Article

A Novel Regression Model for Fractiles: Formulation, Computational Aspects, and Applications to Medical Data

Víctor Leiva [1,*], Josmar Mazucheli [2] and Bruna Alves [2]

1   School of Industrial Engineering, Pontificia Universidad Católica de Valparaíso, Valparaíso 2362807, Chile
2   Department of Statistics, Universidade Estadual de Maringá, Maringá 87020-900, Brazil
*   Correspondence: victor.leiva@pucv.cl or victorleivasanchez@gmail.com

Abstract: Covariate-related response variables that are measured on the unit interval frequently arise in diverse studies when index and proportion data are of interest. A regression on the mean is commonly used to model this relationship. Instead of relying on the mean, which is sensitive to atypical data and less general, we can estimate such a relation using fractile regression. A fractile is a point on a probability density curve such that the area under the curve between that point and the origin is equal to a specified fraction. Fractile or quantile regression modeling has been considered for some statistical distributions. Our objective in the present article is to formulate a novel quantile regression model which is based on a parametric distribution. Our fractile regression is developed reparameterizing the initial distribution. Then, we introduce a functional form based on regression through a link function. The main features of the new distribution, as well as the density, distribution, and quantile functions, are obtained. We consider a brand-new distribution to model the fractiles of a continuous dependent variable (response) bounded to the interval $(0, 1)$. We discuss an R package with random number generators and functions for probability density, cumulative distribution, and quantile, in addition to estimation and model checking. Instead of the original distribution-free quantile regression, parametric fractile regression has lately been employed in several investigations. We use the R package to fit the model and apply it to two case studies using COVID-19 and medical data from Brazil and the United States for illustration.

Keywords: GLM; Marshall–Olkin distribution; maximum likelihood estimation methods; Monte–Carlo simulation methods; quantile function; R statistical software; statistical parameterizations; Weibull distribution

MSC: 60E05; 62J02

1. Introduction, State-of-the-Art, and Motivation

We provide in this section an introduction, the bibliographical review, our objectives, and the plan of this article, as well as a motivating example of the proposed methodology from medicine.

1.1. Bibliographical Review

When information about indices, percentages, proportions, and rates is relevant in an area of research, data derived from continuous random variables constrained to a range of possible values frequently appear. This type of information has become crucial during the COVID-19 pandemic to investigate issues regarding infected and recovered cases, as well as deaths worldwide [1,2]. However, this pandemic also has inducted indirect aspects related to morbidity. For example, the mentioned pandemic produced serious problems of a sedentary lifestyle because of physical inactivity. The medical applications presented in our investigation are associated with these two aspects, that is, COVID-19 pandemic and sedentary lifestyle.

To analyze data constrained to a range of possible values, the regression technique is a statistical method which is popularly used and largely utilized for parameter estimation and hypothesis testing [3]. One estimates the expectation (conditional) of the dependent or response variable employing this technique, while the independent variables (covariates or explanatory variables) are assumed to take a specific value. The response conditional mean is often described by the covariates utilizing a linear or non-linear equation. However, if the response includes extreme atypical values or outliers, common in different types of data, the estimated conditional mean is affected by these outliers, and consequently, the predictions will be inaccurate.

Under the unit mean response modeling, the beta regression [4] comes to mind first when connecting the conditional expectation of the response established over the unit interval with covariates. The beta regression has been proposed by reparameterizing its probability density function (PDF) according to its expected value. This parameterization can be applied to probability distributions with a closed-form expectation, and any model parameter can be transformed by using this expected value.

A fractile is a point on a PDF curve such that the area under the curve between that point and the origin is equal to a specified fraction. The fractile or quantile regression (QR) technique was introduced in [5,6], which can be presented alternatively to a mean regression [7–9]. The QR originally proposed has no assumptions about the distribution of the response and tends to be robust to outlying observations. This property is an advantage of the QR in relation to the mean regression. Similarly to this regression, a QR based on a probability distribution can be postulated via a reparameterized form of any distribution parameter according to its quantile function (QF). The QR models relate a set of covariates to specific quantiles of a response variable, with the median being often considered; see [10] for a study and implementation of this class of models and their applications, including QR for continuous, discrete, and unit responses.

Several parametric QR structures were postulated based on diverse statistical distributions, including the arc-secant hyperbolic Weibull (ASHW), Johnson SB (JOSB), Kumaraswamy (KUMA), log-extended exponential-geometric (LEEG), unit-Birnbaum-Saunders (UBSA), unit-Burr-XII (UBUR), unit-Chen (UCHE), unit-Gompertz (UGOM), unit-Gumbel (UGUM), unit-logistic (ULOG), unit-generalized half-normal-E (UGHE), unit-generalized half-normal-X (UGHX), unit-Weibull-E (UWEE), unit-Weibull-X (UWEX), and Vasicek (VASI) models [11–26]. Some of these distributions are based on classical models [27–29]. In [10], the authors provided a number of parametric QRs, as well as examples and software packages, which can be employed to model rates and indices.

Nevertheless, to our best knowledge, QR structures under the Weibull–Marshall-Olkin distribution have not been established until now. Note that the Weibull–Marshall–Olkin distribution is widely flexible in statistical terms to model diverse types of data, such as shown in the theoretical shape analysis and empirical applications considered in the present study. This flexibility allows the Weibull–Marshall-Olkin distribution to provide a better performance in different classes of statistical modeling in relation to its competing distributions, as confirmed in the mentioned empirical applications here stated.

Therefore, our objective in the present article is to formulate a novel Weibull–Marshall–Olkin QR model. First, we introduce a reparameterization of the Weibull–Marshall–Olkin distribution by adding a quantile parameter and so we formulate the novel regression model, providing high flexibility and allowing us to state a framework such as in GLM—generalized linear models—[30]. We demonstrate that our new model outperforms recently proposed QR models in terms of model fitting.

1.2. An Motivation Case Study from Medicine

Researchers are interested in statistical models for describing data with quantiles that incorporate the effect of covariates. However, the statistical methods for this type of data analysis remain a matter of discussion and are not yet fully disseminated.

A motivation for our investigation comes from a real-world medical study associated with a data set related to the body fat percentage of people who received assistance in a government hospital of the state of Paraná, in the city of Curitiba, Brazil.

The inclusion criteria to recruit people were the following: (i) healthy people (men and women) between 18 and 90 years old having a body mass index (BMI) of 18.5 and 29.9 kg/m^2; (ii) those who do not take supplements or hormone replacement therapy that might change their appearance; and (iii) those who are physically active and are able to walk without the aid of orthoses or prostheses.

The following persons were excluded from the investigation: (i) people with chronic illnesses; (ii) those taking prescription, illicit drugs, or drugs that affect body composition, such as corticosteroids, thyroid hormones in suppressive doses, or insulin-dependent diabetes; and (iii) those who were thin with a BMI under 18.5 kg/m^2 or with a BMI of around 30 kg/m^2.

For the recruited people (participants), all subjects' anthropometric measurements (weight and height) were collected the same day after the total body densitometry of the body fat, lean, and bone masses' total.

All participants filled out the validated Portuguese form of the International Physical Activity Questionnaire (IAPQ) [31], which frequently measures physical activity. The IPAQ-short is made out of eight questions. This questionnaire permits the assessment of week-after-week time spent on proactive tasks of moderate areas of strength for, and in a variety of contexts associated with day-to-day life. These contexts include work, leisure, transportation, and housework, in addition to time spent engaging in straightforward sitting exercises.

Most of the volunteers answered the IPAQ through self-administration or an individual interview. The volunteers were then split into three groups based on how much exercise they had performed. Sedentary individuals do not engage in any form of physical activity for more than ten consecutive minutes each week. At least ten continuous minutes of physical activity at least five days per week or 150 min per week is required to be considered insufficiently active. Assets spend at least 150 min a week doing at least 20 min of vigorous or moderate exercise three times a week, 30 min of walking five times a week, or any additional activity five days a week or more.

The data set was considered in [32], which consists of 298 records about the fat percentage of people who received assistance in a government hospital in Curitiba, Paraná, Brazil. The data set includes two categorical and two continuous covariates, and the five responses correspond to the fat percentages at the android, arms, body, gynaecoid, and legs. The person's BMI (in kg/m^2) and age (in years) are continuous covariates. In contrast, gender (female or male) and IPAQ (sedentary patient, insufficiently active, or active) are the categorical covariates.

We consider the fat percentage at the legs as the response variable mainly, but the other responses are also considered later. We perform a descriptive statistical analysis [33] for continuous variables of a data set related to body fat. Univariate descriptive statistics and correlation analysis for each variable are presented in Table 1. Histograms and scatterplots represent the variables: age, legs, and BMI in Figure 1. The variables age and BMI are substantially correlated. Note that, at 1%, each correlation is statistically significant, encouraging the development of a regression model. When these models are formulated, a collinearity problem may be detected between BMI and age.

Our correlation analysis indicates a Spearman coefficient (ρ) that is significant at 1% when testing H$_0$: $\rho = 0$. Regarding the mentioned variables, we report: legs versus BMI with $r = 0.230$ and *p*-value < 0.001; legs versus age with $r = 0.284$ and *p*-value < 0.001; and BMI versus age with $r = 0.470$ and *p*-value < 0.001, where r is the sample correlation coefficient.

Table 1. Univariate descriptive statistics of data related to the fat percentage at the legs.

Variable Grouped by Level	Mean	SD	Minimum	Maximum	1st Quartile	2nd Quartile	3rd Quartile	CS	CK
legs	0.324	0.11	0.068	0.574	0.235	0.319	0.416	0.021	−0.843
legs (by F)	0.407	0.074	0.203	0.574	0.355	0.406	0.459	−0.205	−0.350
legs (by M)	0.241	0.072	0.068	0.444	0.195	0.238	0.289	0.174	0.021
legs (by A)	0.303	0.117	0.068	0.574	0.213	0.290	0.397	0.223	−0.900
legs (by I)	0.342	0.097	0.133	0.563	0.259	0.350	0.423	−0.098	−0.732
legs (by S)	0.359	0.095	0.174	0.565	0.289	0.357	0.442	0.050	−0.988
age	46.000	19.879	18.000	87.000	25.000	47.000	65.000	0.158	−1.359
bmi	24.716	3.151	18.500	29.900	22.300	24.900	27.200	−0.101	−0.943

Where legs is fat (in %) at legs, male (M), female (F), sedentary (S), insufficiently active (I), active (A); and SD is the standard deviation.
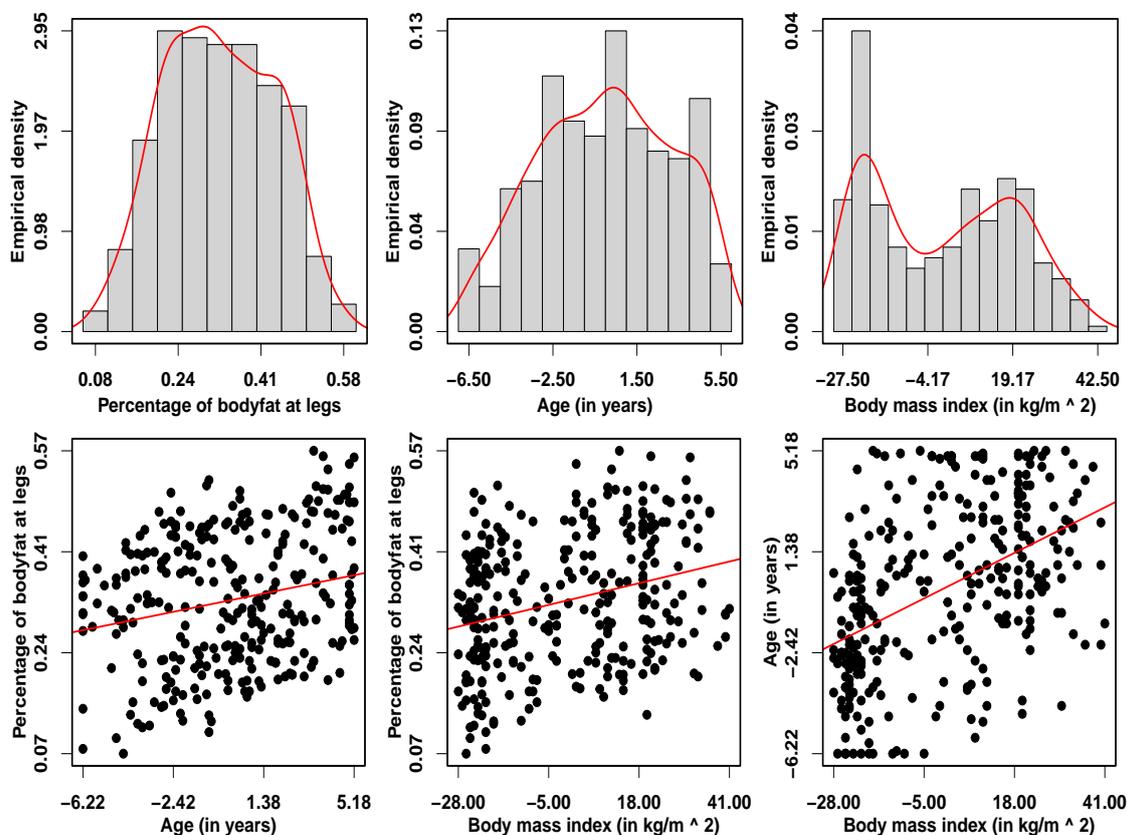
Figure 1. Scatterplots and histograms for the body fat in the legs data set.

　　Before statistically formulating a QR model, we must evaluate any potential collinearity issues in the exploratory data analysis. The variance inflation factor (VIF) is a more formal and often applied method to determine the collinearity between covariates. If VIF > 10, then collinearity can be present. For the covariates under analysis, we report the corresponding VIF: age with VIF = 1.662; BMI with VIF = 1.411; gender/sex with VIF = 1.083; and IPAQ with VIF = 1.280. Thus, we can continue with our QR analysis because every VIF value is less than 10. In summary, derived from our exploratory and VIF analyses, we state that a QR structure with high flexibility may describe these data well and better than its competing models. We recall that the Weibull–Marshall–Olkin distribution is widely flexible in statistical terms to model diverse types of data. Then, the new QR model derived in the present study can provide a better performance in relation to its competitors. Thus, the motivation example inspires our research from an empirical perspective.

1.3. Objectives and Plan of Sections

As mentioned, our objective in the present study is to formulate a new alternative QR model based on a Weibull–Marshall–Olkin distribution to relate covariates and specific quantiles of a response variable. Moreover, we demonstrate how our QR model can be used with real data sets.

Our article is organized into sections. We establish the new model and its methodology in Section 2, whereas Section 3 introduces two case studies as illustration. Section 4 details more information on our study's findings and discusses them. We provide some conclusions, limitations, and ideas for further investigation in Section 4.

2. Methodology

In this section, the new QR model and its statistical methodology is presented.

