University of Brasília
Institute of Exact Sciences
Department of Statistics

Master’s Dissertation

Parametric quantile regression for extreme events

by

Verônica Lelis Bittencourt

Brasília, April 2022

Parametric quantile regression for extreme events

by

Verônica Lelis Bittencourt

A dissertation submitted to the Departament of Statistics at the University of Brasília in partial fulfilment of the requirements for the degree Master in Statistics.

Supervisor: Prof. Dr. Helton Saulo Bezerra dos Santos

Brasília, April 2021

A dissertation submitted to the Departament of Statistics at the University of Brasília in partial fulfilment of the requirements for the degree Master in Statistics.

Approved by:

Prof. Dr. Helton Saulo Bezerra dos Santos
Supervisor, EST/UnB

Prof. Dr. Roberto Vila Gabriel
EST/UnB

Prof. Dr. Jeremias da Silva Leão
DE/UFAM

Prof. Dr. José Augusto Fiorucci
EST/UnB

"The noblest pleasure is the joy of understanding".

(Leonardo da Vinci)

To my parents, without them, nothing would be possible, my partner from all the moments, Lucas, and also, to my supervisor for the support and availability.

Acknowledgments

To God for granting me health, wisdom and strength. For allowing me to make mistakes, learn, grow and persist in the realization of my dreams.

To my supervisor, Prof. Dr. Helton Saulo Bezerra dos Santos, for all the support, availability, patience and commitment.

To the professors of the disciplines taken in the master's degree for shared knowledge.

To my colleagues in the master's degree, for the work and disciplines carried out together, for sharing ideas and for constant support.

To the employees of PPGEST/UnB for the promptness in the clarifications and kindness.

To my parents, who always believed in my ability and didn't put limits on my dreams.

To my boyfriend, for being with me at all times, in the achievements and also in the biggest challenges during this course.

Finally, to all those who contributed, directly or indirectly, to the realization of this dissertation, my sincere thanks.

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível

Superior - Brasil (CAPES) - Finance Code 001.

Resumo

REGRESSÃO QUANTÍLICA PARAMÉTRICA PARA EVENTOS EXTREMOS

A teoria dos valores extremos tem se mostrado muito útil para problemas envolvendo eventos raros e extremos em diversas áreas. Por um longo tempo, as questões ambientais envolvendo principalmente hidrologia e engenharia foram as que mais utilizaram essa teoria. A distribuição de valores extremos generalizados (GEV), desenvolvida por Jenkinson (1955) and Jenkinson (1969), é amplamente utilizada para modelar extremos de eventos naturais, tais como precipitações e temperaturas máximas, e o modelo GEV é de considerável importância para a área ambiental. No entanto, outras áreas como medicina, seguros e setor fianceiro também se beneficiam dessa metodologia, ver, por exemplo Embrechts, Kluppelberg, and Mikosch (1997), Reiss and Thomas (2001), Coles (2001), Haan and Ferreira (2006) e Beirlant, Caeiro, and Gomes (2012). Uma outra distribuição que também vem ganhando atenção em diversas áreas é o modelo Birnbaum-Saunders (BS); ver, por exemplo, Leiva, Sanhueza, and Angulo (2009), Vilca et al. (2010), Saulo et al. (2013), Leiva et al. (2015), Leiva (2016), Balakrishnan and Kundu (2019) e Oliveira et al. (2022). A distribuição BS está relacionada com o modelo normal e é muito útil para modelar dados estritamente positivos e assimétricos. Uma aplicação proeminente dos modelos BS está na teoria dos valores extremos. Ferreira, Gomes, and Leiva (2012) propôs a distribuição de valores extremos BS (EVBS) alterando a normal usual, na re-

presentação estocástica do modelo BS, pela distribuição GEV. Os autores aplicaram o modelo proposto a dados reais atmosféricos. Além disso, Leiva et al. (2016) propuseram um modelo de regressão baseado na distribuição EVBS. O modelo de regressão foi ilustrado através da utilização de dados ambientais e é uma ferramenta muito útil na modelagem de eventos extremos. No entanto, esse modelo é implementado adicionando covariáveis ao parâmetro de localização, o que resulta em uma descrição da variável dependente com base na média. Os modelos de regressão quantílica são uma alternativa à regressão usual baseada na média, pois apresentam como resultados os efeitos das covariáveis sobre a variável resposta nos diferentes quantis de sua distribuição, ou seja, proporciona uma análise ao longo de toda a distribuição condicional; veja Cade, Terrell, and Schroeder (1999), Koenker (2005) e Wei et al. (2006). A média como única medida sumária geralmente é insuficiente para uma avaliação de risco, pois é altamente afetada pela grande variabilidade dos dados e presença de outliers, que podem ser raros, mas suficientes para causar catástrofes na área ambiental e insolvência de seguradoras, por exemplo. Desse modo, dada a importância da regressão quantílica para estimar os efeitos das covariáveis ao longo do espectro da variável resposta, o objetivo principal deste artigo é propor um modelo de regressão quantílica paramétrica baseado na Birnbaum-Saunders de valor extremo. Para tanto, descrevemos a distribuição usual de EVBS (Ferreira, Gomes, and Leiva, 2012) e propomos uma reparametrização dessa distribuição inserindo um parâmetro que representa o quantil da distribuição, chamado Q, onde Q é o $100q$-ésimo quantil de uma variável aleatória T seguindo uma distribuição Birnbaum-Saunders de valor extremo, essa distribuição formulada é denotada por QEVBS. Também estabelecemos alguns resultados relativos à propriedade de uni e bimodalidade, representação estocástica e distribuições relacionadas da distribuição QEVBS. Apresentamos o modelo de regressão QEVBS baseado no quantil, uma abordagem baseada em verossimilhança para estimativa de parâmetros, bem como a função de probabilidade logarítmica, vetor de pontuação, e consideramos dois tipos de resíduos: Cox-Snell generalizado e resíduos quantílicos aleatórios. Um estudo de simulação de Monte Carlo (MC) é realizado para avaliar numericamente o desempenho estatístico dos estimadores de máxima verossimi-

lhança (ML) através do viés e do erro quadrático médio (EQM) e a distribuição empírica dos resíduos. Foram realizadas duas simulação de MC e considerados cinco quantis e três tamanhos amostrais para cada uma delas. Para ilustrarmos a metodologia proposta, foi utilizado um conjunto de dados ambientais reais coletados na região central de Santiago-Chile entre julho e setembro de 2021, que é analisado pela primeira vez aqui. Os dados se referem a concentração máxima diária de ozônio (variável resposta) e a temperatura máxima diária (covariável), assim correspondem a valores máximos, que são intrinsecamente valores extremos, e o modelo de regressão QEVBS pode ser adequado para descrever a relação entre as variáveis. Os resultados do primeiro estudo de simulação MC mostraram bom desempenho dos estimadores de máxima verossimilhança obtendo valores empíricos de viés próximos de zero e, em geral, à medida que o tamanho da amostra aumenta, o EQM diminui. Os resultados da segunda simulação mostraram que os resíduos generalizados de Cox-Snell e quantis aleatórios se comportam de acordo com suas distribuições de referência. Na aplicação do modelo, QQ plots com envelope para Cox-Snell generalizado e resíduos quantílicos aleatórios mostraram um bom ajuste do modelo proposto. Também comparamos o modelo de regressão quantílica de Birnbaum-Saunders de valor extremo proposto com outros modelos comumente usados na literatura de valores extremos (modelos GEV, Weibull e Gumbel) e os resultados se mostram bastante favoráveis ao modelo proposto.

Palavras-chave: Distribuições de valores extremos, regressão quantílica, simulação de Monte Carlo, dados ambientais.

Abstract

The extreme value Birnbaum-Saunders regression model is a very useful tool in the modeling of extreme events. However, this model is implemented by adding covariates to the location parameter. Given the importance of quantile regression to estimate the effects of covariates along the spectrum of the response variable, we introduce an extreme value Birnbaum-Saunders parametric quantile regression model. We implement a likelihood-based approach for parameter estimation and consider two types of residuals. A Monte Carlo simulation study is performed to assess the behavior of the parameter estimation method and the empirical distribution of the residuals. Finally, we illustrate the proposed methodology with the use of a real environmental data set, which is new and is analyzed for the first time here.

Keywords: Extreme value distributions, quantile regression, Monte Carlo simulation, environment data.

Contents

List of Tables

List of Figures

Chapter 1

Parametric quantile regression for extreme events

1.1 Introduction

Extreme value theory has been shown to be very useful for problems involving rare and extreme events. Environmental issues involving hydrology and engineering were the areas that most used this theory for a long time, but medicine, insurance and financial sector, among others, are examples of other areas that benefit from this methodology, see, e.g., Embrechts, Kluppelberg, and Mikosch (1997), Reiss and Thomas (2001), Coles (2001), Haan and Ferreira (2006) and Beirlant, Caeiro, and Gomes (2012). The generalized extreme value (GEV) distribution, developed by Jenkinson (1955) and Jenkinson (1969), is widely used to model extremes of natural events. Particularly, the GEV model is of considerable importance to the environmental area.

A distribution that has been gaining attention in several areas is the Birnbaum-Saunders (BS) model; see, for example, Leiva, Sanhueza, and Angulo (2009), Vilca et al. (2010), Saulo et al. (2013), Leiva et al. (2015), Leiva (2016), Balakrishnan and Kundu (2019), and Oliveira et al. (2022). The BS distribution is related with the normal model and is very useful for modeling strictly positive and asymmetric data. A prominent application of BS models is in extreme value theory. Ferreira, Gomes, and Leiva (2012) proposed the extreme value BS (EVBS) distribution changing the usual normal, in the stochastic representation of the BS model, by the GEV distribution. The authors applied the proposed model to atmospheric data. Further, Leiva et al. (2016) proposed a regression model based on the EVBS distribution. The regression model was illustrated through the used of environmental data.

Quantile regression models are an alternative to the usual mean-based regression, as they present as results the effects of covariates on the response variable in the different quantiles of its distribution, that is, it provides an analysis along the entire conditional distribution; see Cade, Terrell, and Schroeder (1999), Koenker (2005) and Wei et al. (2006). The mean as the only summary measure is generally insufficient for a risk assessment, as it is highly affected by the presence of outliers, which may be rare but sufficient to cause catastrophes in the environmental area and insolvency of insurers, for example.

In this context, the main objective of this paper is to propose a parametric quantile regression model based on the EVBS distribution. To this end, a reparametrization of the EVBS distribution is proposed by inserting a parameter that represents the quantile of the distribution. The rest of this paper proceeds as follows. In Section 1.2, we describe the usual EVBS distribution (Ferreira, Gomes, and Leiva, 2012) and propose a reparameterization of this distribution in terms of a quantile parameter, denoted by QEVBS. In this section, we also establish some results concerning the uni- and bimodality property, stochastic representation, and related distributions, of the QEVBS distribution. In Section 1.3, we introduce the quantile-based QEVBS regression model, as well as the log-likelihood function, score vector, and generalized Cox-Snell and random quantile residuals. In Section 1.4, we carry out Monte Carlo simulation studies for evaluating the performance of the maximum likelihood (ML) estimators and the empirical distribution of residuals. In Section 1.5, an application of the proposed model to a real environmental data set is performed to illustrate the proposed methodology. Finally, In Section 1.6, we mention some concluding remarks.

