# New Lifetime Distribution for Modeling Data on the Unit Interval: Properties, Applications and Quantile Regression 

Suleman Nasiru ${ }^{1(D)}$, Abdul Ghaniyyu Abubakari ${ }^{1(D)}$ and Christophe Chesneau ${ }^{2, *}$<br>1 Department of Statistics, School of Mathematical Sciences, C. K. Tedam University of Technology and Applied Sciences, Kassena-Nankana Navrongo-Kolgo Road, Navrongo, Upper East, Ghana<br>2 Department of Mathematics, LMNO, CNRS-Université de Caen, Campus II, Science 3, 14032 Caen, France<br>* Correspondence: christophe.chesneau@unicaen.fr

Citation: Nasiru, S.; Abubakari, A.G.; Chesneau, C. New Lifetime Distribution for Modeling Data on the Unit Interval: Properties, Applications and Quantile Regression. Math. Comput. Appl. 2022, 27, 105. https://doi.org/10.3390/ mca27060105

Received: 31 October 2022
Accepted: 2 December 2022
Published: 3 December 2022
Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/).


#### Abstract

Probability distributions are very useful in modeling lifetime datasets. However, no specific distribution is suitable for all kinds of datasets. In this study, the bounded truncated Cauchy power exponential distribution is proposed for modeling datasets on the unit interval. The probability density function exhibits desirable shapes, such as left-skewed, right-skewed, reversed J, and bathtub shapes, whereas the hazard rate function displays J and bathtub shapes. For the purpose of modeling dependence between measures in a dataset, a bivariate extension of the proposed distribution is developed. The bivariate probability density function displays monotonic and non-monotonic shapes, making it suitable for modeling complex bivariate relations. Subsequently, the applications of the distribution are illustrated using COVID-19 data. The results revealed that the new distribution provides a better fit to the datasets compared to other existing distributions. Finally, a new quantile regression model is developed and its application demonstrated. The generated quantile regression model offers a decent fit to the data, according to the residual analysis.


Keywords: COVID-19; bounded distribution; estimation methods; Cauchy; regression; bivariate

## 1. Introduction

Disease modeling and prediction are primary tasks of epidemiologists and researchers interested in the estimation of disease occurrences. To perform these tasks, modeling the variability in disease occurrences using probability distributions is essential. With the emergence of the novel coronavirus disease in late 2019 (COVID-19) and its negative impact on humanity, many researchers have proposed new probability distributions (discrete or continuous) for modeling the number of infections, mortality rate, and recovery rates, among others. Some of the proposed probability distributions or families of distributions include: Marshall-Olkin reduced Kies distribution [1], modified inverse Weibull distribution [2], weighted Weibull distribution [3], type I half logistic Burr X-G family [4], unit power Weibull distribution [5], new extended exponentiated Weibull distribution [6], discrete extended odd Weibull exponential distribution [7], odd Weibull inverse Topp-Leone distribution [8], log-logistic tangent distribution [9], discrete-type half-logistic exponential distribution [10], and unit Johnson $S_{U}$ distribution [11].

Among these probability distributions used for modeling diseases, those defined on the unit interval play a major role due to their usefulness in areas such as health, psychology, and epidemiology, among others. For instance, researchers may be interested in modeling mortality or recovery rates. Observations measured on these variables are usually proportions, fractions, or rates, which are defined in the unit interval. Although the beta distribution is the oldest for modeling datasets measured on the unit interval, the intractability of its cumulative distribution function (CDF) and quantile function has called for the development of new distributions with tractable CDFs and quantile functions that are also capable of modeling data on the unit interval. Unit distributions proposed recently
in literature include: unit Gamma/Gompertz distribution [12], bounded odd inverse Pareto exponential distribution [13], bounded shifted Gompertz distribution [14], unit modified Burr-III distribution [15], unit generalized half normal distribution [16], unit Lindley distribution [17], unit Gompertz distribution [18], logit slash distribution [19], unit Weibull distribution [20] and unit inverse Gaussian distribution [21].

Despite the existence of many unit distributions in the literature, no single distribution is capable of modeling all forms of data since the data generating process produces data with different characteristics such as symmetric, skewed, varied degrees of kurtosis, and monotonic and non-monotonic failure rates. This study thus proposes a new unit distribution called the bounded truncated Cauchy power exponential (BTCPE) distribution. The motivations for developing the new distribution are as follows: to provide a model capable of modeling complex data on unit interval that exhibits platykurtic, leptokurtic, reversed J, left-skewed, right-skewed, bathtub, and J shapes; to develop a bivariate distribution for modeling interdependence between random data on unit interval; and to develop a quantile regression model for understanding the relationship between a response variable and given covariates.

The remainder of the paper is organized in nine sections, described as follows: Section 2 presents the development of the BTCPE distribution, Section 3 describes some of its important properties, Section 4 focuses on a special bivariate extension of the BTCPE distribution, Section 5 is devoted to the parametric estimation methods, Section 6 presents the Monte Carlo simulation of nine frequentist estimation methods, Section 7 contains the univariate applications of the BTCPE distribution, Section 8 is about the quantile regression model and its application, and finally the conclusion of the paper is presented in Section 9.

## 2. Bounded Truncated Cauchy Power Exponential Distribution

A random variable $X$ follows the truncated Cauchy power exponential (TCPE) distribution if its CDF and probability density function (PDF), respectively, are defined as

$$
\begin{equation*}
F_{X}(x ; \alpha, \lambda)=\frac{4}{\pi} \arctan \left[\left(1-e^{-\lambda x}\right)^{\alpha}\right], \alpha>0, \lambda>0, x>0 \tag{1}
\end{equation*}
$$

and

$$
\begin{equation*}
f_{X}(x ; \alpha, \lambda)=\frac{4 \alpha \lambda e^{-\lambda x}\left(1-e^{-\lambda x}\right)^{\alpha-1}}{\pi\left[1+\left(1-e^{-\lambda x}\right)^{2 \alpha}\right]}, x>0 \tag{2}
\end{equation*}
$$

The TCPE distribution can be presented as a special case of the TCP Weibull distribution proposed by [22]. Now, we define a new unit distribution, called the BTCPE distribution, corresponding to the distribution of $Y=e^{-X}$. The associated CDF is obtained as follows:

$$
\begin{aligned}
F_{Y}(y ; \alpha, \lambda)=\mathbb{P}\left(e^{-X} \leq y\right) & =\mathbb{P}(-X \leq \log (y)) \\
& =1-\mathbb{P}(X \leq-\log (y)) \\
& =1-F_{X}(-\log (y) ; \alpha, \lambda) .
\end{aligned}
$$

Hence, the CDF of the BTCPE distribution is expressed as

$$
\begin{equation*}
F_{Y}(y ; \alpha, \lambda)=1-\frac{4}{\pi} \arctan \left[\left(1-y^{\lambda}\right)^{\alpha}\right], 0<y<1 \tag{3}
\end{equation*}
$$

and $\alpha>0$ and $\lambda>0$ are the shape parameters that have to be estimated. The associated PDF of the BTCPE distribution is obtained by differentiating Equation (3), and it is given by

$$
\begin{equation*}
f_{Y}(y ; \alpha, \lambda)=\frac{4 \alpha \lambda y^{\lambda-1}\left(1-y^{\lambda}\right)^{\alpha-1}}{\pi\left[1+\left(1-y^{\lambda}\right)^{2 \alpha}\right]}, 0<y<1 . \tag{4}
\end{equation*}
$$

Often, the PDFs are expressed in expanded form for easy derivation of the statistical properties of the proposed distribution. The expanded form of the PDF of the BTCPE
distribution is mainly obtained using the generalized binomial expansion, $(y+a)^{-n}=$ $\sum_{k=0}^{\infty}\binom{-n}{k} y^{k} a^{-n-k},|y|<a$, where $n$ is any real number. Thus, it is given by

$$
\begin{equation*}
f_{Y}(y ; \alpha, \lambda)=\frac{4 \alpha \lambda}{\pi} \sum_{i=0}^{\infty}(-1)^{i} y^{\lambda-1}\left(1-y^{\lambda}\right)^{\alpha(2 i+1)-1}, 0<y<1 . \tag{5}
\end{equation*}
$$

The corresponding hazard rate function (HRF) is given by

$$
\begin{equation*}
h_{Y}(y ; \alpha, \lambda)=\frac{\alpha \lambda y^{\lambda-1}\left(1-y^{\lambda}\right)^{\alpha-1}}{\left[1+\left(1-y^{\lambda}\right)^{2 \alpha}\right] \arctan \left[\left(1-y^{\lambda}\right)^{\alpha}\right]}, 0<y<1 . \tag{6}
\end{equation*}
$$

The shapes of the PDF and HRF for some given parameter values are shown in Figure 1. The PDF exhibits symmetric, bathtub, left-skewed and right-skewed shapes for the given parameter values. The HRF displays bathtub and increasing failure rates.


Figure 1. PDF (left) and HRF (right) of the BTCPE distribution.

## 3. Some Important Properties

This section presents some relevant properties of the BTCPE distribution.

### 3.1. Distribution Inequalities

This subsection investigates some desirable inequalities satisfied by the CDF of the BTCPE distribution. These inequalities are very essential in determining the first order stochastic dominance of random variables [23].

Proposition 1. The CDF of the BTCPE distribution is increasing with respect to the parameter $\alpha$. The CDF of the BTCPE distribution is decreasing with respect to the parameter $\lambda$.

Proof. For the first point, since $\left(1-y^{\lambda}\right)^{\alpha} \log \left(1-y^{\lambda}\right)<0$, for $y \in(0,1)$, we have

$$
\frac{\partial F_{Y}(y ; \alpha, \lambda)}{\partial \alpha}=-\frac{4\left(1-y^{\lambda}\right)^{\alpha} \log \left(1-y^{\lambda}\right)}{\pi\left[1+\left(1-y^{\lambda}\right)^{2 \alpha}\right]} \geq 0 .
$$

This means that $F_{Y}(y ; \alpha, \lambda)$ is increasing with respect to $\alpha$. For the second point, since $y^{\lambda}\left(1-y^{\lambda}\right)^{\alpha-1} \log (y)<0$, for $y \in(0,1)$, we have

$$
\frac{\partial F_{Y}(y ; \alpha, \lambda)}{\partial \lambda}=\frac{4 \alpha y^{\lambda}\left(1-y^{\lambda}\right)^{\alpha-1} \log (y)}{\pi\left[1+\left(1-y^{\lambda}\right)^{2 \alpha}\right]} \leq 0
$$

This implies that $F_{Y}(y ; \alpha, \lambda)$ is decreasing with respect to $\lambda$. This completes the proof of the proposition. From Proposition 1, the following stochastic ordering property follows immediately: if $\alpha_{1} \leq \alpha_{2}$ then $F_{Y}\left(y ; \alpha_{1}, \lambda\right) \leq F_{Y}\left(y ; \alpha_{2}, \lambda\right)$. Also, if $\lambda_{1} \leq \lambda_{2}$ then $F_{Y}\left(y ; \alpha, \lambda_{2}\right) \leq F_{Y}\left(y ; \alpha, \lambda_{1}\right)$.