2.1. A New Weibull–Marshall–Olkin PDF

Let X follow a Weibull model based on $F(x; \varsigma, \nu) = 1 - \exp[-(x/\varsigma)^\nu]$, that is, on the cumulative distribution function (CDF). Also, consider the Marshall–Olkin method to add a new parameter into a distribution proposed in [34]. PDF, CDF, and QF expressions for the extended Weibull Marshall-Olkin (EWMO) model are, respectively, given by

$$f(x; \alpha, \varsigma, \nu) \quad = \quad \frac{\alpha \nu x^{\nu-1} \exp[-(x/\varsigma)^\nu]}{\varsigma^\nu \{1 - \overline{\alpha} \exp[-(x/\varsigma)^\nu]\}^2}, \quad x > 0, \tag{1}$$

$$F(x; \alpha, \varsigma, \nu) \quad = \quad \frac{1 - \exp[-(x/\varsigma)^\nu]}{1 - \overline{\alpha} \exp[-(x/\varsigma)^\nu]}, \quad Q(\tau; \alpha, \varsigma, \nu) = \varsigma \left[-\log\left(\frac{1-\tau}{1-\tau\overline{\alpha}} \right) \right]^{\frac{1}{\nu}}, \tag{2}$$

where $\varsigma > 0$ is a scale parameter, $\nu > 0$ and $\alpha > 0$ are shape parameters, $\overline{\alpha} = (1 - \alpha)$ is a tilt parameter, and $0 < \tau < 1$ is a fractile parameter. As a particular case of the EWMO model, we obtain the extended Marshall–Olkin exponential distribution if $\alpha = 1$. The PDF, CDF, and QF stated in (1) and (2) correspond to the Weibull distribution when $\nu = 1$. Note that starting with a baseline CDF for a random variable X, $F_0(x; \boldsymbol{\theta})$, namely, we can obtain the extended Marshall–Olkin distribution with a CDF defined as

$$F(x; \boldsymbol{\theta}, \alpha) = (1 - \alpha \overline{F}_0(x; \boldsymbol{\theta})) / (1 - \overline{\alpha} \overline{F}_0(x; \boldsymbol{\theta})), \tag{3}$$

where $\overline{F}_0(x; \boldsymbol{\theta})$ defined in (3) is the survival function of X, that is, $\overline{F}_0(x; \boldsymbol{\theta}) = 1 - F_0(x; \boldsymbol{\theta})$, and $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_k)^\top$ is a vector of parameters.

Considering the transformation $X = \exp(-Y)$, we define the unit Weibull–Marshall–Olkin (UWMO) model, with PDF, CDF, and QF being specified as

$$f(y; \alpha, \varsigma, \nu) \quad = \quad \frac{\alpha \nu [-\log(y)]^{\nu-1} \exp\{-[-\log(y)/\varsigma]^\nu\}}{\varsigma^\nu y (1 - \overline{\alpha} \exp\{-[-\log(y)/\varsigma]^\nu\})^2}, \quad 0 < y < 1, \tag{4}$$

$$F(y; \alpha, \varsigma, \nu) \quad = \quad \frac{\alpha}{\exp\left[(-\log(y)/\varsigma)^\nu\right] - \overline{\alpha}}, \quad Q(\tau; \alpha, \varsigma, \nu) = \exp\left\{ -\varsigma \log[(\alpha + \tau\overline{\alpha})/\tau]^{\frac{1}{\nu}} \right\}, \tag{5}$$

with $\alpha > 0$ and $\nu > 0$ being parameters associated with the shape of the distribution, $\overline{\alpha} = (1 - \alpha)$ being a tilt parameter, and $0 < \tau < 1$ being a fractile parameter. Note that now $\varsigma > 0$ is not a parameter related to the scale of the distribution, since $F(y; \alpha, \varsigma, \nu)$ is not equal to $F(y/\varsigma; \alpha, \nu)$. For $\alpha = 1$, we have the unit-Weibull model. In applications, the parameters α, ς, and ν do not have direct interpretations regarding the observed data. However, we can reparameterize the parameter α, taking a one-to-one change into consideration as: $(\alpha, \varsigma, \nu) \to (\mu, \varsigma, \nu)$, where $\mu = Q$ is a quantile defined in (5). Hence, we have that

$$\alpha = h^{-1}(\mu) = \frac{\tau}{(1-\tau)} \left(\exp\left\{ \left[-\frac{\log(\mu)}{\varsigma} \right]^\nu \right\} - 1 \right). \tag{6}$$

From (6), we generate the unit Weibull–Marshall–Olkin quantile (UWMOQ) distribution of parameters $\mu > 0$, $\varsigma > 0$, $\nu > 0$, and $0 < \tau < 1$. Figures 2–4 show plots of potential shapes for the PDF of the UWMOQ distribution for particular values of μ, ς, ν, and τ. The graphical plots display skewness to the left and right as well as bathtub shapes of the UWMOQ distribution, showing the wide shape flexibility of the UWMOQ distribution.
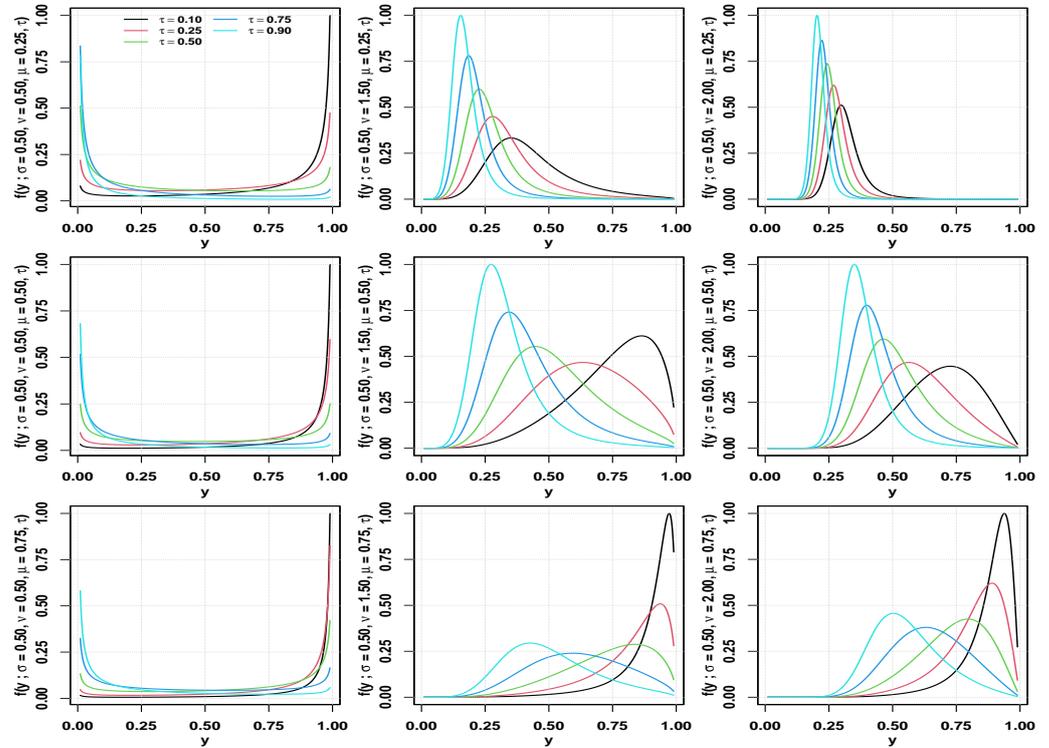

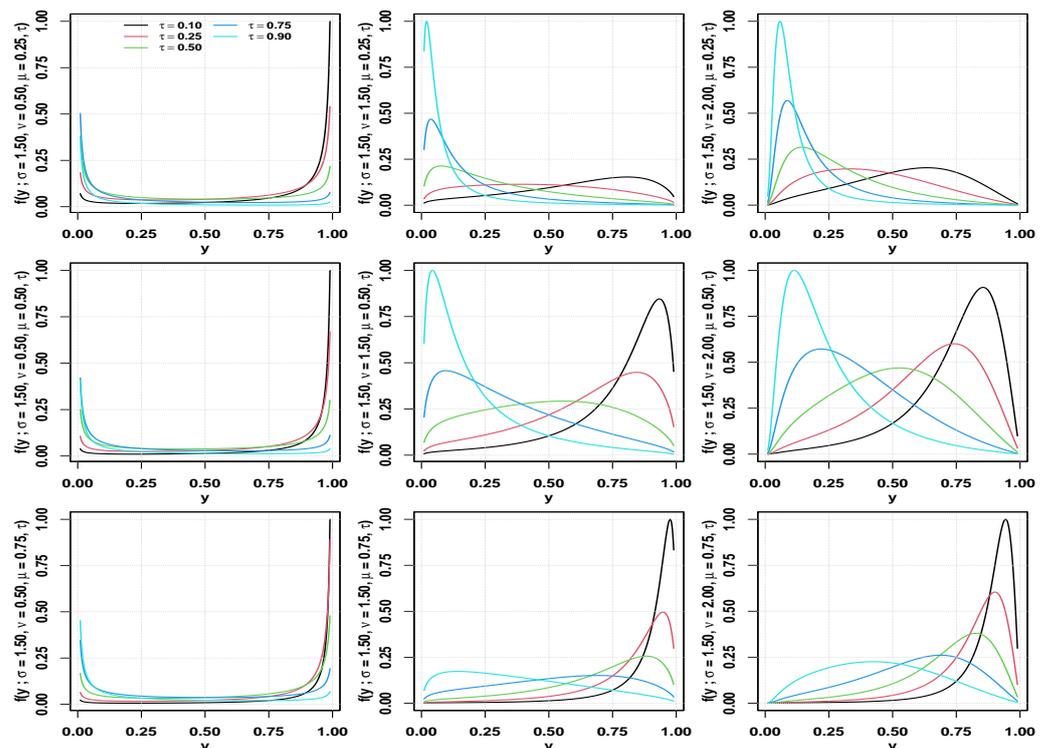**Figure 2.** Plots of the UWMOQ PDF for the indicated values of μ, ν, τ, and $\sigma = \varsigma = 0.5$.

Figure 3. Plots of the UWMOQ PDF for the indicated values of μ, ν, τ, and $\sigma = \varsigma = 1.0$.
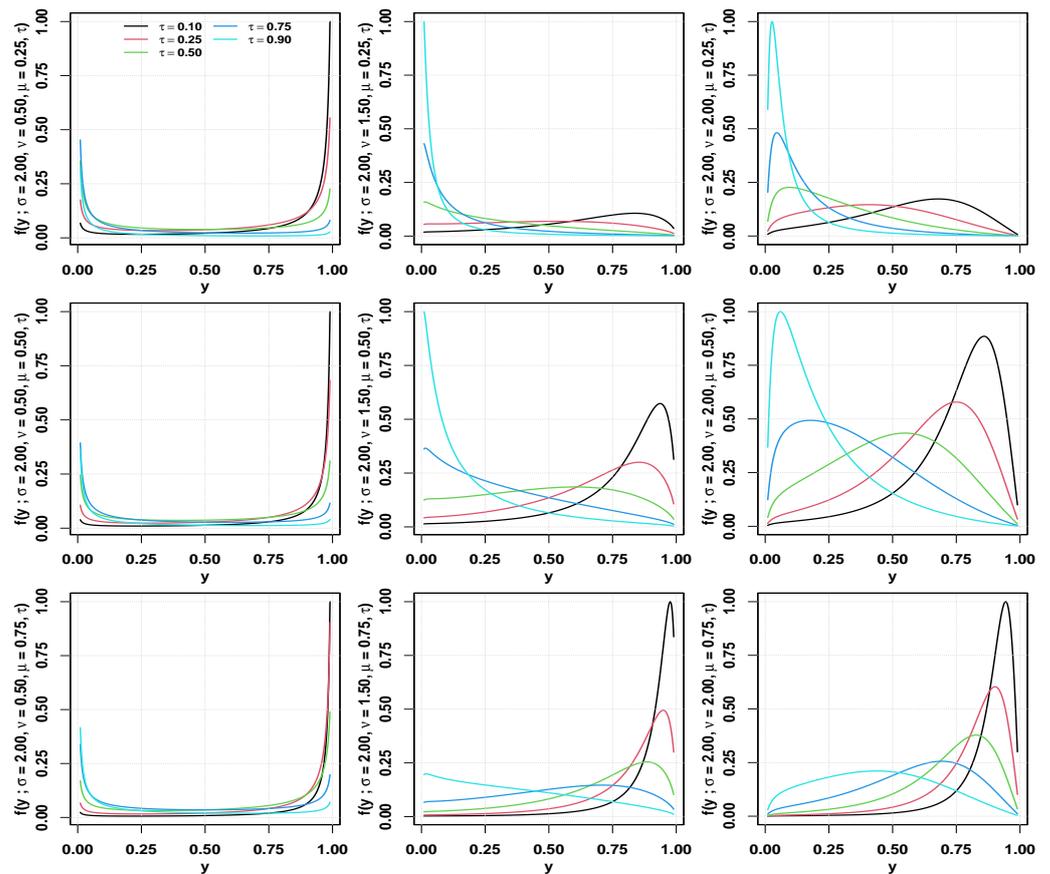
Figure 4. Plots of the UWMOQ PDF for the indicated values of μ, ν, τ, and $\sigma = \varsigma = 2.0$.

2.2. A Novel Quantile Regression Model

Contrasted with regression models that determine the conditional mean of the response variable, we observe that the QR models are a solid substitute. Since they estimate the conditional quantiles of the covariates, these models are also robust regressions. QR models based on probability distributions are generally proposed via a reparameterized form of a distribution parameter according to their QF. Since the QF of the UWMOQ distribution is transformable by its parameters, it is possible to suggest an alternative QR model derived from the UWMOQ PDF.

Now that we have the reparameterized PDF based on (6), we can concentrate on the QR model employing the UWMOQ distribution. Assume observations obtained from the reparameterized UWMOQ PDF with parameters μ_i, $\boldsymbol{\beta}$, and known quantile level τ. Hence, a UWMOQ model for the quantile μ_i is established by

$$g(\mu_i) = \boldsymbol{x}_i \boldsymbol{\beta}^\top, \quad i \in \{1, \ldots, n\}, \tag{7}$$

where $\boldsymbol{\beta} = (\beta_0, \beta_1, \ldots, \beta_p)^\top$ and $\boldsymbol{x}_i = (1, x_{i1}, \ldots, x_{ip})$ represent the QR coefficients vector (unknown) and i-th vector of values (known) of the covariates, respectively. Similarly to (7), we can model ς_i and ν_i. Note that g stated in (7) is the link that connects the covariates to the quantile of the response variable. In addition, the link is determined by the distribution domain. Since the domain of the UWMOQ distribution is bounded on the interval $(0, 1)$, logit or probit links can be used to relate the response variable quantile to covariates. Thus, we utilize a link related to the logit function, which is displayed by

$$g(\mu_i) = \operatorname{logit}(\mu_i) = \log\left(\frac{\mu_i}{1 - \mu_i}\right) = \boldsymbol{x}_i \boldsymbol{\beta}^\top, \quad i \in \{1, \ldots, n\}.$$

2.3. Estimating the Model Parameters

Presently, the estimation of model parameters is carried out with the maximum likelihood (ML) method, considering the parameterizations (α, ς, ν) described. Here, we also propose a diagnostic checking for regression [35] based on a residual analysis.