1.2 Preliminaries

1.2.1 EVBS distribution

Let T be a random variable following a EVBS distribution with shape parameters $\alpha > 0$ and $\xi \in \mathbb{R}$, and scale parameter $\lambda > 0$, denoted by $T \sim \text{EVBS}(\alpha, \lambda, \xi)$. Then, the cumulative distribution function (CDF) associated to T is given by

$$F_T(t; \alpha, \lambda, \xi) = \begin{cases} \exp\big(-(1+\xi a_t)^{-1/\xi}\big), & \text{if } \xi > 0, a_t \geqslant -1/\xi \text{ or } \xi < 0, a_t \leqslant -1/\xi; \\ \exp\big(-\exp(-a_t)\big), & \text{if } \xi = 0, t > 0; \end{cases} \tag{1.1}$$

where

$$a_t = \frac{1}{\alpha}\left(\sqrt{\frac{t}{\lambda}} - \sqrt{\frac{\lambda}{t}}\right).$$

It is well-known that the $100q$-th quantile of $T \sim \text{EVBS}(\alpha, \lambda, \xi)$ is given by

$$Q = Q_T(q; \alpha, \lambda, \xi) = F_T^{-1}(q; \alpha, \lambda, \xi) = \frac{\lambda}{4}\left(\alpha z_q + \sqrt{\alpha^2 z_q^2 + 4}\right)^2, \tag{1.2}$$

with $0 < q < 1$ and z_q is the $100q$-th quantile of the extreme value distribution and it has the following form

$$z_q = \begin{cases} \frac{1}{\xi}\left[(-\log(q))^{-\xi} - 1\right], & \text{if } \xi \neq 0; \\ -\log(-\log(q)), & \text{if } \xi = 0. \end{cases} \tag{1.3}$$

1.2.2 EVBS distribution based on the quantile

Let $0 < q < 1$ be a fixed number and consider a transformation one-to-one $(\alpha, \lambda, \xi) \longmapsto \boldsymbol{\theta}_\xi = (\alpha, Q, \xi)$, where Q is the $100q$-th quantile of T given in the Equation (1.2). Then, reparametrizing the EVBS distribution in function of Q we get the following CDF

$$F_T(t;\boldsymbol{\theta}_\xi) = \begin{cases} \exp\big(-(1+\xi\mathfrak{a}_t)^{-1/\xi}\big), & \text{if } \xi>0, t\geqslant t_\xi \text{ or } \xi<0, t\leqslant t_\xi; \\ \exp\big(-\exp(-\mathfrak{a}_t)\big), & \text{if } \xi=0, t>0. \end{cases} \tag{1.4}$$

Here, $t_\xi = \mathfrak{a}^{-1}_{-1/\xi} = Q[-(\alpha/\xi)+\sqrt{(\alpha/\xi)^2+4}\,]^2/\Lambda_{\alpha,q}$ with $\mathfrak{a}_s^{-1}$ denoting the inverse function of $\mathfrak{a}_t$, $\Lambda_{\alpha,q} = (\alpha z_q + \sqrt{\alpha^2 z_q^2+4}\,)^2$ and

$$\mathfrak{a}_t = \mathfrak{a}_t(\alpha, Q) = \frac{1}{\alpha}\left(\sqrt{\frac{t\Lambda_{\alpha,q}}{4Q}} - \sqrt{\frac{4Q}{t\Lambda_{\alpha,q}}}\right). \tag{1.5}$$

We use the following notation $T \sim \text{QEVBS}(\boldsymbol{\theta}_\xi)$.

Deriving F_T with respect to t, we have that the probability density function (PDF) associated to T is given by

$$f_T(t;\boldsymbol{\theta}_\xi) = \begin{cases} (1+\xi\mathfrak{a}_t)^{(-1/\xi)-1}\exp\big(-(1+\xi\mathfrak{a}_t)^{-1/\xi}\big)\mathfrak{a}_t', & \text{if } \xi>0, t\geqslant t_\xi \text{ or } \xi<0, t\leqslant t_\xi; \\ \exp(-\mathfrak{a}_t)\exp\big(-\exp(-\mathfrak{a}_t)\big)\mathfrak{a}_t', & \text{if } \xi=0,\ t>0; \end{cases} \tag{1.6}$$

for almost all t, with

$$\mathfrak{a}_t' = \frac{1}{2\alpha t}\left(\sqrt{\frac{t\Lambda_{\alpha,q}}{4Q}} + \sqrt{\frac{4Q}{t\Lambda_{\alpha,q}}}\right). \tag{1.7}$$

Since $\lim_{\xi\to 0}(1+\xi x)^{(-1/\xi)-1} = \lim_{\xi\to 0}(1+\xi x)^{-1/\xi} = \exp(-x)$ for all $x\in\mathbb{R}$, it is clear that $\lim_{\xi\to 0} f_T(t;\boldsymbol{\theta}_\xi) = f_T(t;\boldsymbol{\theta}_0)$.

A routine calculation shows that:

- When $\xi>0$; $\lim_{t\to t_\xi^+} f_T(t;\boldsymbol{\theta}_\xi) = \lim_{t\to\infty} f_T(t;\boldsymbol{\theta}_\xi) = 0$.

- When $\xi < 0$; $\lim_{t\to 0^+} f_T(t;\boldsymbol{\theta}_\xi) = 0$ and

$$\lim_{t\to t_\xi^-} f_T(t;\boldsymbol{\theta}_\xi) = \begin{cases} \infty, & \text{if } \xi < -1; \\ \mathfrak{a}'_{t_\xi}, & \text{if } \xi = -1; \\ 0, & \text{if } \xi > -1. \end{cases}$$

- When $\xi = 0$; L'Hôpital's rule gives $\lim_{t\to 0^+} f_T(t;\boldsymbol{\theta}_\xi) = 0$. Moreover, $\lim_{t\to\infty} f_T(t;\boldsymbol{\theta}_\xi) = 0$.

***Proposition* 1.** The point t is a mode of distribution of $T \sim \text{QEVBS}(\boldsymbol{\theta}_\xi)$ if, depending on the parameter ξ, this one is a solution of the following non-linear equation:

$$\begin{aligned} &(1+\xi\mathfrak{a}_t)^{(-1/\xi)-1}\big[1-(\xi+1)(1+\xi\mathfrak{a}_t)^{1/\xi}\big](\mathfrak{a}'_t)^2 + \mathfrak{a}''_t = 0, \quad && \text{when } \xi \neq 0; \\ &\big[\exp(-\mathfrak{a}_t) - 1\big](\mathfrak{a}'_t)^2 + \mathfrak{a}''_t = 0, && \text{when } \xi = 0; \end{aligned}$$

where $\mathfrak{a}_t$ and $\mathfrak{a}'_t$ are given in (1.5) and (1.7), respectively, and

$$\mathfrak{a}''_t = -\frac{1}{4\alpha t^2}\left(\sqrt{\frac{t\Lambda_{\alpha,q}}{4Q}} + 3\sqrt{\frac{4Q}{t\Lambda_{\alpha,q}}}\right). \tag{1.8}$$

Proof. The proof is immediate from a routine differentiation, hence its proof is omitted. □

Modality properties

In this subsection we theoretically characterize the modality shape of QEVBS distribution for the cases: $\xi = 0$, $\xi > 0$, $\xi = -1$, $\xi < -1$ and $-1 < \xi < 0$.

***Theorem* 1.** The QEVBS distribution (1.6) is unimodal when $\xi = 0$.

Proof. Let us denote

$$
\begin{aligned}
g(t) &= \exp(-\mathfrak{a}_t) - 1 = \exp\left(-\frac{1}{\alpha}\left(x - \frac{1}{x}\right)\right) - 1, \\
h(t) &= -\frac{\mathfrak{a}_t''}{(\mathfrak{a}_t')^2} = \frac{x + \dfrac{3}{x}}{\dfrac{1}{\alpha}\left(x + \dfrac{1}{x}\right)^2},
\end{aligned}
\qquad x = \sqrt{\frac{t}{\beta}},\ \beta = \frac{4Q}{\Lambda_{\alpha,q}}.
$$

By Proposition 1, when $\xi = 0$ all mode t satisfies the equation $g(t) = h(t)$. Since (for all $t > 0$)

$$
\begin{aligned}
g'(t) &= -\frac{1}{2\alpha\beta x}\left(\frac{1}{x^2} + 1\right)\exp\left(-\frac{1}{\alpha}\left(x - \frac{1}{x}\right)\right) < 0 \quad \text{and} \\
g''(t) &= \frac{1}{4\alpha^2\beta^2 x^6}\left[\alpha x(x^2+3) + (x^2+1)^2\right]\exp\left(-\frac{1}{\alpha}\left(x - \frac{1}{x}\right)\right) > 0,
\end{aligned}
$$

the function g is strictly decreasing and concave up. Moreover, $g(\beta) = 0$, $\lim_{t\to 0^+} g(t) = \infty$ and $\lim_{t\to\infty} g(t) = -1$.

On the other hand,

$$
h'(t) = -\frac{\alpha(x^4 + 6x^2 - 3)}{2\beta x(x^2+1)^3} \quad \text{and} \quad h''(t) = \frac{3\alpha(x^6 + 9x^4 - 9x^2 - 1)}{4\beta^2 x^3 (x^2+1)^4}.
$$

Note that $h'(t) = 0 \iff x^4 + 6x^2 - 3 = 0 \iff x = \sqrt{2\sqrt{3} - 3} \iff t = t_0 = \beta(2\sqrt{3} - 3)$, in which $h''(t_0) = 3(1.75 + \sqrt{3})(80 - 48\sqrt{3})\alpha/\left[64(2\sqrt{3} - 3)^{3/2}\beta^2\right] < 0$. This implies that t_0 is the unique maximum point of h with maximum value $h(t_0) = \alpha(\sqrt{9 + 6\sqrt{3}})/4$. Further, $h(t) > 0$ for all $t > 0$, and $\lim_{t\to 0^+} h(t) = \lim_{t\to\infty} h(t) = 0$. Notice also that $h''(t) = 0 \iff \alpha(x^2 - 1)/(\beta x) = 0 \iff x = 1 \iff t = \beta$. That is, $t = \beta$ is an inflection point of h.

With the descriptions of the graphs of g and h mentioned previously, it is plausible to expect the plots of g and h to intersect at a single point (Figure 1.1). This ensures the existence of a single positive root of the equation $g(t) = h(t)$. Hence, the QEVBS PDF has a single critical point. Since $\lim_{t\to 0^+} f_T(t;\boldsymbol{\theta}_0) = 0$ and $\lim_{t\to\infty} f_T(t;\boldsymbol{\theta}_0) = 0$, the unimodality follows.

