CHILEAN JOURNAL OF STATISTICS

Edited by Víctor Leiva and Carolina Marchant

A free open-access journal indexed by

INDEXED BY EMERGING SOURCES CITATION (ISI WEB OF SCIENCE) CLARIVATE ANALYTICS

Web of Science Group

Scopus

ELSEVIER

Volume 12 Number 2
December 2021
ISSN: 0718-7912 (print)
ISSN: 0718-7920 (online)

Published by the
Chilean Statistical Society
SOCHE
SOCIEDAD CHILENA DE ESTADÍSTICA

Aims

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

The ChJS covers a broad range of topics in statistics, as well as in artificial intelligence, big data, data science, and machine learning, focused mainly on research articles. However, review, survey, and teaching papers, as well as material for statistical discussion, could be also published exceptionally. Each paper published in the ChJS must consider, in addition to its theoretical and/or methodological novelty, simulations for validating its novel theoretical and/or methodological proposal, as well as an illustration/application with real data.

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

Chilean Journal of Statistics

Volume 12, Number 2

December 2021

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

http://www.soche.cl/chjs

Contents

Chilean Journal of Statistics
Vol. 12, No. 2, December 2021, 189–204

Statistical Modeling
Research Paper

Lomax regression model with varying precision: Formulation, estimation, diagnostics, and application

Moizés da S. Melo[1,*], Laís H. Loose[2], and Jhonnata B. de Carvalho[3]

[1]Department of Statistics, Universidade Federal do Rio Grande do Norte, Natal, Brazil
[2]Department of Statistics, Universidade Federal de Santa Maria, Santa Maria, Brazil
[3]Department of Statistics, Universidade Federal do Amazonas, Manaus, Brazil

(Received: 13 July 2021 · Accepted in final form: 23 November 2021)

Abstract

In this paper, we propose a new regression model with varying precision based on the Lomax distribution with regression structures for both the mean and precision parameters. The structures contain unknown parameters, regressors, and a link function. We discuss methods for parameter estimation, hypothesis testing and diagnostic analysis, along with their asymptotic properties. We also provide the expressions for the score vector as well as for the observed and Fisher information matrices. We conduct a Monte Carlo simulation study to investigate the behavior of the estimators and evaluate their finite sample performance. Finally, we present and discuss an empirical application to illustrate the usefulness of the proposed model.

Keywords: Asymmetrical data · Maximum likelihood method · Monte Carlo simulation · Positive data · Reparametrization.

Mathematics Subject Classification: Primary 62J99 · Secondary 62F10.

1. Introduction

The Lomax distribution, also known as the Pareto type II model, belongs to the class of distributions with decreasing failure rate and was first introduced by Lomax (1954) for modeling business failure data. In the literature, the Lomax distribution has been applied in several fields. For example, Harris (1968) used this distribution for queue problems, Atkinson and Harrison (1978) used it for modeling business failure data, Holland et al. (2006) used it for modeling the distribution of the sizes of computer files on a server, Corbellini et al. (2007) used the Lomax distribution to model firm size distribution, and Chandra and Khan (2013) used it to determine the optimal time for level changes for stress plans in censored samples.

In the context of regression analysis, Beirlant and Goegebeur (2003) presented a regression model for random variables following a Lomax distribution, in which an exponential transformation is used to relate the response variable with covariates. Stasinopoulos and Rigby (2007) developed the package `gamlss` available in the software `R` (R Development Core Team, 2021), in which we can model the parameters using a regression structure and

*Corresponding author. Email: moizes.melo@ufrn.br

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

http://www.soche.cl/chjs

link functions. This package is an innovative proposal that makes it possible to consider regression structures in a wide range of probability distributions. In the approach presented by Stasinopoulos and Rigby (2007), a regression structure using a link function can be considered for modeling each of the distribution parameters. However, the modeling is not performed in terms of the mean of the distribution. This fact can make the interpretation of the parameters difficult, thus limiting the use of the model in practice. A possible approach to interpretation in terms of the mean is to use the invariance property of the estimators. This result can be applied when using some link functions such logarithmic (Das et al., 2010) or the square root. However, this is not possible when using the inverse link function.

When working with regression models for continuous positive variables, one possibility for modeling is to use transformations of the response variable. The most commonly transformation is the logarithmic. For example, Fernández and De Andrade (2020) proposed a log-erf-Frechet regression model and Vigas et al. (2017) proposed a regression model in the location-scale form based on the Poisson-Weibull distribution. In both approaches, the logarithm of the variable of interest is modeled. Nonetheless modeling the mean is the most common approach in regression models (McCullagh and Nelder, 1989; Ferrari and Cribari-Neto, 2004; Fonseca et al., 2016; Palm et al., 2019). Regression models are usually proposed with a focus on constant dispersion or precision parameter. Some extensions for modeling parameters related to the variance of the distribution have been considered in the literature. Among them, we highlight the proposal for modeling the dispersion or precision, such as in generalized linear models (Smyth, 1989) and in the beta regression model (Simas et al., 2010). In more recent proposals, models that address the modeling of the two characteristics, mean and variance, have been introduced in seminal proposals. For example, Santos-Neto et al. (2016) proposed a reparameterized Birnbaum-Saunders regression model with varying precision and Bourguignon and Gallardo (2020) developed the reparameterized inverse gamma regression model with varying precision. In the context of the modal regression, Bourguignon et al. (2020) presented a parametric modal regression with varying precision where the response variable is gamma distributed. Additionally, Altun (2021) introduced a new Lomax regression model, in which the response's mean and shape parameter (α) are modeled by regression structures through the link functions. Recently, Bourguignon and Nascimento (2020) presented a Bayesian approach that considers a new parametrization that is indexed by mean and precision parameters, in which the response variable is a generalized Pareto distribution. The main advantage of this reparametrization is that it allows the mean and precision parameters to be modeled directly, allowing the construction of simple and interpretable models, such as in the context of generalized linear models (McCullagh and Nelder, 1989).

Based on the above discussion, this work has the objective of using the maximum likelihood (ML) approach to make inferences in the regression model with the same reparametrization used in Bourguignon and Nascimento (2020). The methods presented in this article differs from the described in Bourguignon and Nascimento (2020) in one main aspect; the approach to estimate parameters of the models. While Bourguignon and Nascimento (2020) used the Bayesian approach, we presented the ML inference approach. Additionally, in our proposal, the parametric support of the precision parameter is different from the one used by Bourguignon and Nascimento (2020), in this sense, other link functions are suggested and used for precision modeling. In this paper, the estimation of the parameters is performed using the ML method. We obtain analytical expressions for the score vector and Fisher information matrix, and also propose diagnostic measures and tools for model selection. We emphasize that obtaining the Fisher information matrix is possible due to the simplicity of the Lomax probability density function. Such expressions are impossible and/or very costly to obtain in some more complex distributions.

This paper is organized as follows. In Section 2, we present the Lomax distribution, the proposed reparameterization, reparametrized Lomax regression model, and log-likelihood function of the model. In Section 3, we present methods for the estimation and inferences, such as the score vector, the observed and Fisher information matrix, procedures for obtaining confidence intervals and hypothesis tests, additionally we introduce some diagnostic measures to check the goodness-of-fit of the proposed model. Monte Carlo simulation results are presented and discussed in Section 4. We also present and discuss an application. Finally, the conclusions and final remarks are presented in Section 5.

2. Proposed Model

In this section, we introduce the two-parameter Lomax distribution and its main characteristics such as mean, variance, cumulative distribution function, and quantile function. Furthermore, we present the reparametrization in terms of mean and precision parameters, the regression structures for modeling the mean and precision, as well as the log-likelihood function.