Assume a sample of size n, $Y = (Y_1, \ldots, Y_n)^\top$, namely, extracted from a UWMO distribution, which has parameters α, ς, ν, and let its observations be stated as $y = (y_1, \ldots, y_n)^\top$. Using the PDF defined in (4), the proportional log-likelihood function for $(\alpha, \varsigma, \nu)^\top$ is associated with

$$\ell(\alpha, \varsigma, \nu; y) \;\propto\; n\log(\alpha) + n\log(\nu) - n\nu\log(\varsigma) - \nu\sum_{i=1}^{n}\log(y_i) + \sum_{i=1}^{n}\log\left\{\frac{[-\log(y_i)]^{\nu-1}\exp\left\{-\left[-\frac{\log(y_i)}{\varsigma}\right]^{\nu}\right\}}{1-(1-\alpha)\exp\left\{-\left[-\frac{\log(y_i)}{\varsigma}\right]^{\nu}\right\}}\right\}. \tag{8}$$

ML estimates of $(\alpha, \varsigma, \nu)^\top$ are generated by maximizing the logarithmic likelihood function established in (8) and resolving the score elements $U_\alpha = \partial\ell(\alpha, \varsigma, \nu; y)/\partial\alpha = 0$, $U_\varsigma = \partial\ell(\alpha, \varsigma, \nu; y)/\partial\varsigma = 0$, and $U_\nu = \partial\ell(\alpha, \varsigma, \nu; y)/\partial\nu = 0$ accordingly, expressed as

$$U_\alpha = \frac{n}{\alpha} + \sum_{i=1}^{n}\frac{\exp\left\{-\left[-\frac{\log(y_i)}{\varsigma}\right]^{\nu}\right\}}{1-(1-\alpha)\exp\left\{\left[-\frac{\log(y_i)}{\varsigma}\right]^{\nu}\right\}},$$

$$U_\varsigma = \frac{1}{\varsigma}\left\{\nu\sum_{i=1}^{n}\frac{\left[-\frac{\log(y_i)}{\varsigma}\right]^{\nu}}{1-(1-\alpha)\exp\left\{-\left[-\frac{\log(y_i)}{\varsigma}\right]^{\nu}\right\}} - n\nu\right\},$$

$$U_\nu = \frac{n}{\nu} - n\log(\varsigma) - \sum_{i=1}^{n}\log(y_i)$$

$$-\sum_{i=1}^{n}\frac{\log[-\log(y_i)](1-\alpha)\exp\left\{-\left[-\frac{\log(y_i)}{\varsigma}\right]^{\nu}\right\} + \left[-\frac{\log(y_i)}{\varsigma}\right]^{\nu}\log\left[-\frac{\log(y_i)}{\varsigma}\right]}{\cdot}$$

$$+\sum_{i=1}^{n}\frac{\log[-\log(y_i)]}{1-(1-\alpha)\exp\left\{-\left[-\frac{\log(y_i)}{\varsigma}\right]^{\nu}\right\}}.$$

The associated inverse Hessian matrix calculates the asymptotic standard errors (SEs) of $\widehat{\alpha}$, $\widehat{\varsigma}$, and $\widehat{\nu}$. The quantile reparameterization of the UWMO distribution requires considering that

$$\alpha = \frac{\tau}{(1-\tau)}\left\{\exp\left(\left[-\frac{\log(\mu)}{\varsigma}\right]^{\nu}\right) - 1\right\}.$$

Thus, we want to evaluate how the covariates affect $(\alpha, \varsigma, \nu)^\top$ by means of

$$g_1(\mu_i) = \eta_i = \beta_0 + \beta_1 x_{i1} + \cdots + \beta_p x_{ip},$$
$$g_2(\varsigma_i) = \zeta_i = \delta_0 + \delta_1 z_{i1} + \cdots + \delta_q z_{iq},$$
$$g_3(\nu_i) = \rho_i = \gamma_0 + \gamma_1 v_{i1} + \cdots + \gamma_r v_{ir},$$

using functions of the explanatory variables $x_i = (1, x_{i1}, \ldots, x_{ip})^\top$, $z_i = (1, z_{i1}, \ldots, z_{iq})^\top$ and $v_i = (1, v_{i1}, \ldots, v_{ir})^\top$, respectively. We assume that g_1, g_2, and g_3 consist of functions which are strictly monotonic, twice-differentiable, and that map the τ-th quantile μ_i, ς_i, and ν_i to the real line. Suitable choices of g_1 are the Cauchy: Cauchit link; maximum extreme value: log–log link; minimum extreme value: complementary log–log link; the inverse CDF of the logistic: logit link; normal: probit link; distributions. We consider the log link for g_2 and g_3. Observe that x, z, and v may be identical repetitions or subsets one of another.

We use 1st-order and 2nd-order derivatives of the logarithm of the likelihood function to estimate the model parameters. This log-likelihood function is stated as $\ell_i = \ell_i(\Theta) = \log[f(y_i; \Theta, x_i, z_i, v_i, \tau)]$, and then the score vector is constructed with elements defined as

$$\frac{\partial \ell_i}{\partial \beta_j} = \frac{\partial \ell_i}{\partial \mu_i}\frac{\partial \mu_i}{\partial \eta_i}\frac{\partial \eta_i}{\partial \beta_j}, \quad \frac{\partial \ell_i}{\partial \delta_j} = \frac{\partial \ell_i}{\partial \varsigma_i}\frac{\partial \varsigma_i}{\partial \zeta_i}\frac{\partial \zeta_i}{\partial \delta_j}, \quad \frac{\partial \ell_i}{\partial \gamma_j} = \frac{\partial \ell_i}{\partial v_i}\frac{\partial v_i}{\partial \rho_i}\frac{\partial \rho_i}{\partial \gamma_j},$$

where $\Theta = (\beta^\top, \delta^\top, \gamma^\top)^\top$; $\beta = (\beta_0, \ldots, \beta_p)^\top$ is a $(p+1) \times 1$ vector of parameters related to an $n \times (p+1)$ covariates matrix X; $\delta = (\delta_0, \ldots, \delta_q)^\top$ is a $(q+1) \times 1$ vector of parameters related to an $n \times (q+1)$ matrix of covariates Z; and $\gamma = (\gamma_0, \ldots, \gamma_r)^\top$ is an $(r+1) \times 1$ vector of parameters linked to an $n \times (r+1)$ matrix of covariates V. We have the score of the log-likelihood function expressed as $\partial \ell / \partial \beta = X^\top \mathrm{diag}(W_\mu)\dot{\ell}_\mu$, $\partial \ell / \partial \delta = Z^\top \mathrm{diag}(W_\varsigma)\dot{\ell}_\varsigma$, and $\partial \ell \partial \gamma = V^\top \mathrm{diag}(W_v)\dot{\ell}_v$, with "diag" referring to a diagonal matrix of size n and

$$W_\mu = \left[\frac{\partial \mu_1}{\partial \eta_1}, \ldots, \frac{\partial \mu_n}{\partial \eta_n}\right], \quad W_\varsigma = \left[\frac{\partial \varsigma_1}{\partial \zeta_1}, \ldots, \frac{\partial \varsigma_n}{\partial \zeta_n}\right], \quad W_v = \left[\frac{\partial v_1}{\partial \rho_1}, \ldots, \frac{\partial v_n}{\partial \rho_n}\right],$$

$$\dot{\ell}_\mu = \left[\frac{\partial \ell}{\partial \mu_1}, \ldots, \frac{\partial \ell}{\partial \mu_n}\right]^\top, \quad \dot{\ell}_\varsigma = \left[\frac{\partial \ell}{\partial \varsigma_1}, \ldots, \frac{\partial \ell}{\partial \varsigma_n}\right]^\top, \quad \dot{\ell}_v = \left[\frac{\partial \ell}{\partial v_1}, \ldots, \frac{\partial \ell}{\partial v_n}\right]^\top.$$

Then, the corresponding Hessian matrix has elements stated as

$$\frac{\partial^2 \ell}{\partial \beta \partial \beta^\top} = X^\top \mathrm{diag}(\ddot{\ell}_{\mu\mu})\mathrm{diag}\left(W_\mu^2\right)X,$$

$$\frac{\partial^2 \ell}{\partial \delta \partial \delta^\top} = Z^\top \mathrm{diag}(\ddot{\ell}_{\varsigma\varsigma})\mathrm{diag}\left(W_\varsigma^2\right)Z,$$

$$\frac{\partial^2 \ell}{\partial \gamma \partial \gamma^\top} = V^\top \mathrm{diag}(\ddot{\ell}_{vv})\mathrm{diag}\left(W_v^2\right)V,$$

$$\frac{\partial^2 \ell}{\partial \beta \partial \delta^\top} = X^\top \mathrm{diag}(\ddot{\ell}_{\mu\varsigma})\mathrm{diag}(W_\mu)\mathrm{diag}(W_\varsigma)Z,$$

$$\frac{\partial^2 \ell}{\partial \beta \partial \gamma^\top} = X^\top \mathrm{diag}(\ddot{\ell}_{\mu v})\mathrm{diag}(W_\mu)\mathrm{diag}(W_v)V,$$

$$\frac{\partial^2 \ell}{\partial \delta \partial \gamma^\top} = Z^\top \mathrm{diag}(\ddot{\ell}_{\varsigma v})\mathrm{diag}(W_\varsigma)\mathrm{diag}(W_v)V,$$

where

$$\ddot{\ell}_{\mu\mu} = \left[\frac{\partial^2 \ell}{\partial \mu_1 \partial \mu_1}, \ldots, \frac{\partial^2 \ell}{\partial \mu_n \partial \mu_n}\right], \ddot{\ell}_{\varsigma\varsigma} = \left[\frac{\partial^2 \ell}{\partial \varsigma_1 \partial \varsigma_1}, \ldots, \frac{\partial^2 \ell}{\partial \varsigma_n \partial \varsigma_n}\right], \ddot{\ell}_{vv} = \left[\frac{\partial^2 \ell}{\partial v_1 \partial v_1}, \ldots, \frac{\partial^2 \ell}{\partial v_n \partial v_n}\right],$$

$$\ddot{\ell}_{\mu\varsigma} = \left[\frac{\partial^2 \ell}{\partial \mu_1 \partial \varsigma_1}, \ldots, \frac{\partial^2 \ell}{\partial \mu_n \partial \varsigma_n}\right], \ddot{\ell}_{\mu v} = \left[\frac{\partial^2 \ell}{\partial \mu_1 \partial v_1}, \ldots, \frac{\partial^2 \ell}{\partial \mu_n \partial v_n}\right], \ddot{\ell}_{\varsigma v} = \left[\frac{\partial^2 \ell}{\partial \varsigma_1 \partial v_1}, \ldots, \frac{\partial^2 \ell}{\partial \varsigma_n \partial v_n}\right].$$

The ML estimates $(\widehat{\beta}_0, \ldots, \widehat{\beta}_p)^\top$, $(\widehat{\delta}_0, \ldots, \widehat{\delta}_q)^\top$, and $(\widehat{\gamma}_0, \ldots, \widehat{\gamma}_r)^\top$ can be obtained utilizing all-purpose optimization techniques such as conjugate-gradient, quasi-Newton, and Nelder-Mead. These techniques are available in the function `optim` of R [36] or in SAS `proc NLMIXED` [37], and they allow us to attain at least a local maximum of the logarithm of the likelihood function. By definition, an ML estimate is the global maximum point of the log-likelihood function, but this estimate can also be a local maximum point or a saddle point. Directions for identifying such points may be found in [38].

Taking into account certain regularity conditions, we have that

$$\widehat{\Theta} \overset{\cdot}{\sim} N_{p+q+r+3}(\Theta, (\mathcal{I}(\Theta))^{-1}), \tag{9}$$

where $\mathcal{I}(\boldsymbol{\Theta})$ represents the expected information matrix that is written as

$$\mathcal{I}(\boldsymbol{\Theta}) = \mathrm{E}\left[-\frac{\partial^2 \ell(\boldsymbol{\Theta})}{\partial \boldsymbol{\Theta}\,\partial \boldsymbol{\Theta}^\top}\right]. \tag{10}$$

The mentioned regularity conditions include that the true value of the parameter $\boldsymbol{\Theta}$ must be interior to the parameter space, the log-likelihood function must be thrice differentiable, and its third derivatives must be bounded. The interested reader can consult a standard textbook for more technical details and proofs about such conditions, which are general and well-known in statistical inference; see, for example, [39]. By using the expression given in (9), one can arrive at approximate confidence intervals (CIs). However, to obtain the matrix described in (10), we may use the observed information matrix described by $\mathcal{J}(\boldsymbol{\Theta}) = -\partial^2 \ell(\boldsymbol{\Theta})\partial \boldsymbol{\Theta}\,\partial \boldsymbol{\Theta}^\top$, which must be evaluated at $\boldsymbol{\Theta} = \widehat{\boldsymbol{\Theta}}$.

3. Empirical and Computational Aspects

We provide in this section some details of an `R` package developed by the authors for conducting QR with 15 existing distributions. The new UWMO model introduced in the present article is being added to this package. As an alternative to the new in-progress version of such a package, we developed a `SAS` macro to perform our numerical examples [37,40]. In addition to the description of the functionality of the `R` package, in this section, we also present the numerical and empirical applications of our work. First, a Monte Carlo simulation study is conducted. Then, two medical illustrations involving COVID-19 and sports medicine data sets are introduced.

3.1. Computational Package

We employ an `R` package named `unitquantreg`, which was developed in the setting `stats::lm` package. The `unitquantreg()` function of this package is flexible and permits us, through the `ns()` function, to utilize regression splines. This function is available in the `splines` package. Also, we use the `gam()` function of the `mgcv` package. The `unitquantreg` package can be secured from https://github.com/AndrMenezes/unitquantreg (accessed on 30 January 2023) and installed using `devtools::install_github(''AndrMenezes/unitquantreg'')`. Next, we describe the models of the `unitquantreg` package and its commands. QR structures implemented in the `unitquantreg` package use: (i) ASHW; (ii) JOSB; (iii) KUMA; (iv) LEEG; (v) UBSA; (vi) UBUR; (vii) UCHE; (viii) UGOM; (ix) UGUM; (x) ULOG; (xi) UGHE; (xii) UGHX; (xiii) UWEE; (xiv) UWEX; and (xv) VASI distributions. As mentioned, the addition of the UWMO distribution into the `unitquantreg` package is in progress. The 15 models stated in the `R` package named `unitquantreg` follow the standard nomenclature of the `R` software, employing the prefixes `r`, `p`, `d`, and `q` for the random generation, CDF, PDF, and QF, respectively. In the case of a distribution with the unit-Weibull structure, we utilize:

```
duweibull(x, mu, theta, tau = 0.5, log = FALSE)
puweibull(q, mu, theta, tau = 0.5, lower.tail = TRUE, log.p = FALSE)
quweibull(p, mu, theta, tau = 0.5, lower.tail = TRUE, log.p = FALSE)
ruweibull(n, mu, theta, tau = 0.5)
```

The functions `d`, `p`, `q`, and `r` employ `C++` under the `Rcpp` package [41,42]. The `unitquantreg()` function of the `unitquatreg` package operates such as the frameworks `simplexreg()`, `lm()`, `cdfquantreg()`, `betareg()`, `gamlss()`, and `glm()`, following:

```
unitquantreg(formula, data, subset, na.action, tau, family,
            link = c(''logit'', ''probit'', ''cloglog'', ''cauchit''),
            link.theta = c(''log'', ''sqrt'', ''identity''), start = NULL,
            control = unitquantreg.control(),
            model = TRUE, x = FALSE, y = TRUE)
```

Similarly, for example, in the R package named `simplexreg` [43], a regression structure may be specified with the `formula` package of R [44]. Hence, to state formulations for both quantile and shape structures, one may define `formula = y ~ x1 + x2 | z1 + z2`, with `y ~ x1 + x2` establishing the QR, where the explanatory variables `z1` and `z2` are associated with the shape parameter. Information of the generic functions `coef`, `print`, `plot` and `summary` can be extracted from the function `unitquantreg`, that is, from

```
> methods(class = ''unitquantreg'')
 [1] coef confint fitted hnp logLik model.frame model.matrix
 [8] plot predict print residuals summary terms update vcov
```

The QR parameters may be estimated by the ML method employing the R package named `optimx` [45], an optimization function that calls other tools for optimizing in R. Note that `optimx` unifies such tools, permitting some optimization techniques and sanity checks to be considered. The command `unitquantreg.control` enables us to model the process, with its default values being stated as:

method	hessian	gradient	maxit	factr	trace	dowarn	starttests	fnscale
''L-BFGS-B''	''FALSE''	''TRUE''	''5000''	''1e+07''	''0''	''FALSE''	''FALSE''	''1''

with the two most relevant elements being: `hessian` and `gradient`, that allow for controlling whether the `optimx` must employ the analytical Hessian and score elements, respectively, or not. For the 15 distributions that are available in the `unitquantreg` package, the Hessian and score are programmed in C++ to generate accurate estimates and obtain a good performance computationally. Initial values for β, the vector of regression parameters, may be provided by the user. Otherwise, these values for β can be generated from the QR, with the response variable being stated by $y_i^* = g_1(y_i)$, with g_1 being the structure that links with μ, that is, with the quantile parameter. The `quantreg` package [46] has a function named `rq` that is used to generate the initial values. In the case of δ_j, for $j \in \{1, \ldots, q\}$, these values are fixed at 0.1 when assuming a logarithmic link, and at the value 1.1 for links related to the inverse and square root functions. We obtain the SEs by means of the observed version of the Fisher matrix. We employ a Cholesky decomposition to reach the numerical stability of the inverse Hessian matrix.

