Vol. 34, No. 4 (2024) | DOI: 10.37190/ord240407

OPEN ACCESS

Operations Research and Decisions

www.ord.pwr.edu.pl

OPERATIONS RESEARCH AND DECISIONS QUARTERLY

The Opone family of distributions. Beyond the power continuous Bernoulli distribution

Festus C. Opone[1*, 2] Christophe Chesneau[3]

[1] *Department of Statistics, University of Benin, Benin City, Nigeria*
[2] *Department of Statistics, Delta State University of Science and Technology, Ozoro, Nigeria*
[3] *Department of Mathematics, LMNO, University of Caen, 14032 Caen, France*
[*] *Corresponding author, email address: oponef@dsust.edu.ng*

Abstract

Recent developments in applied statistics have given rise to the continuous Bernoulli distribution, a one-parameter distribution with support of $[0, 1]$. In this paper, we use it for a more general purpose: the creation of a family of distributions. We thus exploit the flexible functionalities of the continuous Bernoulli distribution to enhance the modeling properties of well-referenced distributions. We first focus on the theory of this new family, including the quantiles, expansion of important functions, and moments. Then we exemplify it by considering a special baseline: the Topp–Leone distribution. Thanks to the functional structure of the continuous Bernoulli distribution, we create a new two-parameter distribution with support for $[0, 1]$ that possesses versatile shape capacities. In particular, the corresponding probability density function has left-skewed, N-type and decreasing shapes, and the corresponding hazard rate function has increasing and bathtub shapes, beyond the possibilities of the corresponding functions of the Topp–Leone distribution. Its quantile and moment properties are also examined. We then use our modified Topp–Leone distribution from a statistical perspective. The two parameters are supposed to be unknown and then estimated from proportional-type data with the maximum likelihood method. Two different data sets are considered, and reveal that the modified Topp–Leone distribution can fit them better than popular rival distributions, including the unit-Weibull, unit-Gompertz, and log-weighted exponential distributions. It also outperforms the Topp–Leone and continuous Bernoulli distributions.

Keywords: *family of distributions, continuous Bernoulli distribution, moments, quantiles, data fitting.*

1. Introduction

Let us first go through the so-called continuous Bernoulli (CB) distribution, which plays a central role in our study. It can be defined in the following manner [18]:

Received 25 November 2023, accepted 17 October 2024, published online 19 December 2024
ISSN 2391-6060 (Online)/

The costs of publishing this issue have been co-financed by the Department of Operations Research and Business Intelligence at the Faculty of Management, Wrocław University of Science and Technology, Wrocław, Poland

Definition 1. The CB distribution with parameter $\lambda \in [0,1]$ is defined by the following cumulative distribution function (cdf):

$$F(x;\lambda) = \begin{cases} 0, \quad x < 0 \\ x, \quad \lambda = \dfrac{1}{2} \text{ and } x \in [0,1] \\ \dfrac{\lambda^x (1-\lambda)^{1-x} + \lambda - 1}{2\lambda - 1}, \quad \lambda \in (0,1)/\left\{\dfrac{1}{2}\right\} \text{ and } x \in [0,1] \\ 1, \quad x > 1 \end{cases} \tag{1}$$

Thus, like the standard power distribution, the CB distribution has a support of $[0,1]$ and is of a one-parameter continuous type. It has applications in many different domains, with a focus on machine learning, probability theory, and statistics. To learn more about these subjects, we suggest the following references: [14, 18, 29] and [17].

Recent extensions of the CB distribution have been proposed in the literature. We may mention the power CB (PCB) distribution elaborated in [7]. Its main feature is to add a shape parameter to the cdf in equation (1), with the aim of extending the modeling scope of the CB distribution. The basic mathematical properties of the PCB distribution were derived, and a statistical approach was performed. The estimates of the parameters were discussed using the maximum likelihood method. The flexibility of the PCB distribution in real-life data fitting was analyzed using two data sets: a trade share data set, and a tensile strength of polyester fibers data set. Also, fair competitors are considered to highlight the accuracy of the PCB distribution. It is shown that it has better results according to standard statistical criteria. As another efficient extension, the transmuted CB (TCB) distribution was created in [8]. Its main feature is an additional parameter that realizes a linear tradeoff between the min and max of two continuous random variables with the CB distribution. The fundamental mathematical properties of the TCB distribution were deduced, and a statistical method was used. In order to show its efficiency, three proportional data sets were analyzed: a time to infection of kidney dialysis patients data set, records of exceedances of flood peaks data set, and waiting times before service in a bank data set. Empirical findings reveal that the TCB distribution promises more flexibility in fitting these data sets than well-reputed competitors. We also refer to [29, Chapter 9], where a two-dimensional version of the CB distribution is introduced, along with some of its properties. The authors [17] developed a fractile (or quantile) regression model for a response with values in $[0,1]$ using an exponentiated version of the CB distribution. In 2023, there is again a lot of room for work based on the CB distribution, both in theory and in practice.