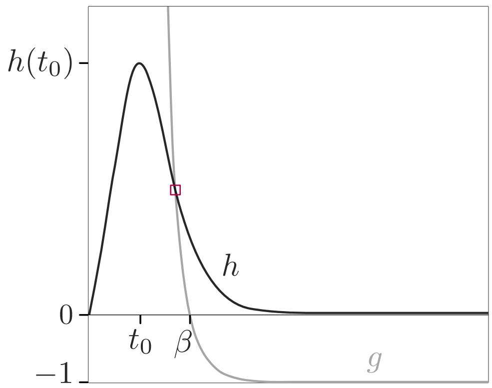


Figure 1.1: Plots of functions g and h for $\xi = 0$.

□

Lemma 1.2.1. *Let* $r(t) = (1+\xi\mathfrak{a}_t)^{(-1/\xi)-1}\big[1-(\xi+1)(1+\xi\mathfrak{a}_t)^{1/\xi}\big]$. *The following is satisfied:*

- *When* $\xi > 0$*;* $\lim_{t\to t_\xi^+} r(t) = \infty$, $\lim_{t\to\infty} r(t) = 0$ *and* $r(t_*) = 0$, *with* $t_* = \mathfrak{a}^{-1}_{[(\xi+1)^{-\xi}-1]/\xi}$.

- *When* $\xi < 0$ *and* $\xi \neq -1$*;*

$$\lim_{t\to 0^+} r(t) = \begin{cases} 0, & \textit{if } \xi < -1; \\ \infty, & \textit{if } \xi > -1; \end{cases} \quad \textit{and} \quad \lim_{t\to t_\xi^-} r(t) = \begin{cases} \infty, & \textit{if } \xi < -1; \\ -\infty, & \textit{if } \xi > -1; \end{cases}$$

  and

$$r(t) \begin{cases} > 0 \textit{ for all } t \in (0, t_\xi), & \textit{if } \xi < -1; \\ = 0 \textit{ in } t_* = \mathfrak{a}^{-1}_{[(\xi+1)^{-\xi}-1]/\xi}, & \textit{if } \xi > -1. \end{cases}$$

- *When* $\xi = -1$*;* $r(t) = 1$.

Proof. The proof immediately follows by applying L'Hôpital's rule. For reasons of space, details are omitted. □

***Theorem* 2.** The QEVBS distribution (1.6) is

1. unimodal when $\xi > 0$;

2. decreasing when $\xi = -1$ and $\alpha < 4/(\sqrt{9+6\sqrt{3}})$;

3. unimodal when $\xi = -1$ and $\alpha = 4/(\sqrt{9+6\sqrt{3}})$;

4. increasing-decreasing-increasing when $\xi = -1$ and $\alpha > 4/(\sqrt{9+6\sqrt{3}})$;

5. increasing-decreasing-increasing when $\xi < -1$ and $\alpha > \tau$, where τ is defined as

$$\tau = \frac{4}{\sqrt{9+6\sqrt{3}}}\,(1+\xi\mathfrak{a}_{\beta(2\sqrt{3}-3)})^{(-1/\xi)-1}\big[1-(\xi+1)(1+\xi\mathfrak{a}_{\beta(2\sqrt{3}-3)})^{1/\xi}\big];$$

6. increasing when $\xi < -1$ and $\alpha < \tau$, where τ is as in Item 5;

7. uni-or bimodal when $-1 < \xi < 0$.

Proof. Let us consider r and h as in Lemma 1.2.1 and Theorem 1, respectively. By Proposition 1, when $\xi \neq 0$ all mode t is a positive root of the equation $r(t) = h(t)$.

Differentiating r gives

$$r'(t) = (\xi+1)(1+\xi\mathfrak{a}_t)^{(-1/\xi)-2}\big[\xi(1+\xi\mathfrak{a}_t)^{1/\xi}-1\big]\mathfrak{a}'_t. \tag{1.9}$$

When $\xi > 0$; $r'(t) = 0 \iff \mathfrak{a}_t = [(1/\xi)^{\xi}-1]/\xi \iff t_1 = \mathfrak{a}^{-1}_{[\xi^{-\xi}-1]/\xi}$ which belongs to the interval (t_ξ, ∞). Since $\lim_{t\to t_\xi^+} r(t) = \infty$ and $\lim_{t\to\infty} r(t) = 0$ (Lemma 1.2.1), it is clear that t_1 is a minimum point of r with minimum value $r(t_1) = -\xi^{\xi}$. Moreover, $r'(t) < 0$ (respectively, > 0) $\iff t < t_1$ (respectively, $t > t_1$) because $1+\xi\mathfrak{a}_t > 0$ and $\mathfrak{a}'_t > 0$. Hence, there is a single point such that $r(t) = h(t)$ (Figure 1.2), and then the QEVBS PDF has a single critical point. Since $\lim_{t\to t_\xi^+} f_T(t;\boldsymbol{\theta}_\xi) = \lim_{t\to\infty} f_T(t;\boldsymbol{\theta}_\xi) = 0$, the unimodality stated in Item 1 follows.

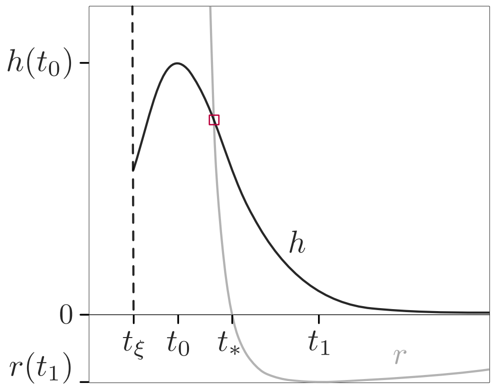


Figure 1.2: Plots of functions r and h for $\alpha/\xi \geqslant (1/\sqrt{2\sqrt{3}-3}) - \sqrt{2\sqrt{3}-3}$ and $\xi > 0$. In this case, $t_\xi \leqslant t_0$.

When $\xi = -1$; $r(t) = 1$ (Lemma 1.2.1). Since h is a positive function with maximum value $h(t_0) = \alpha(\sqrt{9+6\sqrt{3}})/4$ in $t_0 = \beta(2\sqrt{3}-3)$, we have (Figure 1.3)

(i) r and h have no point in common for $h(t_0) < 1$;

(ii) r and h have a single point in common for $h(t_0) = 1$;

(iii) r and h have two points in common for $h(t_0) > 1$.

The three scenarios above ensure that the QEVBS PDF has at most two critical points. But, $\lim_{t\to 0^+} f_T(t;\boldsymbol{\theta}_\xi) = 0$ and $\lim_{t\to t_\xi^-} f_T(t;\boldsymbol{\theta}_\xi) = \mathfrak{a}'_{t_\xi} (> 0)$. So, the proof os Items 2, 3 and 4 follows.

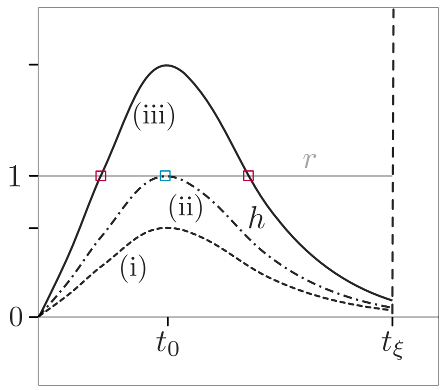


Figure 1.3: Plots of functions r and h for $\xi = -1$. In this case, it is always satisfied that, $t_\xi > t_0$. Three scenarios are highlighted: (i) $h(t_0) < 1$, (ii) $h(t_0) = 1$ and (iii) $h(t_0) > 1$.

Now, let $\xi < -1$ (respectively, $-1 < \xi < 0$). In this case, by using (1.9), see that $r'(t) > 0$ (respectively, $r'(t) < 0$) for all $0 < t < t_\xi$ because $1 + \xi\mathfrak{a}_t > 0$ and $\mathfrak{a}'_t > 0$.

When $\xi < -1$; r is positive, increasing on $(0, t_\xi)$ and $\lim_{t\to 0^+} r(t) = 0$ and $\lim_{t\to t_\xi^-} r(t) = \infty$ (Lemma 1.2.1). Since h is also positive with maximum value $h(t_0)$, we have the following scenarios (Figure 1.4):

(a) r and h have no point in common for $r(t_0) > h(t_0)$;

(b) r and h have a single point in common for $r(t_0) \leqslant h(t_0)$;

(c) r and h have two points in common for $r(t_0) < h(t_0)$.

We claim that scenario (b) cannot occur, otherwise, the QEVBS PDF would have a single critical point. Since $\lim_{t\to 0^+} f_T(t; \boldsymbol{\theta}_\xi) = 0$ and $\lim_{t\to t_\xi^-} f_T(t; \boldsymbol{\theta}_\xi) = \infty$, the QEVBS PDF is forced to have none or at least an even number of critical points, which is a contradiction. This proves the claim. On the other hand, scenarios (a) and (c) ensure that the QEVBS PDF has zero or two critical points. But, $\lim_{t\to 0^+} f_T(t; \boldsymbol{\theta}_\xi) = 0$ and $\lim_{t\to t_\xi^-} f_T(t; \boldsymbol{\theta}_\xi) = \infty$. Then, the proof of Itens 5 and 6 follows.

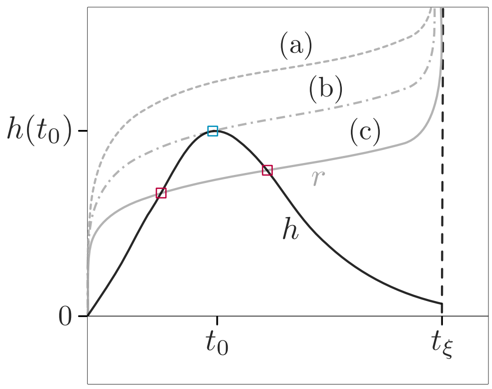


Figure 1.4: Plots of functions r and h for $\xi < -1$. In this case, it is always satisfied that, $t_\xi > t_0$. Three scenarios are highlighted: (a) $r(t_0) > h(t_0)$, (b) $r(t_0) \leqslant h(t_0) = 1$ and (c) $r(t_0) < h(t_0)$.

In the remainder of the proof, we consider the last case $-1 < \xi < 0$. In this case, r is decreasing on $(0, t_\xi)$, crosses the abscissa axis at the point $t_* = \mathfrak{a}^{-1}_{[(\xi+1)^{-\xi}-1]/\xi}$ and $\lim_{t\to 0^+} r(t) = \infty$ and $\lim_{t\to t_\xi^-} r(t) = -\infty$ (Lemma 1.2.1). We have the following scenarios (Figure 1.5):

(d) r and h have a single point in common;

(e) r and h have three points in common.

Both scenarios ensure that the QEVBS PDF has one or three critical points. Since $\lim_{t\to 0^+} f_T(t; \boldsymbol{\theta}_\xi) = 0$ and $\lim_{t\to t_\xi^-} f_T(t; \boldsymbol{\theta}_\xi) = 0$, the uni- or bimodality stated in Item 7 follows.