### 3.2. Quantile Function

The quantile function or the inverse CDF is simply the solution $Q(u ; \alpha, \lambda)$ of the following nonlinear equation: $F_{Y}(Q(u ; \alpha, \lambda) ; \alpha, \lambda)=u$, for all $u \in(0,1)$. Thus, after some algebraic manipulation, we have

$$
\begin{equation*}
Q(u ; \alpha, \lambda)=\left\{1-\left(\tan \left[\frac{\pi}{4}(1-u)\right]\right)^{1 / \alpha}\right\}^{1 / \lambda}, u \in(0,1) \tag{7}
\end{equation*}
$$

The median is obtained by substituting $u=0.5$. The quantile function plays an important role in the generation of random observations from the BTCPE distribution. The quantile function values are also useful in computing measures of skewness and kurtosis. As a classical quantile measure, the MacGillivray measure of skewness [24] is given by

$$
\rho(u ; \alpha, \lambda)=\frac{Q(1-u ; \alpha, \lambda)+Q(u ; \alpha, \lambda)-2 Q(0.5 ; \alpha, \lambda)}{Q(1-u ; \alpha, \lambda)-Q(u ; \alpha, \lambda)}, u \in(0,1) .
$$

In particular, the MacGillivray measure of skewness can be used to efficiently describe the effect of the parameters $(\alpha, \lambda)$ on the skewness. The more the shapes of $\rho(u ; \alpha, \lambda)$ vary according to the parameters, the more flexible the skewness is. Figure 2 shows the plot of this skewness measure for a fixed value of $\lambda$ while $\alpha$ varies and for a fixed value of $\alpha$ while $\lambda$ varies. From Figure 2, the wider variations seen imply that both parameters have a strong influence on the skewness of the BTCPE distribution. In addition, as the values of $\alpha$ or $\lambda$ increase, $\rho(u ; \alpha, \lambda)$ gets closer to the horizontal line. This shows that utilizing higher values of the parameter can result in a symmetrical distribution.


Figure 2. Plots of the MacGillivray skewness.
The kurtosis of the BTCPE distribution can be studied using the Moors kurtosis [25]. The Moors (coefficient of) kurtosis is usually given by

$$
K(\alpha, \lambda)=\frac{Q(7 / 8 ; \alpha, \lambda)-Q(5 / 8 ; \alpha, \lambda)+Q(3 / 8 ; \alpha, \lambda)-Q(1 / 8 ; \alpha, \lambda)}{Q(3 / 4 ; \alpha, \lambda)-Q(1 / 4 ; \alpha, \lambda)} .
$$

Large values of the Moors kurtosis imply that the distribution has a heavy tail, and small values are indications of a light tail. Figure 3 displays the Moors kurtosis for the BTCPE distribution. It can be observed that the BTCPE distribution exhibits various degrees of kurtosis. When the parameters $\alpha$ and $\lambda$ are equal, the distribution displays a platykurtic shape. The overall shapes show how flexible the BTCPE distribution is with regards to modeling datasets having different degrees of kurtosis and skewness.


Figure 3. Plots of Moors kurtosis.

### 3.3. Moments and Moments Generating Function

The $r^{\text {th }}$ moments, incomplete moments and moment generating function of the BTCPE distribution are presented in this subsection.

Proposition 2. If $Y$ is a BTCPE random variable, i.e., a random variable with the BTCPE distribution, then its $r^{\text {th }}$ non-central moment is given by

$$
\begin{equation*}
\mu_{r}^{\prime}=\frac{4 \alpha}{\pi} \sum_{i=0}^{\infty}(-1)^{i} B\left(\frac{r}{\lambda}+1, \alpha(1+2 i)\right), r=1,2, \ldots, \tag{8}
\end{equation*}
$$

where $B(a, b)=\int_{0}^{1} z^{a-1}(1-z)^{b-1} d z$ is the beta integral function.
Proof. The $r^{\text {th }}$ non-central moment of the BTCPE random variable is defined as $\mu^{\prime}{ }_{r}=$ $E\left(Y^{r}\right)=\int_{0}^{1} y^{r} f_{Y}(y ; \alpha, \lambda) d y$. Thus, substituting the expanded form of the PDF given in Equation (5) yields

$$
\mu_{r}^{\prime}=\frac{4 \alpha \lambda}{\pi} \sum_{i=0}^{\infty}(-1)^{i} \int_{0}^{1} y^{r+\lambda-1}\left(1-y^{\lambda}\right)^{\alpha(1+2 i)-1} d y .
$$

Letting $z=y^{\lambda}, y \rightarrow 0, z \rightarrow 0 ; y \rightarrow 1, z \rightarrow 1$ and $d z=\lambda y^{\lambda-1} d y$, we get

$$
\mu_{r}^{\prime}=\frac{4 \alpha}{\pi} \sum_{i=0}^{\infty}(-1)^{i} \int_{0}^{1} z^{\frac{r}{\lambda}}(1-z)^{\alpha(1+2 i)-1} d z
$$

Hence, several algebraic manipulation yield

$$
\mu_{r}^{\prime}=\frac{4 \alpha}{\pi} \sum_{i=0}^{\infty}(-1)^{i} B\left(\frac{r}{\lambda}+1, \alpha(1+2 i)\right) .
$$

This completes the proof.
The non-central moments can be used to derive other important characteristics of the BTCPE distribution such as estimating the variance, coefficient of skewness and kurtosis.

Proposition 3. The $r^{\text {th }}$ incomplete moment of the BTCPE random variable is given by

$$
\begin{equation*}
\varphi_{r}=\frac{4 \alpha}{\pi} \sum_{i=0}^{\infty}(-1)^{i} B\left(y^{\lambda} ; \frac{r}{\lambda}+1, \alpha(1+2 i)\right), r=1,2, \ldots, \tag{9}
\end{equation*}
$$

where $B(q ; a, b)=\int_{0}^{q} z^{a-1}(1-z)^{b-1} d z$ is the incomplete beta integral function.
Proof. By definition, the $r^{\text {th }}$ incomplete moment is given by

$$
\varphi_{r}=E\left(Y^{r} 1\{Y<y\}\right)=\int_{0}^{y} x^{r} f(x ; \alpha, \lambda) d x
$$

Hence, substituting the expanded form of the PDF into the definition yields

$$
\varphi_{r}=\frac{4 \alpha \lambda}{\pi} \sum_{i=0}^{\infty}(-1)^{i} \int_{0}^{y} x^{r+\lambda-1}\left(1-x^{\lambda}\right)^{\alpha(1+2 i)-1} d x
$$

Letting $z=x^{\lambda}, x \rightarrow 0, z \rightarrow 0 ; x \rightarrow y, z \rightarrow y^{\lambda}$ and $d z=\lambda x^{\lambda-1} d x$. Hence, applying similar concepts for proving the incomplete moments yields

$$
\varphi_{r}=\frac{4 \alpha}{\pi} \sum_{i=0}^{\infty}(-1)^{i} B\left(y^{\lambda} ; \frac{r}{\lambda}+1, \alpha(1+2 i)\right) .
$$

This completes the proof.
The moment generating function is useful for deriving the moments of a random variable if only the moment exists.

Proposition 4. The moment generating function of the BTCPE random variable is given by

$$
\begin{equation*}
M_{Y}(t)=\frac{4 \alpha}{\pi} \sum_{r=0}^{\infty} \sum_{i=0}^{\infty} \frac{(-1)^{i} t^{r}}{r!} B\left(\frac{r}{\lambda}+1, \alpha(1+2 i)\right) \tag{10}
\end{equation*}
$$

Proof. By definition and a standard exponential expansion, we have $M_{Y}(t)=E\left(e^{t Y}\right)=$ $\sum_{r=0}^{\infty} \frac{t^{r}}{r!} \mu^{\prime}$. Hence, substituting the $r^{t h}$ non-central moment of the BTCPE distribution into the definition completes the proof.

Table 1 shows the first six moments of the BTCPE distribution and other useful measures, such as the standard deviation (SD), coefficients of variation (CV), skewness (CS) and kurtosis (CK). The SD, CV, CS and CK are, respectively, given by

$$
\begin{gathered}
S D=\sqrt{\mu_{2}^{\prime}-\mu^{2}}, \\
C V=\frac{\sigma}{\mu}=\sqrt{\frac{\mu_{2}^{\prime}}{\mu^{2}}-1},
\end{gathered}
$$

$$
C S=\frac{\mu_{3}^{\prime}-3 \mu \mu_{2}^{\prime}+2 \mu^{3}}{\left(\mu_{2}^{\prime}-\mu^{2}\right)^{3 / 2}}
$$

and

$$
C K=\frac{\mu_{4}^{\prime}-4 \mu \mu_{3}^{\prime}+6 \mu^{2} \mu_{2}^{\prime}-3 \mu^{4}}{\left(\mu_{2}^{\prime}-\mu^{2}\right)^{2}}
$$

Table 1. Values of moment measures, including the SD, CV, CS and CK.

| $\mu_{r}^{\prime}$ | $\alpha=\mathbf{0 . 4}, \boldsymbol{\lambda}=\mathbf{2 . 5}$ | $\alpha=\mathbf{4 . 5}, \boldsymbol{\lambda}=\mathbf{3 . 1}$ | $\boldsymbol{\alpha}=\mathbf{2 0 . 0 , \boldsymbol { \lambda } = \mathbf { 1 . 5 }}$ |
| :---: | :---: | :---: | :---: |
| $\mu_{1}^{\prime}$ | 0.8799 | 0.5602 | 0.1339 |
| $\mu_{2}^{\prime}$ | 0.8021 | 0.3401 | 0.0242 |
| $\mu_{3}^{\prime}$ | 0.7457 | 0.2185 | 0.0053 |
| $\mu_{4}^{\prime}$ | 0.7020 | 0.1465 | 0.0013 |
| $\mu_{5}^{\prime}$ | 0.6667 | 0.1017 | 0.0004 |
| $\mu_{6}^{\prime}$ | 0.6373 | 0.0726 | 0.0001 |
| SD | 0.1668 | 0.1619 | 0.0794 |
| CV | 0.1896 | 0.2890 | 0.5931 |
| CS | -1.9527 | -0.3403 | 0.7713 |
| CK | 6.6850 | 2.7084 | 3.5390 |

From Table 1, the CS is negative for the given parameter values and positive for others. It can be seen that the BTCPE distribution can be leptokurtic or platykurtic depending on the parameter values, since the CK can be lower than 3 or greater than 3, respectively. The coefficient of skewness also reveals that the BTCPE distribution can model both left and right-skewed data.