2.1 The Lomax distribution

Let Y be a random variable with Lomax distribution. Its probability density function is given by

$$f(y;\alpha,\lambda) = \frac{\alpha\lambda^{\alpha}}{(y+\lambda)^{(\alpha+1)}}, \quad y > 0, \tag{1}$$

where $\alpha > 0$ is the shape parameter and $\lambda > 0$ is the scale parameter. The mean and variance of Y are stated, respectively, by $\mathrm{E}(Y) = \lambda/(\alpha-1)$, for $\alpha > 1$, and $\mathrm{Var}(Y) = \alpha\lambda^2/((\alpha-1)^2(\alpha-2))$, for $\alpha > 2$. The cumulative distribution function corresponding to Equation (1) is expressed by

$$F(y;\lambda,\alpha) = 1 - \left(1 + \frac{y}{\lambda}\right)^{-\alpha}.$$

2.2 A reparametrized Lomax distribution

In regression analysis, it is typically more useful and common to model the mean response, as it makes the model parameters easily interpretable. In order to obtain a regression structure for the mean of Y, we consider a new parameterization, which is obtained by taking $\mu = \lambda/(\alpha-1)$ and $\phi = (\alpha-2)/\alpha$ in Equation (1), that is, $\lambda = \mu(\alpha-1)$ and $\alpha = 2/(1-\phi)$. The mean-parametrized Lomax distribution with mean μ and precision ϕ is characterized by the probability density function expressed as

$$f(y;\mu,\phi) = \frac{\frac{2}{1-\phi}\left[\mu\left(\frac{2}{1-\phi}-1\right)\right]^{\frac{2}{1-\phi}}}{\left[y+\mu\left(\frac{2}{1-\phi}-1\right)\right]^{\frac{2}{1-\phi}+1}}, \quad y > 0, \quad \mu > 0, \quad 0 < \phi < 1. \tag{2}$$

The mean and variance are given, respectively, by

$$\mathrm{E}(Y) = \mu \quad \text{and} \quad \mathrm{Var}(Y) = \frac{\mu^2}{\phi}.$$

The new cumulative distribution function is stated as

$$F(y;\mu,\phi) = 1 - \left[1 + \frac{y(1-\phi)}{\mu\,(1+\phi)}\right]^{-\frac{2}{1-\phi}}. \tag{3}$$

2.3 The reparametrized Lomax regression model

Let $Y_1,\ldots,Y_n$ be independent random variables, where each $Y_t, t = 1,\ldots,n$, follows the probability density function stated in Equation (2) with mean μ_t and precision ϕ_t. The regression structures for the mean and precision of Y_t are formulated, respectively, by

$$\eta_{1t} = g_1(\mu_t) = \boldsymbol{x}_t^\top \boldsymbol{\beta} \quad \text{and} \quad \eta_{2t} = g_2(\phi_t) = \boldsymbol{z}_t^\top \boldsymbol{\gamma}, \tag{4}$$

where $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_r)^\top \in \mathrm{R}^{r+1}$ and $\boldsymbol{\gamma} = (\gamma_0, \gamma_1, \ldots, \gamma_q)^\top \in \mathrm{R}^{q+1}$ are vectors of unknown regression parameters assumed to be functionally independent $(r + q + 2 < n)$, $\boldsymbol{x}_t = (1, x_{t1}, \ldots, x_{tr})^\top$ and $\boldsymbol{z}_t = (1, z_{t1}, \ldots, z_{tq})^\top$ are explanatory variables vectors, η_{1t} and η_{2t} are the mean and precision linear predictors, respectively, and g_1 and g_2 are twice-differentiable one-to-one monotonic functions called link functions, where $g_1\colon \mathrm{R}^+ \to \mathrm{R}$ and $g_2\colon (0,1) \to \mathrm{R}$.

The proposed Lomax regression model is defined by Equations (2) and (4). Due to the restriction $\mu_t > 0$, the most common link function and that satisfies the conditions stated for g_1 is the logarithm, $g_1(\mu_t) = \log(\mu_t)$, because it provides non-negative values for $\mu_t = g_1^{-1}(\eta_t) = \exp(\eta_t)$ regardless the values assigned to η_t, is twice-differentiable one-to-one monotonic funcion. Other link functions are usual, but they do not satisfy all the conditions stated for g_1, they are the square root, $g_1(\mu_t) = \sqrt{\mu}_t$, and inverse, $g_1(\mu_t) = 1/\mu_t$ (with special attention to the positivity of the estimates). For the restriction $0 < \phi_t < 1$, we can use the logit $g_2(\phi_t) = \log[\phi_t/(1-\phi_t)]$, probit $g_2(\phi_t) = \Phi^{-1}(\phi_t)$, where Φ is the cumulative distribution function of a standard normal random variable, complementary log-log $g_2(\phi_t) = \log[-\log(1-\phi_t)]$, and log-log $g_2(\phi_t) = -\log[-\log(\phi_t)]$ link functions (Ferrari and Cribari-Neto, 2004; Simas et al., 2010). For more details and a detailed discussion about link functions, see Atkinson (1985, Ch. 7) and McCullagh and Nelder (1989).

2.4 Likelihood function

Let $Y_1,\ldots,Y_n$ be a sample from the proposed Lomax regression model, $y_1,\ldots,y_n$ its observations, and $\boldsymbol{\theta} = (\boldsymbol{\beta}^\top, \boldsymbol{\gamma}^\top)^\top$ the corresponding regression parameter vector. The corresponding log-likelihood function for $\boldsymbol{\theta}$ is given by

$$\ell(\boldsymbol{\theta}) = \sum_{t=1}^{n} \ell_t(\mu_t, \phi_t), \tag{5}$$

where

$$\begin{aligned} \ell_t(\mu_t,\phi_t) = \log(2) - \log(1-\phi_t) + \frac{2}{1-\phi_t}\log\left[\mu_t\left(\frac{2}{1-\phi_t} - 1\right)\right] \\ - \left(\frac{2}{1-\phi_t} + 1\right)\log\left[y_t + \mu_t\left(\frac{1+\phi_t}{1-\phi_t}\right)\right]. \end{aligned} \tag{6}$$

3. Estimation and inference

In this section, we present details for performing point and interval estimation, and hypothesis testing. Initially, we present the score vector, the observed information matrix and Fisher information matrix, next present a test statistic to test hypotheses of interest and the formula for obtaining confidence intervals.

3.1 Score vector

Taking first derivatives of the log-likelihood function with respect to each element of $\boldsymbol{\theta}$, we obtain the score vector $U(\boldsymbol{\theta}) = (\boldsymbol{U}_{\boldsymbol{\beta}}(\boldsymbol{\theta})^\top, \boldsymbol{U}_{\boldsymbol{\gamma}}(\boldsymbol{\theta})^\top)^\top$ given by

$$U_{\beta_i}(\boldsymbol{\theta}) = \frac{\partial \ell(\boldsymbol{\theta})}{\partial \beta_i} = \sum_{t=1}^{n} \frac{\partial \ell_t(\mu_t, \phi_t)}{\partial \mu_t} \frac{\mathrm{d}\mu_t}{\mathrm{d}\eta_{1t}} \frac{\partial \eta_{1t}}{\partial \beta_i},$$

$$U_{\gamma_i}(\boldsymbol{\theta}) = \frac{\partial \ell(\boldsymbol{\theta})}{\partial \gamma_i} = \sum_{t=1}^{n} \frac{\partial \ell_t(\mu_t, \phi_t)}{\partial \phi_t} \frac{\mathrm{d}\phi_t}{\mathrm{d}\eta_{2t}} \frac{\partial \eta_{2t}}{\partial \gamma_i}.$$

