^\top\boldsymbol{\Psi}^{-1}(\boldsymbol{\rho}) + \zeta_i]\boldsymbol{D}(\boldsymbol{\xi}_i)\boldsymbol{D}(\boldsymbol{X}^\top) \\
&\quad + \frac{1}{4}\sum_{i=1}^{n}\boldsymbol{D}(\boldsymbol{X})[2\zeta_i\boldsymbol{D}(\boldsymbol{\phi}_i)\boldsymbol{D}(\boldsymbol{\Psi}^{-1}(\boldsymbol{\rho})\boldsymbol{\phi}_i) - \boldsymbol{D}(\boldsymbol{\xi}_i^{-2})D(\boldsymbol{\phi}_i^2) + \boldsymbol{I}_m]\boldsymbol{D}(\boldsymbol{X}^\top),
\end{aligned}$$

where, for $k = 1, \ldots, l$,

$$\begin{aligned}
\ddot{\ell}_{\alpha\rho_k} &= \frac{2}{\alpha}\sum_{i=1}^{n}[\zeta_i + \zeta_i'\mathrm{MD}_i]\boldsymbol{\phi}_i^\top\boldsymbol{\Psi}^{-1}(\boldsymbol{\rho})\frac{\partial\boldsymbol{\Psi}(\boldsymbol{\rho})}{\partial\rho_k}\boldsymbol{\Psi}^{-1}(\boldsymbol{\rho})\boldsymbol{\phi}_i, \\
\ddot{\boldsymbol{\ell}}_{\boldsymbol{\beta}\rho_k} &= \sum_{i=1}^{n}\boldsymbol{D}(\boldsymbol{X})\boldsymbol{D}(\boldsymbol{\xi}_i)\boldsymbol{\Psi}^{-1}(\boldsymbol{\rho})\left[\zeta_i + \zeta_i'\boldsymbol{\phi}_i\boldsymbol{\phi}_i^\top\boldsymbol{\Psi}^{-1}(\boldsymbol{\rho})\right]\frac{\partial\boldsymbol{\Psi}(\boldsymbol{\rho})}{\partial\rho_k}\boldsymbol{\Psi}^{-1}(\boldsymbol{\rho})\boldsymbol{\phi}_i, \\
\ddot{\ell}_{\rho_k\rho_s} &= -\frac{n}{2}\mathrm{tr}\left(-\boldsymbol{\Psi}^{-1}(\boldsymbol{\rho})\frac{\partial\boldsymbol{\Psi}(\boldsymbol{\rho})}{\partial\rho_k}\boldsymbol{\Psi}^{-1}(\boldsymbol{\rho})\frac{\partial\boldsymbol{\Psi}(\boldsymbol{\rho})}{\partial\rho_s}\right) + \sum_{i=1}^{n}\boldsymbol{\phi}_i^\top\boldsymbol{\Psi}^{-1}(\boldsymbol{\rho})\boldsymbol{\Psi}(\boldsymbol{\rho})_{\rho_k\rho_s}\boldsymbol{\Psi}(\boldsymbol{\rho})^{-1}\boldsymbol{\phi}_i,
\end{aligned}$$

with $\boldsymbol{\Psi}(\boldsymbol{\rho})_{\rho_k\rho_s}$ being given by

$$\zeta_i\frac{\partial\boldsymbol{\Psi}(\boldsymbol{\rho})}{\partial\rho_s}\boldsymbol{\Psi}(\boldsymbol{\rho})^{-1}\frac{\partial\boldsymbol{\Psi}(\boldsymbol{\rho})}{\partial\rho_k} + \zeta_i\frac{\partial\boldsymbol{\Psi}(\boldsymbol{\rho})}{\partial\rho_k}\boldsymbol{\Psi}^{-1}(\boldsymbol{\rho})\frac{\partial\boldsymbol{\Psi}(\boldsymbol{\rho})}{\partial\rho_s} + \zeta_i'\frac{\partial\boldsymbol{\Psi}(\boldsymbol{\rho})}{\partial\rho_s}\boldsymbol{\Psi}(\boldsymbol{\rho})^{-1}\boldsymbol{\phi}_i\boldsymbol{\phi}_i^\top\boldsymbol{\Psi}(\boldsymbol{\rho})^{-1}\frac{\partial\boldsymbol{\Psi}(\boldsymbol{\rho})}{\partial\rho_k},$$

if $k \neq s$; whereas the case $k = s$ conducts to

$$2\zeta_i\frac{\partial\boldsymbol{\Psi}(\boldsymbol{\rho})}{\partial\rho_k}\boldsymbol{\Psi}(\boldsymbol{\rho})^{-1}\frac{\partial\boldsymbol{\Psi}(\boldsymbol{\rho})}{\partial\rho_k} + \zeta_i'\frac{\partial\boldsymbol{\Psi}(\boldsymbol{\rho})}{\partial\rho_k}\boldsymbol{\Psi}(\boldsymbol{\rho})^{-1}\boldsymbol{\phi}_i\boldsymbol{\phi}_i^\top\boldsymbol{\Psi}(\boldsymbol{\rho})^{-1}\frac{\partial\boldsymbol{\Psi}(\boldsymbol{\rho})}{\partial\rho_k},$$

with $\zeta_i' = \zeta'(\mathrm{MD}_i)$, $\boldsymbol{D}(\boldsymbol{\phi}_i) = \mathrm{diag}(\phi_{i1}, \ldots, \phi_{im})$, $\boldsymbol{D}(\boldsymbol{\xi}_i^{-2}) = \mathrm{diag}(\xi_{i1}^{-2}, \ldots, \xi_{im}^{-2})$, and $\boldsymbol{D}(\boldsymbol{\phi}_i^2) = \mathrm{diag}(\phi_{i1}^2, \ldots, \phi_{im}^2)$, whose elements are $\phi_{ij} = [2/\alpha]\sinh([y_{ij} - \mu_{ij}]/2)$.

Under regularity conditions (see Cox and Hinkley, 1974), the estimators $\widehat{\alpha}$, $\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{\rho}}$ are consistent and have a multivariate normal joint asymptotic distribution with asymptotic means α, $\boldsymbol{\beta}$ and $\boldsymbol{\rho}$, respectively, and an asymptotic covariance matrix $\boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}}$ that can be obtained from the corresponding expected Fisher information matrix. Thus, we have that, as $n \to \infty$,

$$\sqrt{n}\,(\widehat{\boldsymbol{\theta}} - \boldsymbol{\theta}) \quad \xrightarrow{\mathcal{D}} \quad \mathrm{N}_{p^*}(\boldsymbol{0}_{p^*\times 1}, \boldsymbol{\Sigma}_{\widehat{\boldsymbol{\theta}}} = \mathcal{J}(\boldsymbol{\theta})^{-1}), \tag{3.6}$$

where $\xrightarrow{\mathcal{D}}$ means "convergence in distribution to", $p^* = p\,m + l + 1$, with $l = m[m-1]/2$, and $\mathcal{J}(\boldsymbol{\theta}) = \lim_{n\to\infty}(1/n)\mathcal{I}(\boldsymbol{\theta})$, with $\mathcal{I}(\boldsymbol{\theta})$ being the corresponding expected Fisher information matrix. Note that $\widehat{\mathcal{I}}(\boldsymbol{\theta})^{-1}$ is a consistent estimator of the asymptotic variance-covariance matrix of $\widehat{\boldsymbol{\theta}}$. In practice, one may approximate the expected Fisher information matrix by its observed version obtained from (3.5), whereas the diagonal elements of its inverse matrix can be used to approximate the corresponding SEs; see Efron and Hinkley (1978) for details about the use of observed versus expected Fisher information matrices. Asymptotic inference for the parameters of the proposed multivariate Birnbaum-Saunders regression models can be conducted by the asymptotic normality defined in (3.6).

3.4 Diagnostics

In this section, we consider the local influence method to produce a diagnostic analysis for multivariate GBS log-linear regression models. Fitting and global influence can be assessed by the MD presented in Section 3.3.

3.4.1 The local influence method

Consider $\ell(\boldsymbol{\theta})$ the log-likelihood function for the parameter $\boldsymbol{\theta}$ of the model defined in (2.8), which we call the non-perturbed model, and the perturbation vector $\boldsymbol{w} \in \mathbb{R}^q$ in the model, for $\boldsymbol{w} \in \Omega$, with Ω being a set of perturbations. Then, $\ell(\boldsymbol{\theta}|\boldsymbol{w})$ is the log-likelihood function of the perturbed model, with $\widehat{\boldsymbol{\theta}}_{\boldsymbol{w}}$ being the ML estimate of $\boldsymbol{\theta}$ obtained from $\ell(\boldsymbol{\theta}|\boldsymbol{w})$. Furthermore, let $\boldsymbol{w}_0 \in \Omega \in \mathbb{R}^q$ be a non-perturbation vector with $\boldsymbol{w}_0 = \mathbf{0}^\top_{q\times 1}$, or $\boldsymbol{w}_0 = \mathbf{1}^\top_{q\times 1}$, or a possible third choice, so that $\ell(\boldsymbol{\theta}) = \ell(\boldsymbol{\theta}|\boldsymbol{w}_0)$. Assuming that $\ell(\boldsymbol{\theta}|\boldsymbol{w})$ is twice continuously differentiable in a neighborhood of $(\widehat{\boldsymbol{\theta}}, \boldsymbol{w}_0)$, we compare the ML estimates $\widehat{\boldsymbol{\theta}}$ and $\widehat{\boldsymbol{\theta}}_{\boldsymbol{w}}$ by the local influence method to investigate how inference is affected by the corresponding perturbation. The likelihood displacement (LD) is given by

$$\mathrm{LD}(\boldsymbol{w}) = 2[\ell(\widehat{\boldsymbol{\theta}}) - \ell(\widehat{\boldsymbol{\theta}}_{\boldsymbol{w}})], \tag{3.7}$$

which is used to detect the influence of $\boldsymbol{w}$. Large values of $\mathrm{LD}(\boldsymbol{w})$ in (3.7) indicate that $\widehat{\boldsymbol{\theta}}$ and $\widehat{\boldsymbol{\theta}}_{\boldsymbol{w}}$ differ considerably in relation to the contours of the non-perturbed log-likelihood function $\ell(\boldsymbol{\theta})$. We study the local behaviour of the influence plot $a(\boldsymbol{w}) = (\boldsymbol{w}^\top, \mathrm{LD}(\boldsymbol{w}))^\top$ around $\boldsymbol{w}_0$. The direction in which the LD locally changes most rapidly is evaluated, that is, the maximum curvature of the surface $a(\boldsymbol{w})$. For $\mathrm{LD}(\boldsymbol{w})$ given in (3.7), the maximum curvature is read to be

$$C_{\max} = \max_{||\boldsymbol{d}||=1} C_{\boldsymbol{d}}, \tag{3.8}$$

where $C_{\boldsymbol{d}} = 2|\boldsymbol{d}^\top \boldsymbol{F}\boldsymbol{d}|$, with the matrix $\boldsymbol{F} \in \mathbb{R}^{n\times n}$ and $\boldsymbol{d}$ being the unit-length direction vector. To compute $C_{\max}$ given in (3.8) and the corresponding direction vector $\boldsymbol{d}_{\max}$, we must calculate

$$\boldsymbol{F} = -\boldsymbol{\Delta}(\widehat{\boldsymbol{\theta}}, \boldsymbol{w}_0)^\top \ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\theta}})^{-1} \boldsymbol{\Delta}(\widehat{\boldsymbol{\theta}}, \boldsymbol{w}_0), \tag{3.9}$$

where $\boldsymbol{\Delta}(\boldsymbol{\theta}, \boldsymbol{w}) \in \mathbb{R}^{p^*\times n}$ is a matrix partitioned accordingly for the perturbed model obtained from (2.8), called perturbation matrix, with elements

$$\Delta_{ij} = \frac{\partial^2 \ell(\boldsymbol{\theta}|\boldsymbol{w})}{\partial \boldsymbol{\theta}_i \partial \boldsymbol{w}_j},\ i = 1, \ldots, n,\ j = 1, \ldots, p^*,$$

evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ and $\boldsymbol{w} = \boldsymbol{w}_0$, where, as mentioned, $p^* = p\,m + l + 1$, with $l = m[m-1]/2$. We recall $-\ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\theta}}) \in \mathbb{R}^{p^*\times p^*}$ is the observed information matrix for the non-perturbed model. Then, $\boldsymbol{d}_{\max}$ is a unit-length eigenvector associated with the largest absolute eigenvalue $C_{\max}$ given in (3.8). If the absolute value of $\boldsymbol{d}_{\max_i}$ is large, it indicates that the case i is potentially influential. However, if the interest is in the subset $\boldsymbol{\theta}_1$ of $\boldsymbol{\theta} = (\boldsymbol{\theta}_1^\top, \boldsymbol{\theta}_2^\top)^\top$, the normal curvature in the direction $\boldsymbol{d}$ for $\boldsymbol{\theta}_1$ is given by

$$C_{\boldsymbol{d}}(\widehat{\boldsymbol{\theta}}_1) = 2|\boldsymbol{d}^\top \boldsymbol{\Delta}(\widehat{\boldsymbol{\theta}}, \boldsymbol{w}_0)^\top [\ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\theta}})^{-1} - \boldsymbol{B}_{22}] \boldsymbol{\Delta}(\widehat{\boldsymbol{\theta}}, \boldsymbol{w}_0)\boldsymbol{d}|,$$

where

$$\boldsymbol{B}_{22} = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{0} & \ddot{\boldsymbol{\ell}}^{-1}_{\boldsymbol{\theta}_2\boldsymbol{\theta}_2} \end{pmatrix}.$$

We may reveal cases influential on $\widehat{\boldsymbol{\theta}}_1$ by using the index-plot of the largest eigenvector of

$$-\boldsymbol{\Delta}(\widehat{\boldsymbol{\theta}}, \boldsymbol{w}_0)^\top [\ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\theta}})^{-1} - \boldsymbol{B}_{22}] \boldsymbol{\Delta}(\widehat{\boldsymbol{\theta}}, \boldsymbol{w}_0).$$

3.4.2 The total local influence method

In addition to $\boldsymbol{d}_{\max_i}$, another direction of interest is $\boldsymbol{d}_i = \boldsymbol{e}_{in}$, which is related to the direction of the case i, where $\boldsymbol{e}_{in} \in \mathbb{R}^n$ is a vector of zeros and a one at the ith position. Thus, the normal curvature is

$$C_i(\boldsymbol{\theta}) = 2|f_{ii}|, \quad \text{for} \quad i = 1, \ldots, n,$$

where f_{ii} is the ith diagonal element of $\boldsymbol{F}$ given in (3.9), evaluated at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$. Case i is considered as potentially influential if

$$C_i(\widehat{\boldsymbol{\theta}}) > 2\overline{C}(\widehat{\boldsymbol{\theta}}), \quad i = 1, \ldots, n,$$

where

$$\overline{C}(\widehat{\boldsymbol{\theta}}) = \frac{1}{n}\sum_{i=1}^{n} C_i(\widehat{\boldsymbol{\theta}}). \tag{3.10}$$

The diagnostic method defined in (3.10) is called total local influence; see Lesaffre and Verbeke (1998) and Verbeke and Molenberghs (2000). Note that is possible to compute the normal curvature for the parameters α, $\boldsymbol{\beta}$ and $\boldsymbol{\rho}$ denoted $C_i(\alpha)$, $C_i(\boldsymbol{\beta})$ and $C_i(\boldsymbol{\rho})$, respectively.

3.4.3 Normal curvatures

By using the model formulated in (2.8) and its perturbed version, we determine normal curvatures for local influence. We compute the observed information matrix $-\ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\theta}})$, derive the perturbation matrix $\boldsymbol{\Delta}(\widehat{\boldsymbol{\theta}}, \boldsymbol{w}_0)$ and then obtain the eigenvector associated with the largest absolute eigenvalue of $\boldsymbol{F}$ given in (3.9) as a local influence measure. Next, we detail the perturbation matrices for different schemes.

Case-weight perturbation (ca). For this scheme, let $\boldsymbol{w} = (w_1, \ldots, w_n)^\top \in \mathbb{R}^n$ be the perturbation vector, where the w_i's are positive values denoting the weight corresponding to the case i, and $\ell_{\rm ca}(\boldsymbol{\theta}|\boldsymbol{w})$ is the perturbed log-likelihood function. Let $\boldsymbol{w}_0 = \boldsymbol{1}_{1\times n}^\top$ be the non-perturbation vector such that $\ell_{\rm ca}(\boldsymbol{\theta}|\boldsymbol{w}_0) = \ell(\boldsymbol{\theta})$. Then, the log-likelihood function for the perturbed model under this scheme is

$$\ell_{\rm ca}(\boldsymbol{\theta}|\boldsymbol{w}) = \sum_{i=1}^{n} w_i \ell_i(\boldsymbol{\theta}), \tag{3.11}$$

with $\ell_i(\boldsymbol{\theta})$ defined from (3.1). Hence, we establish the matrix $\boldsymbol{\Delta}_{\rm ca}(\boldsymbol{\theta}, \boldsymbol{w})$ by taking the derivatives of $\ell_{\rm ca}(\boldsymbol{\theta}|\boldsymbol{w})$ given in (3.11) with respect to $\boldsymbol{\theta}$ and $\boldsymbol{w}$, evaluating them at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ and $\boldsymbol{w} = \boldsymbol{w}_0$. Thus,

$$\begin{aligned}
\left.\frac{\partial^2 \ell_{\rm ca}(\boldsymbol{\theta}|\boldsymbol{w})}{\partial\boldsymbol{\alpha}\partial\boldsymbol{w}_i}\right|_{\boldsymbol{w}=\boldsymbol{w}_0} &= -\frac{1}{\alpha}\left[2\zeta_i \mathrm{MD}_i + m\right], \\
\left.\frac{\partial^2 \ell_{\rm ca}(\boldsymbol{\theta}|\boldsymbol{w})}{\partial\boldsymbol{\beta}\partial\boldsymbol{w}_i}\right|_{\boldsymbol{w}=\boldsymbol{w}_0} &= -\boldsymbol{D}(\boldsymbol{X})[\zeta_i \boldsymbol{D}(\boldsymbol{\xi}_i)\boldsymbol{\Psi}^{-1}(\boldsymbol{\rho}) + \frac{1}{2}\boldsymbol{D}(\boldsymbol{\xi}_i^{-1})]\boldsymbol{\phi}_i, \\
\left.\frac{\partial^2 \ell_{\rm ca}(\boldsymbol{\theta}|\boldsymbol{w})}{\partial\rho_k\partial\boldsymbol{w}_i}\right|_{\boldsymbol{w}=\boldsymbol{w}_0} &= -\frac{1}{2}\mathrm{tr}\left(\boldsymbol{\Psi}^{-1}(\boldsymbol{\rho})\frac{\partial\boldsymbol{\Psi}(\boldsymbol{\rho})}{\partial\rho_k}\right) - \zeta_i\boldsymbol{\phi}_i^\top\boldsymbol{\Psi}^{-1}(\boldsymbol{\rho})\frac{\partial\boldsymbol{\Psi}(\boldsymbol{\rho})}{\partial\rho_k}\boldsymbol{\Psi}^{-1}(\boldsymbol{\rho})\boldsymbol{\phi}_i,
\end{aligned}$$

for $k = 1, \ldots, l$ and $i = 1, \ldots, n$, where $\zeta_i = \zeta(\mathrm{MD}_i)$ is specified as

$$\zeta(\mathrm{MD}_i) = -\frac{1}{2} \quad \text{and} \quad \zeta(\mathrm{MD}_i) = -\frac{[\nu + m]}{2[\nu + \mathrm{MD}_i]},$$

for the multivariate normal and t kernels, respectively, based on expressions given in (3.2) and (3.3).

Perturbation of the correlation matrix (cm). For this scheme, let $\boldsymbol{w} = (w_1, \ldots, w_n)^\top \in \mathbb{R}^n - \{0\}$ be the perturbation vector and $\ell_{\rm cm}(\boldsymbol{\theta}|\boldsymbol{w})$ the corresponding perturbed log-likelihood function. Let $\boldsymbol{w}_0 = \mathbf{1}_{1\times n}^\top$ be the non-perturbation vector such that $\ell_{\rm cm}(\boldsymbol{\theta}|\boldsymbol{w}_0) = \ell(\boldsymbol{\theta})$. Then, the log-likelihood function for the perturbed model under this scheme is

$$\ell_{\rm cm}(\boldsymbol{\theta}|\boldsymbol{w}) = \sum_{i=1}^{n}\left[\log(f_{{\rm EC}_m}(\boldsymbol{\phi}_i; w_i^{-1}\boldsymbol{\Psi}(\boldsymbol{\rho}), g^{(m)})) + \sum_{j=1}^{m}\log(\xi_{ij})\right]. \tag{3.12}$$

Again, we establish the matrix $\boldsymbol{\Delta}_{\rm cm}(\boldsymbol{\theta}, \boldsymbol{w})$ by taking the derivatives now of $\ell_{\rm cm}(\boldsymbol{\theta}|\boldsymbol{w})$ given in (3.12) with respect to $\boldsymbol{\theta}$, and then with respect to $\boldsymbol{w}$, evaluating them at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ and $\boldsymbol{w} = \boldsymbol{w}_0$. Thus,

$$\begin{aligned}
\left.\frac{\partial^2\ell_{\rm cm}(\boldsymbol{\theta}|\boldsymbol{w})}{\partial\boldsymbol{\alpha}\partial\boldsymbol{w}_i}\right|_{\boldsymbol{w}=\boldsymbol{w}_0} &= -\frac{2}{\alpha}\left[\zeta_i'{\rm MD}_i + \zeta_i\right]{\rm MD}_i,\\
\left.\frac{\partial^2\ell_{\rm cm}(\boldsymbol{\theta}|\boldsymbol{w})}{\partial\boldsymbol{\beta}\partial\boldsymbol{w}_i}\right|_{\boldsymbol{w}=\boldsymbol{w}_0} &= -\left[\zeta_i'{\rm MD}_i + \zeta_i\right]\boldsymbol{D}(\boldsymbol{X})\boldsymbol{D}(\boldsymbol{\xi}_i)\boldsymbol{\Psi}^{-1}(\boldsymbol{\rho})\boldsymbol{\phi}_i,\\
\left.\frac{\partial^2\ell_{\rm cm}(\boldsymbol{\theta}|\boldsymbol{w})}{\partial\rho_k\partial\boldsymbol{w}_i}\right|_{\boldsymbol{w}=\boldsymbol{w}_0} &= -\left[\zeta_i'{\rm MD}_i + \zeta_i\right]\phi_i^\top\boldsymbol{\Psi}^{-1}(\boldsymbol{\rho})\frac{\partial\boldsymbol{\Psi}(\boldsymbol{\rho})}{\partial\rho_k}\boldsymbol{\Psi}^{-1}(\boldsymbol{\rho})\boldsymbol{\phi}_i,
\end{aligned}$$

for $k = 1, \ldots, l$, and $i = 1, \ldots, n$, where $\zeta_i' = \zeta'({\rm MD}_i)$, with

$$\zeta'({\rm MD}_i) = 0 \quad \text{and} \quad \zeta'({\rm MD}_i) = \frac{[\nu + m]}{2[\nu + {\rm MD}_i]^2},$$

for the multivariate normal and t kernels, respectively.