In this paper, we use it for a more general objective: we take advantage of the cdf of the CB distribution to modify or enhance the functional properties of any existing distributions. Thus, the cdf of the CB distribution is considered a generator function to create a new family of distributions, called the Opone (Op) family. The use of distributions with support of $[0,1]$ to construct such a family is a well-mastered technique (see, for instance, the beta generated family in [10], the Topp–Leone (TL) generated family in [2], and the Kumaraswamy generated family in [15]) but, to the best of our knowledge, the CB distribution has never been employed in this general family context. In the first part of the paper, we present the Op family by its main functions and illustrate it by some of its important member distributions. Then, we examine the expansion results, which allow us to express complex analytical integral terms related to moments, into a series of manageable functions. After the general study, a focus is put on the special member defined with the TL distribution as the baseline; the OpTL distribution is thus created. It possesses more capacities in terms of statistical modeling in comparison to the former TL distribution. In particular, the corresponding probability density function (pdf) has left-skewed, N-type and decreasing shapes, and the corresponding hazard rate function (hrf) has increasing and bathtub shapes, far beyond the possibilities of the corresponding functions of the TL distribution. Investigations on its moment and quantile features are also conducted. Finally, we statistically apply the OpTL distribution. The maximum likelihood approach is developed to estimate the two involved parameters. Three data sets are considered: one containing trade share data, one about the failure of components, and the last one about the recovery rates of COVID-19 patients in Spain. We show that the OpTL distribution can fit these data sets more accurately than prominent rival distributions like the unit-Weibull, unit-Gompertz, and log-weighted exponential distributions. Additionally, it performed better than the TL and CB distributions. Statistical figures and numerical criteria are used to support this claim.

The organization of the paper is as follows: Section 2 presents the Op family. Section 3 is devoted to the properties and applications of the OpTL distribution in a data fitting scenario. A conclusion is given in Section 4.

2. The Op family of distributions

2.1. Presentation

The basic definition of the Op family is:

Definition 2. The Op family is defined by the cdf given as $F_{\mathrm{Op}}(x;\lambda,\zeta) = F[G(x;\zeta);\lambda]$, where $F(x;\lambda)$ is the cdf of the CB distribution as described in Definition 1, and $G(x;\zeta)$ is the cdf of a continuous univariate distribution.

Table 1. A short list of some Op distributions based on well-referenced distributions and $\lambda \in [0,1] \setminus \{1/2\}$

Distribution	Identity	Support	$G(x;\zeta)$	(λ,ζ)	$F_{\mathrm{Op}}(x;\alpha,\zeta)$
OpTL	Topp–Leone	$[0,1]$	$[x(2-x)]^a$	(λ,a)	$\frac{\lambda^{[x(2-x)]^a}(1-\lambda)^{1-[x(2-x)]^a}+\lambda-1}{2\lambda-1}$
OpK	Kumaraswamy	$[0,1]$	$1-(1-x^a)^b$	(λ,a,b)	$\frac{\lambda^{1-(1-x^a)^b}(1-\lambda)^{(1-x^a)^b}+\lambda-1}{2\lambda-1}$
OpE	Exponential	$[0,+\infty)$	$1-e^{-\theta x}$	(λ,θ)	$\frac{\lambda^{1-e^{-\theta x}}(1-\lambda)^{e^{-\theta x}}+\lambda-1}{2\lambda-1}$
OpW	Weibull	$[0,+\infty)$	$1-e^{-\theta x^\sigma}$	(λ,θ,σ)	$\frac{\lambda^{1-e^{-\theta x^\sigma}}(1-\lambda)^{e^{-\theta x^\sigma}}+\lambda-1}{2\lambda-1}$
OpLom	Lomax	$[0,+\infty)$	$1-(1+\rho x)^{-\theta}$	(λ,ρ,θ)	$\frac{\lambda^{1-(1+\rho x)^{-\theta}}(1-\lambda)^{(1+\rho x)^{-\theta}}+\lambda-1}{2\lambda-1}$
OpPar	Pareto	$[a,+\infty)$	$1-(x/a)^{-b}$	(λ,a,b)	$\frac{\lambda^{1-(x/a)^{-b}}(1-\lambda)^{(x/a)^{-b}}+\lambda-1}{2\lambda-1}$
OpGu	Gumbel	$\mathbb{R}$	$\exp(-e^{-bx})$	(λ,b)	$\frac{\lambda^{\exp(-e^{-bx})}(1-\lambda)^{1-\exp(-e^{-bx})}+\lambda-1}{2\lambda-1}$
OpLog	Logistic	$\mathbb{R}$	$(1+e^{-bx})^{-1}$	(λ,b)	$\frac{\lambda^{(1+e^{-bx})^{-1}}(1-\lambda)^{1-(1+e^{-bx})^{-1}}+\lambda-1}{2\lambda-1}$