From Equation (6), the derivative of $\ell_t(\mu_t, \phi_t)$ with respect to μ_t is defined by

$$\frac{\partial \ell_t(\mu_t, \phi_t)}{\partial \mu_t} = \frac{2}{\mu_t(1-\phi_t)} - \frac{(3-\phi_t)(1+\phi_t)}{(1-\phi_t)^2\,[y_t + c_t]} := b_t, \tag{7}$$

where $c_t = \mu_t(1+\phi_t)/(1-\phi_t)$. Note that $\eta_{1t} = g_1(\mu_t)$, then $\mathrm{d}\mu_t/\mathrm{d}\eta_{1t} = 1/g_1'(\mu_t)$, where g' is the first derivative of function g. We also have that $\partial\eta_{1t}/\partial\beta_i = x_{ti}$. Therefore, it follows that

$$U_{\beta_i}(\boldsymbol{\theta}) = \sum_{t=1}^{n} b_t \frac{1}{g_1'(\mu_t)} x_{ti}, \quad i = 0, 1, \ldots, r,$$

where $x_{t0} = 1$. By taking derivative in Equation (6) with respect to ϕ_t, we define $\partial \ell_t(\mu_t, \phi_t)/\partial \phi_t := a_t$, it follows that

$$a_t = \left[\frac{1}{1-\phi_t} + \frac{2\log(c_t)}{(1-\phi_t)^2} - \frac{2\log(y_t + c_t)}{(1-\phi_t)^2} + \frac{4}{(1+\phi_t)(1-\phi_t)^2} - \frac{2\mu_t(3-\phi_t)}{(1-\phi_t)^3(y_t + c_t)}\right]. \tag{8}$$

For γ_i, we have that $\mathrm{d}\phi_t/\mathrm{d}\eta_{2t} = 1/g_2'(\phi_t)$ and $\partial\eta_{2t}/\partial\gamma_i = z_{ti}$. Therefore, we obtain

$$\boldsymbol{U}_{\gamma_i}(\boldsymbol{\theta}) = \sum_{t=1}^{n} a_t \frac{1}{g_2'(\phi_t)} z_{ti}, \quad i = 0, 1, \ldots, q,$$

where $z_{t0} = 1$. The score vector can be expressed in matrix form as $\boldsymbol{U}_{\boldsymbol{\beta}}(\boldsymbol{\theta}) = \boldsymbol{X}^\top \boldsymbol{M} \boldsymbol{b}$ and $\boldsymbol{U}_{\boldsymbol{\gamma}}(\boldsymbol{\theta}) = \boldsymbol{Z}^\top \boldsymbol{\mathcal{M}} \boldsymbol{a}$, where $\boldsymbol{X}$ is an $n \times r$ matrix with the t-th row given by x_t, $\boldsymbol{Z}$ is an $n \times q$ matrix with t-th row given by z_{ij}, $\boldsymbol{b} = (b_1, \ldots, b_n)^\top$, $\boldsymbol{a} = (a_1, \ldots, a_n)^\top$, $\boldsymbol{M} = \mathrm{diag}\,\{1/g_1'(\mu_1), \ldots, 1/g_1'(\mu_n)\}$, and $\boldsymbol{\mathcal{M}} = \mathrm{diag}\,\{1/g_2'(\phi_1), \ldots, 1/g_2'(\phi_n)\}$.

The ML estimators $\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{\gamma}}$ of the parameters $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are obtained by solving the nonlinear system of equations expressed as $\boldsymbol{U}_{\boldsymbol{\beta}}(\boldsymbol{\theta}) = \boldsymbol{0}$ and $\boldsymbol{U}_{\boldsymbol{\gamma}}(\boldsymbol{\theta}) = \boldsymbol{0}$. Since the above system does not have analytical solution, the use of nonlinear optimization algorithms is required. In this work, we apply the Nelder-Mead simplex method (Nelder and Mead, 1965).

3.2 Observed information matrix

Taking second order derivatives of Equation (5) with respect to each element of $\boldsymbol{\theta}$, we have

$$\begin{aligned}
\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \beta_i \partial \beta_j} &= \sum_{t=1}^{n} \frac{\partial}{\partial \beta_i}\left[\frac{\partial \ell_t(\mu_t,\phi_t)}{\partial \mu_t}\frac{\mathrm{d}\mu_t}{\mathrm{d}\eta_{1t}}\frac{\partial \eta_{1t}}{\partial \beta_j}\right] \\
&= \sum_{t=1}^{n}\left[\frac{\partial^2 \ell_t(\mu_t,\phi_t)}{\partial \mu_t^2}\frac{\mathrm{d}\mu_t}{\mathrm{d}\eta_{1t}} + \frac{\partial \ell_t(\mu_t,\phi_t)}{\partial \mu_t}\frac{\partial}{\partial \mu_t}\left(\frac{\mathrm{d}\mu_t}{\mathrm{d}\eta_{1t}}\right)\right]\frac{\mathrm{d}\mu_t}{\mathrm{d}\eta_{1t}} x_{tj} x_{ti}, \\
\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \beta_i \partial \gamma_j} &= \sum_{t=1}^{n} \frac{\partial}{\partial \gamma_j}\left[\frac{\partial \ell_t(\mu_t,\phi_t)}{\partial \mu_t}\frac{\mathrm{d}\mu_t}{\mathrm{d}\eta_{1t}}\frac{\partial \eta_{1t}}{\partial \beta_i}\right] = \sum_{t=1}^{n}\left[\frac{\partial^2 \ell_t(\mu_t,\phi_t)}{\partial \mu_t \partial \phi_t}\frac{\mathrm{d}\phi_t}{\mathrm{d}\eta_{2t}} z_{tj}\right]\frac{\mathrm{d}\mu_t}{\mathrm{d}\eta_t} x_{ti}, \\
\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial \gamma_i \partial \gamma_j} &= \sum_{t=1}^{n} \frac{\partial}{\partial \gamma_i}\left[\frac{\partial \ell_t(\mu_t,\phi_t)}{\partial \phi_t}\frac{\mathrm{d}\phi_t}{\mathrm{d}\eta_{2t}}\frac{\partial \eta_{2t}}{\partial \gamma_j}\right] \\
&= \sum_{t=1}^{n}\left[\frac{\partial^2 \ell_t(\mu_t,\phi_t)}{\partial \phi_t^2}\frac{\mathrm{d}\phi_t}{\mathrm{d}\eta_{2t}} + \frac{\partial \ell_t(\mu_t,\phi_t)}{\partial \phi_t}\frac{\partial}{\partial \phi_t}\left(\frac{\mathrm{d}\phi_t}{\mathrm{d}\eta_{2t}}\right)\right]\frac{\mathrm{d}\phi_t}{\mathrm{d}\eta_{2t}} z_{tj} z_{ti}.
\end{aligned}$$

In addition, from Equation (7), we have

$$\begin{aligned}
\frac{\partial^2 \ell_t(\mu_t,\phi_t)}{\partial \mu_t^2} &= \frac{\partial}{\partial \mu_t}\left[\frac{2}{\mu_t(1-\phi_t)} - \frac{(3-\phi_t)(1+\phi_t)}{(1-\phi_t)^2\,[y_t+c_t]}\right] \\
&= -\frac{2}{\mu_t^2(1-\phi_t)} + \frac{(3-\phi_t)(1+\phi_t)^2}{(1-\phi_t)^3\,(y_t+c_t)^2} := w_t, \\
\frac{\partial^2 \ell_t(\mu_t,\phi_t)}{\partial \mu_t \partial \phi_t} &= \frac{\partial}{\partial \phi_t}\left[\frac{2}{\mu_t(1-\phi_t)} - \frac{(3-\phi_t)(1+\phi_t)}{(1-\phi_t)^2\,[y_t+c_t]}\right] \\
&= \frac{2}{\mu_t(1-\phi_t)^2} - \frac{2\mu_t(1+\phi_t)^2 + 8y_t(1-\phi_t)}{(1-\phi_t)^4\,(y_t+c_t)^2} := r_t.
\end{aligned}$$

And from Equation (8), we have

$$\begin{aligned}
\frac{\partial^2 \ell_t(\mu_t,\phi_t)}{\partial \phi_t^2} =& \frac{\partial}{\partial \phi_t}\left[\frac{1}{1-\phi_t} + \frac{2\log(c_t) - 2\log(y_t+c_t)}{(1-\phi_t)^2} + \frac{4}{(1+\phi_t)(1-\phi_t)^2}\right. \\
&\left. - \frac{2\mu_t(3-\phi_t)}{(1-\phi_t)^3(y_t+c_t)}\right] \\
=& \frac{5-\phi_t^2}{(1+\phi_t)(1-\phi_t)^3} + \frac{4(1+3\phi_t)}{(1+\phi_t)^2(1-\phi_t)^3} + \frac{4\log(c_t)}{(1-\phi_t)^3} - \frac{4\mu}{(y+c_t)(1-\phi_t)^4} \\
& - \frac{4\log(y+c_t)}{(1-\phi_t)^3} + \frac{4\mu^2(3-\phi_t)}{(y+c_t)^2(1-\phi_t)^5} + \frac{4\mu(\phi_t-4)}{(y+c_t)(1-\phi_t)^4} := s_t.
\end{aligned}$$

Notice also that

$$\frac{\partial}{\partial \mu_t}\left(\frac{\mathrm{d}\mu_t}{\mathrm{d}\eta_{1t}}\right) = -\frac{g''(\mu_{1t})}{[g'(\mu_{1t})]^2} := m_t \quad \text{and} \quad \frac{\partial}{\partial \phi_t}\left(\frac{\mathrm{d}\phi_t}{\mathrm{d}\eta_{2t}}\right) = -\frac{g''(\phi_{2t})}{[g'(\phi_{2t})]^2} := o_t,$$

where g'' is the second derivative of function g.