Perturbation of a covariate (co). In this scheme, we perturb a continuous covariate x_{il} by $x_{il} + w_i$, where $x_{il} \in \mathbb{R}^n$ is the lth column of $\boldsymbol{x}_i$ and $\boldsymbol{w} = (w_1, \ldots, w_n)^\top \in \mathbb{R}^n$ is the perturbation vector. Here, $\boldsymbol{w}$ can be expressed as a proportional value to the standard deviation (SD) of the perturbed covariate and $\boldsymbol{w}_0 = \mathbf{0}_{1\times n}^\top$ is the non-perturbation vector such that $\ell_{\rm ca}(\boldsymbol{\theta}|\boldsymbol{w}_0) = \ell(\boldsymbol{\theta})$. Then, the log-likelihood function for the perturbed model under this scheme is

$$\ell_{\rm co}(\boldsymbol{\theta}|\boldsymbol{w}) = \sum_{i=1}^{n}\left[\log(f_{{\rm EC}_m}(\boldsymbol{\phi}_i(\boldsymbol{w}); \boldsymbol{\Psi}(\boldsymbol{\rho}), g^{(m)})) + \sum_{j=1}^{m}\log(\xi_{ij}(\boldsymbol{w}))\right]. \tag{3.13}$$

Here, we establish $\boldsymbol{\Delta}_{\rm co}(\boldsymbol{\theta}, \boldsymbol{w})$ by taking the derivatives of $\ell_{\rm co}(\boldsymbol{\theta}|\boldsymbol{w})$ given in (3.13) with respect to $\boldsymbol{\theta}$ and $\boldsymbol{w}$, evaluating them at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ and $\boldsymbol{w} = \boldsymbol{w}_0$. Thus, for $k = 1, \ldots, l$ and $i = 1, \ldots, n$,

$$\begin{aligned}
\left.\frac{\partial^2 \ell_{\rm co}(\boldsymbol{\theta}|\boldsymbol{w})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{w}_i}\right|_{\boldsymbol{w}=\boldsymbol{w}_0} &= \frac{2}{\alpha} \boldsymbol{\xi}_i^\top \boldsymbol{D}(\boldsymbol{\beta}_l) \left[\zeta_i' {\rm MD}_i + \zeta_i\right] \boldsymbol{\Psi}^{-1}(\boldsymbol{\rho}) \boldsymbol{\phi}_i, \\
\left.\frac{\partial^2 \ell_{\rm co}(\boldsymbol{\theta}|\boldsymbol{w})}{\partial \boldsymbol{\beta} \partial \boldsymbol{w}_i}\right|_{\boldsymbol{w}=\boldsymbol{w}_0} &= \frac{1}{2} \zeta_i' [\boldsymbol{\xi}_i^\top \boldsymbol{D}(\boldsymbol{\beta}_l) \boldsymbol{\Psi}^{-1}(\boldsymbol{\rho}) \boldsymbol{\phi}_i] \boldsymbol{D}(\boldsymbol{X}) \boldsymbol{D}(\boldsymbol{\xi}_i) \boldsymbol{\Psi}^{-1}(\boldsymbol{\rho}) \boldsymbol{\phi}_i \\
&\quad + \frac{1}{2} \boldsymbol{D}(\boldsymbol{X}) \left[\zeta_i \boldsymbol{D}(\boldsymbol{\phi}_i) \boldsymbol{D}(\boldsymbol{\beta}_l) \boldsymbol{\Psi}^{-1}(\boldsymbol{\rho}) - \frac{1}{2} \boldsymbol{D}(\boldsymbol{\xi}_i^{-2}) \boldsymbol{D}(\boldsymbol{\phi}_i) \boldsymbol{D}(\boldsymbol{\beta}_l)\right] \boldsymbol{\phi}_i \\
&\quad + \frac{1}{2} \boldsymbol{D}(\boldsymbol{X}) \left[\zeta_i \boldsymbol{D}(\boldsymbol{\xi}_i) \boldsymbol{\Psi}^{-1}(\boldsymbol{\rho}) + \frac{1}{2} \boldsymbol{D}(\boldsymbol{\xi}_i^{-1})\right] \boldsymbol{D}(\boldsymbol{\beta}_l) \boldsymbol{\xi}_i \\
\left.\frac{\partial^2 \ell_{\rm co}(\boldsymbol{\theta}|\boldsymbol{w})}{\partial \rho_k \partial \boldsymbol{w}_i}\right|_{\boldsymbol{w}=\boldsymbol{w}_0} &= \boldsymbol{\xi}_i^\top \boldsymbol{D}(\boldsymbol{\beta}_l) \left[\zeta_i' \boldsymbol{\Psi}^{-1}(\boldsymbol{\rho}) \boldsymbol{\phi}_i \boldsymbol{\phi}_i^\top + \zeta_i\right] \boldsymbol{\Psi}^{-1}(\boldsymbol{\rho}) \frac{\partial \boldsymbol{\Psi}(\boldsymbol{\rho})}{\partial \rho_k} \boldsymbol{\Psi}^{-1}(\boldsymbol{\rho}) \boldsymbol{\phi}_i.
\end{aligned}$$

Here, $\boldsymbol{D}(\boldsymbol{\beta}_l)$ is a diagonal matrix with elements β_{lj}, for $j = 1, \ldots, m$.

Perturbation of the response variable (re). In this scheme, we replace $\boldsymbol{Y}_i$ by $\boldsymbol{Y}_i + \boldsymbol{w}_i$, where $\boldsymbol{w}_i = (w_{i1}, \ldots, w_{im})^\top \in \mathbb{R}^m$ denoting the corresponding perturbation to the case i with $\boldsymbol{w}_i > \boldsymbol{0}$. Here $\boldsymbol{w}_i$ can be expressed as a proportional value to the SD of the response and $\boldsymbol{w}_0 = \boldsymbol{0}_{1\times m}^\top$ is the non-perturbation vector such that $\ell_{\rm re}(\boldsymbol{\theta}|\boldsymbol{w}_0) = \ell(\boldsymbol{\theta})$. Then, the log-likelihood function for the perturbed model under this scheme is

$$\ell_{\rm re}(\boldsymbol{\theta}|\boldsymbol{w}) = \sum_{i=1}^{n} \left[\log(f_{{\rm EC}_m}(\boldsymbol{\phi}_i(\boldsymbol{w}); \boldsymbol{\Psi}(\boldsymbol{\rho}), g^{(m)})) + \sum_{j=1}^{m} \log(\xi_{ij}(\boldsymbol{w}))\right]. \tag{3.14}$$

Again, $\boldsymbol{\Delta}_{\rm re}(\boldsymbol{\theta}, \boldsymbol{w})$ is obtained by taking the corresponding derivatives now of $\ell_{\rm re}(\boldsymbol{\theta}|\boldsymbol{w})$ given in (3.14) with respect to $\boldsymbol{\theta}$ and $\boldsymbol{w}$, for $k = 1, \ldots, l$, and $i = 1, \ldots, n$, and evaluating them at $\boldsymbol{\theta} = \widehat{\boldsymbol{\theta}}$ and $\boldsymbol{w} = \boldsymbol{w}_0$. Thus,

$$\begin{aligned}
\left.\frac{\partial^2 \ell_{\rm re}(\boldsymbol{\theta}|\boldsymbol{w})}{\partial \boldsymbol{\alpha} \partial \boldsymbol{w}_i^\top}\right|_{\boldsymbol{w}=\boldsymbol{w}_0} &= -\frac{2}{\alpha} \left[\zeta_i' {\rm MD}_i + \zeta_i\right] \boldsymbol{D}(\boldsymbol{\xi}_i) \boldsymbol{\Psi}^{-1}(\boldsymbol{\rho}) \boldsymbol{\phi}_i, \\
\left.\frac{\partial^2 \ell_{\rm re}(\boldsymbol{\theta}|\boldsymbol{w})}{\partial \boldsymbol{\beta} \partial \boldsymbol{w}_i^\top}\right|_{\boldsymbol{w}=\boldsymbol{w}_0} &= \boldsymbol{D}(\boldsymbol{\xi}_i) [\zeta_i' \boldsymbol{\Psi}^{-1}(\boldsymbol{\rho}) \boldsymbol{\phi}_i \boldsymbol{\phi}_i^\top \boldsymbol{\Psi}^{-1}(\boldsymbol{\rho}) \boldsymbol{D}(\boldsymbol{\xi}_i) - \frac{1}{2} \zeta_i \boldsymbol{\Psi}^{-1}(\boldsymbol{\rho}) \boldsymbol{D}(\boldsymbol{\xi}_i) \\
&\quad - \frac{1}{4} \boldsymbol{D}(\boldsymbol{\xi}_i^{-1})] \boldsymbol{D}(\boldsymbol{X}^\top) \left[\frac{1}{2} \zeta_i \boldsymbol{D}(\boldsymbol{\phi}_i) \boldsymbol{D}(\boldsymbol{\Psi}^{-1}(\boldsymbol{\rho}) \boldsymbol{\phi}_i) - \frac{1}{4} \boldsymbol{D}(\boldsymbol{\xi}_i^{-2}) \boldsymbol{D}(\boldsymbol{\phi}_i^2)\right] \boldsymbol{D}(\boldsymbol{X}^\top), \\
\left.\frac{\partial^2 \ell_{\rm re}(\boldsymbol{\theta}|\boldsymbol{w})}{\partial \rho_k \partial \boldsymbol{w}_i^\top}\right|_{\boldsymbol{w}=\boldsymbol{w}_0} &= -\boldsymbol{D}(\boldsymbol{\xi}_i) \left[\zeta_i' \boldsymbol{\Psi}^{-1}(\boldsymbol{\rho}) \boldsymbol{\phi}_i \boldsymbol{\phi}_i^\top + \zeta_i\right] \boldsymbol{\Psi}^{-1}(\boldsymbol{\rho}) \frac{\partial \boldsymbol{\Psi}(\boldsymbol{\rho})}{\partial \rho_k} \boldsymbol{\Psi}^{-1}(\boldsymbol{\rho}) \boldsymbol{\phi}_i.
\end{aligned}$$

3.4.4 Generalized leverage

The main idea behind the concept of leverage is to evaluate the influence of $\boldsymbol{Y}_i$ on its own predicted value. This influence may be well represented by the derivative $\partial \widehat{\boldsymbol{Y}}_i / \partial \boldsymbol{Y}_i^\top$.

Consider the log-likelihood function for $\boldsymbol{\theta}$ given in (3.1). Let $\widehat{\boldsymbol{\theta}}$ be the ML estimate of $\boldsymbol{\theta}$ and $\boldsymbol{\mu} = \mathrm{E}(\boldsymbol{Y})$. Then, $\widehat{\boldsymbol{Y}} = \boldsymbol{\mu}(\widehat{\boldsymbol{\theta}})$ is the predicted response vector. The matrix of GL may be expressed as

$$\mathrm{GL}(\widehat{\boldsymbol{\theta}}) = \boldsymbol{D}_{\boldsymbol{\theta}}(-\ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\theta}}))^{-1}\ddot{\boldsymbol{\ell}}_{\boldsymbol{\theta}\boldsymbol{Y}}\Big|_{\boldsymbol{\theta}=\widehat{\boldsymbol{\theta}}},$$

where

$$\begin{aligned}\boldsymbol{D}_{\boldsymbol{\theta}} &= \frac{\partial \boldsymbol{\mu}}{\partial \boldsymbol{\theta}^\top} = (\boldsymbol{X}\vdots\boldsymbol{0}) \in \mathbb{R}^{n\times p^*},\\ \ddot{\boldsymbol{\ell}}_{\boldsymbol{\theta}\boldsymbol{Y}} &= \frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \boldsymbol{\theta}\partial \boldsymbol{Y}^\top} = \boldsymbol{\Delta}_{\mathrm{r}}(\boldsymbol{\theta}, \boldsymbol{w}) \in \mathbb{R}^{p^*\times n},\end{aligned}$$

and $\ddot{\boldsymbol{\ell}}(\widehat{\boldsymbol{\theta}}) \in \mathbb{R}^{p^*\times p^*}$ is as given in (3.5), with p^* given in (3.6).

3.5 Application

In this section, we apply the multivariate GBS log-linear regression model and the developed fitting and diagnostic tools implemented by a computational routine in `R` code to real-world data collected by the authors; see Vivanco et al. (2014) and Tables 6.2 and 6.3 of Appendix 6.3. We consider multivariate data useful for regression models based on clinical CT to study the bone quality. As mentioned, note that the multivariate BS distribution is appropriate for modeling strength, which is a mechanical property of bone related to physical properties, as its densities extracted from CT data.

3.5.1 Description of the problem

CT is an important tool in the study of bone quality and a basis for finite element analysis of bone. It provides 3D-images of X-ray attenuation allowing not only the analysis of a 3D geometry but also the measurement of volumetric mineral content. These characteristics allow us to relate mineral density derived from CT to mechanical properties of bone. Thus, one can assess the mechanical properties for fracture risk evaluation or modeling by subject-specific finite element. Several researchers have related CT to elastic modulus (or strength) and bone density for estimation of mechanical properties; see, for example, Duchemin et al. (2008) and Vivanco et al. (2014). We investigate 4 types of densities:
(i) Bulk density, which considers the mass of the intact core, including fat and water;
(ii) Dry density, which excludes fat and water;
(iii) Ash density, which is related to the mineral content; and
(iv) Clinical CT, which is obtained from a calibration equation that is derived from known mineral content phantoms.

Clinical CT scans are used to assess their application in inferring physical properties of human trabecular bone. The prediction of apparent density from ash density allows for estimation of mechanical properties of bone, which can subsequently be used in a finite element model. For example, to determine the utility of clinical CT scans in the prediction of physical properties of human trabecular bone; see Vivanco et al. (2014).

3.5.2 Regression analysis

We consider as responses: (i) bulk density (T_1, in mg/cm^3) and (ii) dry density (T_2, in mg/cm^3). The covariates that could affect these responses are: (i) CT density (X_1, in mg/cm^3) and (ii) ash density (X_2, in mg/cm^3). We illustrate the proposed multivariate models with real-world density data associated with these variables. We work with the log-responses $Y_j = \log(T_j)$, for $j = 1, 2$.

First, we make an exploratory data analysis computing correlations for Y_1, Y_2, X_1, X_2. Figure 3.1 displays the scatter-plots for these variables and their corresponding correlations. From this figure we detect that exist:
(i) Large correlations between (Y_1, Y_2), justifying the use of multivariate distributions;
(ii) Large correlations between X_1, X_2, indicating a possible collinearity problem; and
(iii) Medium correlations between (X_1, Y_1), and (X_1, Y_2), and large correlations between (X_2, Y_1) and (X_2, Y_2), which supports the elimination of X_1.
This must be confirmed by the inferential analysis.

Second, we consider the previous exploratory data analysis, the adequacy of the BS distribution for this problem, and the robustness estimation in BS-t models to propose a multivariate regression model for describing (Y_1, Y_2) in function of X_2 (because X_1 is discarded due to collinearity problem). Therefore, the proposed multivariate log-linear regression model is given by

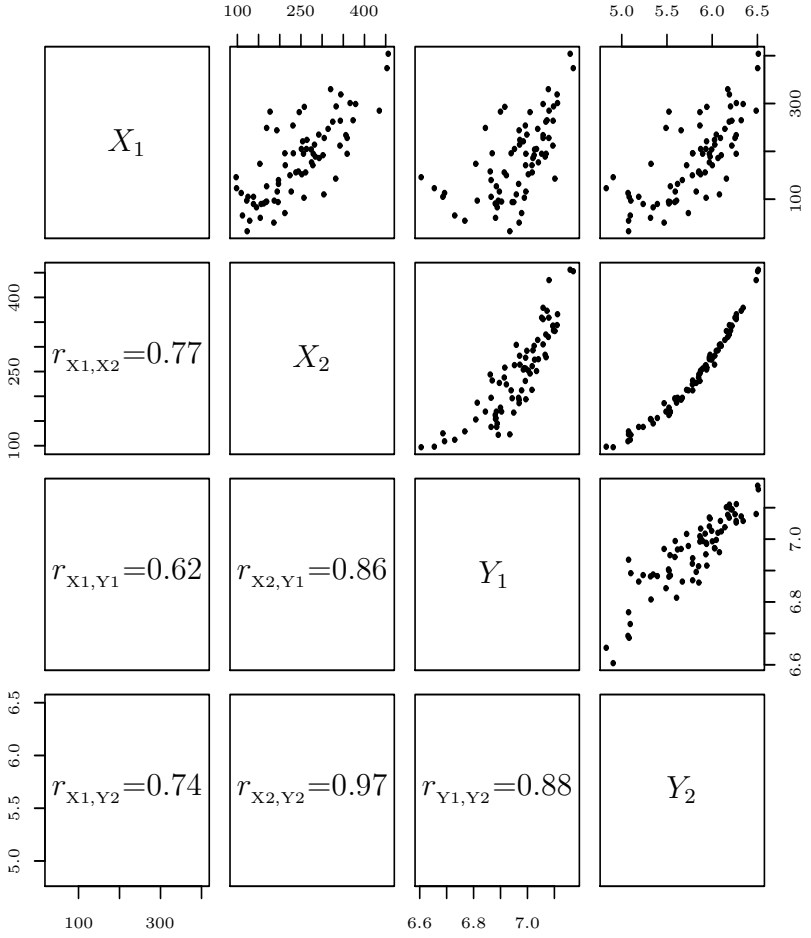
$$\boldsymbol{Y}_i = \boldsymbol{\beta}^\top \boldsymbol{x}_i + \boldsymbol{\varepsilon}_i, \quad i = 1, \ldots, 74,$$

where $\boldsymbol{\varepsilon}_i = (\varepsilon_{i1}, \varepsilon_{i2})^\top \sim \text{log-GBS}_2(\alpha \mathbf{1}_{2\times 1}, \mathbf{0}_{2\times 1}, \boldsymbol{\Psi}(\boldsymbol{\rho})_{2\times 2}, g^{(2)})$. We estimate the parameters of the multivariate BS and BS-t regression models via the ML method, which we have implemented in `R` code. We use the BFGS quasi-Newton method through the `optim` function. Starting values, $\widehat{\boldsymbol{\theta}}^{(0)}$ say, used in the maximization procedure are:

$$\widehat{\alpha}^{(0)} = 0.077963, \quad \widehat{\boldsymbol{\beta}}^{(0)} = \begin{pmatrix} 6.677099 & 4.682944 \\ 0.001157 & 0.004538 \end{pmatrix}, \quad \widehat{\boldsymbol{\Psi}}(\boldsymbol{\rho})^{(0)} = \begin{pmatrix} 1.000000 & 0.378761 \\ 0.378761 & 1.0000 \end{pmatrix}.$$

In addition, we have used the value $\nu = 4$. We have verified that $\nu = 4$ corresponds to the value that maximizes the log-likelihood function within a range of values for ν; see Figure 3.2.

Figure 3.1: Scatter-plots with their corresponding correlations for the indicated variable with bone density data.



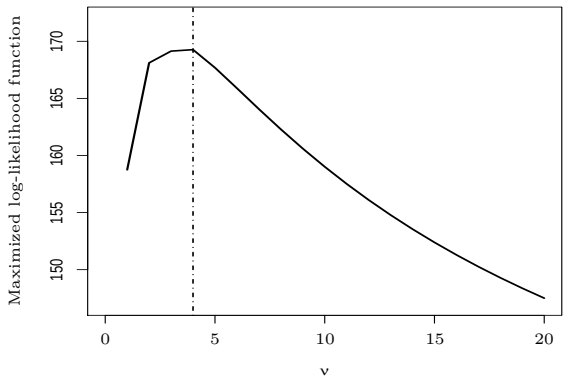
Source: From the author.

Table 3.1 displays the parameter estimates and the maximized log-likelihood value, estimated asymptotic SEs of the corresponding ML estimators for both models, and p-values of each z-test. From this table, and for a 5% significance level, we obtain the following conclusions:

(i) Estimated correlation from the BS_2 and BS-t_2 log-linear models results to be statistically significant, corroborating our conjecture from the exploratory analysis; and

(ii) The regression coefficients β_0 (constant term of the model) and β_1 (slope) must be considered in the prediction of T_1 and T_2. We can also see that the value that maximizes the log-likelihood is greater for BS-t_2 model, indicating a better fit.

Figure 3.2: Profiled maximized log-likelihood in function of ν for $\nu = 1, \ldots, 20$ by 1 with bone density data.



Source: From the author.

Table 3.1: ML estimate of the indicated parameter and model with its corresponding estimated SE, p-value and log-likelihood function.

	BS_2 model			BS-t_2 model		
Parameter	Estimate	SE	p-value	Estimate	SE	p-value
ρ	0.377034	0.050354	< 0.001	0.373313	0.059365	< 0.001
β_{01}	6.676328	0.025193	< 0.001	6.678921	0.029402	< 0.001
β_{02}	4.679009	0.025101	< 0.001	4.687253	0.032067	< 0.001
β_{11}	0.001159	0.000097	< 0.001	0.001151	0.000110	< 0.001
β_{12}	0.004550	0.000097	< 0.001	0.004526	0.000120	< 0.001
α	0.072004	0.003628	< 0.001	0.071836	0.004840	< 0.001
Log-likelihood	165.3831	-	-	169.3681	-	-

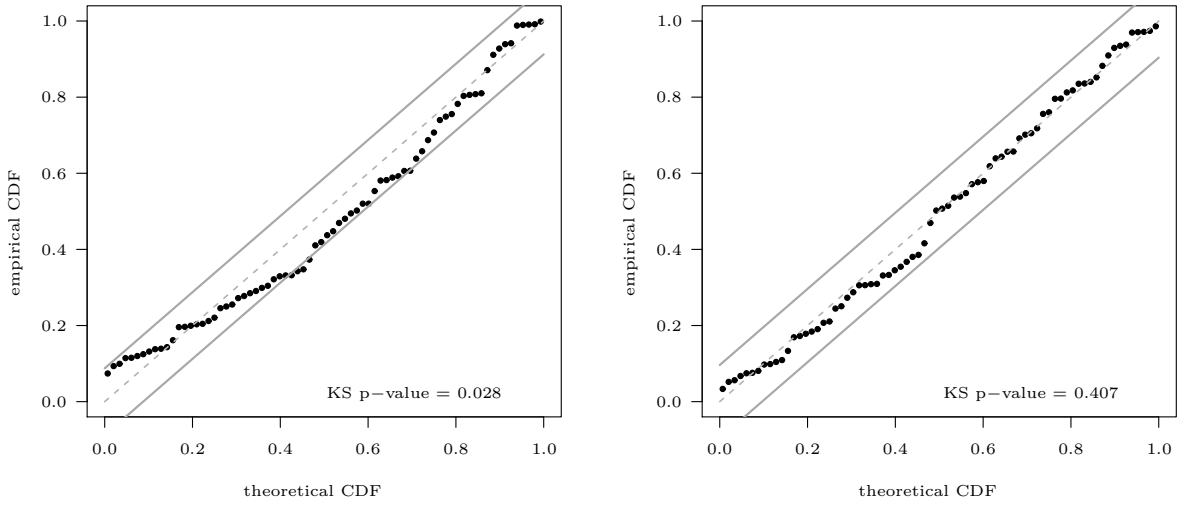
Source: From the author.