Hence, we have

$$F_{\mathrm{Op}}(x;\lambda,\zeta)=\begin{cases}G(x;\zeta), & \lambda=\dfrac{1}{2} \text{ and } x\in\mathbb{R}\\ \dfrac{\lambda^{G(x;\zeta)}(1-\lambda)^{1-G(x;\zeta)}+\lambda-1}{2\lambda-1}, & \lambda\in(0,1)/\left\{\dfrac{1}{2}\right\} \text{ and } x\in\mathbb{R}\end{cases} \quad (2)$$

Thus, $F_{\mathrm{Op}}(x;\lambda,\zeta)$ reduces to the chosen baseline cdf $G(x;\zeta)$ by taking $\lambda=1/2$. Furthermore, if we take $G(x,\zeta)$ as the cdf of the uniform distribution over the interval $[0,1]$, $F_{\mathrm{Op}}(x;\lambda,\zeta)$ corresponds to the cdf of the CB distribution. To the best of our knowledge, the other parameter configurations generate completely new distributions. A small panel of distributions belonging to the Op family are presented in Table 1. To exemplify the interest of the Op family, a focus will be on the OpTL distribution in Section 3.

We end this part by exhibiting some asymptotic properties of $F_{\mathrm{Op}}(x;\lambda,\zeta)$, mainly to show the effect of λ and the baseline cdf at the boundaries. When $G(x;\zeta)\to 0$, we get

$$F_{\mathrm{Op}}(x;\lambda,\zeta)\sim(1-\lambda)c_\lambda G(x;\zeta)$$

where

$$c_\lambda=\frac{2\,\mathrm{arctanh}(1-2\lambda)}{1-2\lambda}\quad\left(\text{or}\quad c_\lambda=\frac{\ln(1-\lambda)-\ln(\lambda)}{1-2\lambda}\right) \quad (3)$$

and when $G(x;\zeta)\to 1$, we have $F_{\mathrm{Op}}(x;\lambda,\zeta)\sim 1-\lambda c_\lambda[1-G(x;\zeta)]$. On these limit bounds, the impact of λ is significant.

2.2. Other crucial functions

In addition to the cdf, other functions are of interest to comprehend the modeling capacities of the Op family. The most crucial of them are described below.

2.2.1. Pdf

To begin, the pdf corresponding to the Op family is obtained as

$$f_{\mathrm{Op}}(x;\lambda,\zeta)=\begin{cases}g(x;\zeta), & \lambda=\dfrac{1}{2} \text{ and } x\in\mathbb{R}\\ c_\lambda g(x;\zeta)\lambda^{G(x;\zeta)}(1-\lambda)^{1-G(x;\zeta)}, & \lambda\in(0,1)/\left\{\dfrac{1}{2}\right\} \text{ and } x\in\mathbb{R}\end{cases} \quad (4)$$

where c_λ is defined in equation (3) and $g(x;\zeta)$ is the pdf associated with $G(x;\zeta)$.

Let us consider the case $\lambda\in(0,1)/\{1/2\}$, and proceed to a basic asymptotic investigation. When $G(x;\zeta)\to 0$, we obtain

$$f_{\mathrm{Op}}(x;\lambda,\zeta)\sim(1-\lambda)c_\lambda g(x;\zeta)$$

and, when $G(x;\zeta)\to 1$, we have

$$f_{\mathrm{Op}}(x;\lambda,\zeta)\sim\lambda c_\lambda g(x;\zeta).$$

These results sketch the importance of the role of the parameter λ in the Op family.

A general mode study is performed below. If a mode exists, say x_m, and $f_{\mathrm{Op}}(x_m;\lambda,\zeta)\neq 0$, it satisfies the equation $\{\ln[f_{\mathrm{Op}}(x;\lambda,\zeta)]\}'|_{x=x_m}=0$, that is, after some developments, if $g(x_m;\zeta)\neq 0$,

$$g'(x;\zeta)|_{x=x_m}-c_\lambda^\dagger g(x_m;\zeta)^2=0, \text{where } c_\lambda^\dagger=2\,\mathrm{arctanh}(1-2\lambda)\ \Big(\text{or } c_\lambda^\dagger=\ln(1-\lambda)-\ln(\lambda)\Big) \tag{5}$$

Depending on the characteristics of $G(x;\zeta)$, this equation may be very complex. In this case, a graphical study will be preferred.

2.2.2. Hrf

The hrf corresponding to the Op family is obtained as

$$h_{\mathrm{Op}}(x;\lambda,\zeta)=\begin{cases}\dfrac{g(x;\zeta)}{1-G(x;\zeta)}, & \lambda=\dfrac{1}{2}\text{ and } x\in\mathbb{R}\\[2mm] c_\lambda^\dagger g(x;\zeta)\dfrac{(1-\lambda)^{1-G(x;\zeta)}}{(1-\lambda)^{1-G(x;\zeta)}-\lambda^{1-G(x;\zeta)}}, & \lambda\in(0,1)/\left\{\dfrac{1}{2}\right\}\text{ and } x\in\mathbb{R}\end{cases} \tag{6}$$

where $c_\lambda^\dagger$ is defined in equation (5). Let us consider the case $\lambda\in(0,1)/\{1/2\}$. Some asymptotic results on $h_{\mathrm{Op}}(x;\lambda,\zeta)$ are presented below. When $G(x;\zeta)\to 0$, we have