Let $\boldsymbol{H} = \mathrm{diag}\{h_1, \ldots, h_n\}$ with $h_t = [w_t/g'(\mu_{1t}) + b_t m_t]/g'(\mu_{1t})$, $\boldsymbol{R} = (r_1, \ldots, r_n)^\top$, and $\boldsymbol{P} = (p_1, \ldots, p_n)^\top$ with $p_t = [s_t/g'(\phi_{2t}) + a_t o_t]/g'(\phi_{2t})$. The joint observed information matrix for $\boldsymbol{\theta}$ is given by

$$\boldsymbol{J}(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{J}_{(\boldsymbol{\beta},\boldsymbol{\beta})} & \boldsymbol{J}_{(\boldsymbol{\beta},\boldsymbol{\gamma})} \\ \boldsymbol{J}_{(\boldsymbol{\gamma},\boldsymbol{\beta})} & \boldsymbol{J}_{(\boldsymbol{\gamma},\boldsymbol{\gamma})} \end{pmatrix},$$

where $\boldsymbol{J}_{(\boldsymbol{\beta},\boldsymbol{\beta})} = -\boldsymbol{X}^\top \boldsymbol{H}\boldsymbol{X}$, $\boldsymbol{J}_{(\boldsymbol{\beta},\boldsymbol{\gamma})} = \boldsymbol{J}^\top_{(\boldsymbol{\gamma},\boldsymbol{\beta})} = -\boldsymbol{X}^\top \boldsymbol{M}\boldsymbol{R}\boldsymbol{\mathcal{M}}\boldsymbol{Z}$, and $\boldsymbol{J}_{(\boldsymbol{\gamma},\boldsymbol{\gamma})} = \boldsymbol{Z}^\top \boldsymbol{P}\boldsymbol{Z}$.

3.3 Information matrix, confidence intervals and hypothesis testing

Before presenting to the important quantities of this subsection, we need some useful results used in obtaining of the Fisher information matrix provided by Lemma below.

Lemma 1: *Let Y_t be a random variable that follows a Lomax distribution with probability density function given in Equation (2). Then,*

$$\mathrm{E}\Big(\frac{1}{Y_t + \mu_t\,(2/(1-\phi_t-1))}\Big) = \frac{2(1-\phi_t)}{\mu_t(1+\phi_t)(3-\phi_t)},$$

$$\mathrm{E}\Big(\frac{1}{(Y_t + \mu_t\,(2/(1-\phi_t-1)))^2}\Big) = \frac{(1-\phi_t)^2}{\mu_t^2(1+\phi_t)^2(2-\phi_t)},$$

$$\mathrm{E}\left(\log\left(Y_t + \mu_t\,(2/(1-\phi_t-1))\right)\right) = \log\left(\frac{\mu_t(1+\phi_t)}{1-\phi_t}\right) + \frac{1-\phi_t}{2}.$$

Proof: Let f be the probability density function of Y_t. Then,

$$\mathrm{E}\left(\frac{1}{Y_t + \mu_t\left(\frac{2}{1-\phi_t} - 1\right)}\right)^k = \int_0^\infty \frac{\frac{2}{1-\phi_t}\left(\mu_t\left(\frac{2}{1-\phi_t}-1\right)\right)^{\frac{2}{1-\phi_t}}}{\left(y_t + \mu_t\left(\frac{2}{1-\phi_t}-1\right)\right)^{\left(\frac{2}{1-\phi_t}+1+k\right)}} \mathrm{d}y_t,$$

where $k > 0$. Making the variable change $x_t = y_t + \mu_t(2/(1-\phi_t-1))$, then the above equation becomes

$$\mathrm{E}\left(\frac{1}{Y_t + \mu_t\,(2/(1-\phi_t-1))}\right)^k = \frac{2(1-\phi_t)^k}{\mu_t^k(1+\phi_t)^k(2+k-\phi_t k)},$$

$$\mathrm{E}\left(\log\left(Y_t + \mu_t\left(\frac{2}{1-\phi_t} - 1\right)\right)\right) = \log\left(\frac{\mu_t(1+\phi_t)}{1-\phi_t}\right) + \frac{1-\phi_t}{2}. \quad \blacksquare$$

The Fisher information matrix is obtained by taking the expected value of the second order derivatives of the log-likelihood function, that is, $\boldsymbol{K}(\boldsymbol{\theta}) = \mathrm{E}[\boldsymbol{J}(\boldsymbol{\theta})]$. Since

$$\begin{aligned}\mathrm{E}\left(\frac{\partial \ell(\mu_t,\phi_t)}{\partial \mu_t}\right) &= \frac{2}{\mu_t(1-\phi_t)} - \frac{(1+\phi_t)(3-\phi_t)}{(1-\phi_t)^2}\mathrm{E}\left(\frac{1}{Y_t + c_t}\right) \\ &= \frac{2}{\mu_t(1-\phi_t)} - \frac{2}{\mu_t(1-\phi_t)} = 0,\end{aligned}$$

the expected value of the derivatives in Section 3.2 are given by

$$\mathrm{E}\left(\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\beta_i\partial\beta_j}\right) = \sum_{t=1}^{n} \mathrm{E}\left(\frac{\partial^2 \ell_t(\mu_t,\phi_t)}{\partial\mu_t^2}\right)\left(\frac{\mathrm{d}\mu_t}{\mathrm{d}\eta_{1t}}\right)^2 x_{tj}x_{ti},$$

$$\mathrm{E}\left(\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\beta_i\partial\gamma_j}\right) = \sum_{t=1}^{n} \mathrm{E}\left(\frac{\partial^2 \ell_t(\mu_t,\phi_t)}{\partial\mu_t\partial\phi_t}\right)\frac{\mathrm{d}\phi_t}{\mathrm{d}\eta_{2t}}\frac{\mathrm{d}\mu_t}{\mathrm{d}\eta_{1t}} z_{tj}x_{ti},$$

$$\mathrm{E}\left(\frac{\partial^2 \ell(\boldsymbol{\theta})}{\partial\gamma_i\partial\gamma_j}\right) = \sum_{t=1}^{n} \mathrm{E}\left(\frac{\partial^2 \ell_t(\mu_t,\phi_t)}{\partial\phi_t^2}\right)\left(\frac{\mathrm{d}\phi_t}{\mathrm{d}\eta_{2t}}\right)^2 z_{tj}z_{ti}.$$