Note that `uniquantreg` verifies whether the optimization algorithm converges or not. In this last case, a message as "optimization failed to converge" warns the user and then she/he should be cautious about the obtained results. Moreover, conditions as Karush–Kuhn–Tucker must be verified by `optimx::optimx` function. If these conditions are not met, a message is similarly generated and then the user must once gain be cautious about the reached results. The `quantreg` package also gives the Moore-Penrose inverse when the Hessian matrix is full rank, and as it is well-known, such a matrix has one negative eigenvalue, at least. We assume a distributional structure for the dependent (response) variable by means of an argument named `family`, and the following labels are utilized for the members of this family:

```
families <- c(''arc-secant hyperbolic Weibull'' = ''ashw'',
              ''Johnson-SB'' = ''johnsonsb'',
              ''Kumaraswamy'' = ''kum'',
              ''log-extended exponential-geometric'' = ''leeg'',
              ''unit-Birnbaum-Saunders'' = ''ubs'',
              ''unit-Burr-XII'' = ''uburrxii'',
              ''unit-Chen'' = ''uchen'',
              ''unit-generalized half-normal-E'' = ''ughne'',
              ''unit-generalized half-normal-X'' = ''ughnx'',
              ''unit-Gompertz'' = ''ugompertz'',
              ''unit-Gumbel'' = ''ugumbel'',
              ''unit-logistic'' = ''ulogistic'',
              ''unit-Weibull'' = ''uweibull'')
```

3.2. Simulation Analyses

To detect performance, our Monte-Carlo study uses empirical values of the root-mean-square error (RMSE) and bias of the ML estimators for the UWMO QR coefficients, $\boldsymbol{\beta}$, and for the parameters ς and ν of the UWMOQ distribution. Also, we calculate the empirical coverage probability (CP) of the 95% CI (CP$_{95\%}$) following the asymptotic normality of the ML estimators.

We consider sample sizes $n \in \{50, 70, 90, 110, 130, 150\}$; $\tau \in \{0.25, 0.50, 0.75\}$; $\varsigma \in \{0.50, 1.50, 2.00\}$, and $\nu \in \{0.50, 1.00, 2.00\}$ inserted on a regression framework stated as $\text{logit}(\mu_i) = \beta_0 + \beta_1 z_{i1}$, for $\beta_0 = 0.50$, $\beta_1 = 0.50$, and $z_{i1} \sim \text{Bernoulli}(p = 0.5)$.

We report that the time spent in the simulation study is less than 30 min. Tables 2–10 indicate the results of the Monte-Carlo experiments for our QR model. According to these results, when estimating $\boldsymbol{\beta}$, ς, and ν, we report a low bias. The value of RMSE is also low and rapidly decreases as the sample size increases. Note that, for each parameter, the CPs approach the nominal level of confidence as n increases. Hence, the results obtained from the simulation study confirm the desirable properties of the ML estimators.

In addition, we report that the results of simulation studies involving two or more covariates are similar to studies with a single covariate. Thus, we have omitted them due to restrictions of space.

Table 2. Empirical RMSE, bias, and 95% CP (true values: $\beta_0 = 0.50$, $\beta_1 = 0.50$, $\varsigma = 0.50$, and $\nu = 0.50$) of the ML estimator for the indicated parameter with simulated data.

		Bias				RMSE				CP$_{95\%}$			
τ	n	β_0	β_1	ς	ν	β_0	β_1	ς	ν	β_0	β_1	ς	ν
	50	0.0638	−0.0043	0.0903	0.0081	0.6356	0.7791	0.5878	0.1281	93.87	95.00	80.65	91.42
	70	0.0478	0.0060	0.1068	0.0039	0.5499	0.6759	0.5684	0.1104	94.30	94.49	83.18	92.54
0.25	90	0.0396	0.0118	0.1131	0.0028	0.5010	0.6077	0.5505	0.0990	94.14	94.34	85.07	93.08
	110	0.0357	0.0105	0.1066	0.0020	0.4427	0.5492	0.5203	0.0898	94.38	94.18	86.17	93.54
	130	0.0305	0.0089	0.1010	0.0017	0.3957	0.4966	0.4920	0.0834	94.42	94.46	87.29	93.46
	150	0.0249	0.0078	0.0978	0.0006	0.3685	0.4619	0.4744	0.0781	94.62	94.94	88.11	94.10
	50	−0.0020	−0.0025	0.1589	0.0333	0.6048	0.7770	0.6468	0.1702	94.16	94.28	81.11	87.76
	70	−0.0006	0.0069	0.1607	0.0201	0.5259	0.6752	0.6222	0.1489	94.08	94.01	82.87	88.64
0.50	90	0.0006	0.0104	0.1643	0.0138	0.4798	0.6037	0.6095	0.1353	94.02	94.12	83.90	89.06
	110	0.0028	0.0099	0.1587	0.0107	0.4249	0.5469	0.5868	0.1243	94.20	93.94	85.47	89.56
	130	0.0022	0.0085	0.1496	0.0090	0.3806	0.4945	0.5544	0.1161	94.42	94.14	86.22	90.28
	150	0.0008	0.0070	0.1401	0.0061	0.3549	0.4596	0.5327	0.1086	94.50	94.70	87.44	91.42
	50	−0.0872	−0.0173	0.2397	0.1019	0.6291	0.7549	0.7244	0.2487	94.12	94.67	82.90	85.40
	70	−0.0655	−0.0074	0.2266	0.0738	0.5453	0.6581	0.7018	0.2114	93.76	94.43	82.38	85.90
0.75	90	−0.0528	−0.0013	0.2236	0.0574	0.4950	0.5902	0.6955	0.1887	93.98	94.28	82.82	86.64
	110	−0.0444	−0.0002	0.2174	0.0449	0.4399	0.5360	0.6791	0.1730	94.24	94.02	82.95	86.00
	130	−0.0382	−0.0010	0.2247	0.0403	0.3964	0.4853	0.6680	0.1614	94.34	94.34	84.52	87.00
	150	−0.0348	−0.0018	0.2190	0.0317	0.3717	0.4520	0.6615	0.1508	93.78	94.72	84.27	87.40

Table 3. Empirical RMSE, bias, and 95% CP (true values: $\beta_0 = 0.50$, $\beta_1 = 0.50$, $\varsigma = 0.50$, and $\nu = 1.00$) of the ML estimator for the indicated parameter with simulated data.

		Bias				RMSE				CP$_{95\%}$			
τ	n	β_0	β_1	ς	ν	β_0	β_1	ς	ν	β_0	β_1	ς	ν
0.25	50	0.0420	0.0012	0.0552	0.0113	0.3212	0.4135	0.3934	0.2461	92.82	93.98	89.02	92.04
	70	0.0305	0.0005	0.0447	0.0058	0.2777	0.3527	0.3429	0.2083	93.72	93.98	90.84	93.22
	90	0.0250	0.0020	0.0382	0.0041	0.2522	0.3144	0.3079	0.1864	93.68	94.12	91.85	93.64
	110	0.0217	0.0022	0.0328	0.0041	0.2231	0.2850	0.2797	0.1660	93.96	94.06	92.84	94.00
	130	0.0183	0.0017	0.0293	0.0033	0.1995	0.2576	0.2594	0.1553	94.48	94.42	93.46	94.22
	150	0.0150	0.0012	0.0233	0.0023	0.1858	0.2399	0.2350	0.1420	94.32	94.88	93.60	94.98
0.50	50	0.0061	−0.0005	0.0977	0.0522	0.3031	0.4097	0.4264	0.3212	93.38	93.46	90.11	88.68
	70	0.0045	−0.0010	0.0724	0.0309	0.2642	0.3500	0.3703	0.2778	93.54	93.66	91.17	89.56
	90	0.0041	0.0004	0.0557	0.0197	0.2409	0.3113	0.3318	0.2546	93.56	93.86	91.57	89.90
	110	0.0042	0.0013	0.0460	0.0147	0.2138	0.2828	0.3027	0.2338	93.70	93.72	92.12	90.50
	130	0.0033	0.0010	0.0401	0.0132	0.1916	0.2559	0.2790	0.2179	94.12	94.16	92.92	90.96
	150	0.0022	0.0006	0.0311	0.0089	0.1788	0.2380	0.2549	0.2028	94.06	94.60	93.16	92.14
0.75	50	−0.0327	−0.0177	0.2026	0.1701	0.3126	0.3983	0.5293	0.4552	93.02	93.52	89.28	86.08
	70	−0.0249	−0.0158	0.1542	0.1181	0.2726	0.3422	0.4644	0.3885	93.04	94.00	87.74	86.12
	90	−0.0199	−0.0124	0.1250	0.0906	0.2479	0.3049	0.4187	0.3477	93.34	94.20	87.03	87.10
	110	−0.0169	−0.0099	0.1012	0.0698	0.2213	0.2780	0.3837	0.3190	93.94	93.56	86.85	86.92
	130	−0.0149	−0.0092	0.0904	0.0621	0.1997	0.2525	0.3568	0.2992	93.74	94.10	86.83	87.34
	150	−0.0136	−0.0086	0.0726	0.0484	0.1870	0.2350	0.3295	0.2795	93.30	94.56	87.25	87.90

Table 4. Empirical RMSE, bias, and 95% CP (true values: $\beta_0 = 0.50$, $\beta_1 = 0.50$, $\varsigma = 0.50$, and $\nu = 2.00$) of the ML estimator for the indicated parameter with simulated data.

		Bias				RMSE				CP$_{95\%}$			
τ	n	β_0	β_1	ς	ν	β_0	β_1	ς	ν	β_0	β_1	ς	ν
0.25	50	0.0236	−0.0052	−0.0046	0.0182	0.1646	0.2210	0.1904	0.4499	92.26	93.74	93.46	93.50
	70	0.0172	−0.0046	−0.0037	0.0105	0.1423	0.1889	0.1620	0.3730	93.28	93.96	94.72	94.30
	90	0.0141	−0.0030	−0.0023	0.0092	0.1291	0.1681	0.1444	0.3287	93.60	93.88	95.44	94.58
	110	0.0121	−0.0021	−0.0014	0.0088	0.1142	0.1527	0.1307	0.2928	93.74	94.26	95.78	94.90
	130	0.0101	−0.0019	−0.0005	0.0086	0.1020	0.1383	0.1208	0.2713	94.54	94.32	96.06	94.96
	150	0.0085	−0.0020	−0.0012	0.0062	0.0951	0.1289	0.1108	0.2485	94.10	94.78	95.94	95.34
0.50	50	0.0035	−0.0046	0.0047	0.0629	0.1545	0.2180	0.1970	0.5780	93.02	93.56	94.34	89.83
	70	0.0024	−0.0041	0.0000	0.0350	0.1348	0.1861	0.1716	0.4984	93.18	93.76	95.12	90.66
	90	0.0021	−0.0026	−0.0019	0.0222	0.1229	0.1651	0.1552	0.4493	93.68	94.08	95.36	91.60
	110	0.0020	−0.0016	−0.0016	0.0186	0.1092	0.1504	0.1418	0.4098	93.68	93.80	95.48	92.38
	130	0.0016	−0.0014	−0.0003	0.0191	0.0979	0.1365	0.1307	0.3785	93.78	94.18	95.42	92.66
	150	0.0010	−0.0014	−0.0018	0.0124	0.0913	0.1268	0.1215	0.3535	94.04	94.54	95.58	93.50
0.75	50	−0.0139	−0.0186	0.0368	0.1896	0.1604	0.2160	0.2182	0.7219	92.30	93.38	89.30	88.70
	70	−0.0106	−0.0156	0.0199	0.1352	0.1401	0.1854	0.1954	0.6536	92.84	93.72	88.64	87.92
	90	−0.0085	−0.0125	0.0115	0.1006	0.1276	0.1648	0.1796	0.5981	93.16	94.18	89.18	87.89
	110	−0.0071	−0.0102	0.0074	0.0808	0.1139	0.1508	0.1668	0.5533	93.70	93.68	88.98	88.46
	130	−0.0065	−0.0095	0.0048	0.0678	0.1028	0.1373	0.1599	0.5273	93.76	94.02	88.68	88.08
	150	−0.0060	−0.0084	0.0015	0.0508	0.0963	0.1276	0.1498	0.4917	93.22	94.40	89.74	88.84

Table 5. Empirical RMSE, bias, and 95% CP (true values: $\beta_0 = 0.50$, $\beta_1 = 0.50$, $\varsigma = 1.50$, and $\nu = 0.50$) of the ML estimator for the indicated parameter with simulated data.

		Bias				RMSE				CP$_{95\%}$			
τ	n	β_0	β_1	ς	ν	β_0	β_1	ς	ν	β_0	β_1	ς	ν
	50	0.0380	−0.0356	−0.4592	0.0023	0.7550	0.8812	0.9959	0.1077	94.46	96.02	72.59	93.82
	70	0.0288	−0.0075	−0.3521	0.0007	0.6545	0.7760	0.9550	0.0910	94.35	95.32	76.19	94.62
0.25	90	0.0243	0.0062	−0.2902	0.0006	0.5970	0.7030	0.9189	0.0807	94.62	94.80	78.18	94.68
	110	0.0232	0.0091	−0.2407	0.0005	0.5275	0.6416	0.8944	0.0722	94.38	94.31	80.00	94.90
	130	0.0198	0.0077	−0.1993	0.0010	0.4727	0.5806	0.8651	0.0648	94.42	94.60	82.10	95.28
	150	0.0149	0.0078	−0.1592	0.0006	0.4404	0.5427	0.8517	0.0596	94.58	94.98	82.90	95.38
	50	−0.0486	−0.0287	−0.3356	0.0133	0.7303	0.8918	1.0055	0.1390	94.77	95.85	77.73	90.32
	70	−0.0350	−0.0017	−0.2630	0.0059	0.6350	0.7837	0.9731	0.1224	94.52	95.00	80.15	91.08
0.50	90	−0.0283	0.0123	−0.1943	0.0046	0.5792	0.7095	0.9485	0.1096	94.40	94.56	82.56	91.98
	110	−0.0214	0.0147	−0.1578	0.0035	0.5121	0.6463	0.9123	0.0993	94.36	94.11	84.45	92.60
	130	−0.0185	0.0130	−0.1156	0.0027	0.4582	0.5854	0.8938	0.0928	94.56	94.44	85.38	92.74
	150	−0.0178	0.0119	−0.0791	0.0013	0.4273	0.5457	0.8837	0.0865	94.68	94.82	86.98	93.68
	50	−0.1576	−0.0327	−0.3069	0.0500	0.7760	0.8854	1.0537	0.1889	95.25	96.08	76.61	87.18
	70	−0.1191	−0.0066	−0.2609	0.0323	0.6698	0.7783	1.0367	0.1640	94.42	95.38	77.88	87.58
0.75	90	−0.0976	0.0073	−0.2285	0.0229	0.6072	0.7026	1.0191	0.1485	94.34	94.88	79.49	88.02
	110	−0.0810	0.0118	−0.1730	0.0179	0.5371	0.6428	1.0023	0.1370	94.38	94.19	80.98	88.24
	130	−0.0712	0.0102	−0.1314	0.0147	0.4821	0.5823	0.9917	0.1287	94.66	94.54	82.26	88.70
	150	−0.0636	0.0097	−0.1057	0.0103	0.4511	0.5438	0.9790	0.1207	94.18	94.80	83.77	89.68

Table 6. Empirical RMSE, bias, and 95% CP (true values: $\beta_0 = 0.50$, $\beta_1 = 0.50$, $\varsigma = 1.50$, and $\nu = 1.00$) of the ML estimator for the indicated parameter with simulated data.