$$h_{\mathrm{Op}}(x;\lambda,\zeta)\sim(1-\lambda)c_\lambda g(x;\zeta)$$

and, when $G(x;\zeta)\to 1$, we find that

$$h_{\mathrm{Op}}(x;\lambda,\zeta)\sim\frac{g(x;\zeta)}{1-G(x;\zeta)}$$

An extremum study of $h_{\mathrm{Op}}(x;\lambda,\zeta)$ is of interest to determine the hazard rate modeling properties of the Op family. Here, if an extremum exists, say x_*, and $h_{\mathrm{Op}}(x_*;\lambda,\zeta)\neq 0$, it satisfies the equation $\{\ln[h_{\mathrm{Op}}(x;\lambda,\zeta)]\}'|_{x=x_*}=0$, that is, after some developments, if $g(x_*;\zeta)\neq 0$,

$$g'(x;\zeta)|_{x=x_*}-g(x_*;\zeta)^2\ln(1-\lambda)+g(x_*;\zeta)^2\frac{(1-\lambda)^{1-G(x_*;\zeta)}\ln(1-\lambda)-\lambda^{1-G(x_*;\zeta)}\ln(\lambda)}{(1-\lambda)^{1-G(x_*;\zeta)}-\lambda^{1-G(x_*;\zeta)}}=0$$

This equation can have a high level of complexity depending on the nature of $G(x;\zeta)$. In this case, a graphical analysis is more convenient.

2.2.3. Quantile function

The quantile function (qf) corresponding to the Op family is indicated as

$$\begin{aligned}Q_{\mathrm{Op}}(x;\lambda,\zeta)&=F_{\mathrm{Op}}^{-1}(x;\lambda,\zeta)\\&=\begin{cases}Q(x;\zeta), & \lambda=\dfrac{1}{2}\text{ and } x\in[0,1],\\[2mm] Q\left\{\dfrac{\ln(1-\lambda)-\ln[(2\lambda-1)x+1-\lambda]}{c_\lambda^\dagger};\zeta\right\}, & \lambda\in(0,1)/\left\{\dfrac{1}{2}\right\}\text{ and } x\in[0,1]\end{cases}\end{aligned} \tag{7}$$

where $c_\lambda^\dagger$ is defined in equation (5) and $Q(x;\zeta)$ is the qf associated with $G(x;\zeta)$.

If we take $x = 1/2$ in equation (7), the median of the Op family is obtained. It can be expressed as

$$M_{ed}(\lambda, \zeta) = Q_{\mathrm{Op}}\left(\frac{1}{2}; \lambda, \zeta\right) = \begin{cases} M^*_{ed}(\zeta), & \lambda = \dfrac{1}{2}, \\ Q\left\{\dfrac{\ln(2) + \ln(1-\lambda)}{c^\dagger_\lambda}; \zeta\right\}, & \lambda \in (0,1)/\left\{\dfrac{1}{2}\right\} \end{cases}$$

where $M^*_{ed}(\zeta)$ is the median associated with $G(x; \zeta)$.

Data generation can be performed with this qf based on the standard inversion approach. To this aim, let us fix a positive integer n and generate n values from the uniform distribution over $[0, 1]$, say $u_1, \ldots, u_n$. Then data from a distribution in the Op family are given by $x_1, \ldots, x_n$, where $x_i = Q_{\mathrm{Op}}(u_i; \lambda, \zeta)$, for $i = 1, \ldots, n$.

The functions described in the above parts, i.e., the cdf, pdf, hrf and qf, will be described in the case of the OpTL distribution in Section 3.

2.3. Some expansion results

Series expansions of the main functions of a distribution are useful to express complex analytical integral terms related to moments into a series of manageable functions. They have computational interests. Indeed, discrete coefficient summations are easy to manipulate, and a high degree of precision can be obtained in comparison to standard integral computation methods.

Two complementary series expansions for the cdf are presented in the next result.

Proposition 1. For $\lambda \in (0, 1)/\{1/2\}$, the two following expansions hold:

Cdf expansion 1. In function of exponentiated versions of the cdf $G(x; \zeta)$, we have

$$F_{\mathrm{Op}}(x; \lambda, \zeta) = \sum_{k,\ell=0}^{+\infty} \sum_{m=0}^{\ell} \alpha_{k,\ell,m}(\lambda) G(x; \zeta)^{k+m} + \beta_\lambda$$

where

$$\alpha_{k,\ell,m}(\lambda) = \frac{1}{2\lambda - 1} \frac{[\ln(\lambda)]^k [\ln(1-\lambda)]^\ell (-1)^m}{k!\ell!} \binom{\ell}{m} \tag{8}$$

and

$$\beta_\lambda = \frac{\lambda - 1}{2\lambda - 1} \tag{9}$$

Cdf expansion 2. In function of exponentiated versions of the survival function (sf) $S(x; \zeta) = 1 - G(x; \zeta)$, we have

$$F_{\mathrm{Op}}(x; \lambda, \zeta) = \sum_{k,\ell=0}^{+\infty} \sum_{m=0}^{k} \alpha^\star_{k,\ell,m}(\lambda) S(x; \zeta)^{\ell+m} + \beta_\lambda$$

where

$$\alpha^\star_{k,\ell,m}(\lambda) = \frac{1}{2\lambda - 1} \frac{[\ln(\lambda)]^k [\ln(1-\lambda)]^\ell (-1)^m}{k!\ell!} \binom{k}{m}$$