3.5.3 Model checking and global influence analysis

As mentioned, m-variate log-GBS model checking can be conducted by using the MD. Here, this distance follows the $\chi^2(m = 2)$ or $2\mathcal{F}(m = 2, \nu = 4)$ distribution if $g^{(2)}$ is the multivariate normal or t_2 kernel, respectively. We substitute the ML estimator of $\boldsymbol{\theta}$ in $\mathrm{MD}_i(\widehat{\boldsymbol{\theta}})$, which has asymptotically the same distribution of $\mathrm{MD}_i(\boldsymbol{\theta})$; see Lange et al. (1989). We use the WH approximations given in (2.18) and (2.19) for transforming this distance, which should follow now a normal distribution; see Ibacache-Pulgar et al. (2014) and references

therein. Then, we check normality of the transformed distances with the WH approximation using GOF techniques. Figure 3.3 shows the corresponding PP-plots with acceptance bands for a significance level of 5%; see Algorithm 4. From this figure, we detect that the BS-t_2 log-linear regression model provides a better fit than the BS_2 model, which is corroborated by the p-values 0.028 and 0.407, respectively, of the Kolmogorov-Smirnov (KS) test associated with these PP-plots; see Barros et al. (2014). Therefore, we can conclude that the multivariate BS-t_2 log-linear regression model fits better to the bone density data.

Figure 3.3: PP-plots with KS acceptance regions at 5% for transformed MDs with BS_2 (a) and BS-t_2 (b) models.

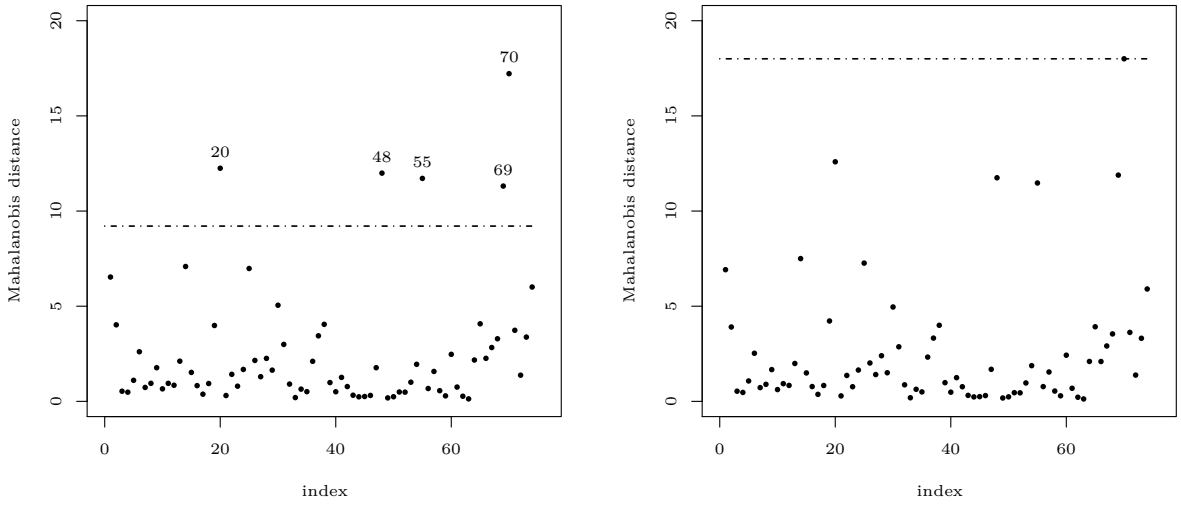


(a) (b)

Source: From the author.

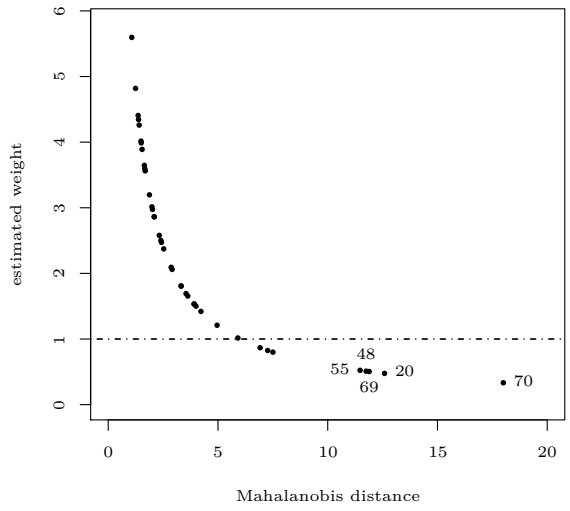
As mentioned, the MD is a global influence measure to detect multivariate outliers. Figures 3.4 (a) and (b) display the index-plots of this distance for BS_2 and BS-t_2 log-linear regression models. In addition, Figure 3.4 (c) presents the plot of estimated weights, say $\widehat{\zeta}_i$, versus MD_i for the BS-t_2 log-linear regression model, with $i = 1, \ldots, 74$. From Figure 3.4 (a) and (b), note that the cases $\{20, 48, 55, 69, 70\}$ appear as possible multivariate outliers in the BS_2 model, but not in the BS-t_2 log-linear regression model. In Figure 3.4 (c), observe that these cases have smaller weight in the BS-t_2 log-linear regression model than the BS_2 log-linear regression model, which confirms the inherent robustness of the ML procedure against possible outlying observations.

Figure 3.4: Index-plots of MDs for the BS_2 (a) and BS-t_2 (b) model; plot of estimated weights of MDs for the BS-t_2 model (c) and BS_2 model –straight line at one–.



(a)

(b)



(c)

Source: From the author.

3.5.4 Local influence analysis

In order to identify possible influential cases under the fitted models, we present some diagnostic graphs for total local influence (C_i). Plots for local influence ($\boldsymbol{d}_{\max}$) are omitted here, but they are similar to total local influence graphs.

Figures 3.5 and 3.6 show the index-plots of C_i under the case-weight perturbation scheme for $\widehat{\boldsymbol{\theta}}$ (a,b), $\widehat{\alpha}$ (c,d), $\widehat{\boldsymbol{\beta}}$ (a,b) and $\widehat{\rho}$ (c,d), respectively, for the indicated model. From this figure, note that cases $\{20, 48, 55, 69, 70\}$ appear with a large influence under the BS_2 log-linear regression model, but not under the BS-t_2 model. These cases have different degrees of influence on $\widehat{\alpha}$ and $\widehat{\boldsymbol{\beta}}$ and coincide with those detected by the MD. Note that there is practically no influence of these cases on $\widehat{\rho}$ for both models according to Figure 3.6(c,d).

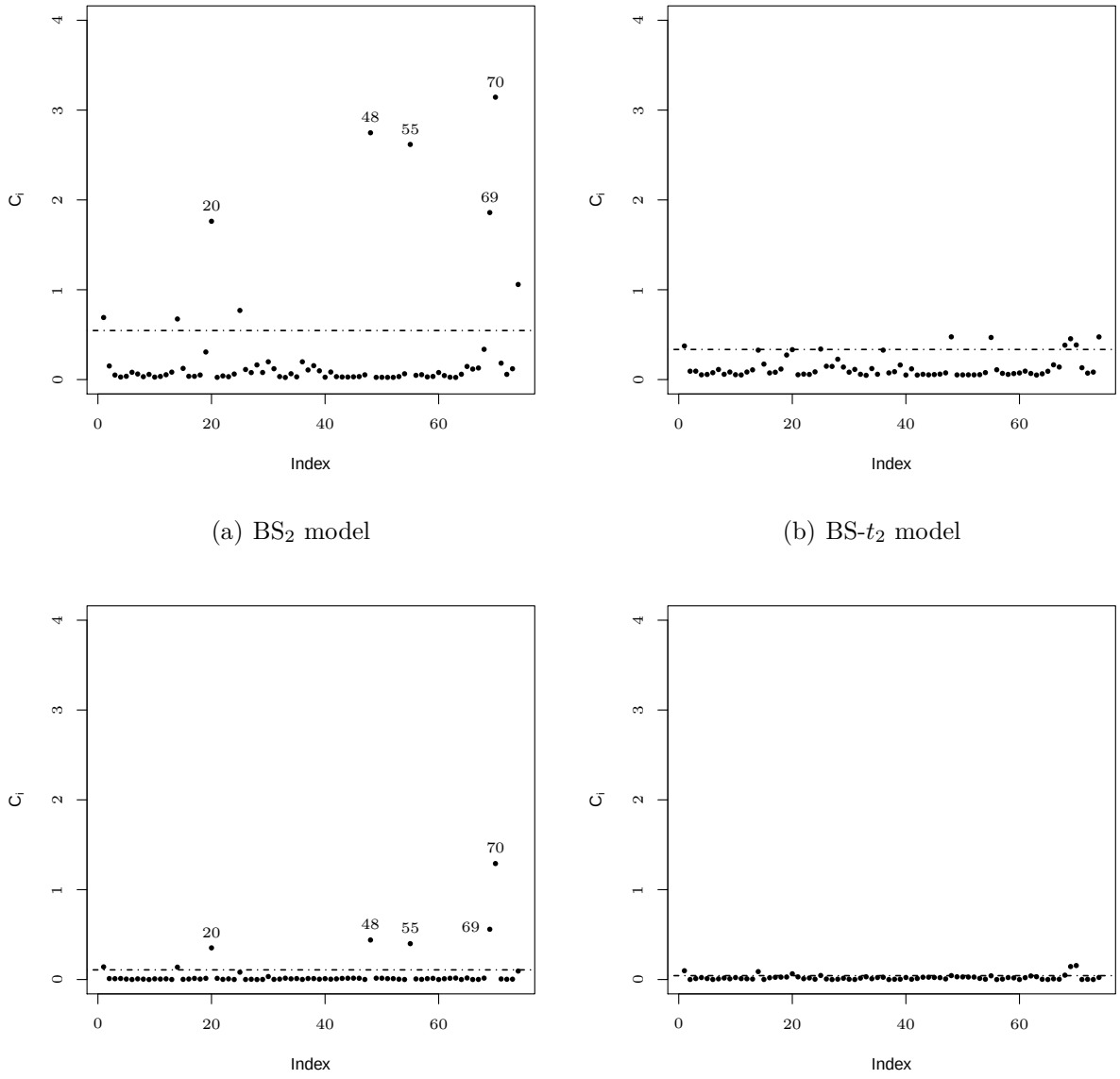
Figures 3.7 and 3.8 display the index-plots of C_i under the correlation perturbation scheme for $C_i(\boldsymbol{\theta})$ (a,b), $C_i(\alpha)$ (c,d), $C_i(\boldsymbol{\beta})$ (a,b) and $C_i(\rho)$ (c,d), respectively, for the indicated model. From this figure, note that again cases $\{20, 48, 55, 69, 70\}$ appear with a large influence under the BS_2 model, but not under the BS-t_2 model. In addition, observe that cases $\{20, 48, 55, 69, 70\}$ have different degrees of influence on $\widehat{\alpha}$ and $\widehat{\boldsymbol{\beta}}$. Again note that there is practically no influence of these cases on $\widehat{\rho}$ for both models according to Figure 3.8(c,d).

Figures 3.9 and 3.10 present the index-plots of C_i under the covariate perturbation scheme for $C_i(\boldsymbol{\theta})$ (a,b), $C_i(\alpha)$ (c,d), $C_i(\boldsymbol{\beta})$ (a,b) and $C_i(\rho)$ (c,d), respectively, for the indicated model. From this figure, note that now cases $\{48, 55, 69, 70, 74\}$ appear with a large influence under the BS_2 model, but not under the BS-t_2 model. In addition, observe that cases $\{69, 70\}$, $\{48, 55, 74\}$ and $\{48, 55, 69, 70\}$ appear as influential on $\widehat{\alpha}$, $\widehat{\boldsymbol{\beta}}$ and $\widehat{\rho}$, respectively.

Figures 3.11 and 3.12 show the index-plots of C_i under the response variable perturbation scheme for $C_i(\boldsymbol{\theta})$ (a,b), $C_i(\alpha)$ (c,d), $C_i(\boldsymbol{\beta})$ (a,b) and $C_i(\rho)$ (a,b), respectively, for the indicated model. From this figure, note that cases $\{20, 48, 55\}$ appear with a large influence under the BS_2 model for variable Y_1, but not under the BS-t_2 model. In addition, cases $\{20\}$, $\{48, 55\}$ and $\{20, 48, 55, 70\}$ have different degrees of influence on $\widehat{\alpha}$, $\widehat{\boldsymbol{\beta}}$ and $\widehat{\rho}$, for the variable Y_1. Furthermore, Figures 3.11 and 3.12 show that cases $\{20, 48, 55, 70\}$ appear with a large influence under the BS_2 model, but not under the BS-t_2 model for variable Y_2. Moreover, observe that cases $\{20, 48, 55, 70\}$, $\{48, 55\}$ and $\{20\}$ appear as influential on $\widehat{\alpha}$, $\widehat{\boldsymbol{\beta}}$ and $\widehat{\rho}$, respectively, for Y_2.

Figure 3.13 show the index-plots of the GL method for the responses variables under study. From Figures 3.13(a, b), note that cases $\{48, 55, 74\}$ appear as possible leverage points for Y_1 under BS_2 model. From Figure 3.13(c, d), cases $\{48, 55, 70, 74\}$ and $\{74\}$ appear now as possible leverage points for Y_2 under the BS_2 and BS-t_2 models, respectively.

Figure 3.5: Total local influence index-plots of case-weight perturbation for $\widehat{\boldsymbol{\theta}}$ (a,b), $\widehat{\alpha}$ (c,d) for the indicated model.



(a) BS_2 model

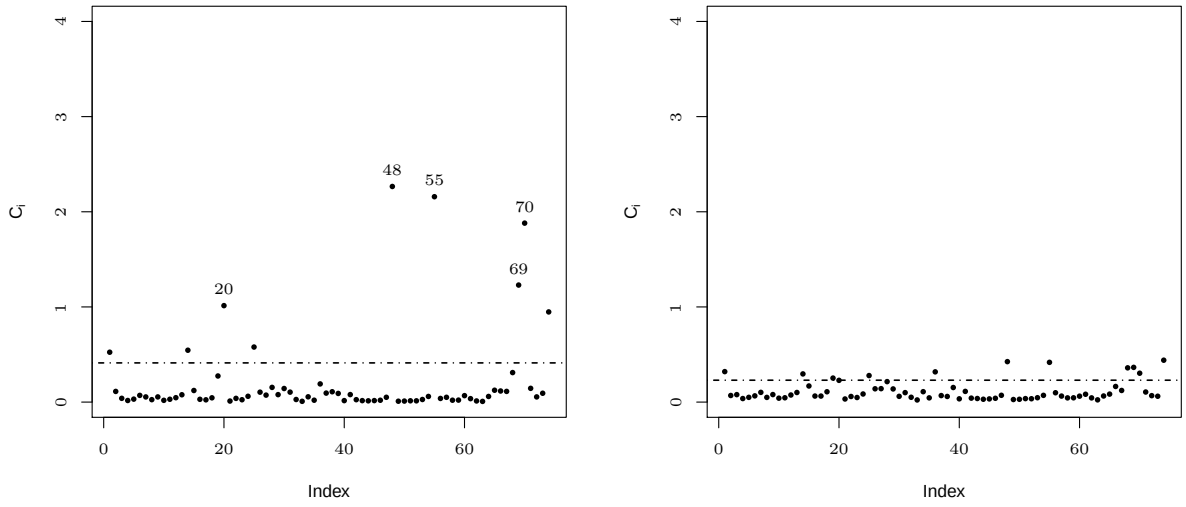
(b) BS-t_2 model

(c) BS_2 model

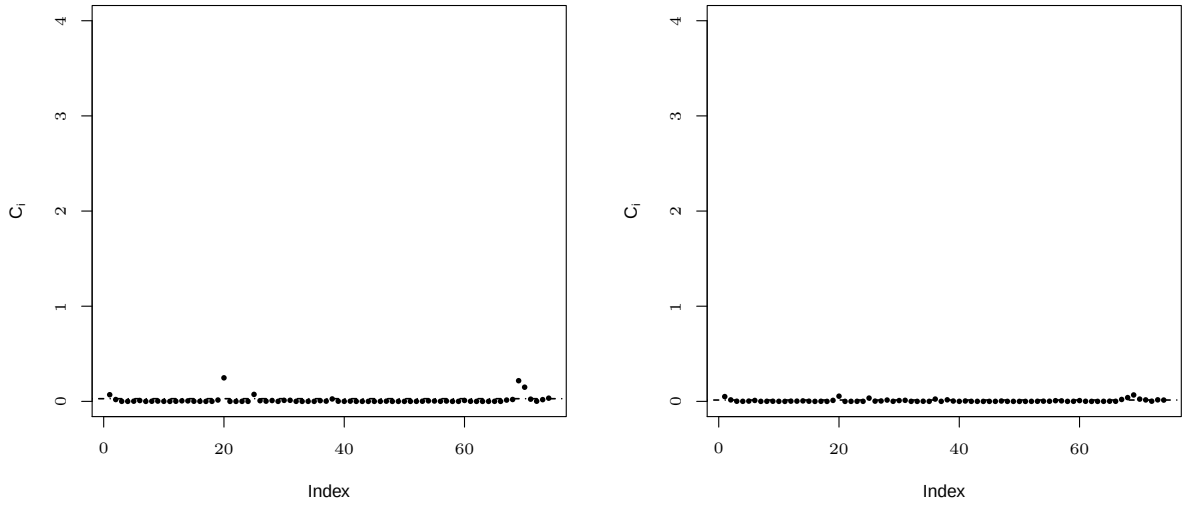
(d) BS-t_2 model

Source: From the author.

Figure 3.6: Total local influence index-plots of case-weight perturbation for $\widehat{\boldsymbol{\beta}}$ (a,b) and $\widehat{\rho}$ (c,d) for the indicated model.



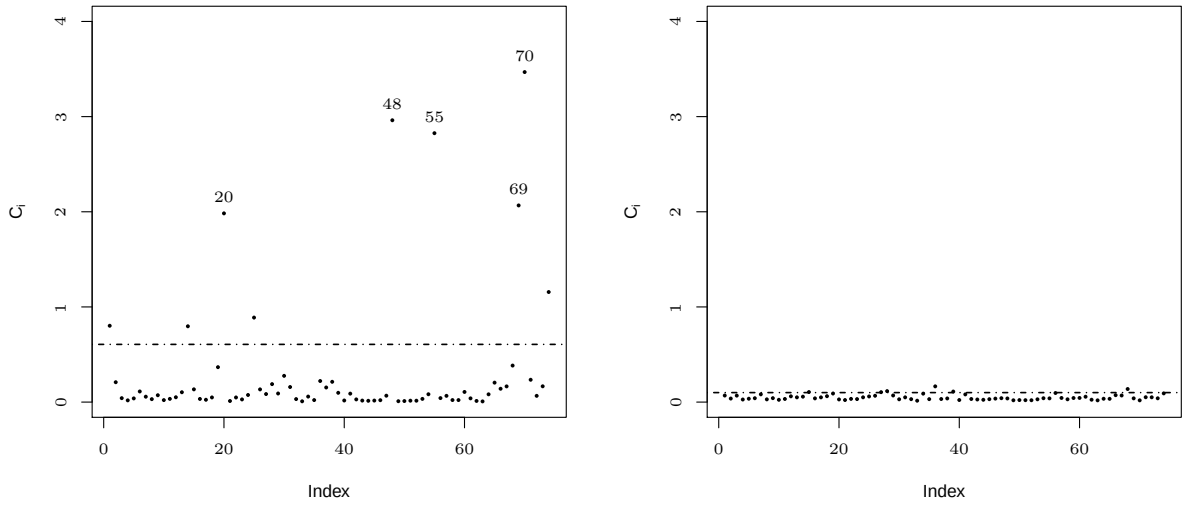
(a) BS_2 model (b) BS-t_2 model



(c) BS_2 model (d) BS-t_2 model

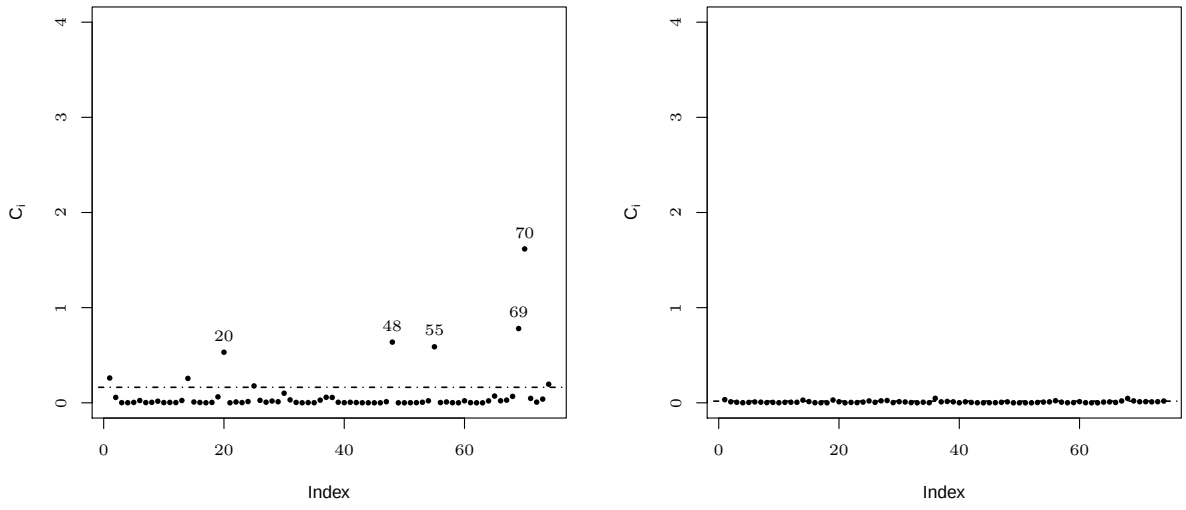
Source: From the author.

Figure 3.7: Total local influence index-plots of correlation perturbation for $\widehat{\boldsymbol{\theta}}$ (a,b) and $\widehat{\alpha}$ (c,d) for the indicated model.