### 3.4. Order Statistics

Order statistics play an imperative role in both statistics and industrial reliability analysis. They can be used to estimate the minimum, maximum, and range of observations. They are used in developing control charts that are useful in industrial quality control analyses. Let $Y_{1: n} \leq Y_{2: n} \leq \ldots \leq Y_{n: n}$ be $n$ order statistics from $n$ BTCPE random variables. Then, the PDF of $Y_{k: n}$ is given by

$$
f_{k: n}(y ; \alpha, \lambda)=\Omega_{k: n}\left[F_{Y}(y ; \alpha, \lambda)\right]^{k-1}\left[1-F_{Y}(y ; \alpha, \lambda)\right]^{n-k} f_{Y}(y ; \alpha, \lambda),
$$

where

$$
\Omega_{k: n}=\frac{n!}{(k-1)!(n-k)!} .
$$

Using the binomial expansion $(1-y)^{\lambda-1}=\sum_{i=0}^{\infty}(-1)^{i}\binom{\lambda-1}{i} y^{i},|y|<1$, we can write

$$
f_{k: n}(y ; \alpha, \lambda)=\Omega_{k: n} \sum_{i=0}^{k-1}(-1)^{i}\binom{k-1}{i}\left[1-F_{Y}(y ; \alpha, \lambda)\right]^{n-k+i} f_{Y}(y ; \alpha, \lambda) .
$$

Thus, we have

$$
\begin{equation*}
f_{k: n}(y ; \alpha, \lambda)=\frac{\Omega_{k: n} 4 \alpha \lambda y^{\lambda-1}\left(1-y^{\lambda}\right)^{\alpha-1}}{\pi\left[1+\left(1-y^{\lambda}\right)^{2 \alpha}\right]} \sum_{i=0}^{k-1}(-1)^{i}\binom{k-1}{i}\left[\frac{4}{\pi} \arctan \left[\left(1-y^{\lambda}\right)^{\alpha}\right]\right]^{n-k+i} . \tag{11}
\end{equation*}
$$

On the other side, the CDF of $Y_{1: n}$ is simply given by

$$
\begin{aligned}
& F_{1: n}(y ; \alpha, \lambda)=1-\left[1-F_{Y}(y ; \alpha, \lambda)\right]^{n} \\
& \quad=1-\left[\frac{4}{\pi} \arctan \left[\left(1-y^{\lambda}\right)^{\alpha}\right]\right]^{n}
\end{aligned}
$$

and the CDF of $Y_{n: n}$ is derived as

$$
F_{n: n}(y ; \alpha, \lambda)=\left[F_{Y}(y ; \alpha, \lambda)\right]^{n}=\left[1-\frac{4}{\pi} \arctan \left[\left(1-y^{\lambda}\right)^{\alpha}\right]\right]^{n}
$$

The distribution of the smallest order statistic represents the lifetime of a system connected in series, and that of the maximum order statistic denotes the lifetime of a system connected in parallel. Hence, they are vital in studying the minimum and maximum time to failure of components in engineering reliability. The minimum and maximum (min$\max$ ) plots of the order statistics can be used to investigate the distributional behavior of observations. The min-max plot captures not only the information in the tails but all the information about the whole distribution. The min-max plots shown in Figure 4 for some parameter values depend on $E\left(Y_{1: n}\right)$ and $E\left(Y_{n: n}\right)$. From the min-max plots, the distribution can exhibit symmetrical, left-skewed, and right-skewed shapes.


Figure 4. Min-max plots for some parameter values.

## 4. Bivariate Extension

Researchers may be interested in modeling the dependence between two (quantitative) measures in a dataset. For instance, one may be interested in modeling the relationship between age and the body mass index of individuals. Bivariate distributions have been used in reliability analysis, queuing theory, finance, and insurance risk analysis, among others, to study interdependency (see [26]). In this section, the bivariate extension of the BTCPE (BEBTCPE) distribution is proposed following the strategy developed in [26,27]. Given a bivariate continuous random vector $(X, Y)$, the CDF of the BEBTCPE distribution with parameters $\alpha, \lambda, \delta_{1}, \delta_{2}, \delta_{3}$, where $\alpha>0, \lambda>0,-1<\delta_{1}+\delta_{3}<1$, $-1<\delta_{2}+\delta_{3}<1,0<x<1$ and $0<y<1$, is given by

$$
\begin{equation*}
F_{X Y}(x, y ; \eta)=\frac{\left(1-\frac{4}{\pi} \arctan \left[\left(1-x^{\lambda}\right)^{\alpha}\right]\right)\left(1-\frac{4}{\pi} \arctan \left[\left(1-y^{\lambda}\right)^{\alpha}\right]\right)}{\left\{1+\left(\delta_{1}+\delta_{3}\right) \frac{4}{\pi} \arctan \left[\left(1-x^{\lambda}\right)^{\alpha}\right]+\left(\delta_{2}+\delta_{3}\right) \frac{4}{\pi} \arctan \left[\left(1-y^{\lambda}\right)^{\alpha}\right]\right\}^{-1}} \tag{12}
\end{equation*}
$$

where $\boldsymbol{\eta}=\left(\alpha, \beta, \delta_{1}, \delta_{2}, \delta_{3}\right)^{\mathrm{T}}$. The parameters $\delta_{1}, \delta_{2}$ and $\delta_{3}$ quantify the dependence between the two variables of a BEBTCPE random vector. The plots of the CDF for the following parameter values are shown in Figure 5:
(a) $\alpha=3.5, \lambda=8.2, \delta_{1}=0.3, \delta_{2}=0.1, \delta_{3}=0.3$;
(b) $\quad \alpha=2.5, \lambda=0.8, \delta_{1}=0.5, \delta_{2}=0.4, \delta_{3}=0.2$ and
(c) $\quad \alpha=0.5, \lambda=4.8, \delta_{1}=-0.3, \delta_{2}=-0.7, \delta_{3}=-0.1$.
(a)
(b)
(c)

Figure 5. CDF plots of the BEBTCPE distribution.
We notice various concave and convex shapes from these plots.
The corresponding bivariate PDF is given by

$$
\begin{equation*}
f_{X Y}(x, y ; \eta)=\frac{(4 \alpha \lambda / \pi)^{2}(x y)^{\lambda-1}\left(1-x^{\lambda}-y^{\lambda}+(x y)^{\lambda}\right)^{\alpha-1}\left[1+\left(1-x^{\lambda}\right)^{2 \alpha}\right]^{-1}\left[1+\left(1-y^{\lambda}\right)^{2 \alpha}\right]^{-1}}{\left\{1+\left(\delta_{1}+\delta_{3}\right) \frac{8}{\pi} \arctan \left[\left(1-x^{\lambda}\right)^{\alpha}\right]+\left(\delta_{2}+\delta_{3}\right) \frac{8}{\pi} \arctan \left[\left(1-y^{\lambda}\right)^{\alpha}\right]\right\}^{-1}} \tag{13}
\end{equation*}
$$

Figure 6 shows the bivariate PDF plots of the BEBTCPE distribution for the following parameter values:
(a) $\quad \alpha=3.5, \lambda=8.2, \delta_{1}=0.3, \delta_{2}=0.1, \delta_{3}=0.3$;
(b) $\quad \alpha=2.5, \lambda=0.8, \delta_{1}=0.5, \delta_{2}=0.4, \delta_{3}=0.2$ and
(c) $\quad \alpha=0.5, \lambda=4.8, \delta_{1}=-0.3, \delta_{2}=-0.7, \delta_{3}=-0.1$.

The first graph displays a non-monotonic shape whereas the other two exhibit monotonic shapes, illustrating the versatility in the bivariate modeling sense.


Figure 6. PDF plots of the BEBTCPE distribution.

## 5. Parameter Estimation Methods

This section presents nine estimation methods for estimating the parameters of the BTCPE distribution. These include the maximum likelihood (ML) estimation (MLE), ordinary least squares (OLS), weighted least squares (WLS), Cramér-von Mises (CVM), percentile (PC) estimation, Anderson-Darling (AD) methods, and maximum and minimum product spacing methods.

### 5.1. Maximum Likelihood Estimation

One of the most common methods used for estimating the parameters of a developed model is the MLE method. Suppose that $Y$ follows the BTCPE distribution, with $\vartheta=(\alpha, \lambda)^{\mathrm{T}}$ as the parameter vector. For a single observation $y$ of $Y$, the log-likelihood function $\ell=\ell(\boldsymbol{\vartheta})$ is

$$
\begin{equation*}
\ell=\log \left(\frac{4 \alpha \lambda}{\pi}\right)+(\lambda-1) \log (y)+(\alpha-1) \log \left(1-y^{\lambda}\right)-\log \left(1+\left(1-y^{\lambda}\right)^{2 \alpha}\right) \tag{14}
\end{equation*}
$$

To obtain the estimates of the parameters for the single observation, the first partial derivative of Equation (14) with respect to the parameters needs to be derive. Here, we obtain

$$
\begin{equation*}
\frac{\partial \ell}{\partial \alpha}=\frac{1}{\alpha}+\log \left(1-y^{\lambda}\right)-\frac{2\left(1-y^{\lambda}\right)^{2 \alpha} \log \left(1-y^{\lambda}\right)}{1+\left(1-y^{\lambda}\right)^{2 \alpha}} \tag{15}
\end{equation*}
$$

and

$$
\begin{equation*}
\frac{\partial \ell}{\partial \lambda}=\frac{1}{\lambda}+\log (y)-\frac{y^{\lambda}(\alpha-1) \log (y)}{1-y^{\lambda}}+\frac{2 \alpha y^{\lambda}\left(1-y^{\lambda}\right)^{2 \alpha-1} \log (y)}{1+\left(1-y^{\lambda}\right)^{2 \alpha}} . \tag{16}
\end{equation*}
$$

Given that $y_{1}, y_{2}, \ldots, y_{n}$ are (independent and identically) observations from $n$ BTCPE random variables, then the total log-likelihood function is given by $\ell_{n}^{*}=\sum_{i=1}^{n} \ell_{i}(\boldsymbol{\vartheta})$, where $\ell_{i}(\boldsymbol{\vartheta}), i=1,2, \ldots, n$ is defined in Equation (14) with $y=y_{i}$. The estimates of the parameters can be obtained by maximizing the total log-likelihood function directly using MATLAB, MATHEMATICA and R software. In this study, the R software is used [28]. Alternatively, the estimates of the parameters can be obtained by equating the first partial derivatives with respect to the parameters to zero and solving the resulting system of equations simultaneously. However, since the resulting system of equations does not have a closed form, the nonlinear system of equations $\left(\frac{\partial \ell_{n}^{*}}{\partial \alpha}, \frac{\partial \ell_{n}^{*}}{\partial \lambda}\right)^{\mathrm{T}}=(0,0)^{\mathrm{T}}$ is solved numerically to obtain the estimates of the parameters.