		Bias				RMSE				CP$_{95\%}$			
τ	n	β_0	β_1	ς	ν	β_0	β_1	ς	ν	β_0	β_1	ς	ν
	50	0.0239	−0.0054	−0.2572	−0.0012	0.4183	0.5146	0.8332	0.1835	93.10	93.96	77.75	95.36
	70	0.0153	−0.0036	−0.1758	−0.0022	0.3631	0.4422	0.7859	0.1512	93.42	93.94	80.87	96.06
0.25	90	0.0120	−0.0003	−0.1236	0.0000	0.3303	0.3959	0.7510	0.1287	93.90	93.98	82.17	95.60
	110	0.0110	0.0002	−0.0921	−0.0002	0.2934	0.3589	0.7225	0.1155	93.84	93.92	83.62	95.70
	130	0.0091	−0.0004	−0.0656	−0.0002	0.2633	0.3243	0.6940	0.1076	94.16	94.40	85.41	95.66
	150	0.0067	−0.0006	−0.0510	−0.0008	0.2457	0.3022	0.6726	0.0988	94.10	94.76	85.77	95.44
	50	−0.0285	0.0084	−0.0992	0.0070	0.4057	0.5268	0.8121	0.2305	93.64	93.56	87.09	92.90
	70	−0.0233	0.0073	−0.0428	0.0033	0.3529	0.4508	0.7612	0.1944	93.72	93.78	89.54	93.94
0.50	90	−0.0193	0.0088	−0.0169	0.0023	0.3217	0.4019	0.7193	0.1736	93.76	93.88	90.60	94.10
	110	−0.0153	0.0083	0.0033	0.0026	0.2852	0.3639	0.6865	0.1537	93.78	93.68	91.70	94.64
	130	−0.0133	0.0067	0.0082	0.0021	0.2554	0.3285	0.6454	0.1437	94.14	94.12	92.01	94.66
	150	−0.0126	0.0058	0.0215	0.0021	0.2384	0.3056	0.6245	0.1285	94.36	94.68	92.28	95.12
	50	−0.0803	0.0102	−0.0126	0.0367	0.4320	0.5312	0.8604	0.2994	93.38	93.66	89.44	89.34
	70	−0.0625	0.0089	0.0196	0.0198	0.3736	0.4548	0.8164	0.2604	93.46	94.02	90.50	90.32
0.75	90	−0.0517	0.0104	0.0343	0.0124	0.3383	0.4049	0.7821	0.2376	93.72	94.12	91.62	90.74
	110	−0.0426	0.0101	0.0381	0.0102	0.2994	0.3672	0.7340	0.2151	93.92	93.84	92.62	91.52
	130	−0.0374	0.0085	0.0436	0.0075	0.2694	0.3318	0.7047	0.2026	94.28	94.14	93.45	91.82
	150	−0.0326	0.0074	0.0415	0.0051	0.2514	0.3082	0.6658	0.1874	93.70	94.62	94.02	92.90

Table 7. Empirical RMSE, bias, and 95% CP (true values: $\beta_0 = 0.50$, $\beta_1 = 0.50$, $\varsigma = 1.50$, and $\nu = 2.00$) of the ML estimator for the indicated parameter with simulated data.

		Bias				RMSE				CP$_{95\%}$			
τ	n	β_0	β_1	ς	ν	β_0	β_1	ς	ν	β_0	β_1	ς	ν
	50	0.0073	−0.0096	−0.2878	−0.0130	0.2325	0.2838	0.7168	0.3056	92.62	93.70	69.63	95.84
	70	0.0036	−0.0069	−0.2310	−0.0131	0.2018	0.2437	0.6651	0.2517	93.40	93.76	72.48	96.14
0.25	90	0.0027	−0.0041	−0.1836	−0.0092	0.1835	0.2182	0.6327	0.2179	93.62	93.82	74.76	95.36
	110	0.0024	−0.0029	−0.1515	−0.0077	0.1630	0.1977	0.6024	0.1944	93.46	93.94	76.40	95.56
	130	0.0016	−0.0026	−0.1236	−0.0063	0.1463	0.1785	0.5773	0.1761	94.22	94.34	78.69	95.40
	150	0.0006	−0.0024	−0.1089	−0.0069	0.1367	0.1664	0.5564	0.1632	93.90	94.74	78.94	95.24
	50	−0.0227	0.0099	−0.1173	−0.0021	0.2268	0.2931	0.6122	0.3425	93.56	93.62	82.63	95.74
	70	−0.0182	0.0083	−0.0875	−0.0040	0.1969	0.2500	0.5490	0.2800	93.74	93.86	84.97	96.16
0.50	90	−0.0149	0.0084	−0.0594	−0.0008	0.1792	0.2229	0.5139	0.2408	93.90	93.98	85.52	95.50
	110	−0.0119	0.0076	−0.0439	−0.0006	0.1584	0.2015	0.4828	0.2151	94.04	93.80	86.84	95.64
	130	−0.0103	0.0063	−0.0355	−0.0006	0.1417	0.1817	0.4498	0.1999	94.20	94.26	88.11	95.60
	150	−0.0094	0.0056	−0.0309	−0.0021	0.1323	0.1689	0.4287	0.1848	94.16	94.72	88.17	95.32
	50	−0.0423	0.0203	−0.0396	0.0083	0.2439	0.3020	0.5576	0.4262	93.24	93.64	91.95	94.06
	70	−0.0317	0.0163	−0.0234	0.0043	0.2099	0.2566	0.4894	0.3511	93.60	93.90	93.06	94.82
0.75	90	−0.0256	0.0150	−0.0137	0.0051	0.1898	0.2281	0.4429	0.3067	93.86	93.98	94.07	94.86
	110	−0.0207	0.0132	−0.0110	0.0038	0.1675	0.2060	0.4051	0.2771	93.74	93.90	94.35	95.00
	130	−0.0176	0.0111	−0.0053	0.0053	0.1501	0.1858	0.3747	0.2522	94.40	94.20	94.64	95.46
	150	−0.0153	0.0098	−0.0062	0.0035	0.1402	0.1722	0.3466	0.2308	94.02	94.74	94.78	95.44

Table 8. Empirical RMSE, bias, and 95% CP (true values: $\beta_0 = 0.50$, $\beta_1 = 0.50$, $\varsigma = 2.00$, and $\nu = 0.50$) of the ML estimator for the indicated parameter with simulated data.

		Bias				RMSE				CP$_{95\%}$			
τ	n	β_0	β_1	ς	ν	β_0	β_1	ς	ν	β_0	β_1	ς	ν
	50	0.0289	−0.0453	−0.7745	0.0012	0.7831	0.9007	1.2837	0.1043	94.66	96.14	69.26	94.16
	70	0.0216	−0.0120	−0.6466	0.0001	0.6784	0.7955	1.2059	0.0873	94.39	95.57	73.20	94.84
0.25	90	0.0185	0.0055	−0.5457	0.0005	0.6187	0.7236	1.1584	0.0764	94.72	94.88	75.44	95.08
	110	0.0185	0.0087	−0.4864	0.0003	0.5465	0.6605	1.1138	0.0685	94.34	94.39	77.55	95.08
	130	0.0160	0.0083	−0.4257	0.0008	0.4898	0.5992	1.0700	0.0616	94.46	94.62	80.02	95.46
	150	0.0116	0.0084	−0.3798	0.0003	0.4564	0.5601	1.0418	0.0572	94.60	95.10	80.43	95.42
	50	−0.0616	−0.0355	−0.6168	0.0100	0.7606	0.9165	1.2559	0.1333	95.05	96.18	76.07	90.82
	70	−0.0444	−0.0062	−0.5178	0.0044	0.6605	0.8047	1.1971	0.1165	94.62	95.32	78.79	91.94
0.50	90	−0.0361	0.0129	−0.4164	0.0036	0.6020	0.7324	1.1593	0.1040	94.50	94.72	81.55	92.58
	110	−0.0278	0.0154	−0.3561	0.0028	0.5318	0.6668	1.1086	0.0938	94.50	94.23	83.58	93.10
	130	−0.0239	0.0144	−0.3036	0.0020	0.4755	0.6051	1.0754	0.0879	94.70	94.48	84.79	93.16
	150	−0.0225	0.0135	−0.2668	0.0009	0.4433	0.5645	1.0484	0.0820	94.74	94.84	85.82	93.94
	50	−0.1752	−0.0398	−0.6059	0.0407	0.8120	0.9129	1.3070	0.1786	95.69	96.28	75.07	87.68
	70	−0.1311	−0.0077	−0.5413	0.0261	0.6985	0.8042	1.2682	0.1555	94.72	95.79	77.11	88.30
0.75	90	−0.1070	0.0098	−0.4848	0.0182	0.6337	0.7281	1.2365	0.1409	94.52	95.09	78.79	88.74
	110	−0.0884	0.0136	−0.3955	0.0142	0.5595	0.6646	1.2127	0.1297	94.62	94.47	80.64	89.08
	130	−0.0778	0.0135	−0.3578	0.0110	0.5015	0.6042	1.1859	0.1227	94.80	94.66	81.67	89.42
	150	−0.0688	0.0127	−0.3159	0.0079	0.4685	0.5640	1.1576	0.1144	94.52	94.88	83.39	90.52

Table 9. Empirical RMSE, bias, and 95% CP (true values: $\beta_0 = 0.50$, $\beta_1 = 0.50$, $\varsigma = 2.00$, and $\nu = 1.00$) of the ML estimator for the indicated parameter with simulated data.

		Bias				RMSE				CP$_{95\%}$			
τ	n	β_0	β_1	ς	ν	β_0	β_1	ς	ν	β_0	β_1	ς	ν
	50	0.0169	−0.0049	−0.5311	−0.0028	0.4356	0.5311	1.0717	0.1741	93.22	93.98	72.97	95.62
	70	0.0096	−0.0026	−0.4125	−0.0034	0.3776	0.4569	1.0000	0.1429	93.46	93.98	76.42	96.16
0.25	90	0.0074	0.0008	−0.3300	−0.0013	0.3436	0.4093	0.9493	0.1224	94.02	94.02	78.52	95.62
	110	0.0071	0.0011	−0.2836	−0.0012	0.3051	0.3709	0.9045	0.1096	93.68	93.98	80.08	95.62
	130	0.0057	0.0004	−0.2272	−0.0009	0.2739	0.3350	0.8730	0.1010	94.24	94.40	82.10	95.52
	150	0.0036	0.0001	−0.2041	−0.0014	0.2556	0.3123	0.8397	0.0928	94.20	94.80	82.56	95.40
	50	−0.0364	0.0119	−0.2897	0.0041	0.4242	0.5473	0.9877	0.2140	93.86	93.64	84.65	94.04
	70	−0.0294	0.0103	−0.1996	0.0013	0.3683	0.4679	0.9169	0.1795	93.84	93.88	87.43	94.76
0.50	90	−0.0243	0.0116	−0.1399	0.0017	0.3354	0.4172	0.8697	0.1572	93.96	94.06	88.52	94.68
	110	−0.0193	0.0105	−0.1033	0.0010	0.2969	0.3773	0.8324	0.1422	94.08	93.78	89.44	94.92
	130	−0.0167	0.0086	−0.0670	0.0021	0.2658	0.3404	0.7947	0.1281	94.34	94.20	90.35	95.34
	150	−0.0154	0.0074	−0.0553	0.0011	0.2481	0.3166	0.7600	0.1181	94.38	94.68	90.49	95.56
	50	−0.0865	0.0177	−0.1758	0.0249	0.4541	0.5560	1.0098	0.2748	93.62	93.78	88.66	90.54
	70	−0.0668	0.0152	−0.1061	0.0117	0.3916	0.4751	0.9578	0.2395	93.58	94.12	90.27	91.44
0.75	90	−0.0544	0.0158	−0.0698	0.0090	0.3540	0.4226	0.9015	0.2143	93.96	94.16	91.88	92.20
	110	−0.0447	0.0145	−0.0337	0.0065	0.3127	0.3823	0.8656	0.1953	93.98	93.84	92.52	92.62
	130	−0.0389	0.0123	−0.0154	0.0051	0.2806	0.3451	0.8285	0.1827	94.46	94.26	93.77	92.78
	150	−0.0339	0.0107	−0.0030	0.0030	0.2619	0.3203	0.7931	0.1692	94.10	94.64	94.29	93.78

Table 10. Empirical RMSE, bias, and 95% CP (true values: $\beta_0 = 0.50$, $\beta_1 = 0.50$, $\varsigma = 2.00$, and $\nu = 2.00$) of the ML estimator for the indicated parameter with simulated data.

		Bias				RMSE				CP$_{95\%}$			
τ	n	β_0	β_1	ς	ν	β_0	β_1	ς	ν	β_0	β_1	ς	ν
	50	0.0034	−0.0084	−0.6018	−0.0170	0.2387	0.2891	1.0116	0.2983	92.78	93.74	61.57	95.40
	70	0.0004	−0.0058	−0.5062	−0.0169	0.2070	0.2482	0.9468	0.2443	93.48	93.76	64.29	95.84
0.25	90	−0.0001	−0.0031	−0.4392	−0.0124	0.1882	0.2223	0.8935	0.2109	93.80	93.80	67.41	95.52
	110	−0.0000	−0.0020	−0.3802	−0.0104	0.1670	0.2013	0.8523	0.1858	93.60	94.02	69.84	95.60
	130	−0.0005	−0.0018	−0.3281	−0.0092	0.1499	0.1817	0.8198	0.1701	94.30	94.30	71.35	95.24
	150	−0.0013	−0.0016	−0.2976	−0.0094	0.1400	0.1693	0.7922	0.1575	94.00	94.82	72.83	95.32
	50	−0.0257	0.0124	−0.3116	−0.0083	0.2331	0.2995	0.8357	0.3178	93.72	93.76	75.51	95.88
	70	−0.0204	0.0102	−0.2419	−0.0096	0.2020	0.2552	0.7673	0.2630	93.84	93.92	78.46	96.12
0.50	90	−0.0166	0.0101	−0.1884	−0.0056	0.1836	0.2275	0.7207	0.2259	94.14	94.06	79.48	95.70
	110	−0.0132	0.0089	−0.1523	−0.0047	0.1621	0.2054	0.6822	0.2023	94.30	93.88	81.48	95.66
	130	−0.0114	0.0074	−0.1222	−0.0037	0.1449	0.1851	0.6500	0.1840	94.34	94.30	82.84	95.42
	150	−0.0104	0.0065	−0.1034	−0.0048	0.1353	0.1721	0.6291	0.1706	94.28	94.76	83.36	95.36
	50	−0.0412	0.0239	−0.1298	0.0013	0.2505	0.3089	0.7296	0.3772	93.70	93.86	87.93	95.06
	70	−0.0307	0.0188	−0.0890	−0.0012	0.2154	0.2618	0.6529	0.3100	93.68	94.00	88.90	95.78
0.75	90	−0.0246	0.0168	−0.0590	0.0026	0.1945	0.2324	0.5991	0.2637	94.00	94.06	90.12	95.48
	110	−0.0196	0.0145	−0.0412	0.0022	0.1714	0.2093	0.5622	0.2360	93.84	93.88	90.62	95.46
	130	−0.0168	0.0121	−0.0303	0.0013	0.1537	0.1886	0.5300	0.2227	94.48	94.16	91.40	95.56
	150	−0.0146	0.0106	−0.0264	−0.0000	0.1436	0.1749	0.5003	0.2036	94.12	94.74	91.16	95.38

3.3. Application to Sports Medicine

Continuing with the motivating example presented in Section 1.2, we utilize several QR models and detect the relationships among the explanatory variables and the fat percentage at the legs. Let us assume that, for a constant quantile value $\tau = 0.5$, the relationship of μ, ς, ν, and the covariates is provided by

$$\begin{aligned}
\mathrm{logit}(\mu_i) &= \beta_0 + \beta_1\,\mathrm{bmi}_i + \beta_2\,\mathrm{age}_i + \beta_3\,\mathrm{sexmale}_i + \beta_4\,\mathrm{ipaqinsufficientlyactive}_i + \beta_5\mathrm{ipaqactive}_i; \\
\log(\varsigma_i) &= \delta_0; \quad \log(\nu_i) = \gamma_0;
\end{aligned}$$

where "age_i = (age_i − 46.00)" with 46.00 being the average age; "bmi_i = (bmi_i − 24.72)" with 24.72 being the average BMI; "sexmale = 1" is employed for "sex = male" and "sexmale = 0" for "sex = female"; "ipaqinsufficientlyactive = 1" is used for "ipaq = insufficiently active" and "ipaqinsufficientlyactive = 0" for "ipaq = sedentary", while "ipaqactive = 1" is utilized for "ipaq = active", and "ipaqactive = 0" for "ipaq = sedentary".