Proof. By using suitable decompositions and the series expansion of the exponential function, we get

$$F_{\mathrm{Op}}(x;\lambda,\zeta)=\frac{1}{2\lambda-1}\lambda^{G(x;\zeta)}(1-\lambda)^{S(x;\zeta)}+\beta_\lambda=\frac{1}{2\lambda-1}e^{G(x;\zeta)\ln(\lambda)}e^{S(x;\zeta)\ln(1-\lambda)}+\beta_\lambda$$

$$=\frac{1}{2\lambda-1}\left\{\sum_{k=0}^{+\infty}\frac{[\ln(\lambda)]^k}{k!}G(x;\zeta)^k\right\}\left\{\sum_{\ell=0}^{+\infty}\frac{[\ln(1-\lambda)]^\ell}{\ell!}S(x;\zeta)^\ell\right\}+\beta_\lambda \qquad (10)$$

On this mathematical basis, let us now distinguish the two kinds of expansions.

For the cdf expansion 1. In function of the exponentiated versions of the cdf $G(x;\zeta)$, by using the standard binomial theorem, we have

$$S(x;\zeta)^\ell=[1-G(x;\zeta)]^\ell=\sum_{m=0}^{\ell}\binom{\ell}{m}(-1)^m G(x;\zeta)^m$$

This decomposition, combined with equation (10), gives the desired result.

For the cdf expansion 2: In function of the exponentiated versions of the sf $S(x;\zeta)$, we have

$$G(x;\zeta)^k=[1-S(x;\zeta)]^k=\sum_{m=0}^{k}\binom{k}{m}(-1)^m S(x;\zeta)^m.$$

This decomposition, combined with equation (10), gives the desired result.

This ends the proof. □

Based on Proposition 1, assuming the interchange of differentiation and summation, we get two different series expansion for the pdf of the Op family upon differentiation of $F_{\mathrm{Op}}(x;\lambda,\zeta)$, that is, for $\lambda\in(0,1)/\{1/2\}$:

Pdf expansion 1. based on the cdf expansion 1 in Proposition 1, we have

$$f_{\mathrm{Op}}(x;\lambda,\zeta)=\sum_{k,\ell=0}^{+\infty}\sum_{m=0}^{\ell}\alpha^{\dagger}_{k,\ell,m}(\lambda)g(x;\zeta)G(x;\zeta)^{k+m-1}$$

where $\alpha^{\dagger}_{k,\ell,m}(\lambda)=(k+m)\alpha_{k,\ell,m}(\lambda)$.

Pdf expansion 2. based on the cdf expansion 2 in Proposition 1, we have

$$f_{\mathrm{Op}}(x;\lambda,\zeta)=\sum_{k,\ell=0}^{+\infty}\sum_{m=0}^{k}\alpha^{\ddagger}_{k,\ell,m}(\lambda)g(x;\zeta)S(x;\zeta)^{\ell+m-1}$$

where $\alpha^{\ddagger}_{k,\ell,m}(\lambda)=-(\ell+m)\alpha^{\star}_{k,\ell,m}(\lambda)$.

Hence, for a random variable X with a distribution into the Op family, and a function $\varphi(x)$, the following moment results hold:

Based on the pdf expansion 1. Using the transfer formula and assuming the interchange of integration and summation, we have

$$E[\varphi(X)] = \int_{-\infty}^{+\infty} \varphi(x) f_{\mathrm{Op}}(x; \lambda, \zeta) dx = \sum_{k,\ell=0}^{+\infty} \sum_{m=0}^{\ell} \alpha^{\dagger}_{k,\ell,m}(\lambda) \gamma_{k,m}$$

where

$$\gamma_{k,m} = \int_{-\infty}^{+\infty} \varphi(x) g(x; \zeta) G(x; \zeta)^{k+m-1} dx$$

Based on the pdf expansion 2. We have

$$E[\varphi(X)] = \int_{-\infty}^{+\infty} \varphi(x) f_{\mathrm{Op}}(x; \lambda, \zeta) dx = \sum_{k,\ell=0}^{+\infty} \sum_{m=0}^{k} \alpha^{\ddagger}_{k,\ell,m}(\lambda) \tau_{\ell,m}$$

where

$$\tau_{\ell,m} = \int_{-\infty}^{+\infty} \varphi(x) g(x; \zeta) S(x; \zeta)^{\ell+m-1} dx$$

The integral terms $\gamma_{k,m}$ and $\tau_{\ell,m}$ can be calculated for most of the standard distributions, i.e., with a well-referenced cdf $G(x; \zeta)$. In particular, this is the case for all the baseline distributions presented in Table 1.