### 5.2. Ordinary and Weighted Least Squares Estimation

Suppose that $y_{(1)}, y_{(2)}, \ldots, y_{(n)}$ are ordered observations from $n$ BTCPE random variables. The OLS estimates of the parameters $\hat{\alpha}_{L S E}$ and $\hat{\lambda}_{L S E}$ are obtained by minimizing the following function:

$$
\begin{equation*}
\operatorname{LSE}(\alpha, \lambda)=\sum_{i=1}^{n}\left[\left(1-\frac{4}{\pi} \arctan \left[\left(1-y_{(i)}^{\lambda}\right)^{\alpha}\right]\right)-\frac{i}{n+1}\right]^{2}, \tag{17}
\end{equation*}
$$

with respect to the parameters $\alpha$ and $\lambda$. On the other hand, the OLS estimates can be obtained by numerically solving the following nonlinear equations:

$$
\begin{equation*}
\sum_{i=1}^{n}\left[\left(1-\frac{4}{\pi} \arctan \left[\left(1-y_{(i)}^{\lambda}\right)^{\alpha}\right]\right)-\frac{i}{n+1}\right] \Delta_{s}\left(y_{(i)} \mid \alpha, \lambda\right)=0, s=1,2 \tag{18}
\end{equation*}
$$

where

$$
\Delta_{1}\left(y_{(i)} \mid \alpha, \lambda\right)=-\frac{8\left(1-y_{(i)}^{\lambda}\right)^{\alpha} \log \left(1-y_{(i)}^{\lambda}\right)}{\pi\left[1+\left(1-y_{(i)}^{\lambda}\right)^{2 \alpha}\right]}
$$

and

$$
\Delta_{2}\left(y_{(i)} \mid \alpha, \lambda\right)=\frac{8 \alpha y_{(i)}^{\lambda}\left(1-y_{(i)}^{\lambda}\right)^{\alpha-1} \log \left(y_{(i)}\right)}{\pi\left[1+\left(1-y_{(i)}^{\lambda}\right)^{2 \alpha}\right]}
$$

The WLS estimates $\hat{\alpha}_{W L S}$ and $\hat{\lambda}_{W L S}$ are obtained by minimizing the following function:

$$
\begin{equation*}
W L S(\alpha, \lambda)=\sum_{i=1}^{n} \frac{(n+1)^{2}(n+2)}{i(n-i+1)}\left[\left(1-\frac{4}{\pi} \arctan \left[\left(1-y_{(i)}^{\lambda}\right)^{\alpha}\right]\right)-\frac{i}{n+1}\right]^{2} \tag{19}
\end{equation*}
$$

with respect to the parameters $\alpha$ and $\lambda$. Alternatively, the WLS estimates can be obtained by numerically solving the following nonlinear equations:

$$
\begin{equation*}
\sum_{i=1}^{n} \frac{(n+1)^{2}(n+2)}{i(n-i+1)}\left[\left(1-\frac{4}{\pi} \arctan \left[\left(1-y_{(i)}^{\lambda}\right)^{\alpha}\right]\right)-\frac{i}{n+1}\right] \Delta_{s}\left(y_{(i)} \mid \alpha, \lambda\right)=0, s=1,2 \tag{20}
\end{equation*}
$$

where $\Delta_{s}\left(x_{(i)} \mid \alpha, \lambda\right), s=1,2$ are defined above.

### 5.3. Cramér-Von Mises Estimation

Let $y_{(1)}, y_{(2)}, \ldots, y_{(n)}$ be ordered observations from $n$ BTCPE random variables. The CVM estimates of the parameters $\hat{\alpha}_{C V M}$ and $\hat{\lambda}_{C V M}$ are obtained by minimizing the following function:

$$
\begin{equation*}
\operatorname{CVM}(\alpha, \lambda)=\frac{1}{12 n}+\sum_{i=1}^{n}\left[\left(1-\frac{4}{\pi} \arctan \left[\left(1-y_{(i)}^{\lambda}\right)^{\alpha}\right]\right)-\frac{2 i-1}{2 n}\right]^{2} \tag{21}
\end{equation*}
$$

with respect to the parameters $\alpha$ and $\lambda$. The estimates of the parameters can also be obtained by numerically solving the following equations:

$$
\begin{equation*}
\sum_{i=1}^{n}\left[\left(1-\frac{4}{\pi} \arctan \left[\left(1-y_{(i)}^{\lambda}\right)^{\alpha}\right]\right)-\frac{2 i-1}{2 n}\right] \Delta_{s}\left(y_{(i)} \mid \alpha, \lambda\right)=0, s=1,2 \tag{22}
\end{equation*}
$$

where $\Delta_{s}\left(y_{(i)} \mid \alpha, \lambda\right), s=1,2$ are given above.

### 5.4. Anderson-Darling Estimation

Another minimum distance estimation method is the AD estimation technique. Let $y_{(1)}, y_{(2)}, \ldots, y_{(n)}$ be ordered observations from $n$ BTCPE random variables. The AD estimates for the parameters of the BTCPE distribution are obtained by minimizing the following function:

$$
\begin{equation*}
A D(\alpha, \lambda)=-n-\frac{1}{n} \sum_{i=1}^{n}(2 i-1)\left[\log \left(1-\frac{4}{\pi} \arctan \left[\left(1-y_{(i)}^{\lambda}\right)\right]\right)+\log \left(\frac{4}{\pi} \arctan \left[\left(1-y_{(i)}^{\lambda}\right)\right]\right)\right], \tag{23}
\end{equation*}
$$

with respect to the parameters $\alpha$ and $\lambda$.

### 5.5. Percentile Estimation

The PC estimation approach is another method of estimating the parameters of a given model. Let $y_{(1)}, y_{(2)}, \ldots, y_{(n)}$ be ordered observations from $n$ BTCPE random variables and
$u_{i}=i /(n+1)$ be an unbiased estimate of $F_{Y}\left(y_{(i)} ; \alpha, \lambda\right)$. The PC estimates of the parameters of the BTCPE distribution are obtained by minimizing the following function:

$$
\begin{equation*}
P C(\alpha, \lambda)=\sum_{i=1}^{n}\left[y_{(i)}-\left\{1-\left(\tan \left[\frac{\pi}{4}\left(1-u_{i}\right)\right]\right)^{1 / \alpha}\right\}^{1 / \lambda}\right]^{2}, \tag{24}
\end{equation*}
$$

with respect to the parameters $\alpha$ and $\lambda$.

### 5.6. Maximum and Minimum Product Spacing Estimation

An alternative parameter estimation technique which is based on the Kullback-Leibler information measure is the maximum product spacing (MPS). Let $y_{(1)}, y_{(2)}, \ldots, y_{(n)}$ be ordered observations from $n$ BTCPE random variables. Consider the uniform spacing

$$
\begin{aligned}
& D_{i}=F_{Y}\left(y_{(i)} ; \alpha, \lambda\right)-F_{Y}\left(y_{(i-1)} ; \alpha, \lambda\right) \\
= & \frac{4}{\pi} \arctan \left[\left(1-y_{(i-1)}^{\lambda}\right)\right]-\frac{4}{\pi} \arctan \left[\left(1-y_{(i)}^{\lambda}\right)\right]
\end{aligned}
$$

where $F_{Y}\left(y_{(0)} ; \alpha, \lambda\right)=0, F_{Y}\left(y_{(n+1)} ; \alpha, \lambda\right)=1$ and $D_{0}(\alpha, \lambda)+D_{1}(\alpha, \lambda)+\ldots+D_{n+1}(\alpha, \lambda)=$ 1. The estimates of the parameters are obtained via the MPS approach by maximizing the logarithm of the geometric mean of the spacing defined by

$$
\begin{equation*}
\operatorname{MPS}(\alpha, \lambda)=\frac{1}{n+1} \sum_{i=1}^{n+1} \log D_{i}(\alpha, \lambda) \tag{25}
\end{equation*}
$$

with respect to the parameters $\alpha$ and $\lambda$.
Additionally, the minimum spacing distance (MSD) estimates for the parameters $\alpha$ and $\lambda$ are obtained by minimizing the following function:

$$
\begin{equation*}
\operatorname{MSD}(\alpha, \lambda)=\sum_{i=1}^{n+1} \vartheta\left(D_{i}(\alpha, \lambda), \frac{1}{n+1}\right) \tag{26}
\end{equation*}
$$

where $\vartheta(x, y)$ is an appropriate distance, with respect to the parameters $\alpha$ and $\lambda$. Although different choices of $\vartheta(x, y)$ exist, in this study the absolute distance $|x-y|$ and the absolute$\log$ distance $|\log x-\log y|$ are utilized. Thus, the minimum spacing absolute distance (MSAD) and minimum spacing absolute-log distance (MSALD) estimates are, respectively, obtained by minimizing the following functions:

$$
\begin{equation*}
\operatorname{MSAD}(\alpha, \lambda)=\sum_{i=1}^{n+1}\left|D_{i}(\alpha, \lambda)-\frac{1}{n+1}\right| \tag{27}
\end{equation*}
$$

and

$$
\begin{equation*}
\operatorname{MSALD}(\alpha, \lambda)=\sum_{i=1}^{n+1}\left|\log D_{i}(\alpha, \lambda)-\log \frac{1}{n+1}\right| \tag{28}
\end{equation*}
$$

where $D_{i}(\alpha, \lambda) \neq \frac{1}{n+1}$ and $\log D_{i}(\alpha, \lambda) \neq \log \frac{1}{n+1}$.

## 6. Simulation

In this section, simulation experiments are carried out to assess how well the proposed parameters of the BTCPE distribution have been estimated. The experiments are carried out with the following two different parameter combinations: $\alpha=4.1, \lambda=2.5$ and $\alpha=$ $3.1, \lambda=8.5$. The experiments are replicated 5000 times with the following different sample sizes: $n=25,75,125,175$ and 225 . The bias ( AB ) and root mean square error (RMSE) of the estimates are then computed and compared.

The AB and RMSE are, respectively, computed using

$$
A B=\frac{1}{R} \sum_{i=1}^{R}\left(\hat{\vartheta}_{i}-\vartheta\right)
$$

and

$$
R M S E=\sqrt{\frac{1}{R} \sum_{i=1}^{R}\left(\hat{\vartheta}_{i}-\vartheta\right)^{2}}
$$

where $\hat{\vartheta}$ is either $\hat{\alpha}$ or $\hat{\lambda}$ and $R=5000$ is used in this study.
From Tables 2 and 3, most of the estimates have their ABs and RMSEs decreasing as the sample size increases. This is an indication that most of the estimates exhibit the consistency property. From Table 2, it can be observed that for sample sizes 25, 75 and 125 the PC estimate is the best for $\alpha$ and, for the sample sizes 175 and 225, the MLE is the best for $\alpha$. For the parameter $\lambda$, the PC estimate is the best for the sample size 25 and the MLE is the best for $75,125,175$ and 225. In Table 3, for sample sizes 25 and 75 the AD estimate is the best for the parameter $\alpha$ and the MLE is the best for 125,175 and 225 . For the parameter $\lambda$, the MLE is the best for sample sizes $25,125,175$ and 225 . The AD estimate is best for $\lambda$ when the sample size is 75 .