As mentioned, we estimate the model parameters with a SAS macro, which is restricted to analyzing data with logit link for μ and log link for ς and ν. Table 11 reports the SEs, 95% CIs, and ML estimates of the corresponding parameters. Tables 12–14 presents values of likelihood-based statistics for model selection considering different QR models. Such statistics correspond to values of the Akaike (AIC) and (BIC) Bayesian information criteria and log-likelihood (LL). Note that smaller values of AIC, BIC, and LL indicate a better concordance between a model and the data set. These statistics report that the UWMOQ model has the best performance to describe the fat percentage at the legs with covariates, as expected, due to the mentioned flexibility of the Weibull–Marshall–Olkin distribution. Such conclusions are supported by the probability plots of Figure 5.

Table 11. ML estimates, SEs, and 95% CI for the indicated τ and parameter with data related to fat percentage at the legs.

				95% CI	
τ	**Parameter**	**Estimate**	**SE**	**Lower**	**Upper**
	β_0	−0.4084	0.0400	−0.4871	−0.3296
	β_1	0.0006	0.0011	−0.0015	0.0027
	β_2	0.0664	0.0062	0.0542	0.0787
0.25	β_3	−0.9202	0.0352	−0.9895	−0.8509
	β_4	−0.0459	0.0484	−0.1411	0.0494
	β_5	−0.1249	0.0454	−0.2144	−0.0355
	δ_0	1.6274	0.3308	0.9763	2.2784
	γ_0	2.3707	0.0481	2.2761	2.4654
	β_0	−0.2537	0.0370	−0.3265	−0.1809
	β_1	0.0006	0.0010	−0.0014	0.0026
	β_2	0.0637	0.0059	0.0520	0.0754
0.50	β_3	−0.8776	0.0335	−0.9435	−0.8117
	β_4	−0.0444	0.0451	−0.1331	0.0443
	β_5	−0.1201	0.0427	−0.2042	−0.0360
	δ_0	1.0004	0.2096	0.5880	1.4128
	γ_0	2.3688	0.0487	2.2730	2.4646
	β_0	−0.1047	0.0361	−0.1759	−0.0336
	β_1	0.0006	0.0010	−0.0014	0.0025
	β_2	0.0612	0.0057	0.0501	0.0724
0.75	β_3	−0.8396	0.0320	−0.9026	−0.7766
	β_4	−0.0431	0.0432	−0.1281	0.0420
	β_5	−0.1155	0.0409	−0.1960	−0.0350
	δ_0	1.0087	0.2280	0.5600	1.4573
	γ_0	2.3695	0.0486	2.2739	2.4651

Table 12. Values of the indicated likelihood-based statistic and distribution for $\tau = 0.25$ with data related to fat percentage at the legs.

Distribution	-2 LL	AIC	AICC	BIC
ASHW	−748.1069	−734.1069	−733.7207	−708.2273
JOSB	−841.6750	−827.6750	−827.2888	−801.7953
KUMA	−823.3698	−809.3698	−808.9836	−783.4902
LEEG	−816.0419	−802.0419	−801.6557	−776.1623
UBSA	−854.4446	−840.4446	−840.0584	−814.5650
UBUR	−667.9945	−653.9945	−653.6083	−628.1148
UCHE	−632.5113	−618.5113	−618.1250	−592.6316
UGHE	−702.0870	−688.0870	−687.7008	−662.2074
UGHX	−819.2782	−805.2782	−804.8920	−779.3986
UGOM	−680.6044	−666.6044	−666.2182	−640.7248
UGUM	−727.1883	−713.1883	−712.8021	−687.3086
ULOG	−849.4922	−835.4922	−835.1060	−809.6125
UWEE	−757.5552	−743.5552	−743.1689	−717.6755
UWEM	−844.5334	−830.5334	−830.1472	−804.6538
UWMO	−859.8022	−843.8022	−843.3039	−814.2254
VASI	−847.9083	−833.9083	−833.5220	−808.0286

Where AIC: Akaike information criterion; AICC: Akaike information corrected criterion; BIC: Bayesian information criterion; LL: log-likelihood. See acronyms of the distributions in Table A1.

Table 13. Values of the indicated likelihood-based statistic and distribution for $\tau = 0.50$ with data related to fat percentage at the legs.

Distribution	-2 LL	AIC	AICC	BIC
ASHW	−749.0414	−735.0414	−734.6551	−709.1617
JOSB	−841.6750	−827.6750	−827.2888	−801.7953
KUMA	−825.4268	−811.4268	−811.0406	−785.5471
LEEG	−818.5204	−804.5204	−804.1342	−778.6408
UBSA	−855.4990	−841.4990	−841.1128	−815.6194
UBUR	−672.1146	−658.1146	−657.7284	−632.2350
UCHE	−630.8162	−616.8162	−616.4300	−590.9365
UGHE	−703.1626	−689.1626	−688.7764	−663.2829
UGHX	−819.2782	−805.2782	−804.8920	−779.3986
UGOM	−684.6376	−670.6376	−670.2514	−644.7580
UGUM	−727.1883	−713.1883	−712.8021	−687.3086
ULOG	−849.4922	−835.4922	−835.1060	−809.6125
UWEE	−758.6517	−744.6517	−744.2655	−718.7720
UWEX	−844.5334	−830.5334	−830.1472	−804.6538
UWMO	−860.9117	−844.9117	−844.4134	−815.3350
VASI	−848.7302	−834.7302	−834.3440	−808.8505

Table 14. Values of the indicated likelihood-based statistic and distribution for $\tau = 0.75$ with data related to fat percentage at the legs.

Distribution	-2 LL	AIC	AICC	BIC
ASHW	−750.1398	−736.1398	−735.7535	−710.2601
JOSB	−841.6750	−827.6750	−827.2888	−801.7953
KUMA	−827.3404	−813.3404	−812.9542	−787.4607
LEEG	−821.4608	−807.4608	−807.0746	−781.5811
UBSA	−856.4866	−842.4866	−842.1004	−816.6069
UBUR	−674.3747	−660.3747	−659.9885	−634.4951
UCHE	−623.7983	−609.7983	−609.4121	−583.9187
UGHE	−704.5493	−690.5493	−690.1631	−664.6696
UGHX	−819.2782	−805.2782	−804.8920	−779.3986
UGOM	−690.5561	−676.5561	−676.1699	−650.6765

Table 14. *Cont.*

Distribution	−2 LL	AIC	AICC	BIC
UGUM	−727.1883	−713.1883	−712.8021	−687.3086
ULOG	−849.4922	−835.4922	−835.1060	−809.6125
UWEE	−759.9440	−745.9440	−745.5578	−720.0644
UWEX	−844.5334	−830.5334	−830.1472	−804.6538
UWMO	−861.7335	−845.7335	−845.2352	−816.1567
VASI	−849.5840	−835.5840	−835.1978	−809.7044
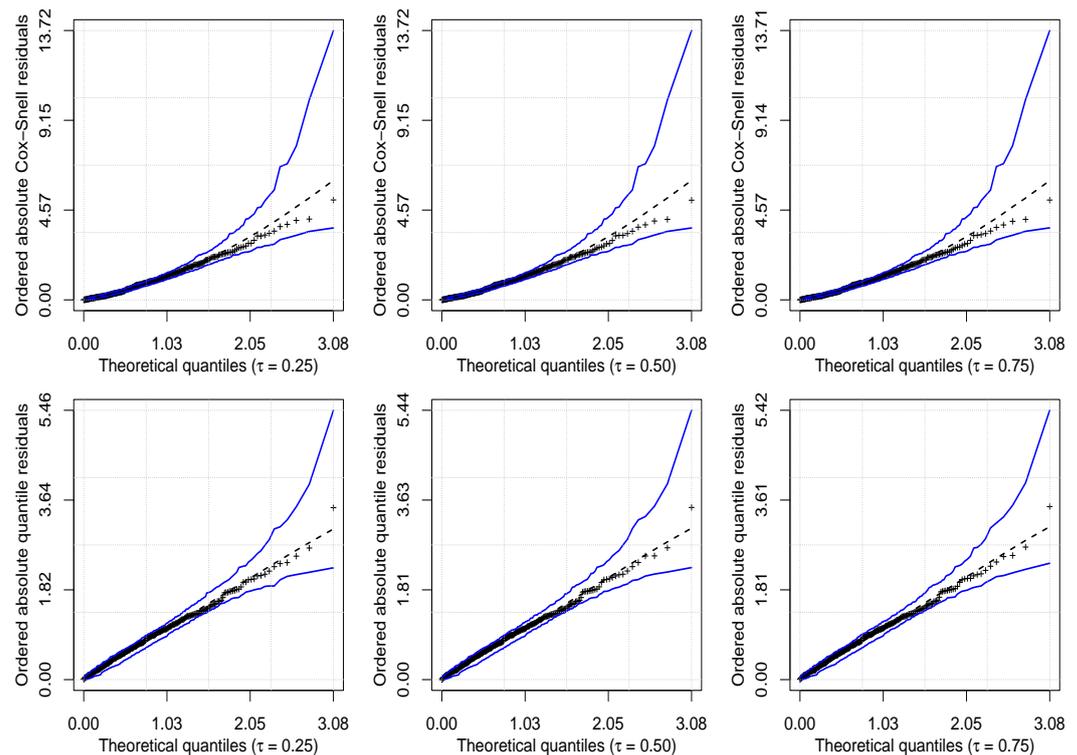
Figure 5. Probability plots with envelope of Cox-Snell (first row) and randomized quantile (second row) residuals for the specified τ quantile level with data related to fat percentage at the legs, where blue lines correspond to the envelope, "- - -" lines are the center of the envelope, and " + + + " symbols represent the empirical data.

One can use a coefficient of determination, R^2, namely, such as in ordinary regression models, to state goodness-of-fit in QR structures [4]. The coefficient of determination is a global indicator of the explained variation for the dependent (response) variable, expressed as

$$R^2 = 1 - \left(\frac{L_0}{L_{p+q}} \right)^{2/n},$$

where L_0 and L_{p+q} are the likelihood functions for statistical formulations considering only the intercept or this plus $p + q$ explanatory variables. Table 15 presents the values of the coefficient of determination R^2 for all distributions with $\tau = 0.5$ (median). Such as in the case of AIC and BIC, the coefficient of determination enables us to compare statistical formulations not necessarily nested, but in this case a larger value states a better model to describe a data set. Based on the results from this table, once again the UWMOQ model presents the best performance.

Table 15. Coefficient of determination (R^2) for the indicated distribution with data related to fat percentage at the legs.

Distribution	R^2
ASHW	0.7641
JOSB	0.7689
KUMA	0.7219
LEEG	0.7678
UBSA	0.7751
UBUR	0.5202
UCHE	0.6967
UGHE	0.7465
UGHX	0.7323
UGOM	0.7558
UGUM	0.7715
ULOG	0.7773
UWEX	0.7578
UWMOQ	0.7808

As mentioned, to detect the adequacy of the model to the data, we also employ probability plots with envelopes based on the Cox-Snell and randomized quantile residuals. Figure 5 shows the good performance of the UWMOQ model with the data set associated with body fat. Hence, we continue with our QR analysis utilizing this model, where the UBSA [47,48] and KUMA [49] models are demonstrated to be competitive.

Note that $\widehat{\beta}_0$ is the estimate of a sedentary female of 46 years old, with a BMI which equals 24.72 kg/m^2. The values of estimated β_1 and β_2 establish that the age and BMI positively affect the fat percentage at the legs. In contrast, the estimates of β_3, β_4, and β_5 are negative, indicating that insufficiently active and active men have less fat percentage at the legs.

For the other response variables available in the data set, that is, the fat percentage at the android, arms, body, and gynecoid, the likelihood-based statistics for model selection and $\tau = 0.5$ are in Tables 16–19. Note that, for the variable body and gynecoid, our proposal is the best.

Table 16. Values of the indicated likelihood-based statistic and distribution for $\tau = 0.50$ with data related to fat percentage at the android.

Distribution	−2 LL	AIC	AICC	BIC
ASHW	−656.6639	−642.6639	−642.2777	−616.7843
JOSB	−733.4400	−719.4400	−719.0538	−693.5603
KUMA	−731.5464	−717.5464	−717.1602	−691.6667
LEEG	−703.3988	−689.3988	−689.0126	−663.5192
UBUR	−804.1466	−790.1466	−789.7604	−764.2669
UBSA	−786.2243	−772.2243	−771.8381	−746.3446
UCHE	−472.4102	−458.4102	−458.0240	−432.5306
UGHE	−635.5863	−621.5863	−621.2001	−595.7066
UGHX	−757.9927	−743.9927	−743.6065	−718.1131
UGOM	−509.3101	−495.3101	−494.9239	−469.4305
UGUM	−601.3444	−587.3444	−586.9581	−561.4647
ULOG	−764.8452	−750.8452	−750.4590	−724.9656
UWEE	−689.5933	−675.5933	−675.2071	−649.7137
UWEX	−772.4956	−758.4956	−758.1094	−732.6159
UWMO	−800.0696	−784.0696	−783.5713	−754.4928
VASI	−751.1870	−737.1870	−736.8008	−711.3073

Table 17. Values of the indicated likelihood-based statistic and distribution for $\tau = 0.50$ with data related to fat percentage at the arms.

Distribution	-2 LL	AIC	AICC	BIC
ASHW	−844.2975	−830.2975	−829.9113	−804.4179
JOSB	−897.9673	−883.9673	−883.5811	−858.0876
KUMA	−859.0303	−845.0303	−844.6441	−819.1507
LEEG	−867.8135	−853.8135	−853.4273	−827.9339
UBSA	−923.3318	−909.3318	−908.9456	−883.4522
UBUR	−661.0975	−647.0975	−646.7113	−621.2179
UCHE	−705.7224	−691.7224	−691.3362	−665.8427
UGHE	−808.1356	−794.1356	−793.7494	−768.2560
UGHX	−858.1305	−844.1305	−843.7443	−818.2509
UGOM	−771.3000	−757.3000	−756.9138	−731.4203
UGUM	−814.9471	−800.9471	−800.5609	−775.0674
ULOG	−902.4015	−888.4015	−888.0153	−862.5218
UWEE	−854.9321	−840.9321	−840.5459	−815.0525
UWEX	−885.2879	−871.2879	−870.9016	−845.4082
UWMO	−923.3566	−907.3566	−906.8583	−877.7798
VASI	−910.9891	−896.9891	−896.6029	−871.1094

Table 18. Values of the indicated likelihood-based statistic and distribution for $\tau = 0.25$ with data related to fat percentage at the body.

Distribution	-2 LL	AIC	AICC	BIC
ASHW	−802.6541	−788.6541	−788.2679	−762.7745
JOSB	−859.1252	−845.1252	−844.7389	−819.2455
KUMA	−837.0105	−823.0105	−822.6243	−797.1309
LEEG	−834.1941	−820.1941	−819.8079	−794.3144
UBSA	−902.1138	−888.1138	−887.7276	−862.2342
UBUR	−864.2166	−850.2166	−849.8304	−824.3369
UCHE	−601.4599	−587.4599	−587.0736	−561.5802
UGHE	−779.5107	−765.5107	−765.1245	−739.6310
UGHX	−851.4477	−837.4477	−837.0615	−811.5680
UGOM	−685.0341	−671.0341	−670.6479	−645.1545
UGUM	−756.7462	−742.7462	−742.3599	−716.8665
ULOG	−879.6604	−865.6604	−865.2742	−839.7808
UWEX	−870.3416	−856.3416	−855.9554	−830.4619
UWEE	−827.4140	−813.4140	−813.0278	−787.5344
UWMO	−909.8991	−893.8991	−893.4008	−864.3223
VASI	−875.2501	−861.2501	−860.8639	−835.3704

Table 19. Values of the indicated likelihood-based statistic and distribution for $\tau = 0.25$ with data related to fat percentage at the gynecoid.