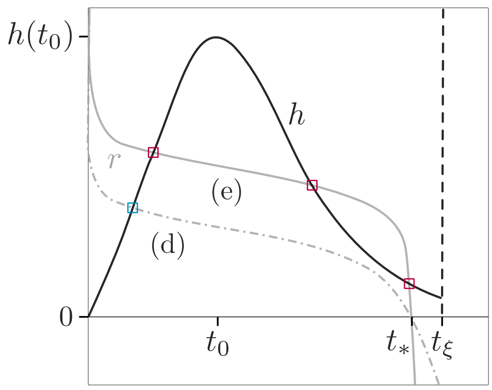


Figure 1.5: Plots of functions r and h for $-1 < \xi < 0$. In this case, it is always satisfied that, $t_\xi > t_0$. The scenarios (d) and (e) are highlighted.

□

Stochastic representation and related distributions

***Proposition* 2.** 1. If $T \sim$ QEVBS$(\boldsymbol{\theta}_\xi)$ then $X = \mathfrak{a}_T \sim GEV(0, 1, \xi)$.

2. If $X \sim GEV(0, 1, \xi)$ then $T = \mathfrak{a}_X^{-1} \sim$ QEVBS$(\boldsymbol{\theta}_\xi)$.

3. If $T \sim$ QEVBS(α, Q, ξ) and $c > 0$ is a constant, then $cT \sim$ QEVBS(α, cQ, ξ).

Proof. If $T \sim$ QEVBS$(\boldsymbol{\theta}_\xi)$ then $\mathbb{P}(X \leqslant x) = \mathbb{P}(T \leqslant \mathfrak{a}_x^{-1}) = F_T(\mathfrak{a}_x^{-1}; \boldsymbol{\theta}_\xi) \overset{(1.4)}{=} F_{\text{GEV}}(x; 0, 1, \xi)$. This proves the first item. The second item is proved similarly.

The proof of Item 3 follows by using the property $\mathfrak{a}_{t/c}(\alpha, Q) = \mathfrak{a}_t(\alpha, cQ)$. □

***Proposition* 3.** 1. If $X \sim \text{Weibull}(\sigma, \mu)$ then $\mathfrak{a}^{-1}_{\mu \log(X/\sigma)} \sim \text{QEVBS}(\boldsymbol{\theta}_0)$.

2. If $T \sim \text{QEVBS}(\boldsymbol{\theta}_0)$ then $\sigma \exp(-\mathfrak{a}_T/\mu) \sim \text{Weibull}(\sigma, \mu)$.

3. If $T_1 \sim \text{QEVBS}(\boldsymbol{\theta}_0)$ and $T_2 \sim \text{QEVBS}(\boldsymbol{\theta}_0)$ are independent random variables, then $\sigma(\mathfrak{a}_{T_1} - \mathfrak{a}_{T_2}) + (\mu_1 - \mu_2) \sim \text{Logistic}(\mu_1 - \mu_2, \sigma)$.

Proof. For $X \sim \text{Weibull}(\sigma, \mu)$ (Weibull distribution) it is well-known that $\mu[1-\sigma \log(X/\sigma)] \sim \text{GEV}(\mu, \sigma, 0)$. Equivalently, $\mu \log(X/\sigma) \sim \text{GEV}(0, 1, 0)$. Then, by using the Item 2 of Proposition 2, the proof of first item follows.

For $T \sim \text{QEVBS}(\boldsymbol{\theta}_0)$, by Item 1 of Proposition 2, $\mathfrak{a}_T \sim \text{GEV}(0, 1, 0)$. By combining this with the known result: If $X \sim \text{GEV}(\mu, \sigma, 0)$ then $\sigma \exp[-(X - \mu)/(\mu\sigma)] \sim \text{Weibull}(\sigma, \mu)$, the proof of Item 2 follows.

Let assume that $T_1 \sim \text{QEVBS}(\boldsymbol{\theta}_0)$ and $T_2 \sim \text{QEVBS}(\boldsymbol{\theta}_0)$ are independent. By Item 1 of Proposition 2, $\mathfrak{a}_{T_1} \sim \text{GEV}(0, 1, 0)$ and $\mathfrak{a}_{T_2} \sim \text{GEV}(0, 1, 0)$. By combining this result with the known fact: If $X_1 \sim \text{GEV}(\mu_1, \sigma, 0)$ and $X_2 \sim \text{GEV}(\mu_2, \sigma, 0)$ are independent then $X_1 - X_2 \sim \text{Logistic}(\mu_1 - \mu_2, \sigma)$ (Logistic distribution), the proof of Item 3 follows. □

1.3 QEVBS regression model

1.3.1 The model

Consider n independent random variables $T_1, \ldots, T_n$ defined over the same measurable space, with $T_i \sim \text{QEVBS}(\boldsymbol{\theta}_{\xi;i})$, $\boldsymbol{\theta}_{\xi;i} = (\alpha, Q_i, \xi)$ for $i = 1, \ldots, n$, and $\mathbf{t} = (t_1, \ldots, t_n)^\top$ its respective observations. Then, the QEVBS regression model can be written as

$$g(Q_i) \;=\; \boldsymbol{x}_i^\top \boldsymbol{\beta}, \quad i = 1, \ldots, n, \tag{1.10}$$

where $\boldsymbol{\beta} = (\beta_0, \ldots, \beta_k)^\top$ denotes an unknown parameter vector to be estimated, $\boldsymbol{x}_i^\top = (1, x_{i1}, \ldots, x_{ik})$ denotes k explanatory variables, $g : \mathbb{R}^+ \to \mathbb{R}$ is an invertible link function, positively supported and twice differentiable.

1.3.2 Estimation

To obtain the ML estimates of the model parameters, $\boldsymbol{\theta}_\xi = (\alpha, \beta_1, \ldots, \beta_k, \xi)$ say, we can maximize the log-likelihood function

$$l(\boldsymbol{\theta}_\xi; \mathbf{t}) = \sum_{i=1}^{n} \log f_{T_i}(t_i; \boldsymbol{\theta}_{\xi;i}), \quad \mathbf{t} = (t_1, \ldots, t_n)^\top \in \mathbb{R}^n, \tag{1.11}$$

where $f_{T_i}(t_i; \boldsymbol{\theta}_{\xi;i})$ is defined in (1.6).

The coordinates of the score vector are given by:

Case $\xi \neq 0$. For $r = 1, \ldots, k$,

$$\frac{\partial l(\boldsymbol{\theta}_\xi; \mathbf{t})}{\partial \alpha} = \sum_{i=1}^{n} \left\{ \frac{\partial \mathfrak{a}_{t_i}}{\partial \alpha} \left[\frac{\xi - 1}{1 + \xi \mathfrak{a}_{t_i}} + \frac{1}{(1 + \xi \mathfrak{a}_{t_i})^{(\xi+1)/\xi}} \right] + \frac{1}{\mathfrak{a}_{t_i}} \frac{\partial \mathfrak{a}'_{t_i}}{\partial \alpha} \right\},$$

$$\frac{\partial l(\boldsymbol{\theta}_\xi; \mathbf{t})}{\partial \beta_r} = \sum_{i=1}^{n} \left\{ \frac{\partial \mathfrak{a}_{t_i}}{\partial Q_i} \left[\frac{\xi - 1}{1 + \xi \mathfrak{a}_{t_i}} + \frac{1}{(1 + \xi \mathfrak{a}_{t_i})^{(\xi+1)/\xi}} \right] + \frac{1}{\mathfrak{a}_{t_i}} \frac{\partial \mathfrak{a}'_{t_i}}{\partial Q_i} \right\} \frac{\partial Q_i}{\partial \beta_r},$$

$$\frac{\partial l(\boldsymbol{\theta}_\xi; \mathbf{t})}{\partial \xi} = \sum_{i=1}^{n} \left\{ \frac{\log(1 + \xi \mathfrak{a}_{t_i})}{\xi^2} - \frac{\mathfrak{a}_{t_i}}{\xi} \left[\frac{\xi + 1}{1 + \xi \mathfrak{a}_{t_i}} - \frac{1}{(1 + \xi \mathfrak{a}_{t_i})^{(\xi+1)/\xi}} \right] \right\},$$

with

$$\frac{\partial \mathfrak{a}_{t_i}}{\partial \alpha} = \frac{z_q \Lambda_{\alpha,q}}{\alpha(\Lambda_{\alpha,q} - \alpha z_q)} \left(\frac{Q_i}{t_i \Lambda_{\alpha,q}} \sqrt{\frac{t_i \Lambda_{\alpha,q}}{Q_i}} + \frac{t_i}{4 Q_i} \sqrt{\frac{Q_i}{t_i \Lambda_{\alpha,q}}} \right) - \frac{1}{\alpha^2} \left(\frac{1}{2} \sqrt{\frac{t_i \Lambda_{\alpha,q}}{Q_i}} - \sqrt{\frac{Q_i}{t_i \Lambda_{\alpha,q}}} \right),$$

$$\frac{\partial \mathfrak{a}_{t_i}}{\partial Q_i} = -\frac{1}{4 \alpha Q_i} \left(\sqrt{\frac{t_i \Lambda_{\alpha,q}}{Q_i}} + 4 \sqrt{\frac{Q_i}{t_i \Lambda_{\alpha,q}}} \right)$$

and

$$\frac{\partial \mathfrak{a}'_{t_i}}{\partial \alpha} = -\frac{z_q \Lambda_{\alpha,q}}{2\alpha(\Lambda_{\alpha,q} - \alpha z_q) t_i} \left(\frac{Q_i}{t_i \Lambda_{\alpha,q}} \sqrt{\frac{t_i \Lambda_{\alpha,q}}{Q_i}} - \frac{t_i}{4Q_i} \sqrt{\frac{Q_i}{t_i \Lambda_{\alpha,q}}} \right) - \frac{1}{2\alpha^2 t_i} \left(\frac{1}{2} \sqrt{\frac{t_i \Lambda_{\alpha,q}}{Q_i}} + \sqrt{\frac{Q_i}{t_i \Lambda_{\alpha,q}}} \right),$$

$$\frac{\partial \mathfrak{a}'_{t_i}}{\partial Q_i} = -\frac{1}{8\alpha Q_i t_i} \left(\sqrt{\frac{t_i \Lambda_{\alpha,q}}{Q_i}} - 4\sqrt{\frac{Q_i}{t_i \Lambda_{\alpha,q}}} \right),$$

and, by using the relation (1.10),

$$\frac{\partial Q_i}{\partial \beta_0} = \frac{1}{g'(Q_i)}; \quad \frac{\partial Q_i}{\partial \beta_r} = \frac{x_{ir}}{g'(Q_i)}.$$

Case $\xi = 0$. For $r = 1, \ldots, k$,

$$\begin{aligned}
\frac{\partial l(\boldsymbol{\theta}_0; \mathbf{t})}{\partial \alpha} &= \sum_{i=1}^{n} \left\{ \frac{\partial \mathfrak{a}_{t_i}}{\partial \alpha} [\exp(-\mathfrak{a}_{t_i}) - 1] + \frac{1}{\mathfrak{a}_{t_i}} \frac{\partial \mathfrak{a}'_{t_i}}{\partial \alpha} \right\}, \\
\frac{\partial l(\boldsymbol{\theta}_0; \mathbf{t})}{\partial \beta_r} &= \sum_{i=1}^{n} \left\{ \frac{\partial \mathfrak{a}_{t_i}}{\partial Q_i} [\exp(-\mathfrak{a}_{t_i}) - 1] + \frac{1}{\mathfrak{a}_{t_i}} \frac{\partial \mathfrak{a}'_{t_i}}{\partial Q_i} \right\} \frac{\partial Q_i}{\partial \beta_r},
\end{aligned}$$

where the partial derivatives of $\mathfrak{a}_{t_i}$, $\mathfrak{a}'_{t_i}$ and of Q_i with respect to parameters are as in Case $\xi \neq 0$.