Table 2. AB and RMSE for $\alpha=4.1$ and $\lambda=2.5$.

| Parameter | $n$ | AB |  |  |  |  |  |  |  |  | RMSE |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | MLE | MPS | MADS | MALDS | OLS | WLS | CVM | AD | PC | MLE | MPS | MADS | MALDS | OLS | WLS | CVM | AD | PC |
| $\alpha$ | 25 | 0.7980 | 2.2327 | -2.3189 | 0.531 | 0.3530 | -2.6423 | 1.1477 | 0.4713 | -0.3155 | 2.2233 | 3.5728 | 2.9322 | 3.3391 | 2.4457 | 2.6729 | 3.5377 | 1.9964 | 1.4972 |
|  | 75 | 0.2140 | 0.7443 | -1.8442 | 0.0634 | 0.1365 | -3.1632 | 0.3415 | 0.1180 | -0.1694 | 0.9157 | 1.3139 | 2.4241 | 1.1330 | 1.1489 | 3.1657 | 1.2820 | 0.9535 | 0.8506 |
|  | 125 | 0.1342 | 0.4372 | $-1.3088$ | -0.0031 | 0.0472 | -3.3268 | 0.1337 | 0.0713 | -0.0783 | 0.6843 | 0.8313 | 1.9860 | 0.7795 | 0.8149 | 3.3279 | 0.8159 | 0.7025 | 0.6641 |
|  | 175 | 0.0914 | 0.2987 | $-0.8738$ | -0.0272 | 0.0460 | -2.1721 | 0.1164 | 0.0657 | -0.0484 | 0.5365 | 0.6791 | 1.5544 | 0.6323 | 0.6845 | 2.1990 | 0.6955 | 0.5941 | 0.5393 |
|  | 225 | 0.0677 | 0.2509 | -0.6976 | 0.0062 | 0.0301 | 3.0791 | 0.1096 | 0.0365 | -0.0623 | 0.4841 | 0.5505 | 1.3266 | 0.5926 | 0.5906 | 3.3377 | 0.6147 | 0.5240 | 0.4860 |
| $\lambda$ | 25 | 0.1871 | 0.5436 | -1.1017 | 0.0300 | $-0.0075$ | $-1.3344$ | 0.2060 | 0.0670 | -0.1737 | 0.6038 | 0.8201 | 1.3862 | 0.7382 | 0.6538 | 1.3687 | 0.7340 | 0.5749 | 0.5401 |
|  | 75 | 0.0478 | 0.2197 | -0.8939 | -0.0079 | 0.0078 | -2.0003 | 0.0802 | 0.1185 | -0.0672 | 0.3089 | 0.4026 | 1.2090 | 0.3996 | 0.3692 | 2.0029 | 0.3886 | 0.3379 | 0.3175 |
|  | 125 | 0.0378 | 0.1293 | -0.6271 | $-0.0160$ | $-0.0055$ | -2.1461 | 0.0305 | 0.0146 | -0.0448 | 0.2407 | 0.2740 | 0.9681 | 0.2987 | 0.2752 | 2.1472 | 0.2798 | 0.2560 | 0.2466 |
|  | 175 | 0.0233 | 0.0866 | $-0.3986$ | -0.0139 | 0.0034 | -1.6394 | 0.0280 | 0.0114 | -0.0267 | 0.1959 | 0.2314 | 0.7264 | 0.2315 | 0.2372 | 1.6455 | 0.2383 | 0.2134 | 0.2008 |
|  | 225 | 0.0208 | 0.0820 | -0.3101 | 0.0021 | 0.0003 | 0.1937 | 0.0230 | 0.0007 | -0.0225 | 0.1810 | 0.1939 | 0.6057 | 0.2196 | 0.2079 | 0.3066 | 0.2124 | 0.1874 | 0.1835 |

Table 3. AB and RMSE for $\alpha=3.1$ and $\lambda=8.5$.

| Parameter | $n$ | AB |  |  |  |  |  |  |  |  | RMSE |  |  |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | MLE | MPS | MADS | MALDS | OLS | WLS | CVM | AD | PC | MLE | MPS | MADS | MALDS | OLS | WLS | CVM | AD | PC |
| $\alpha$ | 25 | 0.5120 | 1.5281 | -1.3158 | 0.3712 | 0.2359 | -1.9121 | 0.7701 | 0.2725 | -0.6748 | 1.4793 | 2.5597 | 2.0817 | 1.9689 | 1.6086 | 1.9366 | 2.4325 | 1.3231 | 1.4204 |
|  | 75 | 0.2081 | 0.5190 | -0.9456 | 0.0477 | 0.0264 | -2.2924 | 0.1973 | 0.0989 | -0.3302 | 0.6848 | 0.8671 | 1.5913 | 0.7865 | 0.7088 | 2.2949 | 0.8122 | 0.6294 | 0.8145 |
|  | 125 | 0.0994 | 0.3218 | -0.6895 | 0.0570 | 0.0153 | -2.4255 | 0.1020 | 0.0757 | -0.2704 | 0.4778 | 0.6228 | 1.2854 | 0.5657 | 0.5532 | 2.4266 | 0.5644 | 0.5109 | 0.6262 |
|  | 175 | 0.0867 | 0.2259 | -0.5478 | 0.0107 | 0.0240 | -1.3857 | 0.0882 | 0.0554 | -0.2187 | 0.4077 | 0.4554 | 1.0538 | 0.4821 | 0.4813 | 1.4810 | 0.4940 | 0.4163 | 0.5242 |
|  | 225 | 0.0461 | 0.1719 | -0.4192 | 0.0007 | 0.0195 | 2.2166 | 0.0555 | 0.0200 | -0.1735 | 0.3331 | 0.3880 | 0.8460 | 0.4021 | 0.4133 | 2.3785 | 0.4184 | 0.3477 | 0.4768 |
| $\lambda$ | 25 | 0.5725 | 1.9602 | -3.5282 | 0.1675 | 0.1107 | -4.6443 | 0.7293 | 0.2368 | -1.6432 | 2.0951 | 3.1555 | 4.8358 | 2.6753 | 2.3883 | 4.7703 | 2.6517 | 2.1203 | 2.6346 |
|  | 75 | 0.2957 | 0.8211 | -2.4563 | -0.0292 | -0.0416 | -6.9243 | 0.2449 | 0.0825 | -0.7220 | 1.1627 | 1.4130 | 3.9050 | 1.3676 | 1.2655 | 6.9330 | 1.3246 | 1.1047 | 1.4877 |
|  | 125 | 0.1098 | 0.4964 | -1.6975 | 0.0128 | -0.0435 | $-7.3960$ | 0.1015 | 0.0784 | -0.5259 | 0.8837 | 1.0321 | 3.0631 | 1.0694 | 0.9899 | 7.3994 | 0.9670 | 0.9327 | 1.1482 |
|  | 175 | 0.1182 | 0.3504 | -1.2348 | -0.0232 | 0.0022 | -5.5678 | 0.1061 | 0.0622 | -0.4011 | 0.7361 | 0.7786 | 2.4368 | 0.8864 | 0.8608 | 5.5937 | 0.8687 | 0.7734 | 0.9365 |
|  | 225 | 0.0631 | 0.2843 | -0.9196 | -0.0034 | 0.0015 | 0.6715 | 0.043 | 0.0249 | -0.3171 | 0.6223 | 0.7091 | 1.9241 | 0.7604 | 0.7452 | 1.0898 | 0.7515 | 0.6590 | 0.8305 |

## 7. Applications

Three applications of the BTCPE distribution are illustrated in this section, and its performance is compared to other competitive distributions defined in the unit interval. The performance of the BTCPE distribution was compared with that of the beta, unit Burr-III (UBIII) [29], bounded M-O extended exponential (BMOEE) [30], unit Gompertz (UG) [18], unit Lindley (UL) [17], unit Weibull (UW) [20] and unit-improved second-degree Lindley (UISDL) [31] distributions. The Akaike information criterion (AIC), Bayesian information criterion (BIC), Anderson-Darling (AD) test, and Cramér-von Mises (CVM) are the model selection techniques employed in arriving at the best model. For these selection techniques, the best model is the one with the smallest test statistic. The datasets represent the mortality rate of COVID-19 patients in Canada and the United Kingdom (UK), and the recovery rate of COVID-19 patients in Spain. The first two datasets were recently reported by [8].

The first dataset is the mortality rate for UK from 1 December 2020 to 29 January 2021. The data are: $0.1292,0.3805,0.4049,0.2564,0.3091,0.2413,0.1390,0.1127,0.3547,0.3126$, $0.2991,0.2428,0.2942,0.0807,0.1285,0.2775,0.3311,0.2825,0.2559,0.2756,0.1652,0.1072$, $0.3383,0.3575,0.2708,0.2649,0.0961,0.1565,0.1580,0.1981,0.4154,0.3990,0.2483,0.1762$, $0.1760,0.1543,0.3238,0.3771,0.4132,0.4602,0.352,0.1882,0.1742,0.4033,0.4999,0.3930$, $0.3963,0.3960,0.2029,0.1791,0.4768,0.5331,0.3739,0.4015,0.3828,0.1718,0.1657,0.4542$, $0.4772,0.3402$.

The second dataset denotes the mortality rate for Canada from 1 November to 26 December 2020. The data are: $0.1622,0.1159,0.1897,0.1260,0.3025,0.2190,0.2075,0.2241$, $0.2163,0.1262,0.1627,0.2591,0.1989,0.3053,0.2170,0.2241,0.2174,0.2541,0.1997,0.3333$, $0.2594,0.2230,0.2290,0.1536,0.2024,0.2931,0.2739,0.2607,0.2736,0.2323,0.1563,0.2677$, $0.2181,0.3019,0.2136,0.2281,0.2346,0.1888,0.2729,0.2162,0.2746,0.2936,0.3259,0.2242$, $0.1810,0.2679,0.2296,0.2992,0.2464,0.2576,0.2338,0.1499,0.2075,0.1834,0.3347,0.2362$.

The third dataset constitutes the recovery rates of COVID-19 patients in Spain from 3 March to 7 May 2020. The dataset can be found in [1] and are: $0.6670,0.5000,0.5000,0.4286$, $0.7500,0.6531,0.5161,0.7895,0.7689,0.6873,0.5200,0.7251,0.6375,0.6078,0.6289,0.5712$, $0.5923,0.6061,0.5924,0.5921,0.5592,0.5954,0.6164,0.6455,0.6725,0.6838,0.6850,0.6947$, $0.7210,0.7315,0.7412,0.7508,0.7519,0.7547,0.7645,0.7715,0.7759,0.7807,0.7838,0.7847$, $0.7871,0.7902,0.7934,0.7913,0.7962,0.7971,0.7977,0.8007,0.8038,0.8289,0.8322,0.8354$, $0.8371,0.8387,0.8456,0.8490,0.8535,0.8547,0.8564,0.8580,0.8604,0.8628,0.6586,0.7070$, 0.7963, 0.8516 .

The ML estimates of the parameters are estimated using the bbmle package in R [32]. The initial values of the parameters of the fitted distributions used for the optimization are obtained using the GenSA package in R [33]. Table 4 displays the descriptive statistics for COVID-19 mortality for the UK and Canada, as well as the recovery rate for Spain. The datasets are platykurtic due to the negative excess kurtosis. The UK mortality is rightskewed and that of Canada is left-skewed. The recovery rate for Spain is also left-skewed. This is affirmed by the boxplot of the datasets shown in Figure 7.

Table 4. Descriptive statistics for datasets.

| Country | Minimum | Maximum | Mean | Skewness | Kurtosis |
| :---: | :---: | :---: | :---: | :---: | :---: |
| UK | 0.0807 | 0.5331 | 0.2888 | 0.0476 | -1.1034 |
| Canada | 0.1159 | 0.3347 | 0.2305 | -0.0850 | -0.4402 |
| Spain | 0.4286 | 0.8628 | 0.7240 | -0.6890 | -0.4761 |

### 7.1. UK COVID-19 Mortality

Table 5 presents ML estimates of the parameters and their corresponding standard errors in brackets, the log-likelihood ( $\ell$ ), AIC, BIC, AD, and CVM for the fitted distributions. Given that it has the lowest values for the AIC, BIC, AD, and CVM and the maximum loglikelihood, the BTCPE distribution offers the best fit to the UK mortality dataset.