Observe that taking the expected value in Equations (7) and (8), and substituting the results from this lemma, we have

$$\begin{aligned}\mathrm{E}\left(\frac{\partial \ell(\mu_t,\phi_t)}{\partial\phi_t}\right) &= \frac{1}{1-\phi_t} + \frac{2\log(c)}{(1-\phi_t)^2} - \frac{2}{(1-\phi_t)^2}\left(\log c + \frac{1-\phi_t}{2}\right) + \frac{4}{(1+\phi_t)(1-\phi_t)^2}\\ &\quad - \frac{2\mu_t(3-\phi_t)}{(1-\phi_t)^3}\frac{2(1-\phi_t)}{\mu_t(1+\phi_t)(3-\phi_t)} = 0,\end{aligned}$$

$$\begin{aligned}\mathrm{E}\left(\frac{\partial^2 \ell(\mu_t,\phi_t)}{\partial\mu_t^2}\right) &= -\frac{2}{\mu_t^2(1-\phi_t)} + \frac{(1+\phi_t)^2(3-\phi_t)}{(1-\phi_t)^3}\mathrm{E}\left(\frac{1}{Y_t+c_t}\right)^2\\ &= -\frac{2}{\mu_t^2(1-\phi_t)} + \frac{(1+\phi_t)^2(3-\phi_t)}{(1-\phi_t)^3}\frac{(1-\phi_t)^2}{\mu_t^2(1+\phi_t)^2(2-\phi_t)} := v_t,\end{aligned}$$

$$\begin{aligned}\mathrm{E}\left(\frac{\partial^2 \ell(\mu_t,\phi_t)}{\partial\mu_t\partial\phi_t}\right) &= \frac{2}{\mu_t(1-\phi_t)^2} - \frac{2\mu_t(1+\phi_t)^2}{(1-\phi_t)^4}\mathrm{E}\left(\frac{1}{(Y_t+c_t)^2}\right) - \frac{8}{(1-\phi_t)^3}\mathrm{E}\left(\frac{Y_t}{(Y_t+c_t)^2}\right)\\ &= \frac{2}{\mu_t(1-\phi_t)^2} - \frac{2}{\mu_t(2-\phi_t)(1-\phi_t)^2} - \frac{8}{\mu_t(3-\phi_t)(2-\phi_t)(1-\phi_t^2)} := d_t,\end{aligned}$$

$$\begin{aligned}\mathrm{E}\left(\frac{\partial^2 \ell(\mu_t,\phi_t)}{\partial\phi_t^2}\right) &= \frac{5-\phi_t^2}{(1+\phi_t)(1-\phi_t)^3} + \frac{4(1+3\phi_t)}{(1+\phi_t)^2(1-\phi_t)^3} + \frac{4\log(c_t)}{(1-\phi_t)^3}\\ &\quad - \frac{4\mu_t}{(1-\phi_t)^4}\mathrm{E}\left(\frac{1}{Y_t+c_t}\right) - \frac{4}{(1-\phi_t)^3}\mathrm{E}(\log(Y_t+c_t)) + \frac{4\mu_t^2(3-\phi_t)}{(1-\phi_t)^5}\mathrm{E}\left(\frac{1}{Y_t+c_t}\right)^2\\ &\quad + \frac{4\mu_t(\phi_t-4)}{(1-\phi_t)^4}\mathrm{E}\left(\frac{1}{Y_t+c_t}\right)\\ &= \frac{5-\phi_t^2}{(1+\phi_t)(1-\phi_t)^3} + \frac{4(1+3\phi_t)}{(1+\phi_t)^2(1-\phi_t)^3} + \frac{4(3-\phi_t)}{(1-\phi_t)^3(1+\phi_t)^2(2-\phi_t)}\\ &\quad + \frac{8(\phi_t-5)}{(1-\phi_t)^3(1+\phi_t)(3-\phi_t)} - \frac{2}{(1-\phi_t)^2} := q_t.\end{aligned}$$

Let $\boldsymbol{V} = \text{diag}\{v_1, \ldots, v_n\}$, $\boldsymbol{D} = (d_1, \ldots, d_n)^\top$, and $\boldsymbol{Q} = \text{diag}\{q_1, \ldots, q_n\}$. The Fisher information matrix for $\boldsymbol{\theta}$ is given by

$$\boldsymbol{K}(\boldsymbol{\theta}) = \begin{pmatrix} \boldsymbol{K}_{(\boldsymbol{\beta},\boldsymbol{\beta})} & \boldsymbol{K}_{(\boldsymbol{\beta},\boldsymbol{\gamma})} \\ \boldsymbol{K}_{(\boldsymbol{\gamma},\boldsymbol{\beta})} & \boldsymbol{K}_{(\boldsymbol{\gamma},\boldsymbol{\gamma})} \end{pmatrix},$$

where $\boldsymbol{K}_{(\boldsymbol{\beta},\boldsymbol{\beta})} = -\boldsymbol{X}^\top \boldsymbol{V}\boldsymbol{M}^2\boldsymbol{X}$, $\boldsymbol{K}_{(\boldsymbol{\beta},\boldsymbol{\gamma})} = \boldsymbol{K}_{(\boldsymbol{\gamma},\boldsymbol{\beta})}^\top = -\boldsymbol{X}^\top \boldsymbol{M}\boldsymbol{D}\boldsymbol{\mathcal{M}}\boldsymbol{Z}$, and $\boldsymbol{K}_{(\boldsymbol{\gamma},\boldsymbol{\gamma})} = -\boldsymbol{X}^\top \boldsymbol{Q}\boldsymbol{\mathcal{M}}^2\boldsymbol{X}$. Note that the parameters $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$ are not orthogonal.

Under usual regularity conditions, the ML estimators $\widehat{\boldsymbol{\theta}}$ of $\boldsymbol{\theta}$ are asymptotically consistent, having approximately normal distribution with mean vector $\boldsymbol{\theta}$ and variance-covariance matrix $\boldsymbol{K}(\boldsymbol{\theta})^{-1}$ in large samples (Pawitan, 2001), that is,

$$\begin{pmatrix} \widehat{\boldsymbol{\beta}} \\ \widehat{\boldsymbol{\gamma}} \end{pmatrix} \sim \mathrm{N}_{r+q+2}\left(\begin{pmatrix} \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{pmatrix}, \boldsymbol{K}(\boldsymbol{\theta})^{-1} \right), \tag{9}$$

where N_{r+q+2} denotes the $(r + q + 2)$-dimensional normal distribution and $\widehat{\boldsymbol{\beta}}$ and $\widehat{\boldsymbol{\gamma}}$ the ML estimators of $\boldsymbol{\beta}$ and $\boldsymbol{\gamma}$, respectively.

Test statistics for hypothesis testing and confidence intervals can be obtained using the asymptotic result presented in Equation (9). Suppose the interest is to test the following hypotheses $\mathcal{H}_0$: $\theta_i = \theta_i^0$ versus $\mathcal{H}_1 : \theta_i \neq \theta_i^0$, where θ_i^0 is a specified value for the unknown parameter θ_i. A useful statistic to test these hypotheses is the signed square root of the Wald statistic, given by $Z = (\widehat{\theta}_i - \theta_i^0)/\sqrt{k^{ii}}$, where k^{ii} is the i-th diagonal element of $\boldsymbol{K}(\widehat{\boldsymbol{\theta}})^{-1}$. This statistic is particularly convenient to test individual parameters (Pawitan, 2001). Under $\mathcal{H}_0$ and for large n, Z has a standard normal distribution. It is also possible to perform more general hypothesis testing inference using the likelihood ratio, Wald, and score statistics.

We can also use the result presented in Equation (9) to construct asymptotic confidence intervals for each parameter θ_i. An approximate $100(1-\alpha)\%$ confidence interval for θ_i is defined as $(\widehat{\theta}_i - z_{1-\alpha/2}\sqrt{k^{ii}}; \widehat{\theta}_i + z_{1-\alpha/2}\sqrt{k^{ii}})$, where $\Phi(z_{1-\alpha/2}) = 1 - \alpha/2$.

3.4 Diagnostic measures

In this subsection we suggest criteria for selecting the Lomax regression model and some diagnostic measures for examining the goodness-of-fit of the proposed model. For model selection, we consider the Akaike Information Criterion (AIC) (Akaike, 1974) and Bayesian Information Criterion (BIC) (Schwarz, 1978) given, respectively, by $\text{AIC} = -2\ell(\widehat{\boldsymbol{\theta}}) + 2(q + r + 2)$ and $\text{BIC} = -2\ell(\widehat{\boldsymbol{\theta}}) + \log(n)(q + r + 2)$.