3. Overview of the OpTL distribution

In this section, we exemplify the Op family by focusing on the theory and application of the OpTL distribution, as sketched in Table 1.

3.1. On the TL distribution

We recall that the cdf of the TL distribution with parameter $a > 0$ is defined by

$$G(x; a) = \begin{cases} 1, \ x > 1 \\ [x(2-x)]^{a}, \ x \in [0,1] \\ 0, \ x < 0 \end{cases} \tag{11}$$

In addition, the corresponding pdf is defined by

$$g(x; a) = \begin{cases} 2a(1-x)\, [x(2-x)]^{a-1}, \ x \in [0,1] \\ 0, \ x \notin [0,1] \end{cases} \tag{12}$$

The hrf of the TL distribution is indicated as

$$h(x;a) = \begin{cases} \dfrac{2a(1-x)\left[x(2-x)\right]^{a-1}}{1-\left[x(2-x)\right]^{a}}, \ x \in [0,1] \\ 0, \ x \notin [0,1] \end{cases}$$

In a nutshell, the TL distribution is of interest in probability and statistics for several reasons, as those described below.

Flexibility. The TL distribution is a flexible distribution that can analyze efficiently a wide range of data types with values of $[0,1]$. This is mostly thanks to the J-shapes of the pdf, and the bathtub shapes of the hrf, which are rare properties of a one-parameter distribution with the support of $[0,1]$.

Computational ease. The TL distribution has a closed-form expression for its cdf, pdf and hrf, which makes it relatively easy to compute compared to other more complex distributions. This makes it an attractive option for analyzing data in many situations where other distributions might be too computationally expensive.

Interpretability. The TL distribution has a clear and intuitive interpretation of its parameters, which makes it simple to understand and interpret the results of statistical analysis. This is particularly important in situations where the distribution is being used to model real-world phenomena, and the results of the analysis will be used to inform decisions.

Applications. The TL distribution has been used in a variety of applications, including finance, actuarial science, biology, and ecology. It has also been used in the analysis of data from surveys, market research, and clinical trials. This broad range of applications demonstrates the versatility and usefulness of the TL distribution in probability and statistics.

For more information on the TL distribution, we refer to [9, 12, 22, 26, 30] and [11].

3.2. The OpTL distribution

Based on the cdf in equation (11) and the cdf of the Op family as described in equation (2), we define the OpTL distribution as follows:

Definition 3. Using equations (2) and (11), the OpTL distribution with parameter λ and a is defined by the following cdf:

$$F_{\mathrm{OpTL}}(x;\lambda,a) = \begin{cases} 1, \ x > 1 \\ \dfrac{\lambda^{[x(2-x)]^{a}}(1-\lambda)^{1-[x(2-x)]^{a}}+\lambda-1}{2\lambda-1}, \ \lambda \in (0,1)/\left\{\dfrac{1}{2}\right\} \text{ and } x \in [0,1] \\ 0, \ x < 0 \end{cases} \quad (13)$$

For the special case $\lambda = 1/2$, it is reduced to the cdf of the TL distribution.

Thus defined, the OpTL distribution presents a new option of two-parameter distribution with support of $[0,1]$. Such distributions have found renewed interest these last few years due to the need to analyze complex proportional or percentage data of various kinds. On this topic, we may refer to [1], [6, 13, 16, 19, 20], [21, 23, 25, 27], and [28].

In the sequel, for the sake of space, we omit the case $\lambda = 1/2$ (so $\lambda \in (0,1)/\{1/2\}$), since it corresponds to the well-known TL distribution.

Based on Proposition 1, the following expansion holds:

$$F_{\mathrm{OpTL}}(x;\lambda,a) = \sum_{k,\ell=0}^{+\infty}\sum_{m=0}^{\ell} \alpha_{k,\ell,m}(\lambda)G(x;a(k+m)) + \beta_\lambda \tag{14}$$

where $\alpha_{k,\ell,m}$ is given in equation (8), β_λ is described in equation (9), and $G(x;a(k+m))$ is defined as in equation (11) with the parameter $a(k+m)$ instead of a.

3.3. Other functions

Based on the cdf in equation (13), the pdf of the OpTL distribution is obtained as

$$f_{\mathrm{OpTL}}(x;\lambda,a) = \begin{cases} c_\lambda 2a(1-x)\left[x(2-x)\right]^{a-1}\lambda^{[x(2-x)]^a}(1-\lambda)^{1-[x(2-x)]^a}, & \lambda \in (0,1)/\left\{\frac{1}{2}\right\} \text{ and } x \in [0,1] \\ 0, \ \lambda \in (0,1)/\left\{\frac{1}{2}\right\} \text{ and } x \notin [0,1] \end{cases} \tag{15}$$

where c_λ is defined in equation (3). Let us now briefly study this function, beginning with its asymptotic properties. When $x \to 0$, we have

$$f_{\mathrm{OpTL}}(x;\lambda,a) \sim (1-\lambda)c_\lambda 2^a a x^{a-1} \to \begin{cases} 0, \ a>1 \\ 2, \ a=1 \\ +\infty, \ a<1 \end{cases} \tag{16}$$

and, when $x \to 1$, we obtain

$$f_{\mathrm{OpTL}}(x;\lambda,a) \sim \lambda c_\lambda 2a(1-x) \to 0$$

From these asymptotic results, we conclude that $f_{\mathrm{OpTL}}(x;\lambda,a)$ can have at least one mode, and is not increasing. For a more deep shape study, in view of the functional complexity of the function, a graphical work will be performed later.