Figure 7. Boxplots of COVID-19 datasets.
Table 5. Parameter estimates and model selection criteria for UK.

| Model | Parameter | $\ell$ | AIC | BIC | AD | CVM |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| BTCPE | $\alpha=16.6904(5.2798)$ <br> $\lambda=2.3884(0.2865)$ | 45.4400 | -86.8726 | -82.6840 | 0.6494 | 0.1049 |
| Beta | $\alpha=4.0502(0.7128)$ <br> $\beta=10.0132(1.8287)$ | 45.4000 | -86.7958 | -82.6071 | 0.7356 | 0.1280 |
| UBIII | $\alpha=0.0757(0.0383)$ <br> $\beta=13.3804(6.5631)$ | 38.9000 | -73.8075 | -69.6188 | 2.8948 | 0.5248 |
| BMOEE | $\alpha=105.2655(59.9004)$ <br> $\beta=3.5949(0.4092)$ | 40.7200 | -77.4396 | -73.2509 | 1.1465 | 0.1698 |
| UW | $\alpha=0.2834(0.0602)$ <br> $\beta=3.1228(0.3047)$ | 42.5600 | -81.1208 | -76.9322 | 1.0656 | 0.1820 |
| UG | $\alpha=686.3600\left(2.2295 \times 10^{-10}\right)$ <br> $\beta=0.0011\left(1.4051 \times 10^{-4}\right)$ | 2.8400 | -1.6760 | 2.5127 | 12.2290 | 2.4707 |
| UL | $\alpha=2.8293(0.3029)$ | 32.3800 | -62.7533 | -60.6590 | 4.4878 | 0.7574 |
| UISDL | $\alpha=3.4259(0.3151)$ | 33.6100 | -65.2142 | -63.1198 | 3.9972 | 0.6545 |

Figure 8 displays the empirical and fitted PDFs and CDFs of the various distributions used to model the UK mortality dataset. The figure gives an indication that the BTCPE distribution provides a good fit to the dataset compared to the other models.

Figure 9 is the probability-probability (P-P) plots of the fitted distributions. Figure 9 once more shows that the BTCPE distribution fits the UK drought mortality well because the points cluster along the diagonal.

The profile log-likelihood plots for the estimated parameter values of the BTCPE distribution for the UK mortality data are shown in Figure 10. From the plots, it can be observed that the estimated values are the true maxima.

### 7.2. Canada COVID-19 Mortality

Table 6 presents ML estimates of the parameters and their corresponding standard errors in brackets and model selection criteria for the fitted distributions. The BTCPE distribution again provides the best fit to the Canada mortality dataset since it has the highest log-likelihood and the lowest values of the AIC, BIC, AD, and CVM.


Figure 8. Empirical and fitted PDFs (left) and CDFs (right) of UK dataset.


Figure 9. P-P plots for UK drought mortality.


Figure 10. Profile log-likelihood plots for estimated parameters of BTCPE for UK.
Table 6. Parameter estimates and model selection criteria for Canada.

| Model | Parameter | $\ell$ | AIC | BIC | AD | CVM |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| BTCPE | $\alpha=622.2064(399.8188)$ <br> $\lambda=4.5085(0.4837)$ | 86.4400 | -168.8806 | -164.8299 | 0.3767 | 0.0689 |
| Beta | $\alpha=14.5128(2.7128)$ <br> $\beta=48.4900(9.1745)$ | 85.9400 | -167.8800 | -163.8293 | 0.4398 | 0.0692 |
| UBIII | $\alpha=0.0080(0.0011)$ <br> $\beta=101.7700\left(8.4127 \times 10^{-8}\right)$ | 30.8900 | -57.7749 | -53.7242 | 14.8770 | 3.1113 |
|  | $\alpha=2822.9776\left(3.3087 \times 10^{-5}\right)$ <br> $\beta=5.4444(0.1439)$ | 80.6700 | -157.3394 | -153.2887 | 1.5514 | 0.2327 |
| UW | $\alpha=0.0552(0.0193)$ <br> $\beta=6.1602(0.5868)$ | 79.9500 | -155.9080 | -151.8573 | 1.4890 | 0.2389 |
|  | $\alpha=628.3885\left(2.4072 \times 10^{-10}\right)$ |  |  |  |  |  |
| $\beta=0.0011\left(1.4212 \times 10^{-4}\right)$ |  |  |  |  |  |  |

Figure 11 shows the empirical and fitted PDFs and CDFs of the various distributions used to model the Canada drought mortality dataset. The figure gives an indication that the BTCPE distribution provides a better fit to the drought mortality for Canada than the other models, as it mimics the empirical PDF and CDF of the dataset better than the other models.


Figure 11. Empirical and fitted PDFs (left) and CDFs (right) of Canada dataset.

Figure 12 shows the P-P plots of the fitted models. Figure 12 gives an indication that the BTCPE distribution provides a good fit to the Canada mortality as the points cluster along the diagonal.


Figure 12. P-P plots for Canada mortality.
Figure 13 displays the profile log-likelihood plots for the estimated parameter values of the BTCPE distribution for the Canada mortality data. It can be observed from the plots that the estimates are unique and represent the true maxima.


Figure 13. Profile log-likelihood plots for estimated parameters of BTCPE for Canada.

### 7.3. Spain COVID-19 Recovery Rate

The ML estimates of the parameters and their corresponding standard errors in brackets and model selection criteria for the fitted distributions are shown in Table 7. Because it has the lowest values for the AIC, BIC, AD, and CVM and the maximum log-likelihood, the BTCPE distribution again offers the best fit to the Spain recovery rate dataset.

Table 7. Parameter estimates and model selection criteria for Canada.

| Model | Parameter | $\ell$ | AIC | BIC | AD | CVM |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| BTCPE | $\alpha=7.1385(1.7764)$ <br> $\lambda=7.1961(0.9033)$ | 58.7500 | -113.4953 | -109.1160 | 0.8770 | 0.1363 |
| Beta | $\alpha=12.7943(2.2291)$ <br> $\beta=4.8994(0.8270)$ | 57.5700 | -111.1489 | -106.7692 | 1.0520 | 0.1783 |
| UBIII | $\alpha=5.4398(0.7948)$ <br> $\beta=2.0613(0.1723)$ | 53.8000 | -103.5927 | -99.2134 | 1.3725 | 0.2209 |
| BMOEE | $\alpha=22.1286(9.9041)$ <br> $\beta=10.0043(1.2381)$ | 51.4600 | -98.9276 | -94.5483 | 1.4958 | 0.2100 |
| UW | $\alpha=8.6445(1.6973)$ <br> $\beta=2.2320(0.2036)$ | 53.9700 | -103.9316 | -99.5523 | 1.3830 | 0.2238 |
| UG | $\alpha=0.2792(0.1059)$ <br> $\beta=3.8482(0.6025)$ | 46.0300 | -88.0569 | -83.6776 | 2.4709 | 0.3691 |
| UL | $\alpha=0.5200(0.0466)$ | 46.1100 | -90.2298 | -88.0402 | 4.2480 | 0.6736 |
| UISDL | $\alpha=0.7403(0.0539)$ | 52.0400 | -102.0717 | -99.8820 | 2.3450 | 0.3194 |

The empirical and fitted PDFs and CDFs of the various distributions used to model the Spain recovery rate dataset are shown in Figure 14. It can be seen that the BTCPE distribution provides a better fit to the recovery rate data than the other models.


Figure 14. Empirical and fitted PDFs (left) and CDFs (right) of Spain dataset.
The P-P plots of the fitted models for the recovery rate data are displayed in Figure 15. The plots indicate that the BTCPE distribution provides a good fit to the recovery rate data as the points cluster along the diagonal.


Figure 15. P-P plots for Spain recovery data.
The profile log-likelihood plots for the estimated parameter values of the BTCPE distribution for the recovery rate data are shown in Figure 16. The plots suggest that the estimates are unique and represent the true maxima.


Figure 16. Profile log-likelihood plots for estimated parameters of BTCPE for Spain.

## 8. Quantile Regression

When the response variable defined in the unit interval is skewed or contaminated with outliers, the beta regression model, which models the conditional mean of the response variable, is no longer reliable. A robust regression model is needed to model
the effects of the covariates on the response variable. In this study, a quantile regression model is proposed for modeling the conditional quantile of the response variable. Given the quantile function of the BTCPE distribution, the PDF of the BTCPE distribution can be re-parameterized in terms of its $u^{\text {th }}$ quantile as $\rho=Q(u ; \alpha, \lambda), \rho \in[0,1]$. If $\lambda=\log \left(1-(\tan [\pi(1-u) / 4])^{1 / \alpha}\right) / \log (\rho)$, then the re-parameterized PDF is

$$
\begin{gather*}
f_{Y}(y ; \alpha, \lambda)= \\
\frac{4 \alpha\left(\log \left(1-(\tan [\pi(1-u) / 4])^{1 / \alpha}\right) / \log (\rho)\right) y^{\left(\log \left(1-(\tan [\pi(1-u) / 4])^{1 / \alpha}\right) / \log (\rho)\right)-1}\left(1-y^{\left(\log \left(1-(\tan [\pi(1-u) / 4])^{1 / \alpha}\right) / \log (\rho)\right)}\right)^{\alpha-1}}{\pi\left[1+\left(1-y^{\left(\log \left(1-(\tan [\pi(1-u) / 4])^{1 / \alpha}\right) / \log (\rho)\right)}\right)^{2 \alpha}\right]} . \tag{29}
\end{gather*}
$$

The parameter $\rho$ is the quantile parameter. The BTCPE quantile regression is defined as

$$
g\left(\rho_{i}\right)=z_{i}^{\prime} \boldsymbol{\theta},
$$

where $\boldsymbol{\theta}=\left(\theta_{0}, \theta_{1}, \ldots, \theta_{p}\right)^{\prime}$ is the vector of unknown parameters, $\rho_{i}$ is the $i^{\text {th }}$ quantile parameter and $z^{\prime}{ }_{i}=\left(1, z_{i 1}, z_{i 2}, \ldots, z_{i p}\right)$ are the known $i^{t h}$ vector of covariates. The link function $g(\cdot)$ is used to link the covariates to the conditional median of the dependent variable $Y$. The logit link function is used to link the covariates to the conditional quantile since $y \in(0,1)$. Hence, we have

$$
g\left(\rho_{i}\right)=\operatorname{logit}\left(\rho_{i}\right)=\log \left(\frac{\rho_{i}}{1-\rho_{i}}\right)
$$

Further, we can write

$$
\rho_{i}=\frac{\exp \left(z_{i}^{\prime} \boldsymbol{\theta}\right)}{1+\exp \left(z^{\prime}{ }_{i} \boldsymbol{\theta}\right)}
$$

Substituting $\rho_{i}$ into the re-parameterized PDF, the log-likelihood for estimating the parameters of the BTCPE quantile regression is given by

$$
\begin{align*}
\ell= & \sum_{i=1}^{n} \log \left((4 \alpha / \pi)\left(\log \left(1-(\tan [\pi(1-u) / 4])^{1 / \alpha}\right) / \log \left(\rho_{i}\right)\right)\right)-\sum_{i=1}^{n} \log \left(1+\left(1-z_{i}\right)^{2 \alpha}\right)+ \\
& \sum_{i=1}^{n}\left[\left(\log \left(1-(\tan [\pi(1-u) / 4])^{1 / \alpha}\right) / \log \left(\rho_{i}\right)\right)-1\right] \log \left(y_{i}\right)+(\alpha-1) \sum_{i=1}^{n} \log \left(1-z_{i}\right), \tag{30}
\end{align*}
$$

where $z_{i}=y_{i}\left(\log \left(1-(\tan [\pi(1-u) / 4])^{1 / \alpha}\right) / \log \left(\rho_{i}\right)\right)$. The estimates of the parameters of the regression equation are obtained by directly maximizing the log-likelihood function. They will be denoted as $\hat{\alpha}$ and $\hat{\boldsymbol{\theta}}=\left(\hat{\theta}_{0}, \ldots, \hat{\theta}_{p}\right)^{\prime}$ of $\alpha$ and $\boldsymbol{\theta}$, respectively.