For validating the proposed model, we perform residual analysis using the randomized quantile residuals (Dunn and Smyth, 1996), defined as $r_t^{(q)} = \Phi^{-1}(F(y_t; \widehat{\mu}_t, \widehat{\phi}_t))$, where $F(y_t; \widehat{\mu}_t, \widehat{\phi}_t)$ is the cumulative distribution function stated in Equation (3). If the model is correctly specified, these residuals should be independent and normally distributed, with zero mean and unit variance.

4. Numerical results

In this section, we provide the simulation study in order to evaluate the performance of the ML estimators of the proposed model under different sample sizes. Also, we present and discuss an empirical application to illustrate the proposed framework.

4.1 Simulation study

We conduct a Monte Carlo simulation study to evaluate the finite sample performance of the likelihood inference for the proposed Lomax regression model. We used 10,000 Monte Carlo replications and considered five sample sizes $n \in \{50, 100, 200, 500, 1000\}$. Performance measures for ML estimator evaluation are the mean, bias, relative bias (RB), standard deviation (SD), and root mean square error (RMSE).

We considered two scenarios with the following true parameter values: (i) Scenario 1: $\boldsymbol{\theta} = (\beta_0, \beta_1, \beta_2, \gamma_0, \gamma_1) = (-3.0, 2.2, 1.5, 0.5, -0.3)$ and (ii) Scenario 2: $\boldsymbol{\theta} = (\beta_0, \beta_1, \beta_2, \gamma_0, \gamma_1) = (0.5, 1.3, -0.7, 0.3, -0.5)$. In both scenarios, the covariates were generated independently from a standard uniform distribution, $\mathcal{U}(0, 1)$, and kept constant during all Monte Carlo replications. We considered the logarithmic and probit link functions for the mean and precision submodels, respectively. This scenario considers the link functions that provided the best fit in the application.

All simulations were performed using the `R` software. The maximization was obtained considering the `optim` function, available in `R`, using the Nelder-Mead method with first analytical derivatives; see Nelder and Mead (1965) for more details. As the iterative optimization algorithm requires a set of initial values for the parameters to be optimized, we suggest to use the following empirical approach to determine these values. The initial values $\boldsymbol{\beta}^{(0)}$ for $\boldsymbol{\beta}$ are obtained by the least squares estimates of $\boldsymbol{\beta}$ from the following linear regression model: $\log(y_t) = \boldsymbol{x}_t^\top \boldsymbol{\beta}$, while the starting values $\boldsymbol{\gamma}^{(0)}$ for $\boldsymbol{\gamma}$ are obtained by $\boldsymbol{\gamma}^{(0)} = (\overline{y}^2/S_y^2)\mathbf{1}_{q+1}^\top$, where $\overline{y}$ and S_y^2 denote the sample mean and variance, respectively, and $\mathbf{1}_{q+1}$ denotes an $(q+1)$-dimensional vector of $1'$s. We have tested the others methods, however, in our study, the Nelder-Mead method provided more robust estimates.

The simulation results are shown in Tables 1 and 2. Based on the results presented, we can verify the good performance of the ML estimators of the Lomax regression model. We observe that the bias and RMSE of the ML estimators of $\boldsymbol{\beta}$ tend toward zero as the sample size increases, indicating the consistency property of the ML estimator. For the vector $\boldsymbol{\gamma}$, the ML estimators are biased in small samples, but the bias decreases as the sample size increases. This suggests that some procedure for inferential improvements can be considered to reduce the problem of biased ML estimator in small samples. We also highlight that this behavior of the ML estimators in precision modeling is recurrent in the literature (Bourguignon and Nascimento, 2020; Simas et al., 2010).

4.2 Empirical application

We illustrate the proposed model using dataset obtained from the United Nations Development Programme (available at http://hdr.undp.org/en/data). The response variable is the carbon dioxide emissions per capita (DEC, measured in tonnes) in 123 countries, including the autonomous territory of Hong-Kong and the United Kingdom collected in 2016. The covariates associated with this response variable are: forest area (FAR, measured in % of total land area), concentration index of exports (CIN, ranging from 0% to 100%, with a larger value denoting a higher concentration of exports), employment in agriculture (EAG, measured in % of total employment), and human development index (HDI). Some summary statistics of the response variable are given in Table 3. Figure 1 shows the dispersion plots between the response variable and covariates. After some adjustments, we consider only the set of regressors statistically significant at the level of 10% in the Lomax regression model. The HDI covariate was not significant for the mean submodel. Also, the FAR, CIN, and EAG covariates were not significant for the precision submodel. We use the observed information matrix obtained numerically using the `optim` function of R software because it provided lower variance estimates than the Fisher information matrix.

Table 1. Monte Carlo simulation results for likelihood inference, evaluation of point estimation, for the Lomax regression model - Scenario 1 - $g_1(\mu_t) = \log(\mu_t)$, $g_2(\phi_t) = \Phi^{-1}(\phi_t)$, $\beta_0 = -3.0$, $\beta_1 = 2.2$, $\beta_2 = 1.5$, $\gamma_0 = 0.5$ and $\gamma_1 = -0.3$.

n	Estimator	Mean	Bias	RB	SD	RMSE
50	$\widehat{\beta}_0$	−3.030	−0.030	0.010	0.556	0.557
	$\widehat{\beta}_1$	2.210	0.010	0.004	0.736	0.736
	$\widehat{\beta}_2$	1.480	−0.020	−0.013	0.714	0.714
	$\widehat{\gamma}_0$	2.161	1.661	3.322	9.454	9.599
	$\widehat{\gamma}_1$	−0.704	−0.404	1.347	55.585	55.586
100	$\widehat{\beta}_0$	−3.013	−0.013	0.004	0.357	0.358
	$\widehat{\beta}_1$	2.198	−0.002	−0.001	0.434	0.434
	$\widehat{\beta}_2$	1.508	0.008	0.005	0.502	0.502
	$\widehat{\gamma}_0$	1.235	0.735	1.471	6.504	6.546
	$\widehat{\gamma}_1$	−0.171	0.129	−0.431	13.708	13.709
200	$\widehat{\beta}_0$	−3.008	−0.008	0.003	0.240	0.240
	$\widehat{\beta}_1$	2.202	0.002	0.001	0.314	0.314
	$\widehat{\beta}_2$	1.499	−0.001	−0.001	0.292	0.292
	$\widehat{\gamma}_0$	0.805	0.305	0.610	2.413	2.432
	$\widehat{\gamma}_1$	−0.214	0.086	−0.286	4.165	4.166
500	$\widehat{\beta}_0$	−3.002	−0.002	0.001	0.135	0.135
	$\widehat{\beta}_1$	2.199	−0.001	0.000	0.179	0.179
	$\widehat{\beta}_2$	1.500	0.000	0.000	0.179	0.179
	$\widehat{\gamma}_0$	0.641	0.141	0.281	0.777	0.789
	$\widehat{\gamma}_1$	−0.368	−0.068	0.225	1.318	1.319
1000	$\widehat{\beta}_1$	−3.001	−0.001	0.000	0.099	0.099
	$\widehat{\beta}_2$	2.199	−0.001	0.000	0.130	0.130
	$\widehat{\beta}_3$	1.501	0.001	0.001	0.129	0.129
	$\widehat{\gamma}_0$	0.552	0.052	0.104	0.468	0.471
	$\widehat{\gamma}_1$	−0.312	−0.012	0.039	0.786	0.786

Table 2. Monte Carlo simulation results for likelihood inference, evaluation of point estimation, for the Lomax regression model - Scenario 2 - $g_1(\mu_t) = \log(\mu_t)$, $g_2(\phi_t) = \Phi^{-1}(\phi_t)$, $\beta_0 = 0.5$, $\beta_1 = 1.3$, $\beta_2 = -0.7$, $\gamma_0 = 0.3$ and $\gamma_1 = -0.5$.