(a) BS_2 model

(b) BS-t_2 model

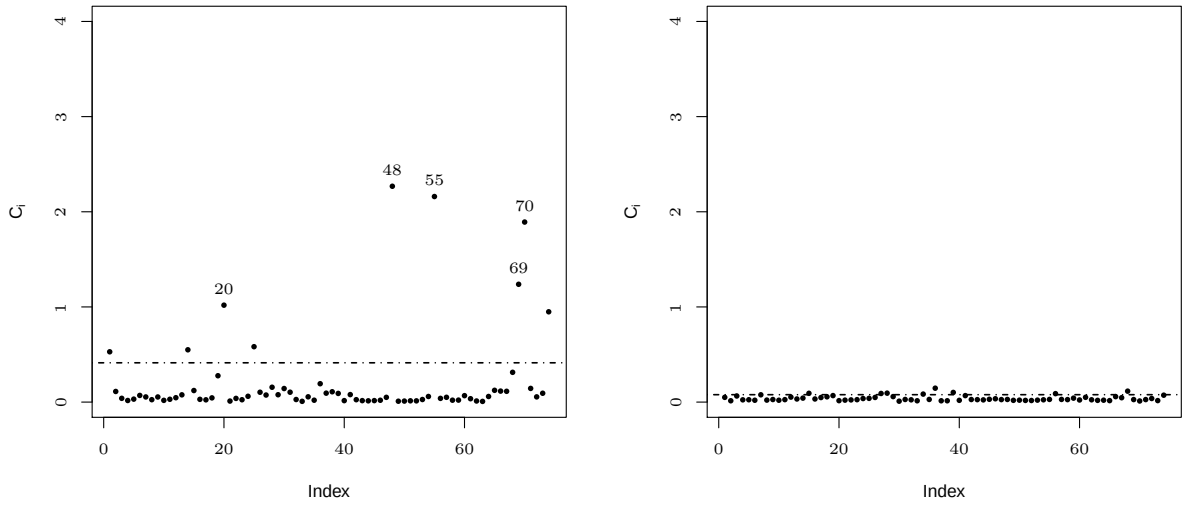


(c) BS_2 model

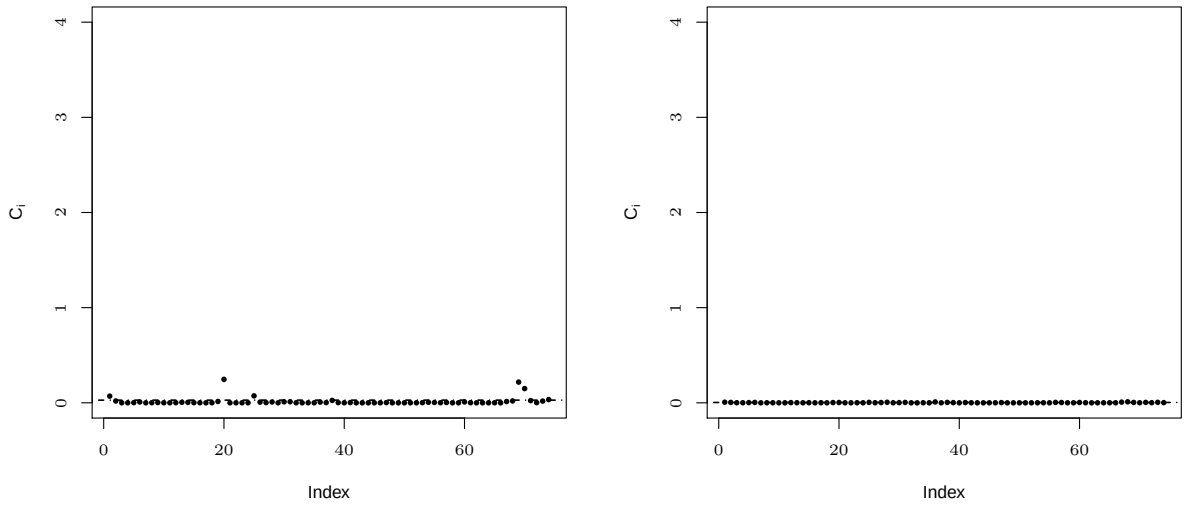
(d) BS-t_2 model

Source: From the author.

Figure 3.8: Total local influence index-plots of correlation perturbation for $\widehat{\boldsymbol{\beta}}$ (a,b) and $\widehat{\rho}$ (c,d) for the indicated model.



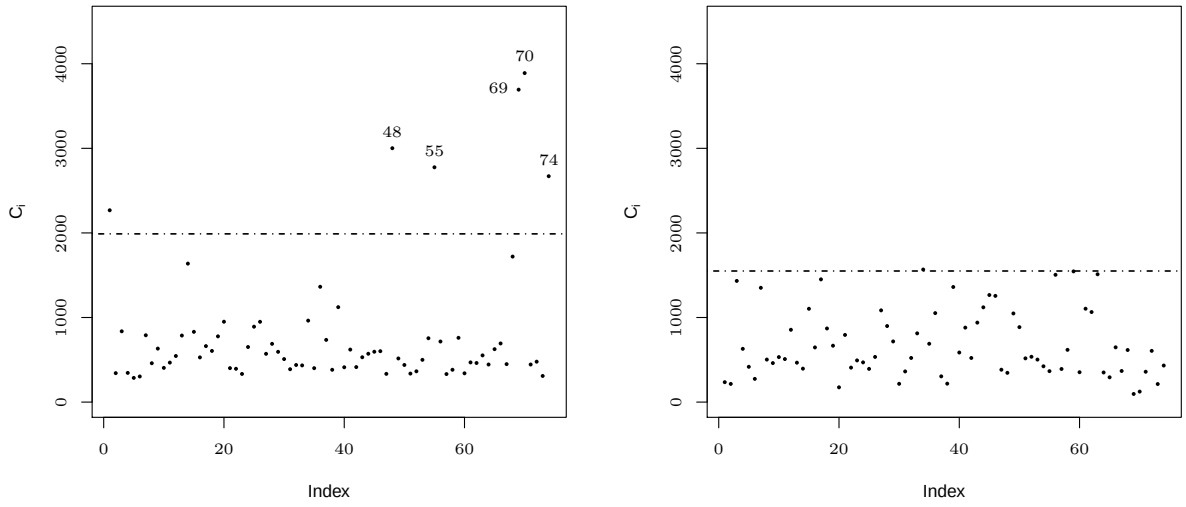
(a) BS_2 model (b) BS-t_2 model



(c) BS_2 model (d) BS-t_2 model

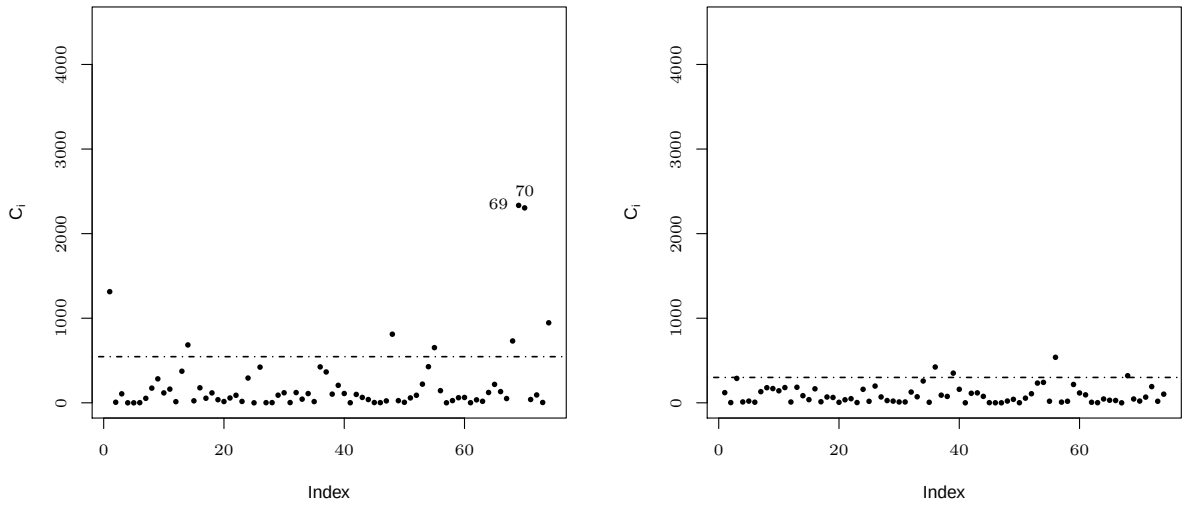
Source: From the author.

Figure 3.9: Total local influence index-plots of covariate perturbation for $\widehat{\boldsymbol{\theta}}$ (a,b) and $\widehat{\alpha}$ (c,d) for the indicated model.



(a) BS_2 model

(b) BS-t_2 model

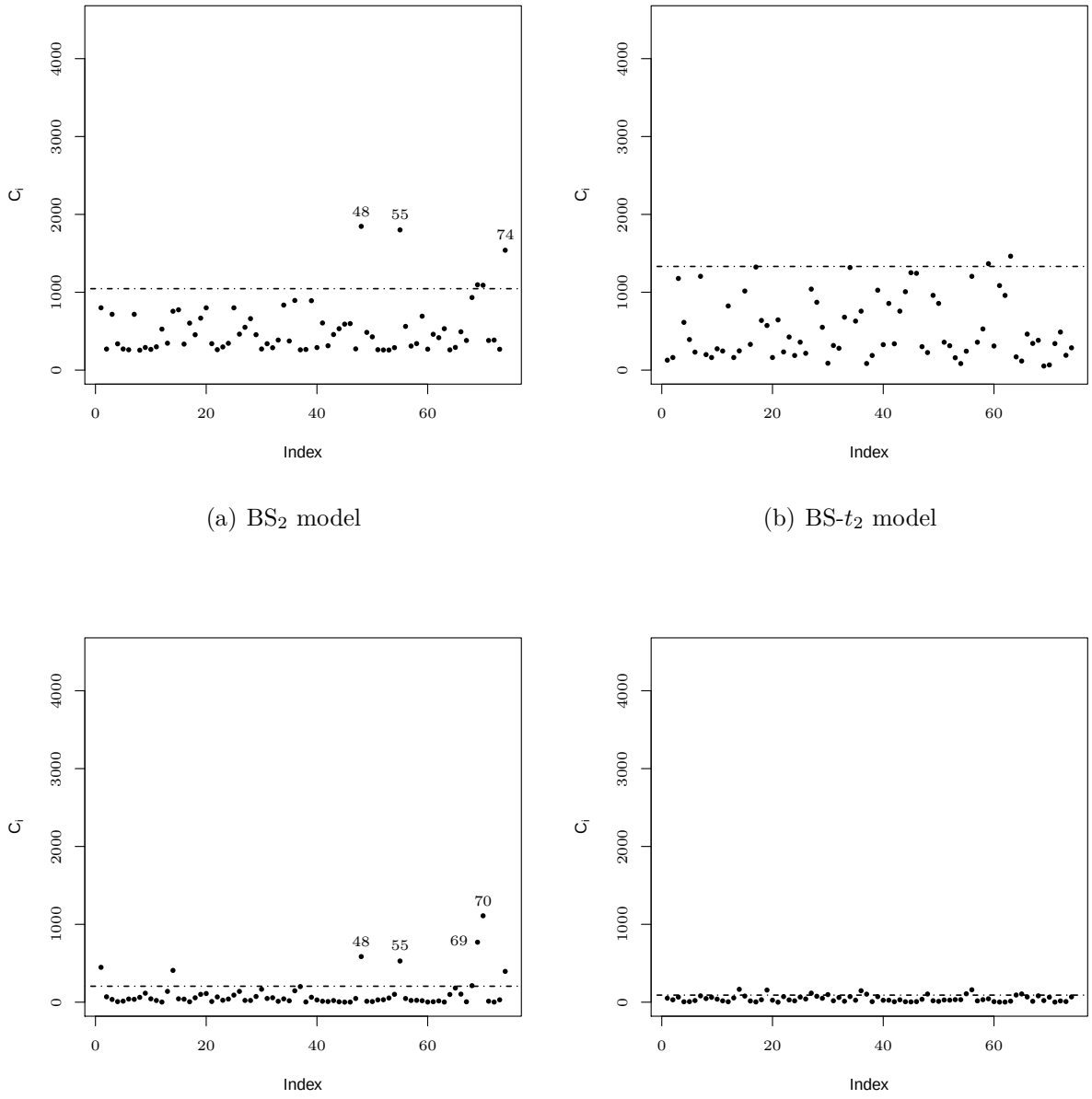


(c) BS_2 model

(d) BS-t_2 model

Source: From the author.

Figure 3.10: Total local influence index-plots of covariate perturbation for $\widehat{\beta}$ (a,b) and $\widehat{\rho}$ (c,d) for the indicated model.



(a) BS_2 model

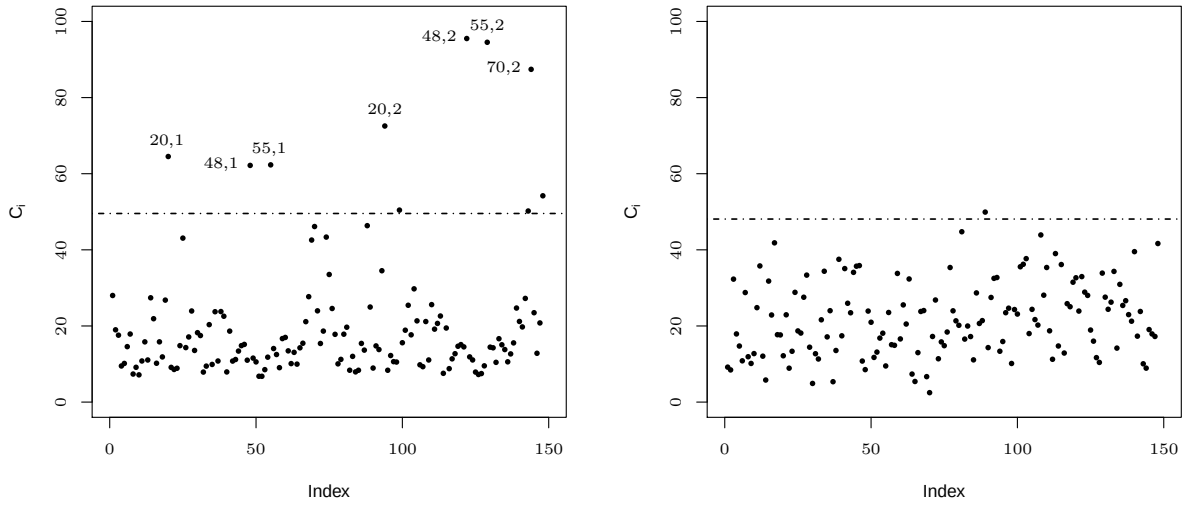
(b) BS-t_2 model

(c) BS_2 model

(d) BS-t_2 model

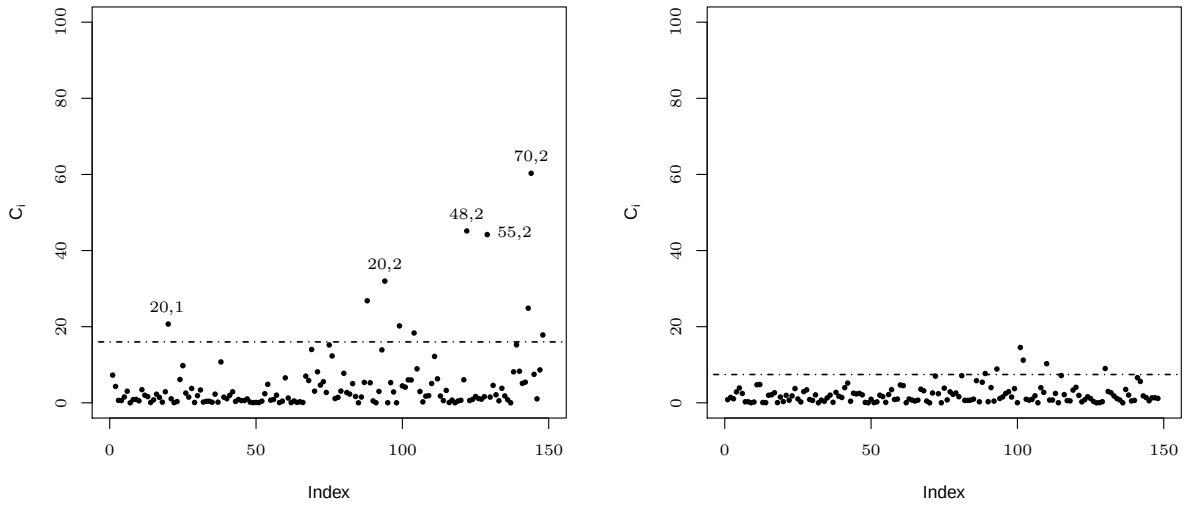
Source: From the author.

Figure 3.11: Total local influence index-plots of response perturbation for $\widehat{\boldsymbol{\theta}}$ (a,b) and $\widehat{\alpha}$ (c,d) for the indicated model.



(a) BS_2 model

(b) BS-t_2 model

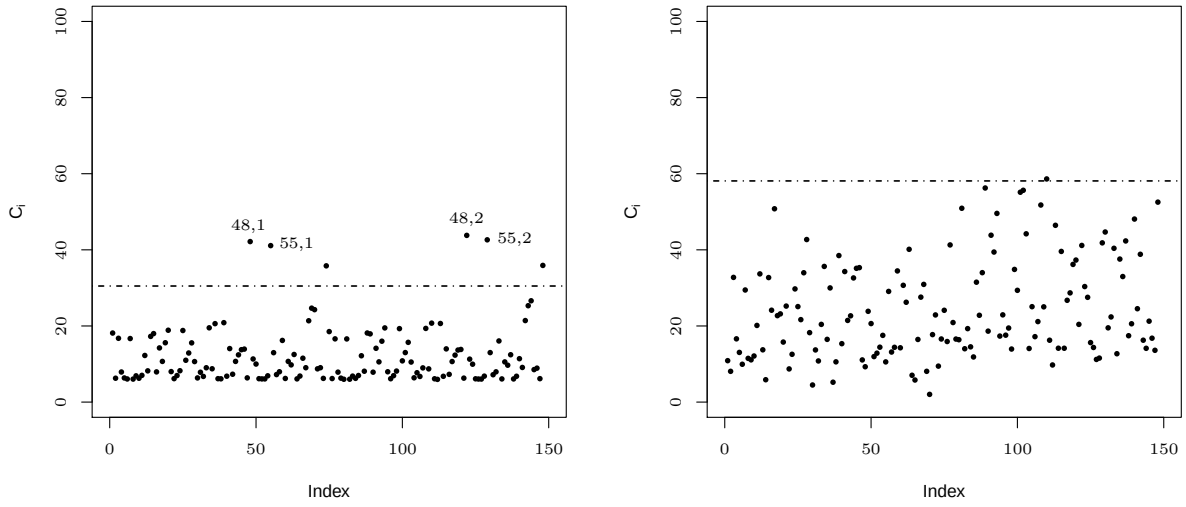


(c) BS_2 model

(d) BS-t_2 model

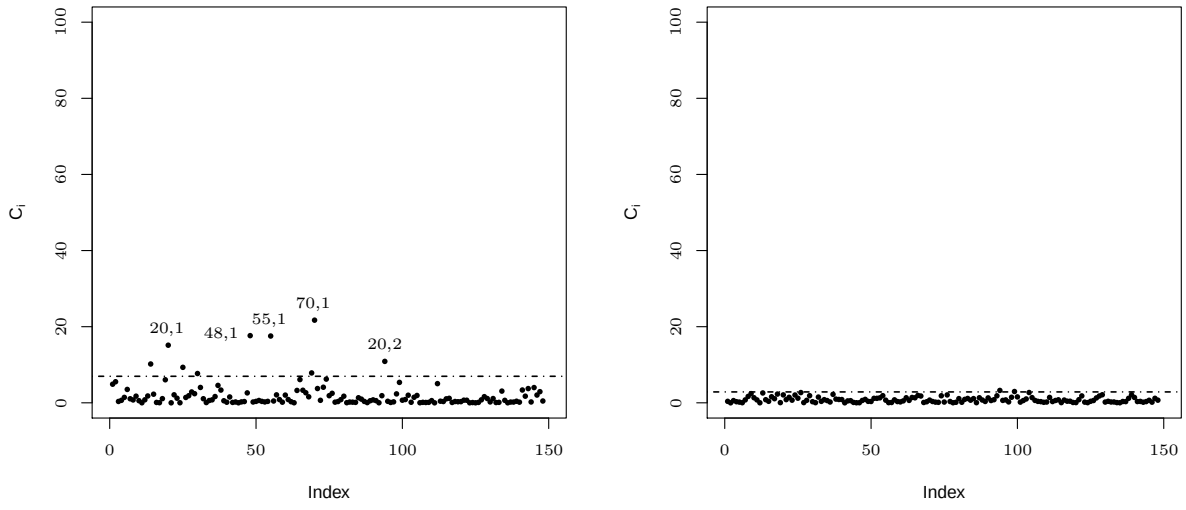
Source: From the author.

Figure 3.12: Total local influence index-plots of response perturbation for $\widehat{\boldsymbol{\beta}}$ (a,b) and $\widehat{\rho}$ (c,d) for the indicated model.