### 8.1. Residual Analysis

Model diagnostics are very essential when fitting a model to a dataset. Often, the behavior of the model residuals is examined to see if the model really provides a good fit to the data. In this study, the randomized quantile residuals are used to assess the adequacy of the regression model. The randomized quantile residuals are defined as

$$
r_{i}=\Phi^{-1}\left(F_{Y}\left(y_{i} ; \hat{\alpha}, \hat{\boldsymbol{\theta}}\right)\right), i=1,2, \ldots, n
$$

where $\Phi^{-1}(\cdot)$ is the quantile of the standard normal distribution. The randomized quantile residuals are expected to be distributed as the standard normal distribution if the models provide a good fit to the data.

### 8.2. Monte Carlo Simulation for Quantile Regression

Monte Carlo simulations are carried out in this section to examine the performance of the ML estimates of the parameters of the BTCPE regression model. The exercise is
performed with two covariates. The following regression structure is adopted for the simulation:

$$
\rho_{i}=\frac{\exp \left(\theta_{0}+\theta_{1} z_{i 1}+\theta_{2} z_{i 2}\right)}{1+\exp \left(\theta_{0}+\theta_{1} z_{i 1}+\theta_{2} z_{i 2}\right)}
$$

The observations for the response variable are generated from the BTCPE distribution using sample sizes $n=50,100,250,350,500,600$ and 700 . The experiments were repeated 5000 times for each sample size. The performance of the ML estimates is examined using AB and RMSE. The simulations were carried out using the median, $u=0.5$. The following parameter combinations were used in the simulation: $I:\left(\alpha, \theta_{0}, \theta_{1}, \theta_{2}\right)=(0.7,0.2,0.8,0.3)$, II : $\left(\alpha, \theta_{0}, \theta_{1}, \theta_{2}\right)=(0.6,0.5,0.4,1.8)$ and III $:\left(\alpha, \theta_{0}, \theta_{1}, \theta_{2}\right)=(0.8,0.4,0.9,0.6)$. From the simulation results shown in Table 8, the ABs and RMSEs of the estimates' decrease as the sample size increases. Hence, the ML estimates for the BTCPE regression parameters are consistent.

Table 8. Simulation results for the quantile regression.

|  |  | I |  |  |  | II | III |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| Parameter | $n$ | AB | RMSE | AB | RMSE | AB | RMSE |
| $\theta_{0}$ | 50 | 0.1949 | 0.2235 | 0.3599 | 0.3753 | 0.2609 | 0.2969 |
|  | 100 | 0.1946 | 0.1961 | 0.3551 | 0.3726 | 0.2178 | 0.2579 |
|  | 250 | 0.1919 | 0.1941 | 0.3465 | 0.3673 | 0.1525 | 0.1926 |
|  | 350 | 0.1898 | 0.1928 | 0.3271 | 0.3544 | 0.1320 | 0.1700 |
|  | 500 | 0.1838 | 0.1927 | 0.3109 | 0.3482 | 0.1101 | 0.1431 |
|  | 600 | 0.1779 | 0.1886 | 0.3051 | 0.3434 | 0.0998 | 0.1318 |
|  | 700 | 0.1761 | 0.1850 | 0.2908 | 0.3333 | 0.0908 | 0.1196 |
| $\theta_{1}$ | 50 | 0.2826 | 0.3067 | 0.3485 | 0.3807 | 0.8194 | 0.8276 |
|  | 100 | 0.2605 | 0.2904 | 0.3181 | 0.3486 | 0.8142 | 0.8238 |
|  | 250 | 0.2290 | 0.2651 | 0.3171 | 0.3363 | 0.8013 | 0.8134 |
|  | 350 | 0.2176 | 0.2539 | 0.3138 | 0.3342 | 0.7872 | 0.8041 |
|  | 500 | 0.2097 | 0.2454 | 0.3083 | 0.3305 | 0.7727 | 0.7945 |
|  | 600 | 0.2079 | 0.2433 | 0.3020 | 0.3272 | 0.7188 | 0.7610 |
|  | 700 | 0.2053 | 0.2389 | 0.2978 | 0.3253 | 0.6862 | 0.7447 |
| $\theta_{2}$ | 50 | 1.5889 | 1.5959 | 1.7104 | 1.7153 | 0.5212 | 0.5338 |
|  | 100 | 1.5835 | 1.5913 | 1.7046 | 1.7102 | 0.5140 | 0.5291 |
|  | 250 | 1.5818 | 1.5910 | 1.6938 | 1.7006 | 0.5073 | 0.5250 |
|  | 350 | 1.5698 | 1.5815 | 1.6751 | 1.6845 | 0.4893 | 0.5130 |
|  | 500 | 1.5566 | 1.5723 | 1.6432 | 1.6578 | 0.4753 | 0.5046 |
|  | 600 | 1.4749 | 1.5132 | 1.5559 | 1.5917 | 0.4601 | 0.4999 |
|  | 700 | 1.3803 | 1.4520 | 1.4593 | 1.5264 | 0.4535 | 0.4921 |
| $\alpha$ | 50 | 0.0792 | 0.0998 | 0.0842 | 0.1110 | 0.1091 | 0.1520 |
|  | 100 | 0.0577 | 0.0745 | 0.0570 | 0.0747 | 0.0872 | 0.1382 |
|  | 250 | 0.0352 | 0.0463 | 0.0339 | 0.0437 | 0.0523 | 0.0859 |
|  | 350 | 0.0295 | 0.0378 | 0.0287 | 0.0366 | 0.0427 | 0.0650 |
|  | 500 | 0.0246 | 0.0316 | 0.0239 | 0.0317 | 0.0340 | 0.0467 |
|  | 600 | 0.0227 | 0.0287 | 0.0217 | 0.0290 | 0.0317 | 0.0449 |
|  | 700 | 0.0210 | 0.0267 | 0.0201 | 0.0259 | 0.0287 | 0.0375 |

### 8.3. Application

The application of the quantile regression model is demonstrated in this section using a real dataset. The data are taken from [34] and are also available at http://www.leg.ufpr. br/doku.php/publications:papercompanions:multquasibeta (accessed on 30 August 2022). The data consist of body fat percentage (response variable) measured in five regions: android, arms, gynoids, legs and trunk. The data are comprised of 298 observations and the independent variables are: age (in years), body mass index (in $\mathrm{kg} / \mathrm{m}^{2}$ ), sex (female or male) and IPAQ (sedentary (S), insufficiently active (I), or active (A)). In this study, the response variable body fat percentage at arms is regressed on age $\left(z_{i 1}\right)$, body mass index $\left(z_{i 2}\right)$ and sex
( $z_{i 3}, 0$ for female and 1 for male). The response variable is regressed on the covariates using the relationship $\operatorname{logit}\left(\rho_{i}\right)=\theta_{0}+\theta_{1} z_{i 1}+\theta_{2} z_{i 2}+\theta_{3} z_{i 3}, i=1,2, \ldots, 298$. Table 9 presents ML estimates, standard errors, and $p$-values for the parameters of the fitted models for the different quantiles. The estimates are all significant at the $5 \%$ level of significance.

Table 9. ML estimates for quantile regression.

| $u$ |  | $\hat{\theta}_{0}$ | $\hat{\boldsymbol{\theta}}_{1}$ | $\hat{\theta}_{2}$ | $\hat{\theta}_{3}$ | $\hat{\alpha}$ |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| 0.10 | Estimates Standard error $p$-value | $\begin{gathered} -3.6699 \\ 0.1681 \\ <0.0001 \end{gathered}$ | $\begin{gathered} 0.0076 \\ 1.1670 \times 10^{-3} \\ <0.0001 \end{gathered}$ | $\begin{gathered} 0.0905 \\ 7.5355 \times 10^{-3} \\ <0.0001 \end{gathered}$ | $\begin{gathered} -1.004 \\ 4.3797 \times 10^{-2} \\ <0.0001 \end{gathered}$ | $\begin{gathered} 308.7724 \\ 9.3305 \times 10^{-5} \\ <0.0001 \end{gathered}$ |
| 0.25 | Estimates Standard error $p$-value | $\begin{gathered} -3.2544 \\ 0.1545 \\ <0.0001 \end{gathered}$ | $\begin{gathered} 0.0071 \\ 1.0687 \times 10^{-3} \\ <0.0001 \end{gathered}$ | $\begin{gathered} 0.0845 \\ 6.9379 \times 10^{-3} \\ <0.0001 \end{gathered}$ | $\begin{gathered} -0.9326 \\ 4.0103 \times 10^{-2} \\ <0.0001 \end{gathered}$ | $\begin{gathered} 325.4705 \\ 4.6137 \times 10^{-5} \\ <0.0001 \end{gathered}$ |
| 0.50 | Estimates Standard error $p$-value | $\begin{gathered} -2.8977 \\ 0.1436 \\ <0.0001 \end{gathered}$ | $\begin{gathered} 0.0067 \\ 9.9065 \times 10^{4} \\ <0.0001 \end{gathered}$ | $\begin{gathered} 0.0792 \\ 6.4570 \times 10^{-3} \\ <0.0001 \end{gathered}$ | $\begin{gathered} -0.8732 \\ 3.7166 \times 10^{-2} \\ <0.0001 \end{gathered}$ | $\begin{gathered} 340.4285 \\ 1.3990 \times 10^{-5} \\ <0.0001 \end{gathered}$ |
| 0.75 | Estimates Standard error $p$-value | $\begin{gathered} -2.6424 \\ 0.1405 \\ <0.0001 \end{gathered}$ | $\begin{gathered} 0.0064 \\ 9.7128 \times 10^{-4} \\ <0.0001 \end{gathered}$ | $\begin{gathered} 0.0766 \\ 6.3363 \times 10^{-3} \\ <0.0001 \end{gathered}$ | $\begin{gathered} -0.8384 \\ 3.6303 \times 10^{-2} \\ <0.0001 \end{gathered}$ | $\begin{gathered} 281.1611 \\ 6.4012 \times 10^{-6} \\ <0.0001 \end{gathered}$ |
| 0.90 | Estimates Standard error $p$-value | $\begin{gathered} -2.4030 \\ 0.1353 \\ <0.0001 \end{gathered}$ | $\begin{gathered} 0.0061 \\ 9.3470 \times 10^{-4} \\ <0.001 \end{gathered}$ | $\begin{gathered} 0.0731 \\ 6.1047 \times 10^{-3} \\ <0.0001 \end{gathered}$ | $\begin{gathered} -0.7987 \\ 3.4900 \times 10^{-2} \\ <0.0001 \end{gathered}$ | $\begin{gathered} 273.9968 \\ 2.9792 \times 10^{-5} \\ <0.0001 \end{gathered}$ |

Table 10 presents the model selection criteria for the different quantiles. It is observed that the 0.90 th quantile provides the best fit for the data as it has the least values of the model selection criteria.