n	Estimator	Mean	Bias	RB	SD	RMSE
50	$\widehat{\beta}_0$	0.443	−0.057	−0.114	0.578	0.581
	$\widehat{\beta}_1$	1.310	0.010	0.008	0.673	0.673
	$\widehat{\beta}_2$	−0.684	0.016	−0.023	0.613	0.614
	$\widehat{\gamma}_0$	1.163	0.863	2.878	16.240	16.263
	$\widehat{\gamma}_1$	−1.372	−0.872	1.744	34.205	34.216
100	$\widehat{\beta}_0$	0.493	−0.007	−0.014	0.332	0.332
	$\widehat{\beta}_1$	1.294	−0.006	−0.005	0.434	0.434
	$\widehat{\beta}_2$	−0.705	−0.005	0.008	0.442	0.442
	$\widehat{\gamma}_0$	0.669	0.369	1.232	6.398	6.408
	$\widehat{\gamma}_1$	−0.774	−0.274	0.547	14.265	14.268
200	$\widehat{\beta}_0$	0.498	−0.002	−0.004	0.200	0.200
	$\widehat{\beta}_1$	1.298	−0.002	−0.001	0.263	0.263
	$\widehat{\beta}_2$	−0.702	−0.002	0.003	0.271	0.271
	$\widehat{\gamma}_0$	0.523	0.223	0.744	2.101	2.113
	$\widehat{\gamma}_1$	−0.732	−0.232	0.465	3.488	3.496
500	$\widehat{\beta}_0$	0.498	−0.002	−0.004	0.152	0.152
	$\widehat{\beta}_1$	1.298	−0.002	−0.002	0.196	0.196
	$\widehat{\beta}_2$	−0.699	0.001	−0.001	0.194	0.194
	$\widehat{\gamma}_0$	0.410	0.110	0.365	0.694	0.702
	$\widehat{\gamma}_1$	−0.612	−0.112	0.223	1.234	1.239
1000	$\widehat{\beta}_0$	0.499	−0.001	−0.001	0.105	0.105
	$\widehat{\beta}_1$	1.297	−0.003	−0.002	0.133	0.133
	$\widehat{\beta}_2$	−0.699	0.001	−0.001	0.138	0.138
	$\widehat{\gamma}_0$	0.344	0.044	0.148	0.397	0.400
	$\widehat{\gamma}_1$	−0.524	−0.024	0.048	0.687	0.687

Table 3. Summary statistics of carbon dioxide emissions.

Min	1st Quantile	Median	Mean	3rd Quantile	Max	Variance
0.100	1.400	3.400	4.953	6.500	29.800	27.597

After testing different combinations of link functions, the link functions that provided the best fit were the logarithm and probit link functions for the mean and precision submodels, respectively, resulting in the regression structures stated as

$$\log(\mu_t) = \beta_1 \mathrm{FAR}_t + \beta_2 \mathrm{CIN}_t + \beta_3 \mathrm{EAG}_t \quad \text{and} \quad \Phi^{-1}(\phi_t) = \gamma_1 \mathrm{HDI}_t, \; t = 1, 2, \ldots, 123.$$

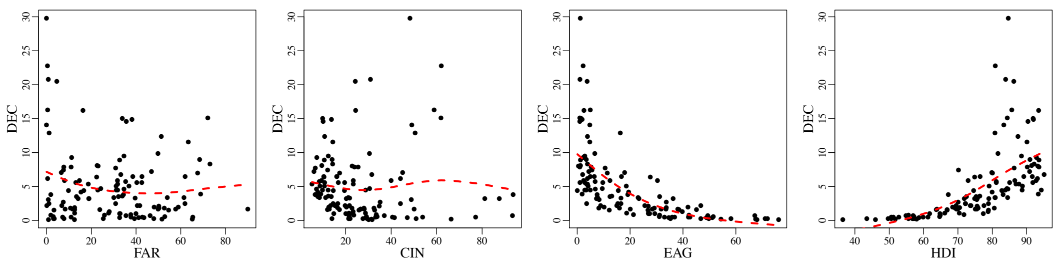


Figure 1. Plot for DEC versus FAR, CIN, EAG and HDI with corresponding smooth curves.

We compare the fitted Lomax regression model with the reparametrized gamma, reparameterized Weibull, and normal linear regression models using the `gamlss` package (Stasinopoulos and Rigby, 2007) in `R`. Some information about the regression structure of these models is summarized in Table 4.

Table 4. Regression structures for the gamma, Weibull, and normal models, with μ and σ^2 representing the mean and variance of the distribution, respectively.

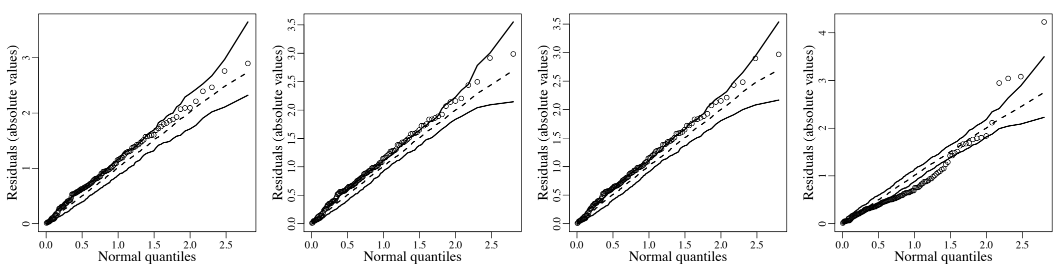
Distribution	Reparametrization	Link function
gamma(θ_1, θ_2)	$\mu = \theta_1\theta_2$ $\sigma = \theta_2\sqrt{\theta_1}$	$g_1(\mu) = \log(\mu)$ $g_2(\sigma) = \log(\sigma)$
Weibull(θ_1, θ_2)	$\mu = \theta_1\Gamma(1 + 1/\theta_2)$ $\sigma = \theta_2$	$g_1(\mu) = \log(\mu)$ $g_2(\sigma) = \log(\sigma)$
normal(θ_1, θ_2^2)	$\mu = \theta_1$ $\sigma = \theta_2$	$g_1(\mu) = \mu$ $g_2(\sigma) = \log(\sigma)$

Table 5 presents the parameter estimates, corresponding standard errors (SE), p-values associated with hypothesis testing based on the Wald square root statistic, and model selection criteria for the four fitted regression models. For comparison purposes, we fitted the reparametrized gamma, reparameterized Weibull, and normal linear regression models considering the same covariates. The two information criteria evaluated indicate that the Lomax regression model presented a better fit when compared to the other models. Considering the nature of the response variable, the Gamma and Weibull distributions are usual models for modeling continuous and positive data and that compete with the Lomax distribution, as their densities can assume a decreasing format. This is confirmed by the observed values of AIC and BIC. The normal model is one of the best known and most widely used in practice, but it is suitable for data with the supported in reals, and its probability density function does not assume a decreasing format. The values of AIC and BIC confirm that the normal model is not a competing model of the proposal presented here.

Table 5. Fit regression models for carbon emissions data.

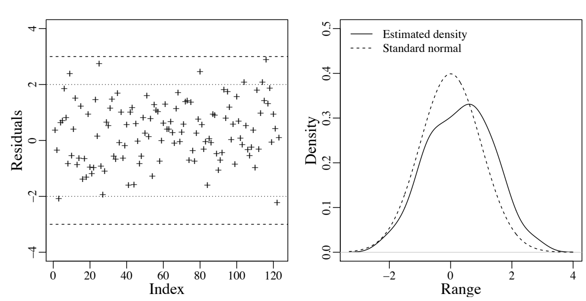
Model	Effect	Parameter	Estimate (SE)	p-value	AIC	BIC
	FAR	β_1	0.0305 (0.0036)	< 0.01		
Lomax	CIN	β_2	0.0466 (0.0045)	< 0.01	650.94	662.19
	EAG	β_3	-0.0424 (0.0048)	< 0.01		
	HDI	γ_1	0.0174 (0.0090)	0.0530		
	FAR	β_1	0.0304 (0.0036)	< 0.01		
Gamma	CIN	β_2	0.0468 (0.0044)	< 0.01	652.01	663.26
	EAG	β_3	-0.0427 (0.0047)	< 0.01		
	HDI	γ_1	0.0002 (0.0007)	0.7658		
	FAR	β_1	0.0304 (0.0036)	< 0.01		
Weibull	CIN	β_2	0.0467 (0.0044)	< 0.01	651.97	663.22
	EAG	β_3	-0.0427 (0.0047)	< 0.01		
	HDI	γ_1	-0.0003 (0.0008)	0.7251		
	FAR	β_1	0.0509 (0.0163)	< 0.01		
Normal	CIN	β_2	0.0959 (0.0189)	< 0.01	742.36	753.60
	EAG	β_2	-0.0724 (0.0172)	< 0.01		
	HDI	γ_1	0.0211 (0.0008)	< 0.01		

Figure 2 presents the half-normal plots with simulated envelopes for the randomized quantile residuals based on 100 replicates for the considered models. From these plots, we note that, except for the normal regression model (Figure 2 (d)), almost all observations appear inside the envelope bands, indicating a good fit of the regression models to the the carbon dioxide emissions per capita (Atkinson, 1981). Figure 3 presents the residuals against the index and estimated probability density function of the residuals in a non-parametric way against the normal standard probability density function. As expected, the residuals seems to be oscillating around zero with constant variance and approximately normally distributed.