The ML estimators of α, β_r with $r = 1, \ldots, k$, and ξ can be determined by solving the following system of equations,

$$\begin{aligned}
&\frac{\partial l(\boldsymbol{\theta}_\xi; \boldsymbol{t})}{\partial \alpha} = \frac{\partial l(\boldsymbol{\theta}_\xi; \boldsymbol{t})}{\partial \beta_r} = \frac{\partial l(\boldsymbol{\theta}_\xi; \boldsymbol{t})}{\partial \xi} = 0, \quad \text{for } \xi \neq 0; \\
&\frac{\partial l(\boldsymbol{\theta}_0; \boldsymbol{t})}{\partial \alpha} = \frac{\partial l(\boldsymbol{\theta}_0; \boldsymbol{t})}{\partial \beta_r} = 0, \quad \text{for } \xi = 0.
\end{aligned}$$

Numerical methods, such as the Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-Newton method, must be used to solve these systems.

1.3.3 Residuals

To assess goodness of fit and departures from the assumptions of the regression model, we use the generalized Cox-Snell (GCS) and random quantile (RQ) residuals, given by

$$\widehat{r}_i^{\mathrm{GCS}} = -\log(1 - F_Y(y_i; \widehat{\boldsymbol{\theta}}_{\xi;i})), \quad i = 1, \ldots, n, \tag{1.12}$$

where F_Y is as in (1.4). If the model is correctly specified, the GCS residual is asymptotically standard exponential distributed. The RQ residual is defined as

$$\widehat{r}_i^{\mathrm{RQ}} = \Phi^{-1}(F_Y(y_i; \widehat{\boldsymbol{\theta}}_{\xi;i})), \quad i = 1, \ldots, n, \tag{1.13}$$

where Φ^{-1} is the inverse function of the standard normal CDF, and $\widehat{\boldsymbol{\theta}}_{\xi;i}$ is an ML estimate of $\boldsymbol{\theta}_{\xi;i}$. The RQ residual follows approximately a standard normal distribution, if the model is correctly specified. For both residuals, the distribution assumption can be assessed through graphical techniques and descriptive statistics.

1.4 Monte Carlo simulation

We present the results of two Monte Carlo simulation studies for the QEVBS regression model. The first study considers the evaluation of the performance of the ML estimation procedure, while the second study evaluates the empirical distribution of the GCS and RQ residuals. Both studies consider simulated data generated from $Q_i = \exp(\beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i})$.

1.4.1 ML estimation

The simulation scenario considers the following setting: quantiles $q \in \{0.1, 0.25, 0.5, 0.75, 0.9\}$, sample sizes $n \in \{50, 150, 400\}$, true parameters $\alpha = 1.5$ and $\xi \in \{-0.25, 0, 0.25\}$, with 1,000 Monte Carlo replications. The covariate values for x_1 and x_2 are obtained from the Uniform (5, 20) and Uniform(10, 50) distributions, respectively.

The performance and recovery of the ML estimators are assessed by the bias and mean squared error (MSE), which are stated respectively by

$$\mathrm{Bias}(\widehat{\varphi}) = \frac{1}{M}\sum_{r=1}^{M}(\widehat{\varphi}^{(r)} - \varphi), \quad \mathrm{MSE}(\widehat{\varphi}) = \frac{1}{M}\sum_{r=1}^{M}(\widehat{\varphi}^{(r)} - \varphi)^2, \tag{1.14}$$

where φ and $\widehat{\varphi}^{(r)}$ are the true parameter value and its respective r-th maximum likelihood estimate, and M is the number of Monte Carlo replicas. The steps of the Monte Carlo simulation study are described in Algorithm 1.

Algoritmo 1: Simulation 1: ML estimation

1: Generate 1000 samples based from the QEVBS regression model.
2: Estimate the model parameters using the ML method for each sample.
3: Compute the bias and MSE using (1.14).

Tables 1.1-1.3 contain the Monte Carlo simulation results. The following sample statistics for the ML estimates are reported: mean, bias and MSE. The results in these tables allow us to conclude that, as expected, in general, as the sample size increases, the bias and MSE decrease, for all values ξ and quantiles considered.

Table 1.1: Mean, Bias and MSE from simulated data in the QEVBS regression with $\alpha = 1.5$ and $\xi = -0.25$.

q		$n = 50$			$n = 150$			$n = 400$		
		Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
0.1	β_0	3.0805	0.0805	1.5007	2.9945	−0.0055	0.4602	2.9909	−0.0091	0.1546
	β_1	0.4750	−0.0250	0.0075	0.4601	−0.0399	0.0134	0.4532	−0.0468	0.0044
	β_2	0.9650	−0.0350	0.0049	0.9501	−0.0499	0.0075	0.9333	−0.0667	0.0058
	α	1.4642	−0.0358	0.0484	1.4588	−0.0412	0.0537	1.4629	−0.0371	0.0384
	ξ	−0.2717	−0.0217	0.0384	−0.2765	−0.0265	0.0243	−0.2263	0.0237	0.0084
0.25	β_0	2.9789	−0.0211	1.3979	2.9787	−0.0213	0.4548	2.9802	−0.0198	0.1514
	β_1	0.4623	−0.0377	0.0113	0.4399	−0.0601	0.0108	0.4455	−0.0545	0.0047
	β_2	0.9477	−0.0523	0.0086	0.9348	−0.0652	0.0101	0.9228	−0.0772	0.0072
	α	1.4378	−0.0622	0.0765	1.4261	−0.0739	0.0783	1.4711	−0.0289	0.0206
	ξ	−0.2599	−0.0099	0.0302	−0.2328	0.0172	0.0120	−0.2036	0.0464	0.0035
0.5	β_0	3.0533	0.0533	1.2073	3.0283	0.0283	0.4134	2.9723	−0.0277	0.1665
	β_1	0.4344	−0.0656	0.0141	0.4297	−0.0703	0.0089	0.4412	−0.0588	0.0042
	β_2	0.9222	−0.0778	0.0132	0.9140	−0.0860	0.0093	0.9103	−0.0897	0.0082
	α	1.3956	−0.1044	0.1218	1.4645	−0.0355	0.0381	1.4797	−0.0203	0.0154
	ξ	−0.2510	−0.0010	0.0200	−0.2107	0.0393	0.0048	−0.2001	0.0499	0.0025
0.75	β_0	2.9633	−0.0367	1.0039	3.0196	0.0196	0.2683	2.9770	−0.0230	0.1036
	β_1	0.4219	−0.0781	0.0155	0.4304	−0.0696	0.0068	0.4357	−0.0643	0.0048
	β_2	0.9074	−0.0926	0.0136	0.8981	−0.1019	0.0110	0.9021	−0.0979	0.0098
	α	1.3614	−0.1386	0.1639	1.4666	−0.0334	0.0300	1.4765	−0.0235	0.0104
	ξ	−0.2293	0.0207	0.0111	−0.2028	0.0472	0.0031	−0.2000	0.0500	0.0025
0.9	β_0	2.9116	−0.0884	0.9729	2.9592	−0.0408	0.2929	2.9601	−0.0399	0.1099
	β_1	0.4101	−0.0899	0.0191	0.4235	−0.0765	0.0084	0.4329	−0.0671	0.0051
	β_2	0.9031	−0.0969	0.0173	0.8944	−0.1056	0.0118	0.8959	−0.1041	0.0109
	α	1.3338	−0.1662	0.1879	1.4629	−0.0371	0.0301	1.4889	−0.0111	0.0097
	ξ	−0.2189	0.0311	0.0087	−0.2006	0.0494	0.0026	−0.2000	0.0500	0.0025

Table 1.2: Mean, Bias and MSE from simulated data in the QEVBS regression with $\alpha = 1.5$ and $\xi = 0$.

q		$n = 50$			$n = 150$			$n = 400$		
		Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
0.1	β_0	3.1311	0.1311	0.9866	3.0736	0.0736	0.3040	3.0239	0.0239	0.1102
	β_1	0.4758	−0.0242	0.0072	0.4687	−0.0313	0.0056	0.4701	−0.0299	0.0035
	β_2	0.9769	−0.0231	0.0058	0.9624	−0.0376	0.0069	0.9476	−0.0524	0.0060
	α	1.4099	−0.0901	0.1268	1.3720	−0.1280	0.1629	1.3514	−0.1486	0.2088
	ξ	−0.0577	−0.0577	0.0508	−0.1334	−0.1334	0.0448	−0.1569	−0.1569	0.0407
0.25	β_0	3.0059	0.0059	1.1992	2.9856	−0.0144	0.3700	3.0016	0.0016	0.1197
	β_1	0.4676	−0.0324	0.0674	0.4589	−0.0411	0.0131	0.4994	−0.0006	1.4514
	β_2	0.9719	−0.0281	0.0162	0.9482	−0.0518	0.0119	0.9203	−0.0797	0.5383
	α	1.3734	−0.1266	0.3904	1.2844	−0.2156	0.2834	1.2844	−0.2156	0.2771
	ξ	−0.0961	−0.0961	0.0487	−0.1311	−0.1311	0.0425	−0.1353	−0.1353	0.0400
0.5	β_0	3.0121	0.0121	1.4879	3.0201	0.0201	0.4530	2.9794	−0.0206	0.1678
	β_1	0.4480	−0.0520	0.0299	0.4313	−0.0687	0.2308	0.4572	−0.0428	0.0123
	β_2	0.9546	−0.0454	0.0202	0.9335	−0.0665	0.0516	0.9326	−0.0674	0.0159
	α	1.2203	−0.2797	0.3620	1.2046	−0.2954	0.3670	1.2689	−0.2311	0.4507
	ξ	−0.0791	−0.0791	0.0449	−0.1086	−0.1086	0.0405	−0.1256	−0.1256	0.0400
0.75	β_0	2.9909	−0.0091	1.7748	2.9797	−0.0203	0.4963	2.9986	−0.0014	0.3386
	β_1	0.4219	−0.0781	0.0418	0.4237	−0.0763	0.1370	0.4010	−0.0990	1.5954
	β_2	0.9316	−0.0684	0.0242	0.9249	−0.0751	0.2018	0.9531	−0.0469	0.7400
	α	1.1719	−0.3281	0.4964	1.1770	−0.3230	0.4663	1.2525	−0.2475	0.4959
	ξ	−0.0832	−0.0832	0.0424	−0.0896	−0.0896	0.0400	−0.1136	−0.1136	0.0400
0.9	β_0	2.9389	−0.0611	4.1137	3.2364	0.2364	16.3833	2.9481	−0.0519	1.2233
	β_1	0.4191	−0.0809	0.0401	0.3631	−0.1369	1.3530	0.3911	−0.1089	2.2122
	β_2	0.9151	−0.0849	0.0353	0.9154	−0.0846	0.2675	0.9344	−0.0656	0.6263
	α	1.1974	−0.3026	0.3824	1.4114	−0.0886	3.0178	1.3825	−0.1175	0.3628
	ξ	−0.0976	−0.0976	0.0411	−0.1336	−0.1336	0.0400	−0.1512	−0.1512	0.0400