(a) BS_2 model

(b) BS-t_2 model

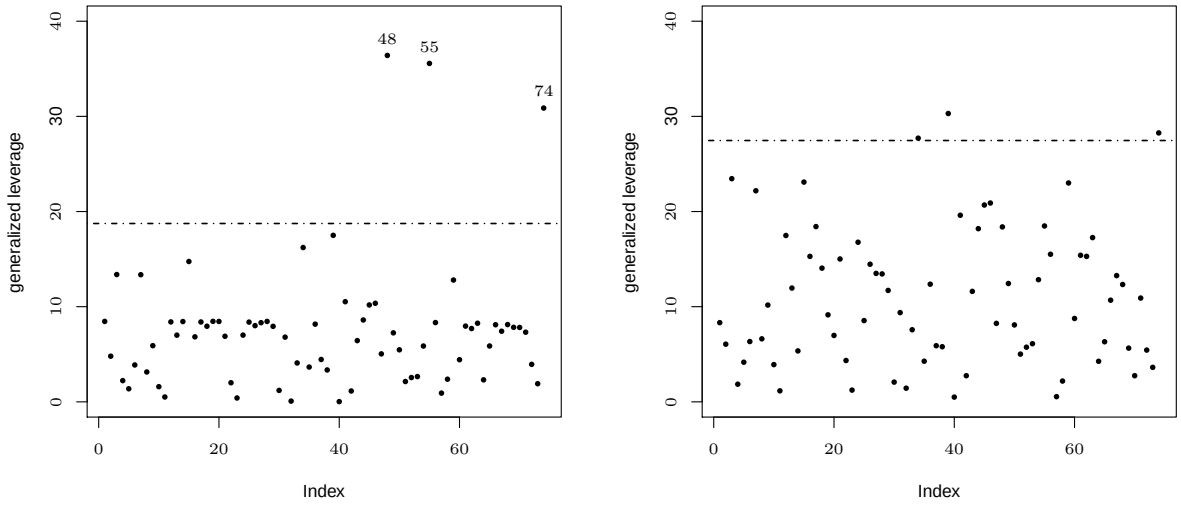


(c) BS_2 model

(d) BS-t_2 model

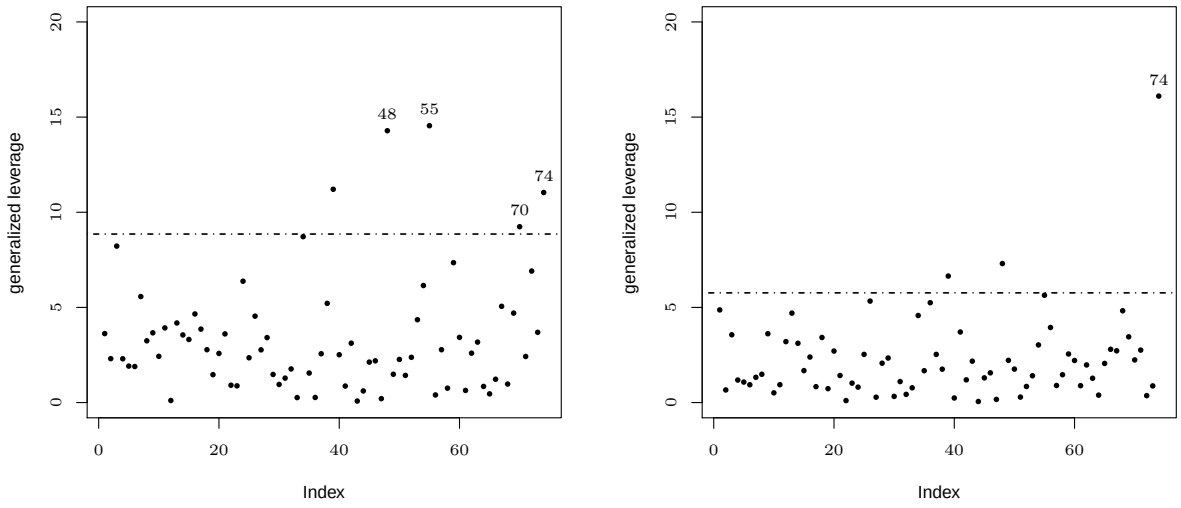
Source: From the author.

Figure 3.13: Index-plots of GL for Y_1 (a,b) and Y_2 (c,d) for the indicated model.



(a) BS_2 model

(b) BS-t_2 model



(c) BS_2 model

(d) BS-t_2 model

Source: From the author.

3.6 Concluding remarks

In this work, we carried out a diagnostic analysis for multivariate generalized Birnbaum-Saunders log-linear regression models. We have considered the Mahalanobis distance for evaluating the suitability of the distributional assumption by transforming this distance with the Wilson-Hilferty approximation and then by using goodness-of-fit techniques. In addition, the Mahalanobis distance has been employed as a global influence measure to detect multivariate outliers. Furthermore, we have obtained appropriate matrices for assessing local influence under perturbation schemes of case-weight, correlation matrix, response variable and a continuous explanatory variable. In addition, we have use the generalized leverage method for detecting effect of the observed response on its own estimated value. We have implemented the obtained results in the `R` software. These results have been applied to real-world multivariate data to illustrate its good performance. We have considered multivariate data useful for regression models based on computed tomography to study the bone quality. The BS distribution is appropriate for modeling physical properties of bone and its densities extracted from CT data, which are related to mechanical properties, as strength, which plays a role as human bone ages affecting its fatigue properties that can be well described by this distribution. Goodness-of-fit tests have concluded that BS-t_2 log-linear regression model was quite appropriate to describe these type of data.

CHAPTER 4

Robust multivariate control charts based on Birnbaum-Saunders distributions

4.1 Resumo

Neste capítulo, desenvolvemos uma metodologia robusta para a construccão de gráficos de controle de qualidade multivariados baseada em distribuições Birnbaum-Saunders generalizadas usando a estatística de Hotelling. Usando o método bootstrap paramétrico, encontramos aproximações da distribuição dessa estatística e obtemos os limites de controle. Realizamos um estudo de simulação de MC para avaliar a metodologia proposta. Os resultados indicam o bom desempenho para fornecer alertas precoces de processos fora de controle. Uma ilustração com dados reais de qualidade do ar de Santiago-Chile é fornecida. Essa ilustração mostra que a metodologia desenvolvida pode ser útil para alertar os episódios de poluição do ar extremo, evitando efeitos adversos na saúde humana.

4.2 Introduction

Multivariate control charts are powerful and simple visual tools for monitoring the quality of a multivariate process by determining whether it is in control or out-of-control. These charts show how several variables jointly influence such a process. Hotelling (1947) was the first to analyze correlated random variables in quality control. He developed a procedure

based on a statistical distance, which generalizes the t statistic. This was later named the Hotelling T^2 statistic in his honor. The T^2 statistic is a useful tool for multivariate normal process control. After this contribution, there was no significant work done in this field until the last few decades, when interested in multivariate statistical quality control was revived due to advances in computing. Since then, a number of authors have done some research in the area of multivariate quality control based on the normal distribution; see, for example, Tracy et al. (1992), Lowry and Montgomery (1995), Liu (1995), Runger et al. (1996), Sullivan and Woodall (1996, 1998), Nedumaran and Pignatiello (2000), Mason et al. (2001), Maravelakis et al. (2002), Noorossana et al. (2002), Yang and Trewn (2004), Bersimis et al. (2007), Yen and Shiau (2010), Chenouri and Variyath (2011) and Yen and Tang (2012).

In traditional control charts, it is usually supposed that the data follow a normal distribution. However, there are many practical applications where the normality assumption is not fulfilled, because the data exhibit heavy tails or skewness. In recent years, some authors have developed multivariate control charts using non-normal distributions. Liu and Tang (1997) proposed that, when we are not sure of the normality of the data, the bootstrap method may be used to determine the control limits. They showed that estimation of the control limits by means of this method is generally better than the normal approach based on the central limit theorem. Chou et al. (2001) introduced a method to determine control limits, working with individual observations, in cases where the data come from a non-normal distribution. Furthermore, Stoumbos and Sullivan (2002) investigated the effects of non-normality on the statistical performance of the exponentially weighted moving average chart, and its special case, the Hotelling chi-squared chart. Phaladiganon et al. (2011) presented a bootstrap-based T^2 control chart to establish its limits when the observed process data are not normally distributed. The limits of bootstrap-based T^2 control charts are calculated based on the percentiles of the T^2 statistic derived from bootstrap samples. Alfaro and Ortega (2012, 2013) proposed robust Hotelling T^2 charts based on the multivariate t distribution.

Traditional control charts use mean and variance-covariance matrix estimators, which are sensitive to outliers in Phase I; see details of Phases I and II of a control chart in Subsection 4.3.2. A univariate outlier is defined as an observation that deviates greatly from other data other points so as to arouse suspicion that it was generated by a different mechanism; see Hawkins (1980). Multivariate outliers are considered to be atypical by not taking the value in a given random variable, but in all the multivariate set of random variables; see Becker and Gather (1999). Multivariate outliers are more difficult to identify than the univariate outliers, since they cannot be considered *outliers* like the situation when you have a single variable under study. Their presence has further detrimental effects than the univariate case, because not only do they distort the position (mean) or dispersion (variance) of the ob-

servations, they also distort the correlations between the variables; see Rocke and Woodruff (1996). Multivariate outliers greatly influence the resulting estimates and cause any out-of-control observations to remain undetected. The identification of outliers in multivariate data is usually based on the MD; see Marchant et al. (2016b). However, sometimes outliers do not have a large MD, which is known as *masking effect*. Such a situation is due to the fact that the estimators based on the model used to generate the MD are statistically non-robust; see, for example, Rocke and Woodruff (1996) and Becker and Gather (1999). Masking effects occur when a group of extreme observations distorts the estimates of the mean and/or variance-covariance matrix, resulting in a small distance from the outlier to the mean.

Jensen et al. (2007), Chenouri et al. (2009), and Alfaro and Ortega (2013) studied the behaviour of different robust alternatives for estimating the process parameters in multivariate control charts. This allowed the researchers to avoid the negative effect of outliers. Alfaro and Ortega (2012) proposed a robust T^2 control chart to protect it in the presence outliers in Phase I, when the data are multivariate t distributed, thus improving the behaviour in Phase II.

The main objectives of this chapter are: (i) to propose a robust methodology for multivariate GBS control charts; (ii) to evaluate its performance by means of MC simulations; and (iii) to apply it to multivariate real-world data. To meet these objectives, we develop the multivariate GBS control charts and determine their limits with the bootstrap method, assuming an in-control status. Then, we estimate the parameters of a multivariate BS and BS-t distributions with the ML method. Fitting tools based on the MD are derived to evaluate the stability and adequacy of these charts. In addition, we consider the MD as an influence measure to detect multivariate outliers. We carry out a MC simulation study to assess the performance of the proposed methodology. We implement the obtained results by a computational routine in the `R` software; see `www.R-project.org`. We employ this routine to carry out an illustration with multivariate air quality data from the city of Santiago, Chile. Note that, the BS distribution was formalized as an adequate model to describe environmental data using the proportionate-effect law; Leiva et al. (2015b).

This chapter is structured as follows. In Section 4.3, we provide some preliminary results, whereas Section 4.4 derives the methodology based on multivariate GBS control charts. In Section 4.5, an MC simulation study is carried out to evaluate the performance of proposed methodology. In Section 4.6, we apply this methodology to multivariate air quality real-world data. Finally, Section 4.7 discusses some conclusions of this chapter.

4.3 Background

In this section, we provide ML estimation aspects of log-GBS distributions. In addition, we discuss some general aspects of traditional multivariate quality control charts.

4.3.1 ML estimation in multivariate log-GBS distributions

Let $\boldsymbol{Y}_1, \ldots, \boldsymbol{Y}_n$ be a random sample from the p-variate log-GBS distribution with $\mathrm{E}(\boldsymbol{Y}_i) = \boldsymbol{\mu}_i$ for $i = 1, \ldots, n$, and let $\boldsymbol{y} = (\boldsymbol{y}_1, \ldots, \boldsymbol{y}_n)^\top$ be their observed values. Then, the log-likelihood function for $\boldsymbol{\theta} = (\boldsymbol{\alpha}^\top, \boldsymbol{\mu}^\top, \mathrm{svec}(\boldsymbol{\Psi})^\top)^\top$, is given by

$$\ell(\boldsymbol{\theta}; \boldsymbol{y}) = \sum_{i=1}^{n} \ell_i(\boldsymbol{\theta}) = \sum_{i=1}^{n} \left[\log(f_{\mathrm{EC}_p}(\boldsymbol{B}_i; \boldsymbol{\Psi}, g^{(p)})) + \sum_{j=1}^{p} \log\left(\frac{1}{\alpha_j} \cosh\left(\frac{y_{ij} - \mu_j}{2} \right) \right) \right], \quad (4.1)$$

where f_{EC_p} is given in (1.11) and $\boldsymbol{B}_i = (B_{i1}, \ldots, B_{ip})^\top$, with elements

$$B_{ij} = \frac{2}{\alpha_j} \sinh\left(\frac{y_{ij} - \mu_j}{2} \right), \quad i = 1, \ldots, n, \quad j = 1, \ldots, p. \quad (4.2)$$

From Table 1.1 and (4.1), if $g^{(p)}$ is the multivariate normal or t kernel, then the log-likelihood functions for $\boldsymbol{\theta}$ are given respectively by

$$\begin{aligned}
\ell_i(\boldsymbol{\theta}) &= -\frac{p}{2}\log(2\pi) - \frac{1}{2}\log(|\boldsymbol{\Psi}|) - \frac{1}{2}\boldsymbol{B}_i^\top \boldsymbol{\Psi}^{-1}\boldsymbol{B}_i + \sum_{j=1}^{p} \log\left(\frac{1}{\alpha_j} \cosh\left(\frac{y_{ij} - \mu_j}{2} \right) \right), \\
\ell_i(\boldsymbol{\theta}) &= -\log\left(\Gamma\left(\frac{\nu}{2} \right) \right) + \log\left(\Gamma\left(\frac{\nu + p}{2} \right) \right) - \frac{p}{2}\log(\nu\pi) + \left[\frac{\nu + p}{2} \right] \log(\nu) \\
&\quad -\frac{1}{2}\log(|\boldsymbol{\Psi}|) - \left[\frac{\nu + p}{2} \right] \log(\nu + \boldsymbol{B}_i^\top \boldsymbol{\Psi}^{-1}\boldsymbol{B}_i) + \sum_{j=1}^{p} \log\left(\frac{1}{\alpha_j} \cosh\left(\frac{y_{ij} - \mu_j}{2} \right) \right),
\end{aligned}$$

where $\boldsymbol{B}_i$ is given in (4.1). In order to obtain the ML estimates of the multivariate log-GBS distribution parameters, the log-likelihood function given in (4.1) must be maximized. As we have already been mentioned in previous chapters, the corresponding likelihood equations must be solved by a non-linear iterative procedure. In our case, initial values for this iterative procedure may be:

(i) $\widehat{\mu}_j^{(0)} = \mathrm{med}(y_{1j}, \ldots, y_{nj})$, for $j = 1, \ldots, p$, where "med" denotes the median of the data.

(ii)

$$\widehat{\alpha}_j^{(0)} = \sqrt{\frac{4}{n} \sum_{i=1}^{n} \left(\sinh\left(\frac{y_{ij} - \widehat{\mu}_j^{(0)}}{2} \right) \right)^2}, \quad j = 1, \ldots, p,$$

where $\widehat{\mu}_j^{(0)} = \mathrm{med}(y_{1j}, \ldots, y_{nj})$.

(iii) $\widehat{\boldsymbol{\Psi}}^{(0)} = \boldsymbol{D}(\widehat{\boldsymbol{\Sigma}}^{(0)})^{-1/2}\widehat{\boldsymbol{\Sigma}}^{(0)}\boldsymbol{D}(\widehat{\boldsymbol{\Sigma}}^{(0)})^{-1/2}$, where $\widehat{\boldsymbol{\Sigma}}^{(0)} = (1/n)\sum_{i=1}^{n}\widehat{\boldsymbol{B}}_i^{(0)}(\widehat{\boldsymbol{B}}_i^{(0)})^{\top}$, with $\widehat{\boldsymbol{B}}_i^{(0)}$ having elements as in (4.2) given by

$$\widehat{B}_{ij}^{(0)} = \frac{2}{\widehat{\alpha_j}^{(0)}}\sinh\left(\frac{y_{ij}-\widehat{\mu}_j^{(0)}}{2}\right), \quad i = 1,\ldots,n, \quad j = 1,\ldots,p,$$

and $\widehat{\alpha_j}^{(0)}$ being computed as in (ii).

4.3.2 Multivariate quality control charts

The construction of a multivariate control chart consists of:
(i) Defining a center line (CL), which represents the expected value of the quality characteristics for all subgroups (samples).
(ii) Establishing lower and upper control limits (UCL and LCL, respectively), which set a distance above and below the CL.
(iii) Plotting points, each of which represents a subgroup of data sampled from the process, representing the mean vector or some other statistic.
UCL and LCL provide a visual display for the expected amount of data dispersion. The control limits are based on the actual behavior of the process, not the desired behavior or specification limits. A process can be in-control and yet not be capable of meeting requirements; see Leiva et al. (2014a).

In the construction of multivariate control charts, Alt (1985) defined the two following phases:

Phase I: An in-control data set is analyzed to estimate the parameters, obtain the control limits and identify multivariate outliers.

Phase II: The estimates and control limits are used to check the data obtained during the process.

Therefore, Phase II consists of using the control chart to detect any departure of the underlying process from a prefixed mean value ($\boldsymbol{\mu}_0$) called the target. In this phase, the idea is that the number of subgroups m collected is greater than 25. In Phase II, note that the data are not taken from an in-control process, unless there is a clear indication of no changes in the process.

The average run length (ARL) is the mean number of points that must be plotted before a point to indicate an out-of-control condition. ARL can be used to evaluate the performance of a control chart and it is calculated as

$$\text{ARL} = \frac{1}{\Pr(\text{one point plotted out of control})}.$$

An in-control ARL is denoted by ARL_0 and expressed as $\mathrm{ARL}_0 = 1/\eta$, where η represents the probability of type-I error. ARL_0 usually takes values in $\{200, 370.4, 500, 1000\}$. On the one hand, the probability that an observation is considered as out of control, when the process is actually in control, indicates a false alarm rate (FAR), which is usually in the set $\{0.005, 0.0027, 0.002, 0.001\}$. On the other hand, the probability of a true out-of-control signal can be obtained from $1 - \Pr(\text{type-II error})$. An out-of-control ARL is denoted by ARL_1 and calculated as $\mathrm{ARL}_1 = 1/[1-\gamma]$, where $\gamma = \Pr(\text{type-II error})$.

4.4 Multivariate GBS quality control charts

In this section, we develop a methodology for multivariate GBS control charts. We use the bootstrap method to determine the control limits in Phase I. Then, we formulate this chart to monitor a process in Phase II.

4.4.1 T^2 statistic for log-GBS distributions

Let $\boldsymbol{Y}_i = (Y_{i1}, \ldots, Y_{ip})^\top \in \mathbb{R}^p$ be a vector that represents p quality characteristics of the ith item in the subgroup for $i = 1, \ldots, n$, where n is the subgroup size and $\boldsymbol{\mu}_0$ is the mean vector of an in-control process (target). Assume that $\boldsymbol{Y}$ follows a p-variate log-GBS distribution, that is, $\boldsymbol{Y} \sim \text{log-GBS}_p(\boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\Psi}, g^{(p)})$, and that there are k subgroups each of size $n > 1$ available from the process. Furthermore, suppose that the vectors $\boldsymbol{Y}_i$ are independent over time. We are interested in testing the hypotheses

$$\mathrm{H}_0\colon\ \boldsymbol{\mu} = \boldsymbol{\mu}_0 = (\mu_{01}, \ldots, \mu_{0p})^\top \quad \text{versus} \quad \mathrm{H}_1\colon\ \boldsymbol{\mu} \neq \boldsymbol{\mu}_0. \tag{4.3}$$

An adaptation of the Hotelling T^2 statistic presented in Gupta et al. (2013, pp. 201-216) can be used for testing the hypotheses in (4.3) as follows. From Property (D1) and considering that

$$\boldsymbol{b}_i = \left(2\sinh\left(\frac{Y_{i1} - \mu_{01}}{2}\right), \ldots, 2\sinh\left(\frac{Y_{ip} - \mu_{0p}}{2}\right)\right)^\top$$

has a p-variate EC distribution, that is, $\boldsymbol{b}_i \sim \mathrm{EC}_p(\boldsymbol{0}_p, \boldsymbol{D}(\boldsymbol{\alpha})\boldsymbol{\Psi}\boldsymbol{D}(\boldsymbol{\alpha}), g^{(p)})$ for $i = 1, \ldots, n$. Then, we obtain a Hotelling T^2 statistic adapted for multivariate log-GBS distributions as

$$T^2 = n[n-1]\overline{\boldsymbol{b}}^\top \boldsymbol{C}^{-1}\overline{\boldsymbol{b}}, \tag{4.4}$$

where $\overline{\boldsymbol{b}} = \sum_{i=1}^n \boldsymbol{b}_i/n$ and $\boldsymbol{C} = \sum_{i=1}^n \boldsymbol{b}_i\boldsymbol{b}_i^\top$. Note that, if $\boldsymbol{Y} \sim \text{log-BS}_p(\boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\Psi})$, T^2 given in (4.4) follows a Fisher distribution with p and $n-p$ DFs, that is, $T^2 \sim \mathcal{F}(p, n-p)$; see Kundu (2015a). However, for the wide family of p-variate log-GBS distributions, this result

is not valid. Particularly, for the multivariate t distribution, Kotz and Nadarajah (2004, pp. 199-200) mentioned that the T^2 statistic has no direct results, which is transferred to p-variate log-BS-t distributions. We propose a multivariate GBS control chart based on the adaptation of the Hotelling T^2 statistic given in (4.4). We discuss a manner to approximate the distribution of this statistic in Subsection 4.4.2. Knowing this distribution, we can obtain its quantiles and then the corresponding control limits.

4.4.2 A bootstrap distribution for the T^2 statistic

As mentioned, for data with a multivariate log-BS-t distribution, the associated T^2 statistic distribution is not known in closed form. We use the parametric bootstrap method to approach this distribution; see Hall (1992). To generate random vectors from p-variate log-BS and log-BS-t distributions, and then to obtain the bootstrap distribution of the T^2 statistic given in (4.4), we use Algorithms 1 and 2, respectively; see Section 2.3.

4.4.3 Phase I

As mentioned in Subsection 4.3.2, control limits must be obtained in Phase I. According to Duncan (1986), Phase I also includes the establishment of the process being statistically in control. Implementation of control charts requires their limits to be generated. Algorithm 5 details how to compute the control limits with the bootstrap distribution of the T^2 statistic defined in (4.4); see Subsection 4.4.2. Note that control charts usually have both LCL and UCL, but sometimes only an UCL is considered; see, for example, Alfaro and Ortega (2012).

4.4.4 Phase II

As mentioned in Subsection 4.3.2, multivariate GBS quality control charts must be used in Phase II to test if the process remains in control when future (new) subgroups, whose statistic is denoted by T^2_{new}, are collected. Then, multivariate GBS control charts are based on the sequence of the T^2_{new} statistic defined in (4.4), for $l = 1, \ldots, m$, where m represent the subgroups size in this phase. Algorithm 6 details how to construct p-variate control charts based on GBS distributions for process monitoring. In Phase II, it is also necessary to check the distributional assumption by using GOF tools and multivariate methods to detect outliers.

Algorithm 5 Estimation and computation of control limits in Phase I.