(a) Lomax (b) Gamma (c) Weibull (d) Normal

Figure 2. Half-normal plot of residuals for the fitted models in this study.



(a) Residuals versus index (b) Probability density function

Figure 3. Residual plots for the proposed Lomax regression model.

The interpretation of the estimated parameters of the Lomax regression model are as follows:

(i) For each 1% that on forest area increases, the mean of the carbon dioxide increases by 3.10% ($e^{\widehat{\beta}_1} = 1.0310$).
(ii) For each 1% that on concentration index of exports increases, the mean of the carbon dioxide increases by 4.77% ($e^{\widehat{\beta}_2} = 1.0477$).
(iii) For each 1% that the employment in agriculture increases, the mean of the carbon dioxide decreases by 4.15% ($e^{\widehat{\beta}_3} = 0.9585$).
(iv) The coefficient of γ_1 is 0.0174, so when the HDI increases, the precision increases.

5. Conclusion

In this paper, we proposed a frequentist approach for the mean-parameterized Lomax regression model with varying precision. The main advantage of this reparametrization is its ability to model the mean directly. This makes the interpretation of the regression coefficients easier in terms of the expectation of the response variable and the proposed model more comparable with other models in the class of generalized linear models. The estimation of the regression model parameters is based on the maximum likelihood approach. We provided closed-form expressions for the score vector, observed information matrix, and Fisher information matrix. Through Monte Carlo simulations, we evaluated the asymptotic properties of maximum likelihood estimators. The simulation results showed that these estimators present a good performance. Finally, we illustrated the practical applicability of the proposed framework through an empirical application.

Supplementary materials

The computational routine implemented in R is available online at https://gist.github.com/moizesmelo/75a365ed957ae1ddbd9da9c3852597f3.

Author contributions Conceptualization, M.S.M., L.H.L.; Methodology, M.S.M., L.H.L.; Software, M.S.M., L.H.L., J.B.C.; Validation, M.S.M., L.H.L., J.B.C.; Formal analysis, M.S.M., L.H.L., J.B.C.; Investigation, M.S.M., L.H.L., J.B.C.; Data Curation, M.S.M., L.H.L., J.B.C.; Visualization, M.S.M., L.H.L., J.B.C.; Supervision, M.S.M., L.H.L., J.B.C.; Writing - original draft, M.S.M., L.H.L., J.B.C.; Writing - review and editing, M.S.M., L.H.L., J.B.C. All authors have read and agreed the published version of the manuscript.

Acknowledgements We thank all the corrections and suggestions pointed out by the Editors-in-Chief and the reviewers.

Funding The authors gratefully acknowledge partial financial support from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

Declaration of Conflict of Interest The authors declare no conflict of interest.

References

Akaike, H., 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control, 19, 716–723

Altun, E., 2021. The Lomax regression model with residual analysis: An application to insurance data. Journal of Applied Statistics, 48, 13–15.

Atkinson, A.B. and Harrison, A.J., 1978. Distribution of personal wealth in Britain. Cambridge University Press, Cambridge, UK.

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

Atkinson, A.C, 1985. Plots,Transformation and Regression. Clarendon Press, Oxford, UK.

Beirlant, J. and Goegebeur, Y., 2003. Regression with response distributions of Pareto-type. Computational Statistics and Data Analysis, 42, 595–619.

Bourguignon, M. and Gallardo, D.I, 2020. Reparameterized inverse gamma regression models with varying precision. Statistica Neerlandica, 74, 611–627.

Bourguignon, M., Leão, J., and Gallardo, D.I., 2020. Parametric modal regression with varying precision. Biometrical Journal, 62, 202–220.

Bourguignon, M. and Nascimento, F.F., 2020. Regression models for exceedance data: A new approach. Statistical Methods and Applications, 30, 1–17.

Chandra, N. and Khan, M.A., 2013. Optimum plan for step-stress accelerated life testing model under type-i censored samples. Journal of Modern Mathematics and Statistics, 7, 58–62.

Corbellini, A., Crosato, L., Ganugi, P., and Mazzoli, M., 2007. Fitting Pareto II distributions on firm size: Statistical methodology and economic puzzles. Advances in Data. Birkhauser, Boston, US, 26, 321–328.

Das, S., Harel, O., Dey, D.K., Covault, J. and Kranzler, H.R., 2010. Analysis of extreme drinking in patients with alcohol dependence using Pareto regression. Statistics in Medicine, 29, 1250–1258.

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

Fernández, L.M.Z. and De Andrade, T.A.N., 2020. The erf-G family: New unconditioned and log-linear regression models. Chilean Journal of Statistics, 11, 3–23.

Ferrari, S.L.P., Cribari-Neto, F., 2004. Beta regression for modelling rates and proportions. Journal of Applied Statistics, 31, 799–815.

Fonseca, R.V., Nobre, J.S., and Farias, R.B.A., 2016. Comparative inference and diagnostic in a reparametrized Birnbaum-Saunders regression model. Chilean Journal of Statistics, 7, 17–30.

Harris, C.M., 1968. The Pareto distribution as a queue service discipline. Operations Research, 16, 307–313.

Holland, O., Golaup, A., and Aghvami, A., 2006. Traffic characteristics of aggregated module downloads for mobile terminal reconfiguration. IEE Proceedings Communications, 153, 683–690.

Lomax, K., 1954. Business failures: Another example of the analysis of failure data. Journal of the American Statistical Association, 49, 847–852.

McCullagh, P. and Nelder, J., 1989. Generalized Linear Models. Chapman and Hall, Boca Raton, FL, US.

Nelder, J.A. and Mead, R., 1965. A simplex method for function minimization. The Computer Journal, 7, 308–313.

Palm, B.G., Bayer, F.M., Cintra, R.J., Pettersson, M.I., and Machado, R., 2019. Rayleigh regression model for ground type detection in SAR imagery. IEEE Geoscience and Remote Sensing Letters, 16, 1660–1664.

Pawitan, Y., 2001. In all likelihood: Statistical modelling and inference using likelihood. Oxford University Press, Oxford, UK.

R Development Core Team, 2021. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna, Austria.

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, 10, 2825–2855.

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

Simas, A.B., Barreto-Souza, W., and Rocha, A.V., 2010. Improved estimators for a general class of beta regression models. Computational Statistics and Data Analysis, 54, 348–366.

Smyth, G.K., 1989. Generalized linear models with varying dispersion. Journal of the Royal Statistical Society B, 51, 47–60.

Stasinopoulos, D.M. and Rigby, R.A., 2007. Generalized additive models for location scale and shape (GAMLSS) in R. Journal of Statistical Software, 23, 1–46.

Vigas, V.P., Silva, G.O. and Louzada, F., 2017. The Poisson-Weibull regression model. Chilean Journal of Statistics, 8, 25–51.

Information for authors

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

Preparation of accepted manuscripts

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

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

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

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

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

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

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

Copyright