In a last remark on this pdf, based on equation (14), the following expansion holds:

$$f_{\mathrm{OpTL}}(x;\lambda,a) = \sum_{k,\ell=0}^{+\infty}\sum_{m=0}^{\ell} \alpha_{k,\ell,m}(\lambda)g(x;a(k+m)) \tag{17}$$

where $g(x;a(k+m))$ is the pdf related to the TL distribution, i.e., defined in equation (12) with the parameter $a(k+m)$.

The expansion of $f_{\mathrm{OpTL}}(x;\lambda,a)$ finds an interest in approximation of various moments quantities, for computation purposes. The hrf of the OpTL distribution is obtained as

$$h_{\mathrm{OpTL}}(x;\lambda,a) = \begin{cases} c_\lambda^\dagger 2a(1-x)\left[x(2-x)\right]^{a-1} \dfrac{(1-\lambda)^{1-[x(2-x)]^a}}{(1-\lambda)^{1-[x(2-x)]^a}-\lambda^{1-[x(2-x)]^a}}, \ \lambda \in (0,1)/\left\{\dfrac{1}{2}\right\} \text{ and } x\in[0,1] \\ 0, \ \lambda \in (0,1)/\left\{\dfrac{1}{2}\right\} \text{ and } x \notin [0,1] \end{cases} \quad (18)$$

where $c_\lambda^\dagger$ is defined in equation (5). Let us now briefly examine this function, beginning with its asymptotic properties.

When $x \to 0$, we have

$$h_{\mathrm{OpTL}}(x;\lambda,a) \sim (1-\lambda)c_\lambda 2^a a x^{a-1} \to \begin{cases} 0, \ a>1 \\ 2, \ a=1 \\ +\infty, \ a<1 \end{cases}$$

and, when $x \to 1$, we get

$$h_{\mathrm{OpTL}}(x;\lambda,a) \sim \frac{2}{1-x} \to +\infty$$

These results demonstrate the flexibility of $h_{\mathrm{OpTL}}(x;\lambda,a)$. For a more deep shape study, in view of the functional complexity of the function, we chose to do a graphical work in the next part.

With regard to equation (7), the qf of the OpTL distribution is obtained as

$$Q_{\mathrm{OpTL}}(x;\lambda,a) = 1-\sqrt{1-\left\{\frac{\ln(1-\lambda)-\ln[(2\lambda-1)x+1-\lambda]}{c_\lambda^\dagger}\right\}^{1/a}}, \ \lambda \in (0,1)/\left\{\frac{1}{2}\right\} \text{ and } x\in[0,1]$$

The closed form expression of $Q_{\mathrm{OpTL}}(x;\lambda,a)$ is a clear plus; diverse quantile measures can be expressed and computed without problem.

3.4. Graphical work

For comparison purposes, and as a visual benchmark, Figure 1 displays the pdf and hrf of the TL distribution for varying values of $a \in (0,1)$.

Clearly, for $a \in (0,1)$, the pdf the TL distribution is strictly a decreasing function, while the hrf accommodates only a bathtub-shaped property.

Now, let us investigate the shapes of the pdf and hrf of the OpTL distribution. Figure 2 displays them for several values of λ and $a \in (0,1)$.

From this figure, we notice that the pdf of the OpTL distribution accommodates left-skewed, N-type and decreasing shapes, while the hrf exhibits increasing and bathtub shapes.

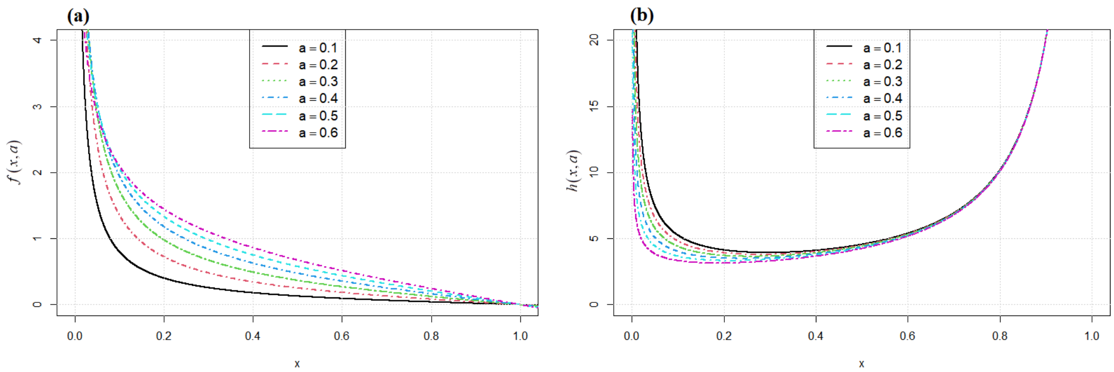

Figure 1. The pdf (a) and hrf (b) plots of the TL distribution for varying values of a