1: Collect k subgroups $(\boldsymbol{y}_{1h},\ldots,\boldsymbol{y}_{nh})^\top$ of size $n > 1$ for an in-control process, with $h = 1,\ldots,k$, assuming that the logarithm of the data follows a log-GBS$_p(\boldsymbol{\alpha},\boldsymbol{\mu},\boldsymbol{\Psi},g^{(p)})$ distribution.
2: Compute the ML estimates of $\boldsymbol{\alpha}$, $\boldsymbol{\mu}$ and $\boldsymbol{\Psi}$ using the data of the pooled sample of size $N = k \times n$ collected in Step 1 and check the distributional assumption using GOF tools.
3: Generate a parametric bootstrap sample $(\boldsymbol{y}_1^*,\ldots,\boldsymbol{y}_n^*)^\top$ of size n from a p-variate log-GB distribution using the ML estimates obtained in Step 2 as the distribution parameters.
4: Compute T^2 defined in (4.4) with $(\boldsymbol{y}_1^*,\ldots,\boldsymbol{y}_n^*)^\top$, which is denoted by T^{2*}, assuming a target $\boldsymbol{\mu}_0$.
5: Repeat Steps 3-4 a large number of times (for example, $B = 10,000$) and obtain B bootstrap statistics of T^2, denoted by $T_1^{2*},\ldots,T_{\mathrm{B}}^{2*}$.
6: Fix η as the desired FAR of the chart.
7: Use the B bootstrap statistics obtained in Step 5 to find the $100(\eta/2)$th and $100(1-\eta/2)$th quantiles of the distribution of T^2, which are the LCL and UCL for the chart of FAR η, respectively.

Algorithm 6 Process monitoring using the multivariate GBS control chart in Phase II.

1: Take a subgroup of size n, $(\boldsymbol{y}_1,\ldots,\boldsymbol{y}_n)^\top$ say, from the process.
2: Calculate the T^2_{new} statistic from the sample obtained in Step 1.
3: Declare the process as in control if T^2_{new} falls between LCL and UCL obtained in Algorithm 5; otherwise, that is, if the T^2_{new} falls below the LCL or above the UCL, the chart signals a out-of-control condition.
4: Repeat Steps 1-3 for m subgroups taken at regular time intervals.

4.5 Simulation study

In this section, we evaluate the performance of the proposed methodology in Phases I and II. We use the `R` software in all of our calculations for this study.

4.5.1 Phase I

For Phase I, we consider the following simulation scenario. Using Algorithms 1 and 2, we generate B = 10000 bootstrap samples for $k = 20$ and subgroup sizes $n \in \{5, 10, 25, 50, 100\}$ from p-variate log-BS and log-BS-t distributions. For each bootstrap sample, we compute its T^2 statistic with the formula defined in (4.4), obtaining $T_1^{2*},\ldots,T_{10000}^{2*}$. We use an overall FAR $\eta = 0.0027$ to obtain the LCL and UCL based on the 0.27th and 99.73th quantiles of

the 10000 statistics. Table 4.1 shows the UCL and LCL obtained for p-variate log-BS and log-BS-t distributions with $p \in \{2, 3, 4\}$ with true values for their parameters established as:

(a) $\boldsymbol{\alpha} = (0.4, 0.5)^\top$, $\boldsymbol{\mu} = (2, 1)^\top$ and $\psi = 0.8$ for $p = 2$.

(b) $\boldsymbol{\alpha} = (0.4, 0.5, 0.4)^\top$, $\boldsymbol{\mu} = (2, 1, 5)^\top$ and $\boldsymbol{\Psi} = \begin{pmatrix} 1.0 & 0.8 & 0.5 \\ 0.8 & 1.0 & 0.2 \\ 0.5 & 0.2 & 1.0 \end{pmatrix}$ for $p = 3$.

(c) $\boldsymbol{\alpha} = (0.4, 0.5, 0.4, 0.5)^\top$, $\boldsymbol{\mu} = (2, 1, 5, 3)^\top$ and $\boldsymbol{\Psi} = \begin{pmatrix} 1.0 & 0.8 & 0.5 & 0.2 \\ 0.8 & 1.0 & 0.7 & 0.3 \\ 0.5 & 0.7 & 1.0 & 0.5 \\ 0.2 & 0.3 & 0.5 & 1.0 \end{pmatrix}$ for $p = 4$.

These values are chosen with the following criteria: (i) $\alpha_j \leq 0.5$ according to Marchant et al. (2016a); (ii) $\mu_j \in \{1, 2, 3, 5\}$ based on Alfaro and Ortega (2012); and (iii) $\boldsymbol{\psi}_{rs}$ with small, medium and large correlations. We use $\nu = 4$ DF for the log-BS-t distribution according to Marchant et al. (2016a) and references therein for a justification about this value, which often maximizes the log-likelihood function. From Table 4.1, as n increases, the control limits become narrower for both distributions, whereas the UCL decreases progressively. Note that the LCLs are very close to zero, which is a reason this limit is often not calculated and set as zero; see 4.4.3. For a fixed n, the limits of the log-BS-t distribution are narrower than those for the log-BS distribution, which can be attributed to the non-robustness to outliers of the ML estimation of its parameters, possibly affecting the detection of out-of-control conditions in Phase II; see Section 4.2.

4.5.2 Phase II

For Phase II, we generate $m = 30$ new subgroups using the same scenario of Phase I. We calculate T^2_{new} with algorithm 6. We generate M = 5000 MC replications. We perturb one (1) observation in each subgroup and consider three perturbation levels (low, moderate and high), corresponding to one (1), five (5) and ten (10) perturbed subgroups. When a subgroup is not perturbed, we denote it with zero. The performance of the control charts is judged in terms of the detection rate of the new subgroup, which is obtained as the proportion of statistic values that are above the UCL. The results are shown in Table 4.2.

From this table, note that the multivariate BS-t control chart performs well in detecting out-of-control conditions in Phase II. Observe that the low and moderate perturbations are well detected by the control chart. Also, notice that the BS-t chart has perform better as

Table 4.1: Simulated control limits for data following p-variate log-GBS distributions with $\eta = 0.0027$.

			UCL			
	$p = 2$		$p = 3$		$p = 4$	
n	BS	BS-t	BS	BS-t	BS	BS-t
5	11.6027	6.7079	14.8541	8.2944	16.0207	9.3704
10	11.1691	6.0518	13.6852	7.4770	15.9451	8.8128
25	10.7939	5.8128	13.7052	7.1973	15.7863	8.6813
50	10.9625	5.6462	13.5828	6.8775	15.5876	7.9276
100	10.9106	5.3870	13.4448	6.6943	15.5746	7.8409
			LCL			
	$p = 2$		$p = 3$		$p = 4$	
n	BS	BS-t	BS	BS-t	BS	BS-t
5	0.0041	0.0031	0.0472	0.0265	0.1810	0.0578
10	0.0054	0.0031	0.0504	0.0246	0.1727	0.0795
25	0.0050	0.0024	0.0480	0.0218	0.1509	0.0813
50	0.0052	0.0030	0.0530	0.0245	0.1676	0.0672
100	0.0059	0.0031	0.0596	0.0252	0.1576	0.0775

Source: From the author.

the number of variables increases. In addition, as the level of perturbation increases, the detection of out-of-control condition decreases due to the masking effect, especially when ten subgroups are perturbed. However, this effect is attenuated in the case of the BS-t chart due to the robustness of its estimation method in relation to the BS chart (omitted here). Finally, in general, when the subgroup size n increases, the performance for detecting out-of-control conditions also increases.

4.6 Data analysis

In this section, we apply the multivariate GBS control charts and GOF tools implemented by a computational routine in `R` code to real-world air quality data.

4.6.1 Description of the problem and data

Multiple studies show an association between pollutant concentrations such as particulate matter (PM), ozone (O_3), nitrogen dioxide (NO_2) and sulfur dioxide (SO_2). These pollutants produce premature deaths and several cardio-respiratory diseases in children and adults. PM is a contaminant highly associated with mortality and morbidity. PM is classified according

Table 4.2: Out-of-control detection rate in Phase II for indicated p and target in the BS-t distribution.

		$p = 2$	$p = 3$	$p = 4$
n	Perturbation level	$\boldsymbol{\mu}_0 = (2, 1)^\top$	$\boldsymbol{\mu}_0 = (2, 1, 5)^\top$	$\boldsymbol{\mu}_0 = (2, 1, 5, 3)^\top$
5	0	0.4320	0.9790	1.0000
	1	0.4960	0.9810	1.0000
	5	0.2278	0.8526	1.0000
	10	0.1597	0.5922	0.9875
10	0	0.8210	0.9940	0.9950
	1	0.8960	0.9980	0.9960
	5	0.7706	0.9768	0.9704
	10	0.6235	0.9145	0.9037
25	0	0.8500	0.9900	0.9880
	1	0.9040	0.9960	0.9930
	5	0.6736	0.9784	0.9406
	10	0.5983	0.9189	0.8158
50	0	0.8540	0.9970	0.9980
	1	0.8880	0.9980	0.9990
	5	0.5538	0.9680	0.9636
	10	0.4558	0.8525	0.8424
100	0	0.8580	1.0000	1.0000
	1	0.8720	1.0000	1.0000
	5	0.4794	0.9768	0.9796
	10	0.3793	0.8345	0.8536

Source: From the author.

to its diameter, because particle size determines sites of deposition within the respiratory tract. Coarser particles (those with a diameter over 10 μm) do not penetrate into airways. Instead, these particles are deposited in the upper respiratory tract and are cleared by cilia action. Inhalable particles measuring less than 10 μm are called PM10. Those smaller than 2.5 μm are called PM2.5. As size decreases, there is a higher possibility for PM to penetrate deeper into smaller alveoli and airways. In particular, various effects are produced from exposure to PM, but the nature of those induced effects vary according to the PM composition. Indeed, there is evidence of an increase in the risk of cardiovascular diseases and mortality from exposure to PM2.5, which occurs even after short time periods, such as hours or weeks. For more details about PM concentrations, the interested reader is referred to Marchant et al. (2013b).

Because of a combination of meteorological and topographic factors, Santiago, the capital of Chile, endures bad atmospheric ventilation in both winter and summer periods. During

winter, there is an accumulation of PM and gaseous contaminants, whereas increased solar radiation in summer favors ozone-producing photochemical reactions. Gaseous contaminants such as carbon monoxide (CO), NO_2, SO_2, and PM are the main contributors to air quality problems in Santiago, Chile. In addition, NO_2 and SO_2 are precursors of PM; see Marchant et al. (2013b). In this illustration, we use the following variables: (i) PM2.5 in mg/normalized cubic meters (m^3N)–X_1–; and (ii) PM10 in mg/m^3N–X_2–. The data to be considered were collected by the Metropolitan Environmental Health Service. We utilize these data for our analysis below, which are available at `http://www.mma.gob.cl`. The data were collected in 2003 as 1 h (hourly) average values, at (a) Las Condes and (b) Pudahuel monitoring stations, located in Santiago. We select data from these stations mainly because they are more suitable to conditions of low and high stability, allowing us to analyze different pollution patterns. Chilean guidelines for air quality are established by the Ministry of the Environment. The maximum concentrations (in mg/m^3N) according these guidelines are 50 and 150, during 24 hours for PM2.5 and PM10, respectively. These values are considered as the targets in this illustration.

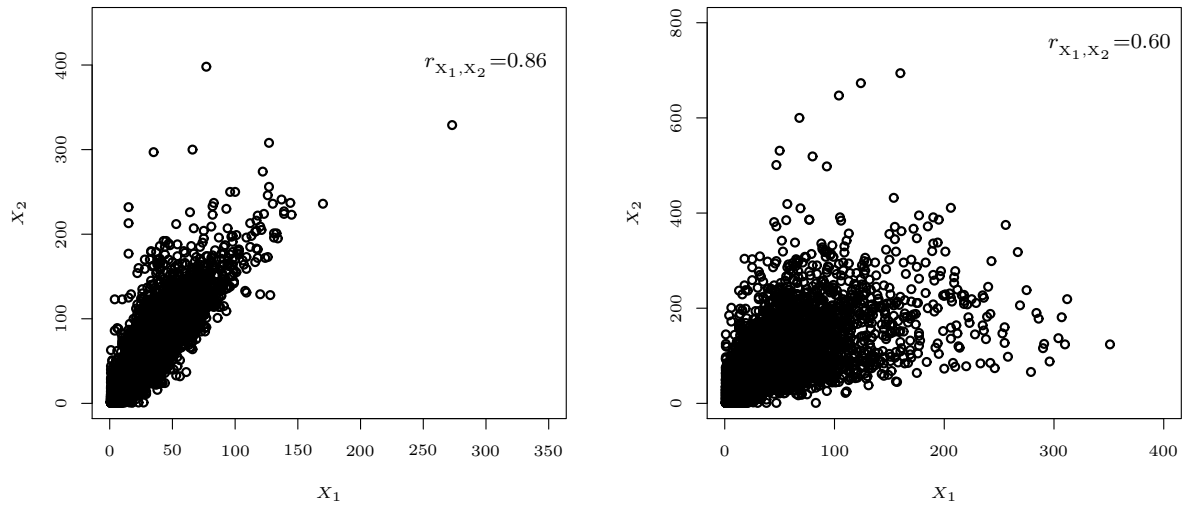
4.6.2 Exploratory data analysis

First, we conduct an exploratory data analysis by computing correlations between X_1 and X_2 in both stations. Figures 4.1 (a) and (b) display the scatter-plots for these variables and their corresponding correlations. From this figure, we detect that there are large and medium correlations between X_1 and X_2 for Las Condes and Pudahuel stations, respectively. Exploratory data analysis for each separate variable at the two stations was conducted and marginal BS distributions seem good candidates for describing these data.

4.6.3 Phase I

We initially use concentrations of the months of January and February to calculate the control limits according to Algorithm 5 with $k = 59$, $n = 24$, $N = 1416$, B = 10000 and FAR $\eta = 0.0027$. We use these months since the air quality is stable (considered under control), because the meteorological and topographical conditions favor no saturation of PM concentrations. We employ the transformed MD with Wilson-Hilferty approximation for obtaining a normal distribution; see Subsection 2.6. Then, a GOF technique is utilized to check step 2 of Algorithm 5; see Marchant et al. (2016b). Figure 4.2 shows the corresponding PP-plots with acceptance bands for a significance level of 5% in Pudahuel station based on BS and BS-t distributions (for Las Condes station, the results are similar). From this figure, we corroborate the good fit of the BS and BS-t distributions to the data in Phase I, which

Figure 4.1: Scatter-plots with their corresponding correlations for the indicated variable for (a) Las Condes and (b) Pudahuel stations.



(a) Las Condes (b) Puhahuel

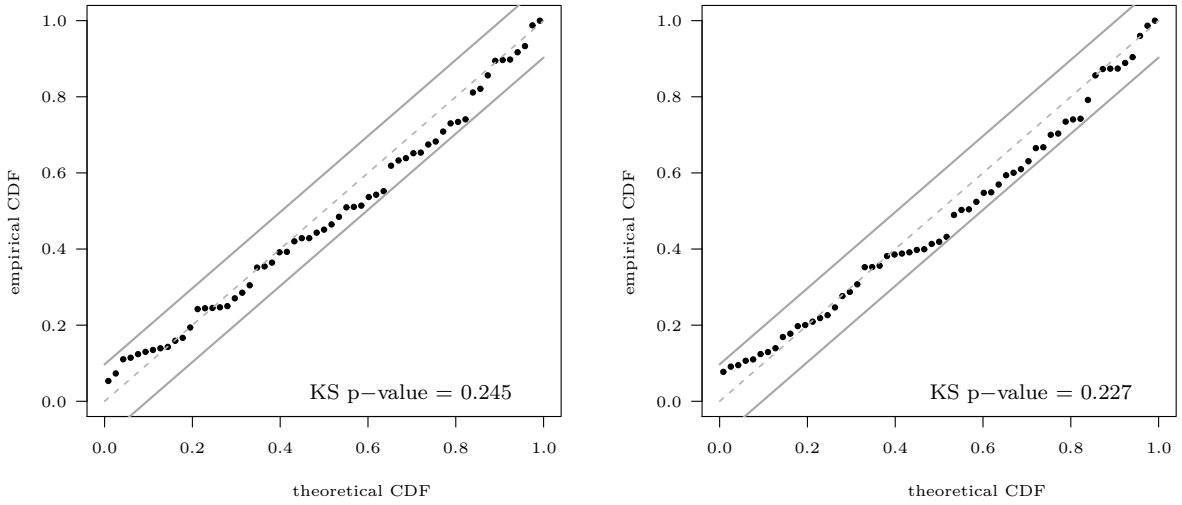
Source: From the author.

is supported by the p-values 0.245 and 0.227, respectively, of the Kolmogorov-Smirnov (KS) test associated with these PP-plots; see Algorithm 4 and Marchant et al. (2016a).

4.6.4 Phase II

We use the control limits obtained in Phase I (Subsection 4.6.3) to monitor April of 2003; see Marchant et al. (2013b). For the control chart of this month, the number of subgroups and the subgroup size are $m = 30$ days and $n = 24$, respectively, giving a total of 720 observations. Once again, we employ the transformed MD to assess the goodness of fit of the most appropriate distribution to these data. Figure 4.3 displays the PP-plots with acceptance bands for a significance level of 5%. From this figure, we detect that the BS-t distribution provides a better fit than the BS distribution for both stations, which is corroborated by the p-values 0.901 (Las Condes/BS-t) and 0.520 (Pudahuel/BS-t) versus 0.095 (Las Condes/BS) and 0.238 (Pudahuel/BS) of the KS test associated with these PP-plots.

Figure 4.2: PP-plots with KS acceptance regions at 5% for transformed MDs with BS (a) and BS-t (b) distributions based on Pudahuel data.

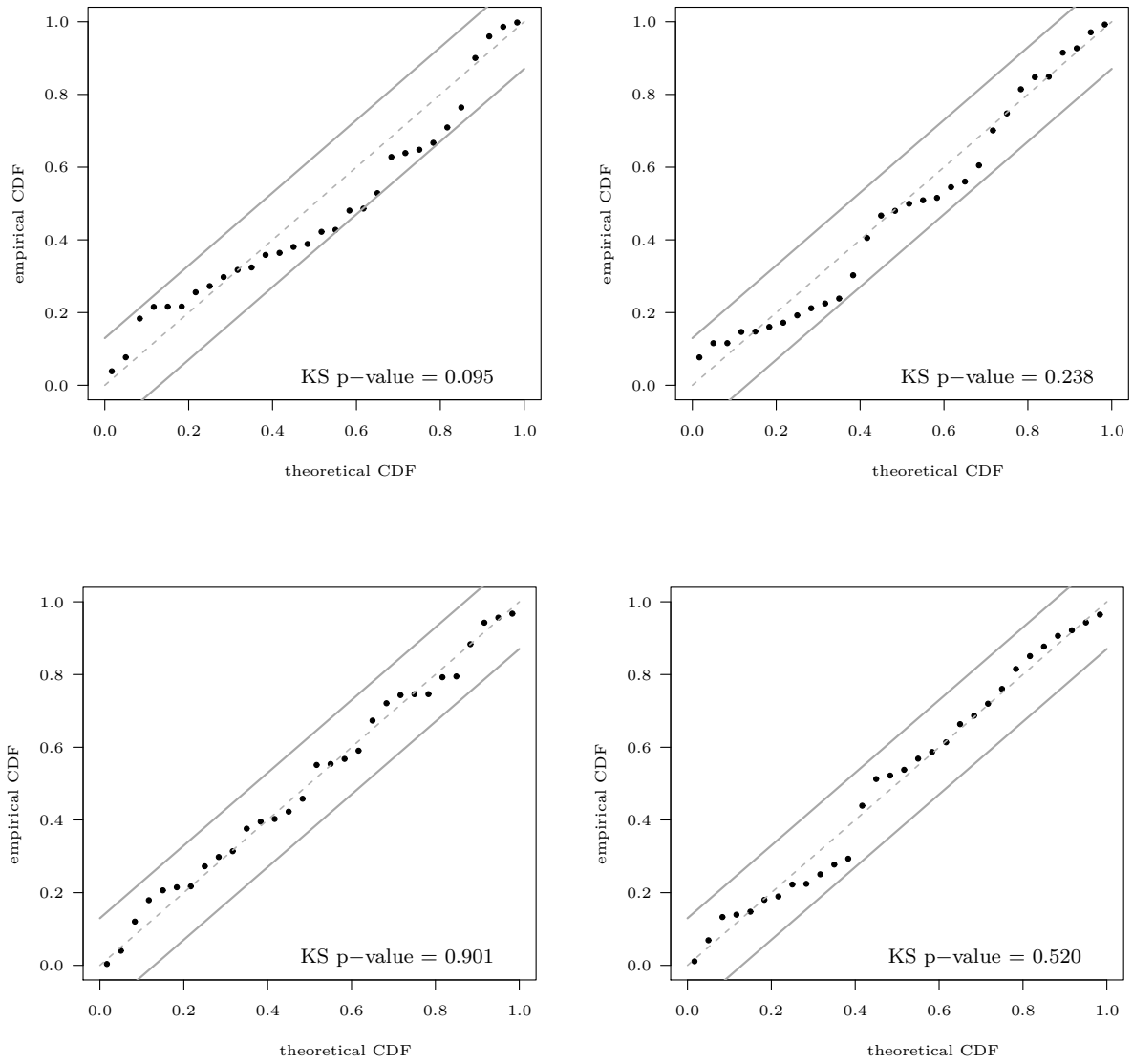


(a) BS (b) BS-t

Source: From the author.

In addition, we utilize the MD as a measure to detect multivariate outliers. Figures 4.4 (a) and (b) depict graphical plots for both stations with the BS-t distribution. From this figure, note that 19 April (labelled 19) is detected as a multivariate outlier. However, this observation does not influence the control charts shown in Figure 4.5 (a) and (b) for both stations due to the robustness of the ML estimation for the BS-t distribution. It is known that the Las Condes station is less contaminated than the other 7 stations (in the year 2003), due to better ventilation at its high altitude; see Marchant et al. (2013b). This can be the reason why such a station does not have points outside of the limits. On the contrary, 10 April (labelled 10) exceeds the limit in the Pudahuel station. Therefore, an environmental alert must be declared as an out-of-control condition for the next day. Note that, if at least one of the air quality monitoring stations presents a dangerous PM level for human health in Santiago, then an out-of-control condition must be declared. The interested reader is referred to CONAMA (1998) for details of the official decree of the Ministry of Environment (CONAMA) of the Chilean government that indicates this regulation. Observe that our criterion is coherent with the official information provided by the Chilean Ministry of Health, which established environmental alerts for the day 11-April-2003; see `www.seremisaludrm.cl/sitio/pag/aire/indexjs3airee001.asp`. Finally, we

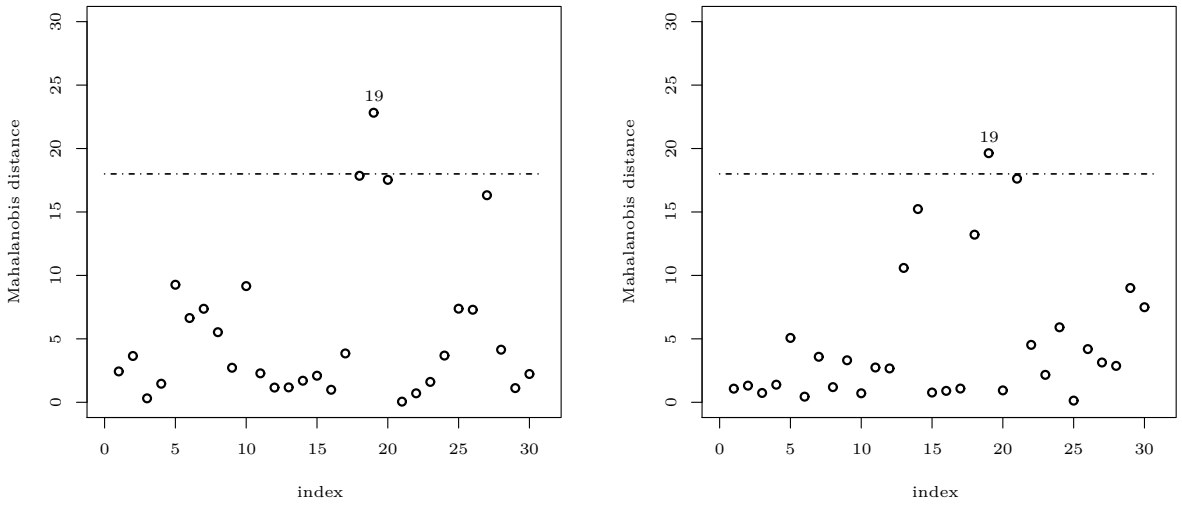
Figure 4.3: PP-plots with KS acceptance regions at 5% for transformed MDs with BS_2 (first panel) and BS-t_2 (second panel) distributions for Las Condes (left) Pudahuel (right) stations based on pollutant data.