Distribution	-2 LL	AIC	AICC	BIC
ASHW	−782.8493	−768.8493	−768.4631	−742.9697
JOSB	−874.0575	−860.0575	−859.6713	−834.1779
KUMA	−859.6809	−845.6809	−845.2947	−819.8013
LEEG	−847.9210	−833.9210	−833.5348	−808.0414
UBSA	−893.3770	−879.3770	−878.9908	−853.4973
UBUR	−778.2260	−764.2260	−763.8398	−738.3463
UCHE	−626.8613	−612.8613	−612.4750	−586.9816
UGHE	−739.2528	−725.2528	−724.8666	−699.3731
UGHX	−861.7670	−847.7670	−847.3808	−821.8873
UGOM	−705.0838	−691.0838	−690.6976	−665.2041

Table 19. *Cont.*

Distribution	−2 LL	AIC	AICC	BIC
UGUM	−760.1826	−746.1826	−745.7964	−720.3030
ULOG	−884.1771	−870.1771	−869.7909	−844.2975
UWEE	−795.4313	−781.4313	−781.0451	−755.5516
UWEX	−883.5176	−869.5176	−869.1314	−843.6379
UWMO	−898.7473	−882.7473	−882.2490	−853.1706
VASI	−880.3849	−866.3849	−865.9986	−840.5052

3.4. COVID-19 Application with Data from the US

Our second illustration uses a data set related to COVID-19. These data can be secured from https://github.com/tatianefribeiro/RUBXII_Regression_COVID-19/tree/master (accessed on 30 January 2023). For purposes of comparison, we consider the distributions of the `unitquantreg` package. For details about the characteristics of these distributions, see Table A1 in Appendix A.

We define the response variable as: recovery rate $= 1 -$ mortality rate, throughout all 50 USA states, and the subsequent model is formulated as

$$\text{logit}(\mu_i) = \beta_0 + \beta_1 \text{PD}_i + \beta_2 \text{GINI}_i + \beta_3 \text{BEDS}_i + \beta_4 \text{SR}_i + \beta_5 \text{PR}_i + \beta_6 \text{LE}_i + \beta_7 \text{T90} + \beta_8 \text{T180}_i, \quad (11)$$

with GINI being the GINI indicator in 2017; PD relating to the population density in 2020 (people/m^2); BEDS being the number of hospital beds per 100,000 people in 2018; PR corresponding to the level of poverty in 2020; SR being the state-by-state smoking rate in 2020; LE indicating the life expectancy in 2018; T90 being a dummy variable that is set as one if the response reflects the recovery rate after 90 days for the 10th confirmed case, and it is set as zero, otherwise. In contrast, T180 also is a dummy variable that is set as one if the response reflects the recovery rate after 180 days for the 10th confirmed case; otherwise, it is zero. Additional information about the covariates related to the indicators GINI, BEDS, and LE in 2020 is available in [50], where the mortality rate, defined as mortality rate $= (1 -$ recovery rate), was considered as the response variable.

Note that, in the three time periods, the response presents a positive coefficient of skewness (CS), and the average is far from the median. At 30 and 90 days, the coefficient of kurtosis (CK) is CK > 3, informing a leptokurtic distribution. Histograms and scatter plots are shown in Figure 6 for the variables BEDS, GINI, PD, PR, RR, SR, and LE. The empirical distribution of the response is leptokurtic, asymmetric, and ranges from 0.820 to 0.999. Several models implemented in our R package work well to model the data. In addition, note that the response variable only has a statistically significant correlation at 1% with the covariates PD and GINI, allowing the model to leave out the other covariates. However, when the regression models are specified, substantial correlations between some covariates may point to multicollinearity problems.

As opposed to what occurred with the body fat data, the results acquired for the COVID-19 data set in terms of significance at 5% are somewhat different among the models. We use the measure R^2 once again as a criterion for comparing all models, with the UGOM QR having the best fit closely followed by the UWMOQ model.

Correlation coefficients reveal pairwise correlations. The VIF is employed to identify which explanatory variables are collinear and then they must be removed before the analysis of the model established in (11), which we call M0. The VIF values for model M0 are:

```
> car::vif(fitcovidULOG)
PD        GINI      BEDS      SR        PR        LE        T90       T180
2.136667  2.551950  1.740564  6.988683  4.176044  6.791217  1.423759  1.419650
```
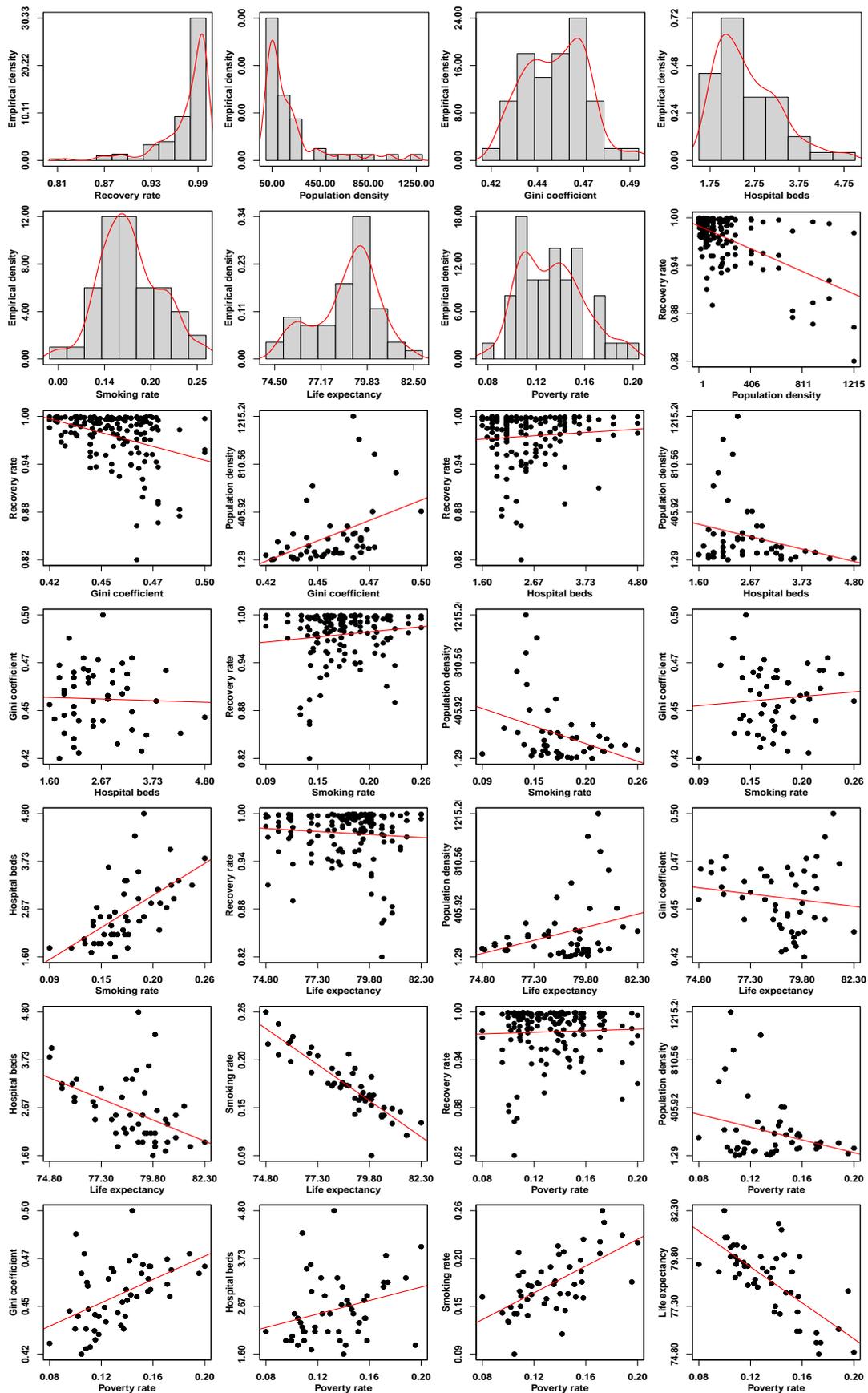
Figure 6. Scatterplots and histograms about the COVID-19 data set.

As previously stated, VIF > 10 suggests potential collinearity. Here, the covariates SR and LE have VIF scores of 6.989 and 6.791, respectively. Thus, these covariates could be considered, but the models considering such covariates must be carefully analyzed. If one covariate is removed, the VIF values must be recalculated, and the procedure is repeated until all VIF values are sufficiently small (less than five, for example) to ensure no collinearity issues. Then, we state three QR models to describe the recovery rate:

(M1 –reduced 1–) with covariates PD, GINI, BEDS, PR, LE, T90, and T180;

(M2 –reduced 2–) with covariates PD, GINI, BEDS, SR, PR, T90, and T180;

(M1 –reduced 3–) with covariates PD, GINI, BEDS, PR, T90, and T180.

The corresponding VIF values for models M1, M2, and M3 are:

```
> car::vif(reduced.1)
PD        GINI      BEDS      PR        LE        T90       T180
2.068567  2.516064  1.183617  4.169128  2.701867  1.393607  1.374571
> car::vif(reduced.2)
PD        GINI      BEDS      SR        PR        T90       T180
2.016583  2.253713  1.637897  2.888653  3.165383  1.433448  1.422845
> car::vif(reduced.3)
PD        GINI      BEDS      PR        T90       T180
2.005546  2.224827  1.120871  2.155601  1.401250  1.379976
```

Now, observe that the VIF is small for all reduced models (M1, M2, and M3). We can conclude from these results that the difference between the four models (M0, M1, M2, and M3) is not significant at 5%. M3 is chosen because of the parsimony principle. Essential details about the final model are the following: the model is congruent with the hypotheses established from the descriptive statistical analysis, which recommended to only include the explanatory variables T180, T90, GINI, and PD at a significant level of 1%. Observe that $\widehat{\beta}_0$ represents the percentage of recovered COVID-19 cases when all covariates take a value equal to zero. According to the estimate of the parameter β_3, hospital beds have a favorable impact on the recovery rates, meaning that recovery rates are higher in states with several hospital beds in the USA. The estimates of the parameters β_1, β_2, and β_4 have an unfavorable consequence on the rate of recovery, showing that the recovered COVID-19 cases are less in USA states with higher levels of population, GINI indicator, and level of poverty. Moreover, $\widehat{\beta}_5$ and $\widehat{\beta}_6$ are also negative, with $\widehat{\beta}_6$ being less than $\widehat{\beta}_5$. This suggests that the period after the 10th confirmed case has a negative impact on the recovery rate.

4. Concluding Remarks

The fractile or quantile regression models use conditional quantile functions and provide a method for stating the relationship between a response variable or outcome and covariates or explanatory variables. This method permits a more in-depth investigation of the response conditional distribution for various quantiles and provides a robust alternative for estimating the response central tendency. Although quantile regression structures were applied in various fields, parametric approaches have only been investigated in the past ten years. As mentioned, several parametric quantile regressions were postulated based on the arc-secant hyperbolic Weibull, Johnson SB, Kumaraswamy, log-extended exponential-geometric, unit-Birnbaum-Saunders, unit-Burr-XII, unit-Chen, unit-Gompertz, unit-Gumbel, unit-logistic, unit-generalized half-normal-E, unit-generalized half-normal-X, unit-Weibull-E, unit-Weibull-X, and Vasicek distributions [11–21,23–25]. In [10], the authors provided a number of parametric quantile regressions, as well as examples and software packages, which can be employed to model rates and indices. These indicators are particularly of interest when analyzing COVID-19 data, as we did in one of our applications. Therefore, new and more precise information based on our methodology was obtained, due to our model fitting the COVID-19 data in a better form than in other existing models. Our quantile regression formulation under the Weibull–Marshall–Olkin distribution had not been established until now.

We use two medical data sets and a computational library written in the R statistical software. By reparameterizing a distribution as a quantile function, we made an updated examination of the parametric quantile regression developed. To describe continuous variables with unit interval bounds, several distributions were used. We summarized their essential characteristics in this work. The mentioned R library was introduced for parameter estimation and model validation with unit interval distributions. Through the mentioned medical applications, we have demonstrated how to use the techniques and capabilities of the package.

A future work is related to the inclusion of quantile regression for inflation levels of zero, one, or zero-one. Considering that in such a case the likelihood function factors in two parts, one exclusively dependent on the discrete element and another solely dependent on the continuous element, it must be noted that the computation execution of zero or one is evidently enhanced. Therefore, the discrete elements can be estimated using the function of the R statistical software: `stats::glm()`, while the continuous elements may be determined employing the `uniquantreg::uniquantreg()` function. In future releases of the `uniquantreg` package, we will intend to add a wrapper function for estimation and hypothesis testing in other quantile regression models, as well as the model presented in this investigation.

Explanatory variables may affect the quantiles and other parameters. Then, one can study the joint effect that the explanatory variables can have on both quantiles and another parameters in regression formulations. Moreover, we want to study statistical structures generated from the current research under settings related to partial least squares, temporal, spatial, and multivariate frameworks [51–56]. Similarly, considering censored data may also be of interest to study [57], as well as control charts for quantiles based on covariates [58]. We are planning to conduct studies on these issues in the future.

Author Contributions: Conceptualization, V.L., J.M. and B.A. Data curation, J.M. and B.A. Formal analysis, V.L., J.M. and B.A. Investigation, V.L., J.M. and B.A. Methodology, V.L., J.M. and B.A. Writing—original draft, J.M. and B.A. Writing—review and editing, V.L. All authors have read and agreed to the published version of the manuscript.

Conflicts of Interest: The authors declare that they have no conflict of interest.

Appendix A. Existing Distributions

Table A1. PDF and reparameterization of the indicated distribution and its acronym.

Distribution	Acronym	PDF and Reparameterization
Arc-secant hyperbolic Weibull	ASHW	$f(z; \alpha, \beta) = \frac{\alpha\beta}{z\sqrt{1-z^2}} \mathrm{arcsech}(z)^{\beta-1} \exp\left[-\alpha \, \mathrm{arcsech}(z)^\beta\right]$, where $\alpha = -\log(\tau)/\mathrm{arcsech}(\mu)^\beta$
Johnson–SB	JOSB	$f(z; \alpha, \beta) = \frac{\beta}{\sqrt{2\pi}} \frac{1}{z(1-z)} \exp\left\{-\frac{1}{2}\left[\alpha + \beta \log\left(\frac{z}{1-z}\right)\right]^2\right\}$, where $\alpha = \Phi^{-1}(\tau) - \beta \log(\mu/(1-\mu))$
Kumaraswamy	KUMA	$f(z; \alpha, \beta) = \alpha\beta z^{\beta-1}(1-z^\beta)^{\alpha-1}$, where $\alpha = \log(1-\tau)/\log(1-\mu^\beta)$

Table A1. Cont.