Table 10. Model selection criteria for quantile regression.

| $\boldsymbol{u}$ | $-2 \ell$ | AIC | BIC |
| :---: | :---: | :---: | :---: |
| $\mathbf{0 . 1 0}$ | -885.3517 | -875.3517 | -856.8663 |
| $\mathbf{0 . 2 5}$ | -887.4067 | -877.4067 | -858.9212 |
| $\mathbf{0 . 5 0}$ | -889.1990 | -879.1990 | -860.7136 |
| $\mathbf{0 . 7 5}$ | -889.8634 | -879.8634 | -861.3779 |
| $\mathbf{0 . 9 0}$ | -890.8307 | -880.8307 | -862.3453 |

Figure 17 shows the rate of change of the regression coefficients for the different quantile levels and the corresponding $95 \%$ confidence interval (CI). It can be observed that all the coefficients approach zero as the quantile level increases, suggesting that they are more important in explaining smaller quantiles.

Figures 18 and 19 show the P-P plots and half-normal plots with simulated envelopes, respectively, for the randomized quantile residuals. These figures display good fits of the BTCPE quantile regression model to the $u^{t h}$ percentage of body fat in arms for $u \in$ (0.10, 0.25, 0.50, 0.75, 0.90).


Figure 17. Rate of change of regression coefficients for different quantiles.


Figure 18. P-P plots for randomized quantile residuals.


Figure 19. Half-normal plots with simulated envelopes for randomized quantile residuals.

## 9. Conclusions

In this study, the BTCPE distribution is proposed for modeling datasets that are defined on the unit interval. The PDF of this distribution exhibits left-skewed, right-skewed, reversed J, and approximately symmetric shapes. The HRF displays increasing and bathtub shapes. This makes the distribution a suitable candidate for modeling datasets that exhibit these traits. Nine estimation methods were proposed for estimating the parameters of the distribution, and simulation results revealed that most of these estimates were consistent when it came to the estimation of the parameters of the distribution. The applications of the BTCPE distribution were illustrated using datasets on the mortality rate and recovery rates of COVID-19. The results revealed that for the three datasets, the BTCPE model provided a better fit than the other competing models. A quantile regression model for studying the relationship between the conditional quantiles of a bounded response variable and a set of covariates was proposed. The application of the regression model was illustrated using real data. The study only defined the cumulative distribution and probability density functions of the bivariate distribution. Our future research will study the detailed properties of the bivariate distribution, estimate its parameters, and illustrate its applications.

Author Contributions: Conceptualization, S.N., A.G.A., and C.C.; Data curation, S.N., A.G.A., and C.C.; Methodology, S.N., A.G.A., and C.C.; Supervision, S.N., and C.C.; Validation, S.N., and C.C.; Visualization, S.N., and A.G.A.; Writing, S.N., and A.G.A.; Review \& editing, S.N., and C.C. All authors read and agreed to the published version of the manuscript.

Funding: This research received no external funding.
Conflicts of Interest: The authors declare no conflict of interest.

## References

1. Afify, A.Z.; Nassar, M.; Kumar, D.; Cordeiro, G.M. A new unit distribution: Properties and applications. Electron. J. Appl. Stat. Anal. 2022, 15, 460-484.
2. Almazah, M.M.A.; Ullah, K.; Hussam, E.; Hossain, M.; Aldallal, R.; Riad, F.H. New Statistical Approaches for Modeling the COVID-19 Data Set: A Case Study in the Medical Sector. Complexity 2022, 2022, 1325825. [CrossRef]
3. Alahmadi, A.A.; Alqawba, M.; Almutiry, W.; Shawki, A.W.; Alrajhi, S.; Al-Marzouki, S.; Elgarhy, M. A New Version of Weighted Weibull Distribution: Modelling to COVID-19 Data. Discret. Dyn. Nat. Soc. 2022, 2022, 3994361. [CrossRef]
4. Algarni, A.; Almarashi, A.M.; Elbatal, I.; Hassan, A.S.; Almetwally, E.M.; Daghistani, A.M.; Elgarhy, M. Type I Half Logistic Burr X-G Family: Properties, Bayesian, and Non-Bayesian Estimation under Censored Samples and Applications to COVID-19 Data. Math. Probl. Eng. 2021, 2021, 5461130. [CrossRef]
5. Bantan, R.A.R.; Shafiq, S.; Tahir, M.H.; Elhassanein, A.; Jamal, F.; Almutiry, W.; Elgarhy, M. Statistical Analysis of COVID-19 Data: Using a New Univariate and Bivariate Statistical Model. J. Funct. Spaces 2022, 2022, 2851352. [CrossRef]
6. Arif, M.; Khan, D.M.; Aamir, M.; Khalil, U.; Bantan, R.A.R.; Elgarhy, M. Modeling COVID-19 Data with a Novel Extended Exponentiated Class of Distributions. J. Math. 2022, 2022, 1908161. [CrossRef]
7. Nagy, M.; Almetwally, E.M.; Gemeay, A.M.; Mohammed, H.S.; Jawa, T.M.; Sayed-Ahmed, N.; Muse, A.H. The New Novel Discrete Distribution with Application on COVID-19 Mortality Numbers in Kingdom of Saudi Arabia and Latvia. Complexity 2021, 2021, 7192833. [CrossRef]
8. Almetwally, E.M. The Odd Weibull Inverse Topp-Leone Distribution with Applications to COVID-19 Data. Ann. Data Sci. 2021, 9, 121-140. [CrossRef]
9. Muse, A.H.; Tolba, A.H.; Fayad, E.; Abu Ali, O.A.; Nagy, M.; Yusuf, M. Modelling the COVID-19 Mortality Rate with a New Versatile Modification of the Log-Logistic Distribution. Comput. Intell. Neurosci. 2021, 2021, 8640794. [CrossRef]
10. Haq, M.A.U.; Babar, A.; Hashmi, S.; Alghamdi, A.S.; Afify, A.Z. The Discrete Type-II Half-Logistic Exponential Distribution with Applications to COVID-19 Data. Pak. J. Stat. Oper. Res. 2021, 17, 921-932. [CrossRef]
11. Gündüz, S.; Korkmaz, M. A New Unit Distribution Based on the Unbounded Johnson Distribution Rule: The Unit Johnson SU Distribution. Pak. J. Stat. Oper. Res. 2020, 16, 471-490. [CrossRef]
12. Bantan, R.; Jamal, F.; Chesneau, C.; Elgarhy, M. Theory and Applications of the Unit Gamma/Gompertz Distribution. Mathematics 2021, 9, 1850. [CrossRef]
13. Nasiru, S.; Abubakari, A.G.; Angbing, I.D. Bounded Odd Inverse Pareto Exponential Distribution: Properties, Estimation, and Regression. Int. J. Math. Math. Sci. 2021, 2021, 9955657. [CrossRef]
14. Jodrá, P. A bounded distribution derived from the shifted Gompertz law. J. King Saud Univ.-Sci. 2020, 32, 523-536. [CrossRef]
15. Haq, M.A.U.; Hashmi, S.; Aidi, K.; Ramos, P.L.; Louzada, F. Unit Modified Burr-III Distribution: Estimation, Characterizations and Validation Test. Ann. Data Sci. 2020, 99, 1-26. [CrossRef]
16. Korkmaz, M.Ç. The unit generalized half normal distribution: A new bounded distribution with inference and application. $U$. P. B. Sci. Bull. Ser. A 2020, 82, 133-140.
17. Mazucheli, J.; Menezes, A.F.B.; Chakraborty, S. On the one parameter unit-Lindley distribution and its associated regression model for proportion data. J. Appl. Stat. 2019, 46, 700-714. [CrossRef]
18. Mazucheli, J.; Menezes, A.F.; Dey, S. Unit-Gompertz distribution with applications. Statistica 2019, 79, 25-43.
19. Korkmaz, M. A new heavy-tailed distribution defined on the bounded interval. J. Appl. Stat. 2019, 47, 2097-2119. [CrossRef]
20. Mazucheli, J.; Menezes, A.F.; Ghitany, M.E. The unit Weibull distribution and associated inference. J. Appl. Probab. Stat. 2018, 13, 1-22.
21. Ghitany, M.E.; Mazucheli, J.; Menezes, A.F.B.; Alqallaf, F. The unit-inverse Gaussian distribution: A new alternative to twoparameter distributions on the unit interval. Commun. Stat.-Theory Methods 2018, 48, 3423-3438. [CrossRef]
22. Aldahlan, M.A.; Jamal, F.; Chesneau, C.; Elgarhy, M.; Elbatal, I. The Truncated Cauchy Power Family of Distributions with Inference and Applications. Entropy 2020, 22, 346. [CrossRef] [PubMed]
23. Shaked, M.; Shanthikumar, J.G. Stochastic Orders; Wiley: New York, NY, USA, 2007.
24. MacGillivray, H.L. Skewness and Asymmetry: Measures and Orderings. Ann. Stat. 1986, 14, 994-1011. [CrossRef]
25. Moors, J.J. A quantile alternative for kurtosis. J. R. Stat. Soc. Ser. D 1988, 37, 25-32. [CrossRef]
26. Elhassanein, A. On Statistical Properties of a New Bivariate Modified Lindley Distribution with an Application to Financial Data. Complexity 2022, 2022, 2328831. [CrossRef]
27. Ganji, M.; Bevrani, H.; Hami, N. A New Method For Generating Continuous Bivariate Families. J. Iran. Stat. Soc. 2018, 17, 109129. [CrossRef]
28. R Core Team. R: A Language and Environment for Statistical Computing; R Foundation for Statistical Computing: Vienna, Austria, 2019; Available online: https://www.R-project.org/ (accessed on 30 October 2022).
29. Modi, K.; Gill, V. Unit Burr-III distribution with application. J. Stat. Manag. Syst. 2019, 23, 579-592. [CrossRef]
30. Ghosh, I.; Dey, S.; Kumar, D. Bounded M-O Extended Exponential Distribution with Applications. Stoch. Qual. Control 2019, 34, 35-51. [CrossRef]
31. Altun, E.; Cordeiro, G.M. The unit-improved second-degree Lindley distribution: Inference and regression modeling. Comput. Stat. 2019, 35, 259-279. [CrossRef]
32. Bolker, B. Tools for General Maximum Likelihood Estimation; R Development Core Team: Vienna, Austria, 2014. Available online: https://github.com/bbolker/bbmle (accessed on 30 October 2022).
33. Xiang, Y.; Gubian, S.; Suomela, B.; Hoeng, J. Generalized simulated annealing: GenSA package. R J. 2013, 5, 13-29. [CrossRef]
34. 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, 17, 39-53. [CrossRef] [PubMed]