Table 1.3: Mean, Bias and MSE from simulated data in the QEVBS regression with $\alpha = 1.5$ and $\xi = 0.25$.

q		$n = 50$			$n = 150$			$n = 400$		
		Mean	Bias	MSE	Mean	Bias	MSE	Mean	Bias	MSE
0.1	β_0	3.0526	0.0526	0.8202	3.0634	0.0634	0.2321	3.0512	0.0512	0.3911
	β_1	0.5053	0.0053	0.0177	0.4820	−0.0180	0.0216	0.4956	−0.0044	0.3281
	β_2	1.0106	0.0106	0.0143	1.0036	0.0036	0.0142	0.9991	−0.0009	0.3784
	α	1.2691	−0.2309	1.4077	0.9336	−0.5664	0.8062	0.5016	−0.9984	1.4274
	ξ	0.3038	0.0538	0.0495	0.2791	0.0291	0.0423	0.1987	−0.0513	0.0309
0.25	β_0	3.0927	0.0927	1.3013	3.0249	0.0249	0.3308	3.1654	0.1654	1.4116
	β_1	0.5121	0.0121	0.1136	0.4754	−0.0246	0.6886	0.4551	−0.0449	1.5110
	β_2	0.9883	−0.0117	0.0335	0.9819	−0.0181	0.2124	1.0048	0.0048	0.4855
	α	1.1847	−0.3153	0.5828	0.6924	−0.8076	1.2689	0.2952	−1.2048	2.0500
	ξ	0.3002	0.0502	0.0468	0.2346	−0.0154	0.0315	0.1659	−0.0841	0.0212
0.5	β_0	3.0008	0.0008	1.5623	3.1282	0.1282	2.0887	3.0816	0.0816	0.3589
	β_1	0.5487	0.0487	1.0291	0.5555	0.0555	3.2426	0.4221	−0.0779	1.3033
	β_2	0.9295	−0.0705	0.5195	0.9332	−0.0668	0.2953	0.9382	−0.0618	0.3851
	α	0.8358	−0.6642	1.3379	0.5459	−0.9541	4.5417	0.1589	−1.3411	2.0752
	ξ	0.2208	−0.0292	0.0338	0.1800	−0.0700	0.0185	0.1618	−0.0882	0.0143
0.75	β_0	3.1031	0.1031	3.0027	3.0983	0.0983	1.0571	2.9180	−0.0820	4.3098
	β_1	0.4294	−0.0706	0.1311	0.4446	−0.0554	1.3651	0.3763	−0.1237	3.8415
	β_2	0.9357	−0.0643	0.0893	0.8946	−0.1054	0.6312	0.9330	−0.0670	2.1309
	α	0.5570	−0.9430	1.3430	0.2957	−1.2043	2.0860	0.3191	−1.1809	3.9285
	ξ	0.1615	−0.0885	0.0273	0.1507	−0.0993	0.0190	0.1553	−0.0947	0.0144
0.9	β_0	2.9009	−0.0991	4.2848	2.9096	−0.0904	32.1892	3.0049	0.0049	1.8110
	β_1	0.4674	−0.0326	0.4604	0.4110	−0.0890	2.6626	0.5477	0.0477	5.4679
	β_2	0.8447	−0.1553	0.2199	0.7668	−0.2332	1.3460	0.6520	−0.3480	1.7127
	α	0.4326	−1.0674	1.6137	0.1392	−1.3608	2.1962	0.1963	−1.3037	3.1735
	ξ	0.1184	−0.1316	0.0310	0.0895	−0.1605	0.0335	0.0773	−0.1727	0.0369

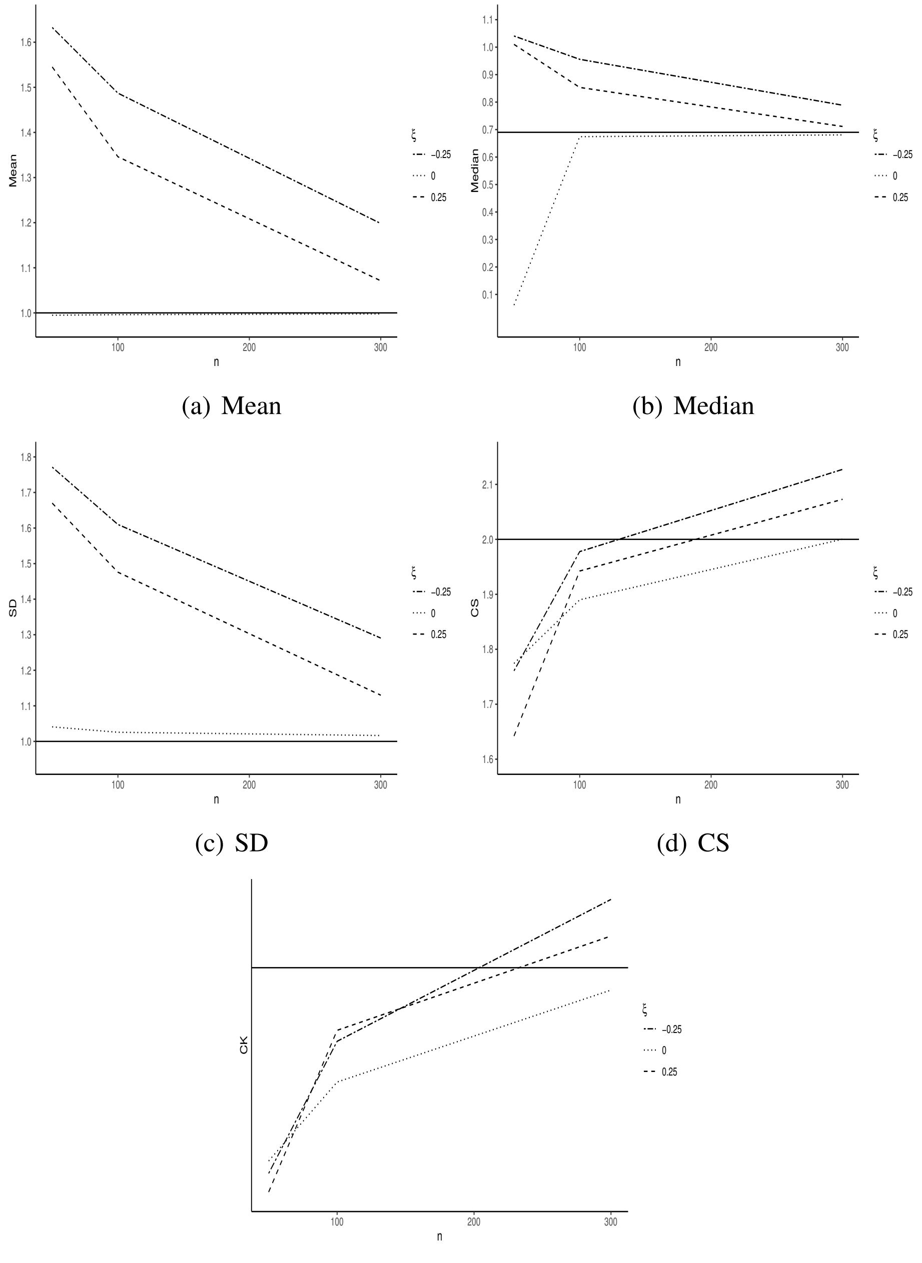
1.4.2 Empirical distribution of the residuals

We present a Monte Carlo simulation study that evaluates the performance of the GCS and RQ residuals. The simulation scenario considers the following setting: quantiles $q \in \{0.1, 0.25, 0.5, 0.75, 0.9\}$, sample sizes $n \in \{50, 100, 300\}$, true parameters $\alpha = 1.5$ and $\xi \in \{-0.25, 0, 0.25\}$, with 1,000 Monte Carlo replications. The covariate values for x_1 and x_2 are obtained from the Uniform (1, 10) and Uniform(1, 5) distributions, respectively. The steps of the Monte Carlo simulation study are described in Algorithm 2.

Algoritmo 2: Simulation 1: Residuals

1: Generate 1000 samples based from the QEVBS regression model.
2: Estimate the model parameters using the ML method for each sample.
3: Calculate, for each replication, the GCS and RQ residuals and the respective descriptive statistics: mean, median, standard deviation (SD), coefficient of skewness (CS) and coefficient of (excess) kurtosis (CK).
4: Determine the means of the descriptive statistics obtained in Step 3 from the 1000 replications.

Tables 1.4-1.5 and Figures 1.6-1.7 ($q = 0.50$) report the Monte Carlo simulation results for the GCS and RQ residuals. The objective here is to verify whether the GCS and RQ residuals behave according to their reference distributions. Then, the mean, median, SD, CS and CK are expected to be 1, 0.69, 1, 2 and 6, respectively, for the GCS residuals, and 0, 0, 1, 0 and 0, respectively, for the RQ residuals. From these tables and figures note that, in general, the considered residuals conform well with their reference distributions. Nevertheless, note that the CK values for the RQ residuals are persistently large with $\xi \in \{-0.25, 0.25\}$ and $q \in \{0.50, 0.75, 0.90\}$.



(a) Mean

(b) Median

(c) SD

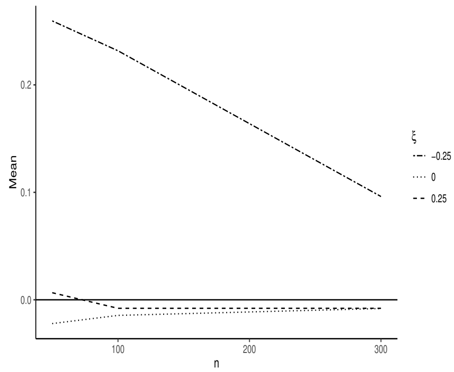
(d) CS

(e) CK

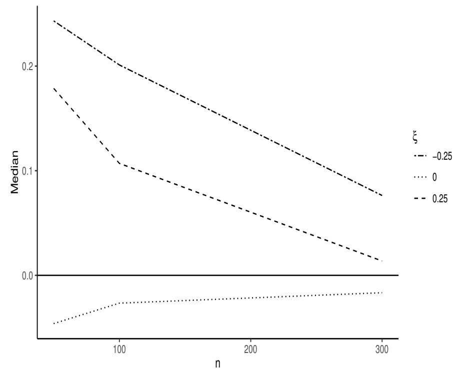
Figure 1.6: Statistics values for GCS residuals ($q = 0.50$).