Source: From the author.

recommend this methodology based on multivariate GBS control charts because these can be useful for alerting citizens about episodes of extreme air pollution. Thus, pollutants must be monitored to prevent adverse effects on human health for the population of Santiago, Chile. In addition, we show the coherence between our criterion and when Chilean health authority rules environmental alerts in the real-world.

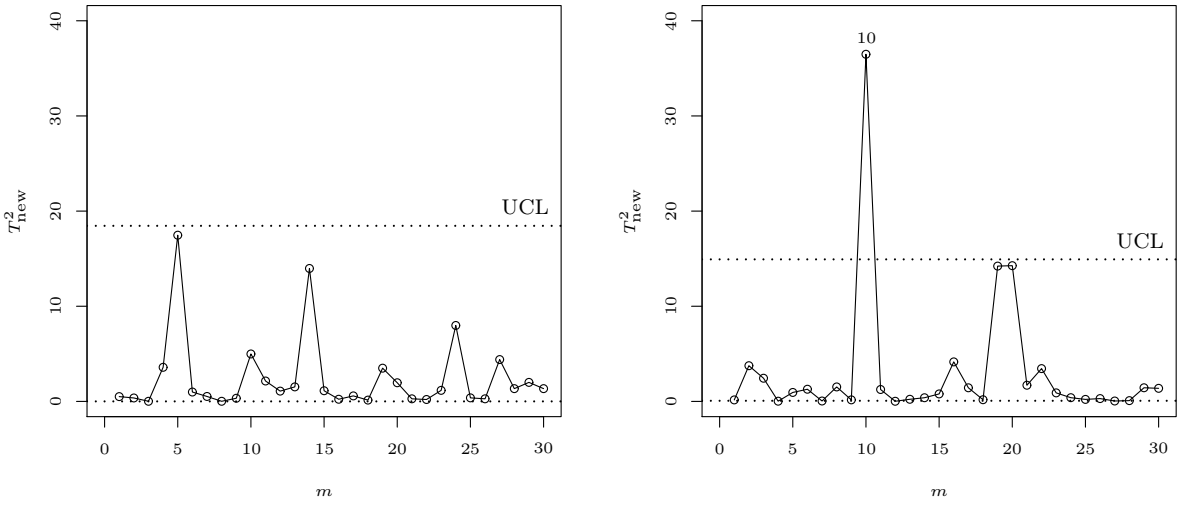
Figure 4.4: Index-plots of MDs for the BS$_t$ Las Condes (a) and BS-t Pudahuel (b) stations based on pollutant data.



(a) Las Condes (b) Pudahuel

Source: From the author.

Figure 4.5: Bivariate BS-t control charts for Las Condes (a) and Pudahuel (b) stations for April 2003 based on pollutant data.



(a) Las Condes (b) Pudahuel

Source: From the author.

4.7 Concluding remarks

In this chapter, we proposed a robust methodology for multivariate control charts based on generalized Birnbaum-Saunders distributions. This methodology estimates the multivariate Birnbaum-Saunders and Birnbaum-Saunders-t parameters with the maximum likelihood method. It also considers the Mahalanobis distance to detect multivariate outliers and evaluate the adequacy of the distributional assumption. For data following a multivariate Birnbaum-Saunders-t model, the distribution of the associated Hotelling statistic is not known in closed form. Then, the methodology uses the bootstrap method to obtain an approximation of this distribution. Once such a distribution is known, the proposed methodology obtains its quantiles to construct the control limits of the multivariate chart. A Monte Carlo simulation study was conducted to evaluate the proposed methodology in Phases I and II. We concluded by means of this simulation study that the multivariate Birnbaum-Saunders-t control chart performs well in the detection of out-of-control conditions in Phase II and behave better than the multivariate Birnbaum-Saunders control chart in both phases. We illustrated the proposed methodology with real-world air quality data of Santiago, Chile, through two monitoring stations with different climatic conditions and pollution levels. This illustration showed that our methodology is useful for alerting citizens about episodes of extreme air pollution and for preventing adverse effects on human health for the population of Santiago, Chile. Specifically, the control charts based on the Birnbaum-Saunders-t distribution performs well for early prediction of extreme situations of contamination. In addition, we showed the coherence between our criterion and real-world situations in which Chilean health authority ruled environmental alerts.

CHAPTER 5

Discussion

5.1 Resumo

Neste capítulo são apresentadas as conclusões gerais deste trabalho intitulado "Ensaios sobre métodos Birnbaum-Saunders generalizada multivariada". Neste trabalho podemos destacar quatro importantes resultados a saber: (i) propomos distribuições Birnbaum-Saunders generalizadas logarítmicas multivariadas, (ii) derivamos modelos de regressão multivariados Birnbaum-Saunders generalizados, (iii) propomos métodos de diagnóstico para modelos de regressão multivariados Birnbaum-Saunders generalizados e, finalmente, (iv) desenvolvemos gráficos de controle de qualidade multivariados para distribuições Birnbaum-Saunders generalizadas. Além disso, neste capítulo são apresentados alguns problemas para trabalhos futuros que podem ser desenvolvidos. Podemos destacar a possibilidade de explorar o efeito de considerar diferentes parâmetros de forma em cada resposta, derivar uma análise de resíduos para modelos de regressão multivariados Birnbaum-Saunders generalizados e considerar efeitos aleatórios nos modelos propostos neste trabalho. Além disso, na área de gráficos de controle de qualidade multivariados, podemos desenvolver um gráfico de controle para observações individuais ou correlacionadas.

5.2 Conclusion

In this work, we presented essays on multivariate generalized Birnbaum-Saunders methods. Initially, in Chapter 2, we proposed multivariate log-GBS distributions and derived new multivariate generalized Birnbaum-Saunders regression models, including their maximum likelihood estimation by means of the EM algorithm. In Chapter 3, we carried out a diagnostic analysis for multivariate generalized Birnbaum-Saunders log-linear regression models by goodness-of-fit techniques and global and local influence methods. In Chapter 4, we proposed a robust methodology for multivariate generalized Birnbaum-Saunders control charts, whose limits were based on the distribution of the Hotelling statistic obtained with the bootstrap method. In each chapter of this work an application with real-world data was provided to illustrate the good performance of the proposed methodology.

5.3 Future research

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

(i) It is possible to explore the effect of considering different shape parameters in each response.

(ii) Heterogeneity problems presented in the data could also be considered in the type of models derived in the thesis.

(iii) Other estimation procedures can be investigated.

(iv) Residual analysis may be derived for these regression models. Such a topic is somewhat complex in multivariate regression, still more for asymmetrical distributions; see Cysneiros and Paula (2004) and Manghi et al. (2016) for some ideas on residual analysis in elliptical multivariate models.

(v) Distributional assumption can be evaluated by multivariate GOF methods.

(vi) Random effects can be added in the models proposed in this work.

(vii) A multivariate GBS quality control charts for individual or correlated observations can be developed.

All the possible future works provide to us challenging aspects to be studied.

CHAPTER 6

Appendix

6.1 The EM algorithm

The EM algorithm was proposed by Dempster et al. (1977). They pointed out that the method had been proposed many times in special circumstances by earlier authors. In particular, a very detailed treatment of the EM method for exponential families was published by Sundberg (1976). Dempster et al. (1977) generalized the method and sketched a convergence analysis for a wider class of problems. Regardless of earlier inventions, the innovative Dempster et al. (1977) paper in the Journal of the Royal Statistical Society received an enthusiastic discussion at the Royal Statistical Society meeting with Sundberg calling the paper "brilliant". The Dempster et al. (1977) paper established the EM method as an important tool of statistical analysis.

The EM iteration alternates between performing an expectation (E) step, which creates a function for the expectation of the log-likelihood evaluated using the current estimate for the parameters, and a maximization (M) step, which computes parameters maximizing the expected log-likelihood found on the E-step. These estimates are then used to determine the distribution of the latent variables in the next E-step. According to Ferreira et al. (2011), the EM algorithm has a large popularity due to (i) its computational simplicity in the M-step, because it involves only complete data ML estimation; and (ii) its stable and straightforward implementation, because the iterations converge monotonically and there are no need for second derivatives.

Let $\boldsymbol{Y}$ and $\boldsymbol{Z}$ be the vector of observed and missing data, respectively, thereby complete data vector $\boldsymbol{Y}_c$ is composed of $\boldsymbol{Y}$ and $\boldsymbol{Z}$ as $\boldsymbol{Y}_c = (\boldsymbol{Y}^\top, \boldsymbol{Z}^\top)^\top$. Consider $\ell(\boldsymbol{\theta}, \boldsymbol{Y}_c)$ the log-likelihood function of complete data for the vector parameters $\boldsymbol{\theta} \in \boldsymbol{\Theta}$. The EM algorithm addresses problems with incomplete data indirectly through the substitution of not observable part in $\ell(\boldsymbol{\theta}, \mathbf{Y_c})$ by its conditional expectation given $\boldsymbol{Y}$, using the adjustment to current $\boldsymbol{\theta}$. That is, consider the function Q defined as

$$Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}) = \mathbb{E}[\ell(\boldsymbol{\theta}, \boldsymbol{Y}_c|\boldsymbol{Y}, \widehat{\boldsymbol{\theta}})].$$

The two steps of the EM approach are summarized for a general setting in Algorithm 7.

Algorithm 7 EM approach for a general setting

E-step. Given $\widehat{\boldsymbol{\theta}} = \widehat{\boldsymbol{\theta}}^{(r-1)}$, compute $Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(r-1)})$, for $r = 1, 2, \ldots$; and
M-step. Find $\widehat{\boldsymbol{\theta}}^{(r)}$ such that $Q\Big(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(r)}\Big) = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(r-1)})$, for $r = 1, 2, \ldots$, where $Q(\boldsymbol{\theta}|\widehat{\boldsymbol{\theta}}^{(r-1)})$ is the expected value of the complete data log-likelihood function conditional on the observed data and evaluated at $(r-1)$th estimation of $\boldsymbol{\theta}$.

Algorithm 7 must be iterated until to reach convergence. Each iteration of the EM algorithm increases the likelihood function of observed data, say $L(\boldsymbol{\theta}, \boldsymbol{Y})$, and under appropriate conditions the EM algorithm presents monotonous convergence to the global or local maximum of $L(\boldsymbol{\theta}, \boldsymbol{Y})$; see Wu (1983) and McLachlan and Krishnan (1997). Note that the convergence analysis developed in Dempster et al. (1977) was flawed and a correct convergence analysis was published by Wu (1983). This author established the EM methods convergence outside of the exponential family.

6.2 Proof of Theorem 1

Let $\boldsymbol{Y} \sim \text{log-GBS}_m(\boldsymbol{\alpha}, \boldsymbol{\mu}, \boldsymbol{\Psi}, g^{(m)})$ be partitioned as in (2.4). $\boldsymbol{W} = B(\boldsymbol{Y}; \boldsymbol{\alpha}, \boldsymbol{\mu}) \sim \text{EC}_m(\mathbf{0}, \boldsymbol{\Psi}, g^{(m)})$, with $B(\boldsymbol{Y}; \boldsymbol{\alpha}, \boldsymbol{\mu})$ given in (2.2) and $\boldsymbol{\Psi} > \mathbf{0}$. Thus,

(a) if $\boldsymbol{W} \sim \text{EC}_m(\mathbf{0}, \boldsymbol{\Psi}, g^{(m)})$ and $\boldsymbol{D} \in \mathbb{R}^{q\times m}$ is a constant matrix with $\text{rk}(\boldsymbol{D}) = q < m$, then $\boldsymbol{DW} \sim \text{EC}_q(\mathbf{0}, \boldsymbol{D\Psi D}^\top, g^{(q)})$; see Fang et al. (1990, pp. 43). Let $\boldsymbol{D} = [\boldsymbol{I}_q|\mathbf{0}_{q\times[m-q]}]$, so that $\boldsymbol{DY} = \boldsymbol{Y}_1$ and $\boldsymbol{DW} \sim \text{EC}_q(\mathbf{0}, \boldsymbol{D\Psi D}^\top = \boldsymbol{\Psi}_{11}, g^{(q)})$, then

$$\boldsymbol{Y}_1 \sim \text{log-GBS}_q(\boldsymbol{\alpha}_1, \boldsymbol{\mu}_1, \boldsymbol{\Psi}_{11}, g^{(q)}).$$

Similarly if $\boldsymbol{D} = [\mathbf{0}_{(m-q)\times q}|\boldsymbol{I}_{(m-q)}] \in \mathbb{R}^{[m-q]\times m}$, then $\boldsymbol{Y}_2 \sim \text{log-GBS}_{m-q}(\boldsymbol{\alpha}_2, \boldsymbol{\mu}_2, \boldsymbol{\Psi}_{22}, g^{(m-q)})$;

(b) if $\boldsymbol{W} = (\boldsymbol{W}_1, \boldsymbol{W}_2)^\top \sim \mathrm{EC}_m(\boldsymbol{0}, \boldsymbol{\Psi}, g^{(m)})$, with $\boldsymbol{W}_1 \in \mathbb{R}^q$, $\boldsymbol{W}_2 \in \mathbb{R}^{m-q}$ and $\boldsymbol{\Psi}$ being partitioned as in (2.4), then the CDF of $\boldsymbol{W}_1|\boldsymbol{W}_2$ is

$$F_{\boldsymbol{W}_1|\boldsymbol{W}_2=z_2}(z_1) = F_{\mathrm{EC}_q}(\boldsymbol{z}_1 - \boldsymbol{\Psi}_{12}\boldsymbol{\Psi}_{22}^{-1}\boldsymbol{z}_2; \boldsymbol{\Psi}_{11.2}, g^{(q)}_{(\boldsymbol{z}_2^\top \boldsymbol{\Psi}_{22}^{-1}\boldsymbol{z}_2)}),$$

where $\boldsymbol{z}_1 = (z_1, \ldots, z_q)^\top$, $\boldsymbol{z}_2 = (z_{q+1}, \ldots, z_m)^\top$ and $\boldsymbol{\Psi}_{11.2} = \boldsymbol{\Psi}_{11} - \boldsymbol{\Psi}_{12}\boldsymbol{\Psi}_{22}^{-1}\boldsymbol{\Psi}_{21}$; see Fang et al. (1990, pp. 45-46). In our case, $\boldsymbol{W} = (\boldsymbol{W}_1, \boldsymbol{W}_2)^\top = \boldsymbol{B}(\boldsymbol{Y}; \boldsymbol{\alpha}, \boldsymbol{\mu}) \sim \mathrm{EC}_m(\boldsymbol{0}, \boldsymbol{\Psi}, g^{(m)})$, with $\boldsymbol{W}_1 \in \mathbb{R}^q$ and $\boldsymbol{W}_2 \in \mathbb{R}^{m-q}$.

(c) the demonstration of this item follows immediately from (b). ∎

6.3 Data Sets

Fatigue Data Set

Table 6.1: Fatigue data for the indicated variable used in Chapter 2.

X_1	X_2	X_3	T_1	T_2	T_3	T_4
0.07	23.00	581.08	1850	1.260	144000	6420
0.07	23.00	818.92	470	1.349	36700	33700
0.07	31.96	581.08	1830	1.532	156000	9430
0.07	31.96	818.92	523	1.614	39900	36600
0.13	23.00	581.08	2030	1.801	181000	12100
0.13	23.00	818.92	581	1.824	46900	32000
0.13	31.96	581.08	2230	1.939	203000	13200
0.13	31.96	818.92	632	1.928	52300	32100
0.05	27.50	700.00	889	1.275	78600	19900
0.15	27.50	700.00	1410	1.921	125000	15000
0.10	20.00	700.00	1060	1.692	92100	20900
0.10	35.00	700.00	1390	1.888	111000	21200
0.10	27.50	500.00	2430	1.666	213000	9170
0.10	27.50	900.00	243	1.685	19500	74800
0.10	27.50	700.00	1130	1.651	96900	19900

Source: Lepadatu et al. (2005).

Biomaterial data set

CT scan data and different densities expressed in mg/cm^3 of human trabecular bone specimens used in Chapter 3. Here, **Voxels** represents the number of voxels, **HU core** the mean of the voxel Hounsfield units (HUs) within a bone core, **SD HU** the standard deviation of the voxel HUs within a bone core, **CV HU** the coefficient of variation of the voxel HUs within a bone core, $\rho_{\mathbf{ct}}$ the computer tomography density derived from HUs by means of a calibration equation from mineral standards, $\rho_{\mathbf{bulk}}$ the bulk density defined as the mass of the intact core, including fat and water, over specimen bulk volume, $\rho_{\mathbf{dry}}$ the dry apparent density considered mass of the intact core excluding fat and water, over specimen bulk volume and $\rho_{\mathbf{ash}}$ the ash density considered mineral mass over bulk volume.

Table 6.2: CT scan data to study the bone quality used in Chapter 3.

Age	Gender	ID core	Voxels	HU core	SD HU	CV HU	Density (mg/cm^3)			
							ρ_{ct}	ρ_{bulk}	ρ_{dry}	ρ_{ash}
68	F	213	2311	144	153	1.060	105	801	161	125
68	F	231	1811	154	177	1.150	116	988	338	227
68	F	233	2214	260	262	1.010	195	1156	526	359
68	F	234	2447	327	205	0.626	235	1094	422	292
68	F	235	2026	272	227	0.832	205	1045	376	263
87	F	411	1902	230	132	0.573	156	1006	347	238
87	F	412	1902	329	121	0.368	228	1187	519	359
87	F	413	2059	401	132	0.330	282	1107	352	246
87	F	421	1376	104	122	1.173	71	1073	310	212
87	F	422	1791	232	115	0.497	156	1116	373	261
87	F	423	1862	284	123	0.433	195	1175	391	279
87	F	424	1761	213	111	0.520	143	1214	472	332
87	F	425	2267	349	123	0.352	244	1063	286	193
87	F	426	1614	88	89	1.010	55	869	160	129
87	F	427	1833	425	159	0.374	301	1226	527	366
86	F	511	4250	180	157	0.870	132	1061	275	197
86	F	512	3231	121	117	0.967	83	980	210	145
86	F	521	3047	129	140	1.080	95	996	249	169
86	F	522	3083	148	136	0.922	105	958	179	138
86	F	523	3618	45	95	2.130	33	1027	160	123
86	F	524	3750	131	136	1.040	94	1036	267	196
86	F	525	3623	140	142	1.020	103	1084	377	257
86	F	526	3256	232	179	0.770	171	1089	404	278
86	F	531	4972	149	167	1.120	116	1090	268	194
86	F	532	4740	128	155	1.210	97	984	164	122
86	F	534	5141	125	153	1.220	94	1042	253	167
86	F	535	4170	74	138	1.860	61	974	204	154
86	F	536	4413	116	151	1.300	90	978	188	138

Source: Vivanco et al. (2014).

Table 6.3: (continued) CT scan data to study the bone quality.

Age	Gender	ID core	Voxels	HU core	SD HU	CV HU	Density (mg/cm^3)			
							ρ_{ct}	ρ_{bulk}	ρ_{dry}	ρ_{ash}
81	F	611	2487	193	102	0.529	127	977	254	169
81	F	613	2891	324	116	0.360	224	1066	414	264
81	F	621	1448	211	110	0.521	140	958	290	197
81	F	622	2688	297	136	0.457	204	1126	401	274
81	F	624	3210	329	125	0.380	228	1162	442	305
77	F	712	3109	377	161	0.428	265	1179	556	373
77	F	713	2462	280	116	0.415	192	1124	446	302
77	F	714	2319	102	107	1.052	66	837	163	112
77	F	722	3817	361	160	0.442	254	1088	352	231
77	F	724	2922	234	136	0.580	158	955	348	244
77	F	726	3112	422	195	0.463	299	1162	568	379
79	M	311	1668	260	118	0.454	177	1142	391	275
79	M	312	1952	450	125	0.278	319	1224	488	344
79	M	321	1554	278	101	0.364	189	1171	394	284
79	M	322	1612	465	147	0.317	330	1185	479	320
79	M	323	1601	417	146	0.349	294	1208	494	333
79	M	324	1542	309	104	0.336	212	1205	501	342
79	M	325	1755	377	137	0.365	264	1206	501	343
79	M	326	1520	223	112	0.501	150	1014	324	224
79	M	327	1625	564	187	0.332	404	1285	672	456
79	M	331	1900	374	174	0.466	262	1174	487	325
79	M	332	1781	355	122	0.344	247	1139	464	314
79	M	333	1486	284	145	0.512	195	1090	363	256
79	M	334	1368	227	124	0.549	152	1100	355	252
79	M	335	1448	299	106	0.353	205	1134	355	251
79	M	336	1225	252	145	0.577	171	1115	304	213
79	M	337	1333	524	164	0.313	374	1300	667	453
66	M	811	2394	258	129	0.500	174	905	205	153
66	M	812	3029	310	162	0.522	214	1065	415	282
66	M	821	2270	273	93	0.339	186	1119	430	293
66	M	822	2207	337	98	0.290	234	1162	526	356
66	M	823	2307	287	99	0.384	196	962	325	232
66	M	824	2838	358	117	0.328	249	938	242	169
66	M	825	2168	403	91	0.227	283	992	250	177
66	M	826	1893	146	64	0.441	90	975	220	156
66	M	827	2001	320	124	0.387	221	1080	379	254
66	M	831	2045	287	120	0.418	196	1032	325	212
66	M	832	2029	147	87	0.589	91	973	250	162
66	M	833	885	93	53	0.573	51	1063	237	186
66	M	834	2500	177	84	0.477	113	806	159	109
66	M	835	3000	220	88	0.401	146	739	135	97
66	M	836	3334	190	75	0.396	123	776	125	98
66	M	837	2172	155	81	0.523	97	910	272	187
66	M	838	1942	172	88	0.514	110	1052	437	304
66	M	839	3428	416	182	0.439	293	1008	380	258
66	M	8310	3477	405	127	0.313	285	1188	656	435

Source: Vivanco et al. (2014).