Distribution	Acronym	PDF and Reparameterization
Log-extended exponential–geometric	LEEG	$f(z;\alpha,\beta) = \frac{\beta(1+\alpha)z^{\beta-1}}{(1+\alpha z^{\beta})^2}$ where $\alpha = -(1-\tau\mu^{-\beta})/(1-\tau)$
Unit–Birnbaum–Saunders	UBSA	$f(z;\alpha,\beta) = \frac{1}{2z\alpha\beta\sqrt{2\pi}}\left[\left(-\frac{\alpha}{\log(z)}\right)^{\frac{1}{2}} + \left(-\frac{\alpha}{\log(z)}\right)^{\frac{3}{2}}\right]\exp\left[\frac{1}{2\beta^2}\left(2+\frac{\log(z)}{\alpha}+\frac{\alpha}{\log(z)}\right)\right],$ where $\alpha = \log(\mu)\{\beta\Phi^{-1}(1-\tau)(4+\beta\Phi^{-1}(1-\tau))^{\frac{1}{2}} - 2 - [\beta\Phi^{-1}(1-\tau)]^2\}/2)$
Unit–Burr-XII	UBUR	$f(z;\alpha,\beta) = \frac{\alpha\beta}{z}[-\log(z)]^{\alpha-1}\{1+[-\log(z)]^{\alpha}\}^{(-\beta-1)},$ where $\alpha = \log(\tau^{-1/\beta}-1)/\log[-\log(\mu)]$
Unit–Chen	UCHE	$f(z;\alpha,\beta) = \frac{\alpha}{z}\beta[-\log(z)]^{\beta-1}\exp\left\{[-\log(z)]^{\beta}\right\}\exp\left[\alpha\left(1-\exp\left\{[-\log(z)]^{\beta}\right\}\right)\right],$ where $\alpha = \log(\tau)/(1-\exp\{[-\log(\mu)]^{\beta}\})$
Unit-generalized half-normal-E	UGHE	$f(z;\alpha,\beta) = \sqrt{\frac{2}{\pi}}\frac{\beta}{z[-\log(z)]}\left[-\frac{\log(z)}{\alpha}\right]^{\beta}\exp\left\{-\frac{1}{2}\left[-\frac{\log(z)}{\alpha}\right]^{2\beta}\right\},$ where $\alpha = -\log(\mu)[\Phi^{-1}(\tau/2)]^{-1/\beta}$
Unit-generalized half-normal-X	UGHX	$f(z;\alpha,\beta) = \sqrt{\frac{2}{\pi}}\frac{\beta}{z(1-z)}\left[\frac{z}{\alpha(1-z)}\right]^{\beta}\exp\left\{-\frac{1}{2}\left[\frac{z}{\alpha(1-z)}\right]^{2\beta}\right\},$ where $\alpha = [\mu/(1-\mu)]\{\Phi^{-1}[(\tau+1)/2]\}^{1/\beta}$
Unit–Gompertz	UGOM	$f(z;\alpha,\beta) = \alpha\beta z^{-(1+\beta)}\exp\left[\alpha\left(1-z^{-\beta}\right)\right],$ where $\alpha = \log(\tau)/(1-\mu^{-\beta})$
Unit–Gumbel	UGUM	$f(z;\alpha,\beta) = \frac{\beta}{z(1-z)}\exp\left\{-\alpha - \beta\log\left(\frac{z}{1-z}\right) - \exp\left[-\alpha - \beta\log\left(\frac{z}{1-z}\right)\right]\right\},$ where $\alpha = \beta\log[(1-\mu)/\mu] + \log[-1/\log(\tau)]$
Unit-logistic	ULOG	$f(z;\alpha,\beta) = \frac{\beta\exp(\alpha)\left(\frac{z}{1-z}\right)^{\beta-1}}{\left[1+\exp(\alpha)\left(\frac{z}{1-z}\right)^{\beta}\right]^2},$ where $\alpha = \log[\tau/(1-\tau)] - \beta\log[\mu/(1-\mu)]$
Unit–Weibull–E	UWEE	$f(z;\alpha,\beta) = \frac{1}{z}\alpha\beta[-\log(z)]^{\beta-1}\exp\left\{-\alpha[-\log(z)]^{\beta}\right\},$ where $\alpha = -\log(\tau)/[-\log(\mu)]^{\beta}$
Unit–Weibull–X	UWEX	$f(z;\alpha,\beta) = \frac{1}{(1-z)^2}\alpha\beta\left(\frac{z}{1-z}\right)^{\beta-1}\exp\left[-\alpha\left(\frac{z}{1-z}\right)^{\beta}\right],$ where $\alpha = -\log(1-\tau)[\mu/(1-\mu)]^{-\beta}$
Vasicek	VASI	$f(z;\alpha,\beta) = \sqrt{\frac{1-\beta}{\beta}}\exp\left\{\frac{1}{2}\left[\Phi^{-1}(z)^2 - \left(\frac{\Phi^{-1}(z)\sqrt{1-\beta}-\Phi^{-1}(\alpha)}{\sqrt{\beta}}\right)^2\right]\right\},$ where $\alpha = \Phi[\Phi^{-1}(\mu)\sqrt{1-\beta}-\Phi^{-1}(\tau)\sqrt{\beta}]$

References

1. Shahin, A.I.; Almotairi, S. A deep learning BiLSTM encoding-decoding model for COVID-19 pandemic spread forecasting. *Fractal Fract.* **2021**, *5*, 175. [CrossRef]
2. Ospina, R.; Leite, A.; Ferraz, C.; Magalhaes, A.; Leiva, V. Data-driven tools for assessing and combating COVID-19 out-breaks based on analytics and statistical methods in Brazil. *Signa Vitae* **2022**, *18*, 18–32.
3. Li, S.; Chen, J.; Li, B. Estimation and testing of random effects semiparametric regression model with separable space-time filters. *Fractal Fract.* **2022**, *6*, 735. [CrossRef]
4. Ferrari, S.; Cribari Neto, F. Beta regression for modelling rates and proportions. *J. Appl. Stat.* **2004**, *31*, 799–815. [CrossRef]
5. Koenker, R.; Bassett, G.J. Regression quantiles. *Econometrica* **1978**, *46*, 33–50. [CrossRef]
6. Koenker, R. *Quantile Regression*; Cambridge University Press: Cambridge, UK, 2005.
7. Haupt, H.; Fritsch, M. Quantile trend regression and its application to central England temperature. *Mathematics* **2022**, *10*, 413. [CrossRef]
8. Shin, K.; You, S. Quantile regression analysis between the after-school exercise and the academic performance of Korean middle school students. *Mathematics* **2021**, *10*, 58. [CrossRef]
9. Yu, K.; Moyeed, R.A. Bayesian quantile regression. *Stat. Prob. Lett.* **2001**, *54*, 437–447. [CrossRef]

10. Mazucheli, M.; Alves, B.; Menezes, A.F.B.; Leiva, V. An overview on parametric quantile regression models and their computational implementation with applications to biomedical problems including COVID-19 data. *Comput. Methods Programs Biomed.* **2022**, *221*, 106816. [CrossRef]

11. Galarza, C.E.; Zhang, P.; Lachos, V.H. Logistic quantile regression for bounded outcomes using a family of heavy-tailed distributions. *Sankhya B* **2021**, *83*, 325–349. [CrossRef]

12. Jodrá, P.; Jiménez-Gamero, M.D. A quantile regression model for bounded responses based on the exponential-geometric distribution. *REVSTAT Stat. J.* **2020**, *4*, 415–436.

13. Korkmaz, M.Ç.; Chesneau, C. On the unit Burr-XII distribution with the quantile regression modeling and applications. *Comput. Appl. Math.* **2021**, *40*, 29. [CrossRef]

14. Korkmaz, M.Ç.; Chesneau, C.; Korkmaz, Z.S. A new alternative quantile regression model for the bounded response with educational measurements applications of OECD countries. *J. Appl. Stat.* **2023**, *50*, 131–154. [CrossRef]

15. Korkmaz, M.Ç.; Chesneau, C.; Korkmaz, Z.S. On the arcsecant hyperbolic normal distribution. Properties, quantile regression modeling and applications. *Symmetry* **2021**, *13*, 117. [CrossRef]

16. Korkmaz, M.Ç.; Emrah, A.; Chesneau, C.; Yousof, H.M. On the unit-Chen distribution with associated quantile regression and applications. *Math. Slovaca* **2021**, *72*, 765–786. [CrossRef]

17. Korkmaz, M.Ç.; Chesneau, C.; Korkmaz, Z.S. Transmuted unit Rayleigh quantile regression model: Alternative to beta and Kumaraswamy quantile regression models. *Univ. Politeh. Buchar. Sci. Bull. A Appl. Math. Phys.* **2021**, *83*, 149–158.

18. Korkmaz, M.Ç.; Korkmaz, Z.S. The unit log–log distribution: A new unit distribution with alternative quantile regression modeling and educational measurements applications. *J. Appl. Stat.* 2023, *in press*. [CrossRef]

19. Saulo, H.; Dasilva, A.; Leiva, V.; Sánchez, L.; Fuente-Mella, H.L. Log-symmetric quantile regression models. *Stat. Neerl.* **2021**, *76*, 124–163. [CrossRef]

20. Sánchez, L.; Leiva, V.; Saulo, H.; Marchant, C.; Sarabia, J.M. A new quantile regression model and its diagnostic analytics for a Weibull distributed response with applications. *Mathematics* **2021**, *9*, 2768. [CrossRef]

21. Mazucheli, J.; Leiva, V.; Alves, B.; Menezes, A.F.B. A new quantile regression for modeling bounded data under a unit Birnbaum-Saunders distribution with applications in medicine and politics. *Symmetry* **2021**, *13*, 682. [CrossRef]

22. Mazucheli, M.; Alves, B.; Korkmaz, M.C.; Leiva, V. Vasicek quantile and mean regression models for bounded data: New formulation, mathematical derivations, and numerical applications. *Mathematics* **2022**, *10*, 1389. [CrossRef]

23. Mazucheli, M.; Korkmaz, M.C.; Menezes, A.F.B.; Leiva, V. The unit generalized half-normal quantile regression model: Formulation, estimation, diagnostics, and numerical applications. *Soft Comput.* **2023**, *27*, 279–295. [CrossRef]

24. Saulo, H.; Vila, R.; Borges, G.V.; Bourguignon, M.; Leiva, V.; Marchant, C. Modeling income data via new parametric quantile regressions: Formulation, computational statistics, and application. *Mathematics* **2023**, *11*, 448. [CrossRef]

25. Saulo, H.; Vila, R.; Bittencourt, V.L.; Leao, J.; Leiva, V.; Christakos, G. On a new extreme value distribution: Characterization, parametric quantile regression, and application to extreme air pollution events. *Stoch. Environ. Res. Risk Assess.* 2023, *in press*. [CrossRef]

26. Sánchez, L.; Leiva, V.; Galea, M.; Saulo, H. Birnbaum-saunders quantile regression and its diagnostics with application to economic data. *Appl. Stoch. Model. Bus. Ind.* **2021**, *37*, 53–73. [CrossRef]

27. Johnson, N.L.; Kotz, S.; Balakrishnan, N. *Continuous Univariate Distributions*; Wiley: New York, NY, USA, 1994; Volume 1.

28. Johnson, N.L.; Kotz, S.; Balakrishnan, N. *Continuous Univariate Distributions*; Wiley: New York, NY, USA, 1995; Volume 2.

29. Kotz, S.; Leiva, V.; Sanhueza, A. Two new mixture models related to the inverse Gaussian distribution. *Methodol. Comput. Appl. Probab.* **2010**, *12*, 199–212. [CrossRef]

30. Lindsey, J.K. *Applying Generalized Linear Models*; Springer: New York, NY, USA, 2000.

31. Benedetti, T.R.B.; Antunes, P.d.C.; Rodriguez-Añez, C.R.; Mazo, G.Z.; Petroski, É.L. Reproducibility and validity of the International Physical Activity Questionnaire (IPAQ) in elderly men. *Rev. Bras. Med. Esp.* **2007**, *13*, 11–16. [CrossRef]

32. Petterle, R.R.; Bonat, W.H.; Scarpin, C.T.; Jonasson, T.; Borba, V.Z.C. Multivariate quasi-beta regression models for continuous bounded data. *Int. J. Biostat.* **2020**, *1*, 39–53. [CrossRef]

33. Rousseeuw, P.; Croux, C.; Todorov, V.; Ruckstuhl, A.; Salibian-Barrera, M.; Verbeke, T.; Koller, M.; Maechler, M. Robustbase: Basic Robust Statistics. R Package Version 0.92-6; 2016. Available online: https://CRAN.R-project.org/package=robustbase (accessed on 30 January 2023).

34. Marshall, A.W.; Olkin, I. A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika* **1997**, *84*, 641–652. [CrossRef]

35. Zeileis, Z.; Hothorn, T. Diagnostic checking in regression relationships. *R J.* **2002**, *2*, 7–10.

36. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2019.

37. SAS. *SAS/STAT® 14.1 Users Guide*; SAS Institute: Cary, NC, USA, 2015.

38. Jiang, J. *Linear and Generalized Linear Mixed Models and Their Applications*; Springer: New York, NY, USA, 2006.

39. Cox, D.R.; Hinkley, D.V. *Theoretical Statistics*; CRC Press: Boca-Raton, FL, USA, 1979.

40. Korosteleva, O. *Advanced Regression Models with SAS and R*; CRC Press: Boca Raton, FL, USA, 2019.

41. Eddelbuettel, D.; Balamuta, J.J. Extending R with C++: A brief introduction to Rcpp. *Amer. Stat.* **2018**, *72*, 28–36. [CrossRef]

42. Eddelbuettel, D.; Francois, R. Rcpp: Seamless R and C++ integration. *J. Stat. Soft.* **2011**, *40*, 1–18. [CrossRef]

43. Zhang, P.; Qiu, Z.; Shi, C. simplexreg: An R package for regression analysis of proportional data using the simplex distribution. *J. Stat. Soft.* **2016**, *71*, 1–21. [CrossRef]
44. Zeileis, A.; Croissant, Y. Extended model formulas in R: Multiple parts and multiple responses. *J. Stat. Soft.* **2010**, *34*, 1–13. [CrossRef]
45. Nash, J.C.; Varadhan, R. Unifying optimization algorithms to aid software system users: Optimx for R. *J. Stat. Soft.* **2011**, *43*, 1–14. [CrossRef]
46. Koenker, R. quantreg: Quantile Regression. R Package Version 5.86. 2021. Available online: https://CRAN.R-project.org/package=quantreg (accessed on 30 January 2023).
47. Mazucheli, J.; Menezes, A.F.; Dey, S. The unit-Birnbaum-Saunders distribution with applications. *Chil. J. Stat.* **2018**, *9*, 47–57.
48. Marchant, C.; Leiva, V.; Cavieres M.F.; Sanhueza, A. Air contaminant statistical distributions with application to PM10 in Santiago, Chile. *Rev. Environ. Contam. Toxicol.* **2013**, 223, 1–31.
49. Mitnik, P.A.; Baek, S. The Kumaraswamy distribution: Median-dispersion reparameterizations for regression modeling and simulation-based estimation. *Stat. Pap.* **2013**, *54*, 177–192. [CrossRef]
50. Ribeiro, T.F.; Cordeiro, G.M.; Peña-Ramírez, F.A.; Guerra, R.R. A new quantile regression for the COVID-19 mortality rates in the United States. *Comput. Appl. Math.* **2022**, *40*, 255. [CrossRef]
51. Huerta, M.; Leiva, V.; Liu, S.; Rodriguez, M.; Villegas, D. On a partial least squares regression model for asymmetric data with a chemical application in mining. *Chemom. Intell. Lab. Syst.* **2019**, *190*, 55–68. [CrossRef]
52. Leiva, V.; Saulo, H.; Leao, J.; Marchant, C. A family of autoregressive conditional duration models applied to financial data. *Comput. Stat. Data Anal.* **2014**, *79*, 175–191. [CrossRef]
53. Marchant, C.; Leiva, V.; Cysneiros F.J.A.; Liu, S. Robust multivariate control charts based on Birnbaum-Saunders distributions. *J. Stat. Comput. Simul.* **2018**, *88*, 182–202. [CrossRef]
54. Martinez, S.; Giraldo, R.; Leiva, V. Birnbaum-Saunders functional regression models for spatial data. *Stoch. Environ. Res. Risk Assess.* **2019**, *33*, 1765–1780. [CrossRef]
55. Saulo, H.; Leao, J.; Leiva, V.; Aykroyd, R.G. Birnbaum-Saunders autoregressive conditional duration models applied to high-frequency financial data. *Stat. Pap.* **2019**, *60*, 1605–1629. [CrossRef]
56. Sanchez, L.; Leiva, V.; Galea, M.; Saulo, H. Birnbaum-Saunders quantile regression models with application to spatial data. *Mathematics* **2020**, *8*, 1000. [CrossRef]
57. Leao, J.; Leiva, V.; Saulo, H.; Tomazella, V. Incorporation of frailties into a cure rate regression model and its diagnostics and application to melanoma data. *Stat. Med.* **2018**, *37*, 4421–4440. [CrossRef]
58. Leiva, V.; dos Santos, R.A.; Saulo, H.; Marchant, C.; Lio, Y. Bootstrap control charts for quantiles based on log-symmetric distributions with applications to monitoring of reliability data. *Qual. Reliab. Eng. Int.* **2023**, *39*, 1–24. [CrossRef]