Table 1.4: Summary statistics of the GCS residuals with simulated data.

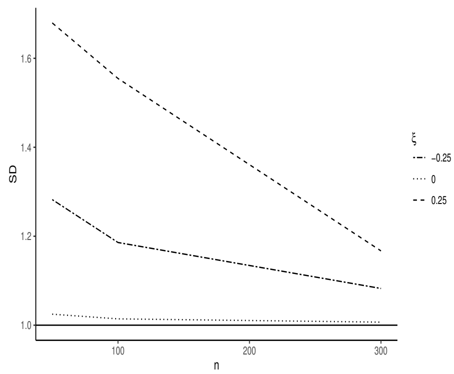
q	*Statistic*	$n=50$			$n=100$			$n=300$		
		ξ=-0.25	ξ=0	ξ=0.25	ξ=-0.25	ξ=0	ξ=0.25	ξ=-0.25	ξ=0	ξ=0.25
0.1	Mean	1.2973	0.9954	1.0638	1.2221	0.9961	1.0472	1.2278	0.9979	0.9993
	Median	0.8319	0.6604	0.7498	0.8055	0.6721	0.7351	0.8249	0.6825	0.6964
	SD	1.4051	1.0354	1.0409	1.3046	1.0288	1.0290	1.2886	1.0146	0.9950
	CS	1.6022	1.7169	1.4964	1.7945	1.8954	1.7023	1.9853	1.9836	1.9027
	CK	2.7144	3.2538	2.4149	4.0036	4.5727	3.5691	5.5664	5.5593	5.1284
0.25	Mean	1.3743	0.9950	1.0535	1.2020	0.9956	1.0272	1.2219	0.9976	0.9987
	Median	0.8994	0.6599	0.7400	0.7795	0.6699	0.7152	0.8246	0.6817	0.6941
	SD	1.4506	1.0389	1.0387	1.2931	1.0273	1.0163	1.2641	1.0164	0.9922
	CS	1.6700	1.7489	1.5439	1.8595	1.8781	1.7015	2.0508	1.9891	1.8695
	CK	3.0162	3.4487	2.6035	4.3389	4.4670	3.5878	6.1273	5.5788	4.8831
0.5	Mean	1.6327	0.9948	1.5452	1.4871	0.9961	1.3461	1.1981	0.9978	1.0712
	Median	1.0406	0.0611	1.0101	0.9557	0.6742	0.8534	0.7887	0.6806	0.7113
	SD	1.7713	1.0409	1.6701	1.6100	1.0257	1.4758	1.2906	1.0165	1.1299
	CS	1.7608	1.7742	1.6421	1.9777	1.8900	1.9424	2.1273	2.0003	2.0729
	CK	3.4821	3.6341	3.2538	5.0989	4.6001	5.2333	6.8359	5.7262	6.3877
0.75	Mean	1.8050	0.9941	1.4887	2.1657	0.9962	1.1856	1.6200	0.9976	1.1125
	Median	1.0850	0.6563	0.8825	1.4257	0.6732	0.7631	1.0502	0.68126	0.6999
	SD	2.0906	1.0418	1.7639	2.2812	1.0244	1.2898	1.7397	1.0167	1.2282
	CS	1.9845	1.7522	1.8133	2.1839	1.8761	1.9547	2.2846	1.9971	2.1424
	CK	4.7185	3.4277	4.0314	6.5064	4.5055	5.2573	7.9629	5.6668	6.9411
0.9	Mean	1.7956	0.9913	1.2745	2.1073	0.9924	1.1053	1.6406	0.9963	1.0171
	Median	1.0784	0.6563	0.7470	1.3663	0.6627	0.6947	1.0627	0.6786	0.6616
	SD	2.0470	1.0415	1.5193	2.2554	1.0329	1.2323	1.7631	1.0211	1.1138
	CS	1.9645	1.7579	1.8595	2.3405	1.9195	2.0603	2.3418	2.0226	2.2971
	CK	4.5059	3.4701	4.0941	7.6897	4.7936	5.8199	8.3939	5.8572	8.2580



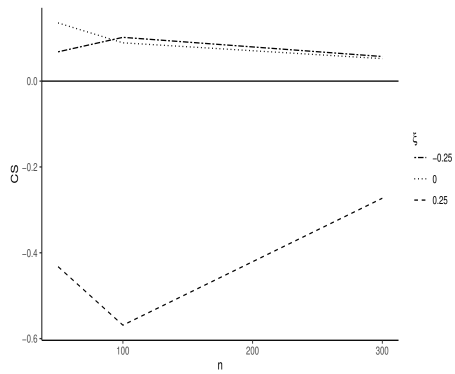
(a) Mean



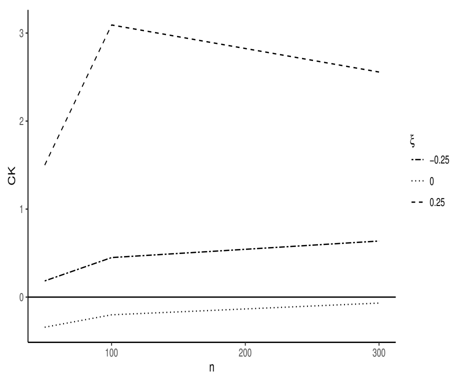
(b) Median



(c) SD



(d) CS



(e) CK

Figure 1.7: Statistics values for RQ residuals ($q = 0.50$).

Table 1.5: Summary statistics of the RQ residuals with simulated data.

q	$Statistic$	$n = 50$			$n = 100$			$n = 300$		
		ξ=-0.25	ξ=0	ξ=0.25	ξ=-0.25	ξ=0	ξ=0.25	ξ=-0.25	ξ=0	ξ=0.25
0.1	Mean	0.0806	−0.0214	0.0201	0.0722	−0.0150	0.0145	0.1035	−0.0075	−0.0006
	Median	0.1052	−0.0472	0.0422	0.0795	−0.0290	0.0238	0.1000	−0.0143	0.0034
	SD	1.2337	1.0249	1.0375	1.1418	1.0147	1.0294	1.0910	1.0064	1.0015
	CS	−0.0754	0.1278	−0.0532	−0.0539	0.0939	−0.0547	−0.0281	0.0479	−0.0047
	CK	−0.1049	−0.3861	−0.3798	0.1548	−0.1803	−0.0488	0.3885	−0.0728	−0.0866
0.25	Mean	0.1505	−0.0217	0.0170	0.0629	−0.0159	0.0069	0.1052	−0.0086	−0.0020
	Median	0.1421	−0.0476	0.0334	0.0665	−0.0322	0.0162	0.0920	−0.0154	0.0005
	SD	1.1799	1.0247	1.0381	1.1457	1.0149	1.0232	1.0712	1.0070	1.0029
	CS	0.0192	0.1312	−0.0417	−0.0158	0.0955	−0.0472	0.0169	0.0547	−0.0088
	CK	−0.0162	−0.3509	-0.3544	0.1406	−0.2057	−0.1126	0.6008	−0.0806	−0.1028
0.5	Mean	0.2595	−0.0221	0.0066	0.2317	−0.0145	−0.0079	0.0961	−0.0081	−0.0079
	Median	0.2432	−0.0461	0.1787	0.2011	−0.0266	0.1070	0.0763	−0.0166	0.0137
	SD	1.2824	1.0247	1.6796	1.1859	1.0142	1.5547	1.0825	1.0068	1.1670
	CS	0.0682	0.1356	−0.4323	0.1020	0.0892	−0.5687	0.0573	0.0524	−0.2729
	CK	0.1837	−0.3423	1.4997	0.4492	−0.2012	3.0921	0.6371	−0.0680	2.5572
0.75	Mean	0.3897	−0.0239	−0.1864	0.6104	−0.0143	−0.0586	0.3504	−0.0084	0.0066
	Median	0.3198	−0.0525	0.0491	0.5160	−0.0278	0.0228	0.2859	−0.0159	0.0005
	SD	1.3340	1.0256	1.9642	1.2543	1.0140	1.3871	1.1420	1.0069	1.1821
	CS	0.2476	0.1428	−0.4976	0.3065	0.0873	−0.3619	0.2267	0.0538	−0.3578
	CK	0.7512	−0.3579	1.9586	1.0904	−0.2072	2.2437	1.0433	−0.0722	4.1542
0.9	Mean	0.4317	−0.0274	−0.2083	0.6218	−0.0208	−0.0584	0.3663	−0.0110	−0.0182
	Median	0.3229	−0.0530	−0.0525	0.4961	−0.0418	−0.0338	0.2898	−0.0194	−0.0428
	SD	1.2489	1.0253	1.6791	1.2086	1.0155	1.2405	1.1343	1.0078	1.0550
	CS	0.3575	0.1537	−0.1440	0.4642	0.1164	0.0030	0.2599	0.0686	0.0580
	CK	0.6437	−0.3645	0.8340	1.4506	−0.1808	0.7781	1.2029	−0.0693	1.3067

1.5 Application to real data

To illustrate the applicability of the proposed model we use an environmental data set obtained from the Chilean Directorate of Meteorology and the Chilean National Information System of Air Quality; see `https://climatologia.meteochile.gob.cl` and `https://sinca.mma.gob.cl/index.php/estacion/index/id/273`, respectively. These data were collected between 01-July-2021 and 22-September-2021 in the central region of Santiago, Chile. The response variable ($ozone$) is the daily maximum ozone concentration in the environment, measured in parts-per-billion (ppb), whereas the covariate ($temp$) is the daily maximum temperature (measured in °C); see Appendix A. Thus, the data under analysis correspond to maximum values, which are intrinsically extreme values, and the QEVBS regression model can be suitable for describing appropriately the relationship between the response variable and the covariate.

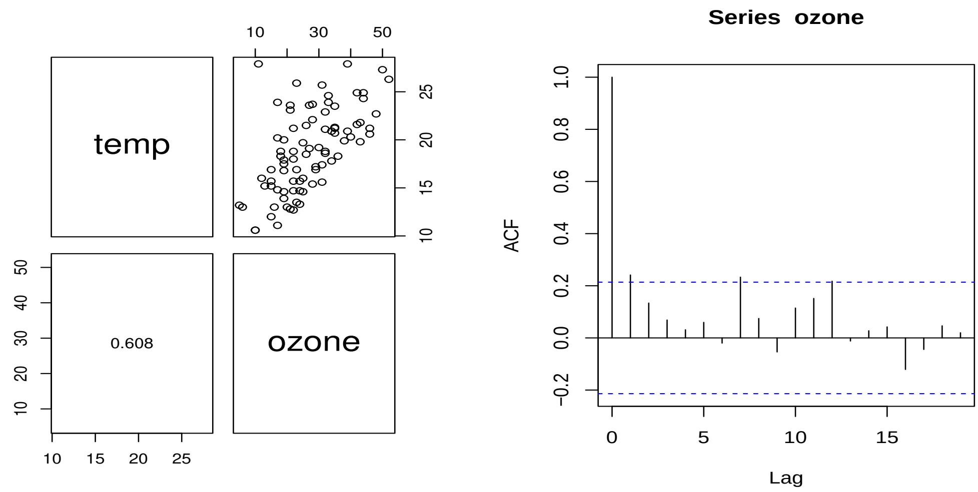
Table 1.6 provides descriptive statistics of the daily maximum ozone concentration in the central region of Santiago, which includes the sample, mean, median, minimum, maximum, SD,CS, CK, and coefficient of variation (CV) values. We observe that the median and mean are respectively 25.00 and 26.87, namely, the mean is greater than the median indicating a positively skewed feature in the data distribution. Moreover, the CV is 39.4541, which reports a high level of dispersion around the mean. We also observe that the CS value confirms the skewed nature and the CK value indicates a high kurtosis feature in the data distribution.