References

Achcar, J. A. (1993). Inference for the Birnbaum-Saunders fatigue life model using Bayesian methods. *Computational Statistics and Data Analysis*, 15:367–380.

Alfaro, J. L. and Ortega, J. F. (2012). Robust Hotelling's T^2 control charts under non-normality: the case of t-Student distribution. *Journal of Statistical Computation and Simulation*, 83:1437–1447.

Alfaro, J. L. and Ortega, J. F. (2013). A new control chart in contaminated data of t-Student distribution for individual observations. *Applied Stochastic Models in Business and Industry*, 29:79–91.

Alt, F. B. (1985). Multivariate quality control. In Kotz, S., Johnson, N. L., and Read, C. B., editors, *The Encyclopedia of Statistical Sciences*, volume 6, pages 110–112. Wiley, New York, US.

Assumpção, R., Uribe-Opazo, M., and Galea, M. (2014). Analysis of local influence in geostatistics using student-t distribution. *Journal of Applied Statistics*, 41:2323–2341.

Atkinson, A. (2009). Econometric applications of the forward search in regression: robustness, diagnostics, and graphics. *Econometric Reviews*, 28:21–39.

Azevedo, C., Leiva, V., Athayde, E., and Balakrishnan, N. (2012). Shape and change point analyses of the Birnbaum-Saunders-t hazard rate and associated estimation. *Computational Statistics and Data Analysis*, 56:3887–3897.

Balakrishnan, N., Gupta, R., Kundu, D., Leiva, V., and 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., Leiva, V., Sanhueza, A., and Vilca, F. (2009). Estimation in the Birnbaum-Saunders distribution based on scale-mixture of normals and the EM-algorithm. *Statistics and Operations Research Transactions*, 33:171–192.

Barros, M., Galea, M., Gonzalez, M., and Leiva, V. (2010). Influence diagnostics in the tobit censored response model. *Statistical Methods and Applications*, 19:379–397.

Barros, M., Leiva, V., Ospina, R., and Tsuyuguchi, A. (2014). Goodness-of-fit tests for the Birnbaum-Saunders distribution with censored reliability data. *IEEE Transactions on Reliability*, 63:543–554.

Barros, M., Paula, G. A., and Leiva, V. (2008). A new class of survival regression models with heavy-tailed errors: robustness and diagnostics. *Lifetime Data Analysis*, 14:316–332.

Becker, C. and Gather, U. (1999). The masking breakdown point of multivariate outlier identification rules. *Journal of the American Statistical Association*, 94:947–955.

Bersimis, S., Psarakis, S., and Panaretos, J. (2007). Multivariate statistical process control charts: An overview. *Quality and Reliability Engineering International*, 23:517–543.

Bhatti, C. (2010). The Birnbaum-Saunders autoregressive conditional duration model. *Mathematics and Computers in Simulation*, 80:2062–2078.

Billor, N. and Loynes, R. (1999). An application of the local influence approach to ridge regression. *Journal of Applied Statistics*, 26:177–183.

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

Caro-Lopera, F., Leiva, V., and 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.

Castro-Kuriss, C., Kelmansky, D., Leiva, V., and Martínez, E. (2009). A new goodness-of-fit test for censored data with an application in monitoring processes. *Communications in Statistics – Simulation and Computation*, 38:1161–1177.

Castro-Kuriss, C., Kelmansky, D., Leiva, V., and Martínez, E. (2010). On a goodness-of-fit test for normality with unknown parameters and type-II censored data. *Journal of Applied Statistics*, 37:1193–1211.

Castro-Kuriss, C., Leiva, V., and Athayde, E. (2014). Graphical tools to assess goodness-of-fit in non-location-scale distributions. *Colombian Journal of Statistics*, 37:341–365.

Chatterjee, S. and Hadi, A. S. (1988). *Sensitivity Analysis in Linear Regression*. Wiley, New York, US.

Chenouri, S. and Variyath, A. M. (2011). A comparative study of phase II robust multivariate control charts for individual observations. *Quality and Reliability Engineering International*, 27:857–865.

Chenouri, S., Variyath, A. M., and Steiner, S. H. (2009). A multivariate robust control chart for individual observations. *Journal of Quality Technology*, 41:259–271.

Chou, Y. M., Mason, R. L., and Young, J. C. (2001). The control chart for individual observations from a multivariate non-normal distribution. *Communications in Statistics – Theory and Methods*, 30:1937–1949.

CONAMA (1998). Establishment of primary quality guideline for PM10 that regulates environmental alerts. Technical Report Decree 59, Ministry of Environment (CONAMA) of the Chilean Government, Santiago, Chile.

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

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

Cysneiros, A., Cribari-Neto, F., and Araujo, C. G. J. (2008). On Birnbaum-Saunders inference. *Computational Statistics and Data Analysis*, 52:4939–4950.

Cysneiros, F. J. A. and Paula, G. A. (2004). One-sided tests in linear models with multivariate t-distribution. *Communications in Statistics – Simulation and Computation*, 33:747–771.

D'Agostino, C. and Stephens, M. A. (1986). *Goodness-of-Fit Techniques*. Marcel Dekker, New York, US.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximun likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society*, 39:1–38.

Desmond, A., Rodríguez-Yam, G., and Lu, X. (2008). Estimation of parameters for a Birnbaum-Saunders regression model with censored data. *Journal of Statistical Computation and Simulation*, 78:983–997.

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

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

Duchemin, L., Bousson, V., Raossanaly, C., Bergot, C., Laredo, J., Skalli, W., and Mitton, D. (2008). Prediction of mechanical properties of cortical bone by quantitative computed tomography. *Medical Engineering and Physics*, 30:321–328.

Duncan, A. (1986). *Quality Control and Industrial Statistics*. Irwin, Homewood, IL, US.

Dupuis, D. J. and Mills, J. E. (1998). Robust estimation of the Birnbaum-Saunders distribution. *IEEE Transactions on Reliability*, 1:88–95.

Efron, B. and Hinkley, D. (1978). Assessing the accuracy of the maximum likelihood estimator: observed vs. expected Fisher information. *Biometrika*, 65:457–487.

Fang, K. T., Kotz, S., and Ng, K. W. (1990). *Symmetric Multivariate and Related Distributions*. Chapman and Hall, London, UK.

Fang, K. T. and Zhang, Y. T. (1990). *Generalized Multivariate Analysis*. Springer, Berlin, Germany.

Ferreira, C. S., Bolfarine, H., and Lachos, V. H. (2011). Skew scale mixtures of normal distributions: properties and estimation. *Statistical Methodology*, 8:154–171.

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

Galea, M., Leiva, V., and Paula, G. A. (2004). Influence diagnostics in log-Birnbaum-Saunders regression models. *Journal of Applied Statistics*, 31:1049–1064.

Galea, M., Paula, G. A., and Bolfarine, H. (1997). Local influence in elliptical linear regression models. *The Statistician*, 46:71–79.

Garcia-Papani, F. M., Uribe-Opazo, M. A., Leiva, V., and Aykroyd, R. G. (2016). Birnbaum-Saunders spatial modelling and diagnostics applied to agricultural engineering data. *Stochastic Environmental Research and Risk Assessment*, pages in press available at `http://dx.doi.org/10.1007/s00477--015--1204--4`.

Gupta, A. K., Varga, T., and Bodnar, T. (2013). *Elliptically Contoured Models in Statistics and Portfolio Theory*. Springer, New York, US.

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*. Springer-Verlag, New Jersey, US.

Haslett, J. and Dillane, D. (2004). Application of 'delete = replace' to deletion diagnostics for variance component estimation in the linear mixed model. *Journal of The Royal Statistical Society B*, 66:131–143.

Hawkins, D. M. (1980). *Identification of Outliers*. Chapman and Hall, New York, US.

Hotelling, H. (1947). Multivariate quality control. In Eisenhart, C., Hastay, M., and Wallis, W. A., editors, *Techniques of Statistical Analysis*, pages 111–184. McGraw-Hill, New York, US.

Ibacache-Pulgar, G., Paula, G., and Galea, M. (2014). On influence diagnostics in elliptical multivariate regression models with equicorrelated random errors. *Statistical Modelling*, 16:14–21.

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

Jensen, W. A., Birch, J. B., and Woodall, W. H. (2007). High breakdown estimation methods for phase I multivariate control charts. *Quality and Reliability Engineering International*, 23:615–629.

Johnson, N. L., Kotz, S., and Balakrishnan, N. (1995). *Continuous Univariate Distributions*, volume 2. Wiley, New York, US.

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

Khosravi, M., Leiva, V., Jamalizadeh, A., and Porcu, E. (2016). A nonlinear Birnbaum-Saunders model based on a bivariate construction and its characteristics. *Communications in Statistics – Theory and Methods*, 45:772–793.

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

Kotz, S. and Nadarajah, S. (2004). *Multivariate t-Distributions and their Applications*. Cambridge University Press, New York, US.

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.

Kundu, D., Balakrishnan, N., and Jamalizadeh, A. (2010). Bivariate Birnbaum-Saunders distribution and associated inference. *Journal of Multivariate Analysis*, 101:113–125.

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

Lange, K. (2000). *Numerical Analysis for Statisticians*. Springer, New York, US.

Lange, K., Little, J., and Taylor, M. (1989). Robust statistical modeling using the t distribution. *Journal of the American Statistical Association*, 84:881–896.

Lange, K. and Sinsheimer, J. S. (1993). Normal/independent distributions and their applications in robust regression. *Journal of Computational and Graphical Statistics*, 2:175–198.

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

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

Leiva, V., Barros, M., Paula, G. A., and Galea, M. (2007). Influence diagnostics in log-Birnbaum-Saunders regression models with censored data. *Computational Statistics and Data Analysis*, 51:5694–5707.

Leiva, V., Barros, M., Paula, G. A., and Sanhueza, A. (2008a). Generalized Birnbaum-Saunders distribution applied to air pollutant concentration. *Environmetrics*, 19:235–249.

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

Leiva, V., Liu, S., Shi, L., and Cysneiros, F. J. A. (2015a). Diagnostics in elliptical regression models with stochastic restrictions applied to econometrics. *Journal of Applied Statistics*, 43:627–642.

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

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

Leiva, V., Ponce, M., Marchant, C., and Bustos, O. (2012). Fatigue statistical distributions useful for modeling diameter and mortality of trees. *Colombian Journal of Statistics*, 35:349–367.

Leiva, V., Rojas, E., Galea, M., and Sanhueza, A. (2014b). Diagnostics in Birnbaum-Saunders accelerated life models with an application to fatigue data. *Applied Stochastic Models in Business and Industry*, 30:115–131.

Leiva, V., Ruggeri, F., Saulo, H., and Vivanco, J. F. (2017). A methodology based on the Birnbaum-Saunders distribution for reliability analysis applied to nano-materials. *Reliability Engineering and System Safety*, page in press available at `http://dx.doi.org/10.1016/j.ress.2016.08.024`.

Leiva, V., Sanhueza, A., and 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., Sanhueza, A., Kotz, S., and Araneda, N. (2010). A unified mixture model based on the inverse Gaussian distribution. *Pakistan Journal of Statistics*, 26:445–460.

Leiva, V., Sanhueza, A., Sen, P. K., and Paula, G. A. (2008b). Random number generators for the generalized Birnbaum-Saunders distribution. *Journal of Statistical Computation and Simulation*, 78:1105–1118.

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

Leiva, V., Santos-Neto, M., Cysneiros, F. J. A., and 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, E., Leão, J., and Marchant, C. (2014d). A family of autoregressive conditional duration models applied to financial data. *Computational Statistics and Data Analysis*, 79:175–191.

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

Leiva, V., Soto, G., Cabrera, E., and Cabrera, G. (2011b). New control charts based on the Birnbaum-Saunders distribution and their implementation. *Colombian Journal of Statistics*, 34:147–176.

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

Leiva, V. and Vivanco, J. F. (2016). Fatigue models. *Wiley StatsRef: Statistics Reference Online*, pages 1–10.

Lemonte, A., Simas, A., and Cribari-Neto, F. (2008). Bootstrap-based improved estimators for the two-parameter Birnbaum-Saunders distribution. *Journal of Statistical Computation and Simulation*, 78:37–49.

Lemonte, A. J. (2011). Covariance matrix formula for Birnbaum-Saunders regression models. *Journal of Statistical Computation and Simulation*, 81:899–908.

Lemonte, A. J. (2013). Multivariate Birnbaum-Saunders regression model. *Journal of Statistical Computation and Simulation*, 83:2244–2257.

Lemonte, A. J. and Cordeiro, G. M. (2009). Birnbaum-Saunders nonlinear regression models. *Computational Statistics and Data Analysis*, 53:4441–4452.

Lemonte, A. J., Cribari-Neto, F., and Vasconcellos, K. L. P. (2007). Improved statistical inference for the two-parameter Birnbaum-Saunders distribution. *Computational St atistics and Data Analysis*, 51:4656–4681.

Lemonte, A. J. and Ferrari, S. L. P. (2011a). Signed likelihood ratio tests in the Birnbaum-Saunders regression model. *Journal of Statistical Planning and Inference*, 141:1031–1040.

Lemonte, A. J. and Ferrari, S. L. P. (2011b). Small-sample corrections for score tests in Birnbaum-Saunders regressions. *Communications in Statistics – Theory and Methods*, 40:232–243.

Lemonte, A. J., Ferrari, S. L. P., and Cribari-Neto, F. (2010). Improved likelihood inference in Birnbaum-Saunders regressions. *Computational Statistics and Data Analysis*, 54:1307–1316.

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

Lepadatu, D., Kobi, A., Hambli, R., and Barreau, A. (2005). Lifetime multiple response optimization of metal extrusion die. *Proceedings of the Annual Reliability and Maintainability Symposium, Institute of Electrical and Electronics Engineers*, pages 37–42.

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

Li, A. P., Chen, Z. X., and Xie, F. C. (2012). Diagnostic analysis for heterogeneous log-Birnbaum-Saunders regression models. *Statistics and Probability Letters*, 89:1690–1698.

Lio, Y. L. and Park, C. (2008). A bootstrap control chart for Birnbaum-Saunders percentiles. *Quality and Reliability Engineering International*, 24:585–600.

Liu, C. and Rubin, D. B. (1995). ML estimation of the t distribution using the EM and its extensions, ECM and ECME. *Statistica Sinica*, 5:19–39.

Liu, R. (1995). Control charts for multivariate processes. *Journal of the American Statistical Association*, 90:1380–1387.

Liu, R. Y. and Tang, J. (1997). Control charts for dependent and independent measurements based on boostrap methods. *Journal of the American Statistical Association*, 91:1694–1700.

Liu, S., Leiva, V., Ma, T., and Welsh, A. (2016). Influence diagnostic analysis in the possibly heteroskedastic linear model with exact restrictions. *Statistical Methods and Applications*, 5:227–249.

Lowry, C. A. and Montgomery, D. C. (1995). A review of multivariate control charts. *IEE Transactions*, 27:800–810.

Lucas, A. (1997). Robustness of the student t based M-estimator. *Communications in Statistics – Theory and Methods*, 41:1165–1182.

MacDonald, I. L. (2014). Numerical maximization of likelihood: A neglected alternative to EM? *Intern Statistical Review*, 82:296–308.

Manghi, R. F., Cysneiros, F. J. A., and Paula, G. A. (2016). On elliptical multilevel models. *Journal of Applied Statistics*, 43:2150–2171.

Maravelakis, P. E., Bersimis, S., Panaretos, J., and Psarakis, S. (2002). On identifying the out-of-control variable in a multivariate control chart. *Communications in Statistics – Theory and Methods*, 31:2391–2408.

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

Mason, R. L., Chou, Y. M., and Young, J. C. (2001). Applying Hotellings T^2 statistic to batch processes. *Journal of Quality Technology*, 33:466–479.

McLachlan, G. and Krishnan, T. (1997). *The EM Algorithm and Extensions*. Wiley, New York, US.

Meng, X. L. and Rubin, D. B. (1993). Maximum likelihood estimation via the ecm algorithm: A general framework. *Biometrika*, 80:267–278.

Mittelhammer, R. C., Judge, G. G., and Miller, D. J. (2000). *Econometric Foundations*. Cambridge University Press, New York, US.

Nedumaran, G. and Pignatiello, J. J. (2000). On constructing T^2 control charts for retrospective examination. *Communications in Statistics – Simulation and Computation*, 29:621–632.

Ng, H. K. T., Kundu, D., and Balakrishnan, N. (2003). Modified moment estimation for the two-parameter Birnbaum-Saunders distribution. *Computational Statistics and Data Analysis*, 43:283–298.

Nocedal, J. and Wright, S. J. (1999). *Numerical Optimization*. Springer, New York, US.

Noorossana, R., Woodall, W. H., and Amiriparian, S. (2002). On the economic design of multivariate control charts. *Communications in Statistics – Theory and Methods*, 31:1665–1673.

Osorio, F., Paula, G. A., and Galea, M. (2007). Assessment of local influence in elliptical linear models with longitudinal structure. *Computational Statistics and Data Analysis*, 51:4354–4368.

Owen, W. and Padgett, W. (1999). Accelerated test models for system strength based on Birnbaum-Saunders distribution. *Lifetime Data Analysis*, 5:133–147.

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

Owen, W. J. and Padgett, W. J. (2000). A Birnbaum-Saunders accelerated life model. *IEEE Transactions on Reliability*, 49:224–229.

Paula, G. A. (1993). Assessing local influence in restricted regression models. *Communications in Statistics – Theory and Methods*, 16:63–79.

Paula, G. A., Leiva, V., Barros, M., and Liu, S. (2012). Robust statistical modeling using the Birnbaum-Saunders-t distribution applied to insurance. *Applied Stochastic Models in Business and Industry*, 28:16–34.

Phaladiganon, P., Kim, S. B., Chen, V., Baek, J. G., and Park, S. K. (2011). Bootstrap-based T^2 multivariate control charts. *Communications in Statistics – Simulation and Computation*, 40:645–662.

R-Team (2016). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rieck, J. R. (2003). A comparison of two random number generators for the Birnbaum-Saunders distribution. *Communications in Statistics – Theory and Methods*, 32:929–934.

Rieck, J. R. and Nedelman, J. R. (1991). A log-linear model for the Birnbaum-Saunders distribution. *Technometrics*, 3:51–60.

Rocke, D. M. and Woodruff, D. L. (1996). Identification of outliers in multivariate data. *Journal of the American Statistical Association*, 91:1047–1061.

Rojas, F., Leiva, V., Wanke, P., and Marchant, C. (2015). Optimization of contribution margins in food services by modeling independent component demand. *Colombian Journal of Statistics*, 38:1–30.

Runger, G. C., Alt, F. A., and Montgomery, D. C. (1996). Contributors to a multivariate statistical process control signal. *Communications in Statistics – Theory and Methods*, 25:2103–2213.

Sánchez, L., Leiva, V., Caro-Lopera, F. J., and 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., and 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., and 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., and 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., and Barros, M. (2016). Reparameterized Birnbaum-Saunders regression models with varying precision. *Electronic Journal of Statistics*, pages in press available at `http://dx.doi.org/10.1214/16--EJS1187`.

Saulo, H., Leiva, V., and 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., and 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.

Shewhart, W. A. (1931). *Economic Control of Quality of Manufactured Product*. D. Van Nostrand Company, New York, US.

Shi, L. (1997). Local influence in principal components analysis. *Biometrika*, 84:175–186.

Stoumbos, G. Z. and Sullivan, J. (2002). Robustness to non-normality of the multivariate EWMA control charts. *Journal of Quality Technology*, 34:304–315.

Sullivan, J. H. and Woodall, W. H. (1996). A comparison of multivariate quality control charts for individual observations. *Journal of Quality Technology*, 28:398– 408.

Sullivan, J. H. and Woodall, W. H. (1998). Adapting control charts for the preliminary analysis of multivariate observations. *Communications in Statistics – Simulation and Computation*, 27:953–979.

Sundberg, R. (1976). An iterative method for solution of the likelihood equations for incomplete data from exponential families. *Communications in Statistics – Simulation and Computation*, 5:55–64.

Tracy, N. D., Young, J. C., and Mason, R. L. (1992). Multivariate control charts for individual observations. *Journal of Quality Technology*, 24:88–95.

Tsionas, E. (2001). Bayesian inference in Birnbaum-Saunders regression. *Communications in Statistics – Theory and Methods*, 30:179–193.

Uribe-Opazo, M. A., Borssoi, J. A., and Galea, M. (2012). Influence diagnostics in gaussian spatial linear models. *Journal of Applied Statistics*, 39:615–630.

Vanegas, L. H., Rondon, L. M., and Cysneiros, F. J. A. (2012). Diagnostic procedures in Birnbaum-Saunders nonlinear regression models. *Computational Statistics and Data Analysis*, 56:1662–1680.

Verbeke, G. and Molenberghs, G. (2000). *Linear Mixed Models for Longitudinal Data*. Springer, New York.

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

Vilca, F., Balakrishnan, N., and Zeller, C. (2014b). The bivariate sinh-elliptical distribution with applications to Birnbaum-Saunders distribution and associated regression and measurement error models. *Computational Statistics and Data Analysis*, 80:1–16.

Vilca, F., Sanhueza, A., Leiva, V., and 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.

Vilca, F., Santana, L., Leiva, V., and Balakrishnan, N. (2011). Estimation of extreme percentiles in Birnbaum-Saunders distributions. *Computational Statistics and Data Analysis*, 55:1665–1678.

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

Vivanco, J. F., Burgers, T. A., and et al (2014). Estimating the density of femoral head trabecular bone from hip fracture patients using computed tomography scan data. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 228:616–626.

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

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

Wu, C. (1983). On the convergence properties of the em algorithm. *The Annals of Statistics*, 11:95–103.

Xiao, Q., Liu, Z., Balakrishnan, N., and Lu, X. (2010). Estimation of the Birnbaum-Saunders regression model with current status data. *Computational Statistics and Data Analysis*, 54:326–332.

Xie, F. C. and Wei, B. C. (2007). Diagnostics analysis for log-Birnbaum-Saunders regression models. *Computational Statistics and Data Analysis*, 51:4692–4706.

Yang, K. and Trewn, J. (2004). *Multivariate Statistical Methods in Quality Management*. New York, McGraw Hill.

Yen, C. L. and Shiau, J. J. H. (2010). A multivariate control chart for detecting increases in process dispersion. *Statistica Sinica*, 20:1683–1707.

Yen, C. L. and Tang, J. (2012). Methods for identifying influential variables in an out-of-control multivariate normal process. *Statistica Sinica*, 222:847–868.