Table 1.6: Summary statistics for the the daily maximum ozone concentration in the environment, in ppb.

n	Min.	Median	Mean	Max.	SD	CS	CK	CV
84	5.00	25.00	26.87	52.00	10.6090	0.3261	2.5063	39.4541

Figure 1.8 shows the scatterplot (with correlation) between the daily maximum ozone concentrations and maximum temperatures, as well as the autocorrelation function (ACF) for the ozone concentrations. From this figure, note that a positively relation between the maximum

level of ozone and the maximum temperature exists, whose sample correlation coefficient is 0.608. Moreover, ACF plot indicates to the absence of dependence in the observations.



(a) scatterplot matrices (b) ACF

Figure 1.8: Scatterplot (left) and ACF plot (right) for the daily maximum ozone data.

We consider the following regression model

$$Q_i \quad = \quad \exp\left(\beta_0 + \beta_1 temp_i\right), \quad i = 1, \ldots, 84, \tag{1.15}$$

to describe the environmental data from Santiago, Chile.

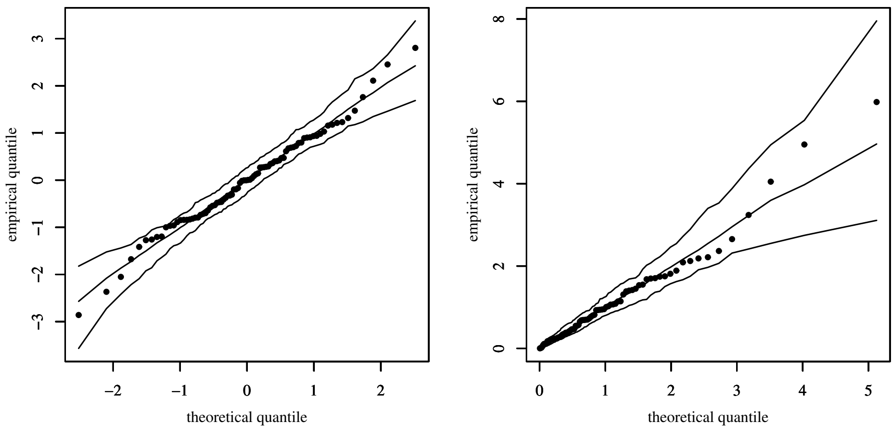
The ML estimates of the parameters of the model, with approximate standard errors in parenthesis, are presented in Table 1.7, for different values of q. Note that the estimates of the β_0 (intercept) vary across q, indicating the importance to consider a quantile approach. Figure 1.9 displays the quantile versus quantile (QQ) plots with simulated envelope of the GCS and RQ residuals. This figure indicates that the GCS and RQ residuals in the QEVBS regression model show good agreements with the expected standard exponential and standard normal distributions, respectively.

We compare the QEVBS regression model ($q = 0.50$) given in 1.15 with other models commonly used in the extreme value literature. In particular, we consider the generalized

Table 1.7: ML estimates (standard errors in parentheses) for the QEVBS regression model across different values of q.

	$q = 0.01$	$q = 0.10$	$q = 0.25$	$q = 0.50$	$q = 0.75$	$q = 0.95$	$q = 0.99$
$\beta_0\ (intercept)$	0.5041*	0.8181*	0.9757*	1.1169*	1.2178*	1.2961*	1.3185*
	(0.1511)	(0.1425)	(0.1392)	(0.1369)	(0.1358)	(0.1349)	(0.1353)
$\beta_1\ (temp)$	0.0746*	0.0746*	0.0746*	0.0746*	0.0746*	0.0747*	0.0746*
	(0.0068)	(0.0068)	(0.0068)	(0.0068)	(0.0068)	(0.0068)	(0.0068)
α	0.4095*	0.4095*	0.4095*	0.4095*	0.4095*	0.4094*	0.4095
	(0.0340)	(0.0340)	(0.0340)	(0.0340)	(0.0340)	(0.0340)	(0.0340)
ξ	-0.73	-0.73	-0.73	-0.73	-0.73	-0.73	-0.73

* significant at 5% level. ** significant at 10% level.



(a) RQ (b) GCS

Figure 1.9: QQ plot and its envelope for the GCS and RQ residuals for the ozone data ($q = 0.50$, other values of q show similar results).

extreme value (GEV), Gumbel and Weibull regression models. We report the Akaike (AIC) and Bayesian (BIC) information criteria in Table 1.8. Note that AIC $= -2\ell(\widehat{\boldsymbol{\theta}}) + 2m$ and BIC $= -2\ell(\widehat{\boldsymbol{\theta}}) + m\log(n)$, where $\ell(\widehat{\boldsymbol{\theta}})$ is the log-likelihood function of the model with vector of parameters $\boldsymbol{\theta}$ evaluated at $\widehat{\boldsymbol{\theta}}$, n is the sample size, and m is the number of model parameters. From Table 1.8, we observe that the QEVBS regression model provides the best adjustment compared to the other models based on the values of AIC and BIC.

Table 1.8: AIC and BIC values for the indicated regression models.

	QEVBS	GEV	Weibull	Gumbel
$\ell(\widehat{\boldsymbol{\theta}})$	-289.571	-298.7502	-291.6617	-301.0489
AIC	587.142	605.5003	589.3235	610.0978
BIC	596.865	615.2236	596.6159	619.8210

1.6 Concluding remarks

We introduced and analyzed a new class of quantile regression models based on the extreme value Birnbaum-Saunders distribution. This model is formulated from a reparametrization of the distribution extreme value Birnbaum-Saunders as a function of Q, where Q is the $100q$-th quantile of a random variable T following an extreme value Birnbaum-Saunders distribution. Two Monte Carlo simulations were carried out to numerically evaluate the statistical performance of the maximum likelihood estimators and the empirical distribution of the generalized Cox-Snell and random quantile residuals. Results from the first simulation study showed good performance of the maximum likelihood estimators obtaining empirical values of bias near to zero, as shown in Tables 1.1-1.3, and the results from the second simulation study showed that the generalized Cox-Snell and random quantile residuals behave according to their reference distributions, as shown in Tables 1.4-1.5. In the application to real data, QQ plots with envelope for the generalized Cox-Snell and random quantile residuals provided in Figure 1.9 showed a good fit of the proposed model. We also compared the proposed extreme value Birnbaum-Saunders quantile regression model with other models commonly used in the extreme value literature, the results are seen to be quite favorable to the proposed model.

Appendix A

Data set

Daily maximum ozone concentration data (T, in ppb): 15, 15, 22, 39, 28, 20, 18, 22, 5, 27, 34, 10, 19, 22, 22, 31, 32, 33, 17, 11, 23, 31, 46, 17, 29, 22, 19, 21, 21, 27, 35, 50, 39, 33, 32, 12, 26, 10, 23, 25, 25, 35, 15, 23, 43, 22, 19, 17, 15, 19, 6, 19, 34, 19, 29, 28, 31, 13, 28, 52, 17, 32, 18, 42, 16, 35, 40, 24, 25, 30, 42, 44, 21, 24, 24, 26, 32, 44, 48, 38, 36, 43, 35, 46.

Daily maximum temperature data (X, in °C): 15.7, 16.9, 18.8, 20.9, 22.1, 13.0, 18.8, 21.2, 13.2, 19.1, 20.9, 10.6, 13.9, 14.7, 18.0, 15.6, 18.6, 24.6, 23.9, 27.9, 25.9, 25.7, 21.2, 11.1, 17.2, 15.7, 17.9, 23.1, 23.6, 23.6, 23.5, 27.3, 27.9, 23.9, 21.1, 16.0, 18.5, 10.6, 13.5, 16.0, 19.7, 20.7, 12.0, 16.9, 19.8, 12.7, 17.5, 20.2, 15.2, 14.6, 13.0, 20.0, 17.8, 16.8, 16.9, 15.4, 17.4, 15.2, 23.7, 26.3, 14.8, 18.8, 18.3, 21.6, 13.0, 21.2, 20.3, 15.7, 14.6, 19.2, 24.9, 24.3, 12.8, 14.7, 13.3, 21.5, 22.9, 24.9, 22.7, 19.9, 18.3, 21.8, 21.3, 20.6.

References

Balakrishnan, N. and Kundu, Debasis (2019). "Birnbaum-Saunders distribution: A review of models, analysis, and applications". *Applied Stochastic Models in Business and Industry* 35, 4–132 (with discussions).

Beirlant, J., Caeiro, F., and Gomes, M. (2012). "An overview and open research topics in statistics of univariate extremes". *Revstat Stat*, 10:1–31.

Cade, B. S., Terrell, J. W., and Schroeder, R. L. (1999). "Estimating effects of limiting factors with regression quantiles". *Ecology*, 80:311–323.

Coles, S. (2001). *An introduction to statistical modeling of extreme values*. London: Springer.

Embrechts, P., Kluppelberg, C., and Mikosch, T. (1997). *Modelling Extremal Events for Insurance and Finance*. Vol. 2. Springer-Verlag.

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

Haan, L. de and Ferreira, A. (2006). *Extreme value theory: an introduction.* New York: Springer.

Jenkinson, A.F. (1955). "The frequency distribution of the annual maximum (or minimum) values of meteorological elements". *Quarterly Journal of Royal Meteorological Society* 81, pp. 58–171.

— (1969). "Estimation of maximum floods". *World Meteorological Organizationy Technical Note 98 (5)*, pp. 183–227.

Koenker, R. (2005). *Quantile Regression*. Cambridge: Cambridge University Press.

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

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

Leiva, V et al. (2015). “A criterion for environmental assessment using Birnbaum-Saunders attribute control charts”. *Environmetrics* 2, pp. 463–476.

Leiva, V. et al. (2016). “Extreme value Birnbaum-Saunders regression models applied to environmental data”. *Stochastic Environmental Research and Risk Assessment* 30, pp. 1045–1058.

Oliveira, Kessys L. P. et al. (2022). “On a length-biased Birnbaum-Saunders regression model applied to meteorological data”. *Communications in Statistics - Theory and Methods* 0.0, pp. 1–20.

Reiss, R. D. and Thomas, M. (2001). *Satistical Analysis of Ext reme Values*. 2nd ed. Boston Berling: Birkhauser Verlag.

Saulo, H et al. (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, pp. 147–149.

Vilca, F et al. (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, pp. 771–782.

Wei, Y. et al. (2006). “Quantile regression methods for reference growth charts.” *Statistics in Medicine*, 25:1369–1382.