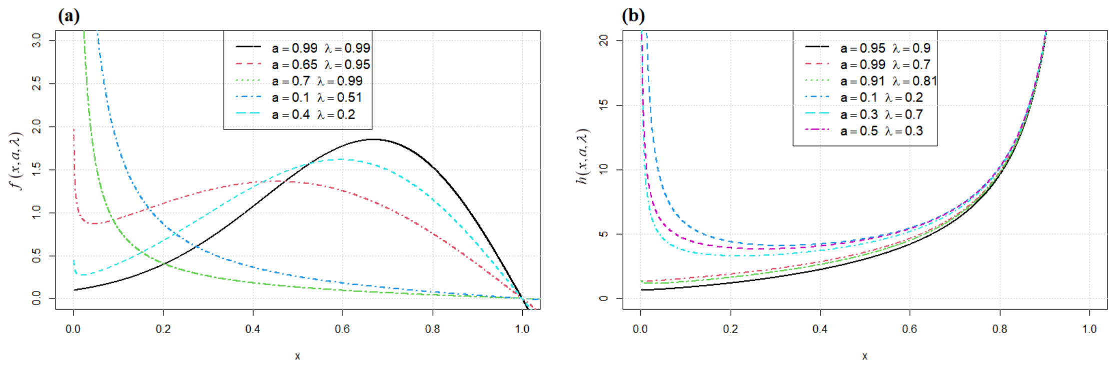

Figure 2. The pdf (a) and hrf (b) plots of the OpTL distribution for varying values of λ and a

If we compare Figures 1 and 2, it is clear that the functional capabilities of the TL distribution are thus considerably enhanced by our CB distribution generator. The OpTL distribution is thus preferable from a statistical modeling viewpoint.

3.5. Moment analysis

Let X be a random variable that follows the OpTL distribution with parameters λ and a (with $\lambda \in (0,1)/\{1/2\}$ to lighten the developments), and r be a non-negative integer. Then the rth moment of X is obtained as

$$v_r = E(X^r) = \int_{-\infty}^{+\infty} x^r f_{\mathrm{OpTL}}(x; \lambda, a) dx.$$

Based on equation (16), we get

$$v_r = c_\lambda 2a \int_0^1 x^r (1-x) \left[x(2-x)\right]^{a-1} \lambda^{[x(2-x)]^a} (1-\lambda)^{1-[x(2-x)]^a} dx.$$

This integral is complicated to manage from an analytical viewpoint, but it can be calculated easily with the help of any mathematical software. Furthermore, based on the series expansion in equation (17), we have

$$v_r = \sum_{k,\,\ell=0}^{+\infty} \sum_{m=0}^{\ell} \alpha_{k,\,\ell,\,m}(\lambda) w_r(a(k+m))$$

where $w_r(a(k+m))$ denotes the rth moment of a random variable Y that follows the TL distribution with parameter $a_* = a(k+m)$. See [22] for the exact and technical expression of this moment. In the expression above, by replacing the infinite upper bound by a finite integer, we obtain an approximation of v_r; the larger the chosen integer, the more precise the approximation.

Classical moment measures can be derived to v_r. In particular, the standard deviation of X follows from the standard formula: $\sigma = \sqrt{v_2 - v_1^2}$. In addition, the rth central moment of X around 0 is given by

$$v_r^* = E[(X - v_1)^r] = \sum_{k=0}^{r} \binom{r}{k} (-1)^{r-k} v_1^{r-k} v_k$$

Based on these ingredients, the moment skewness and kurtosis coefficients of X are, respectively, obtained as the following ratio expressions:

$$S = \frac{v_3^*}{\sigma^3}, \quad K = \frac{v_4^*}{\sigma^4}$$

Numerical computations of the mean, variance, moment skewness, and moment kurtosis for the OpTL distribution are shown in Table 2 for varying parameter values.

Table 2. Numerical computation of the moments of the OpTL distribution

λ	a	v_1^*	σ^2	S	K
0.1	1	0.2038	0.0361	1.3391	4.4503
	3	0.4232	0.0350	0.4226	2.7973
	6	0.5617	0.0248	0.1283	2.5943
	9	0.6327	0.0189	-0.0509	3.3990
0.3	1	0.2786	0.0494	0.8563	2.9095
	3	0.4941	0.0402	0.1534	2.2405
	6	0.6198	0.0269	-0.1370	2.6439
	9	0.6827	0.0200	-0.2463	2.7134
0.6	1	0.3606	0.0574	0.4399	2.2178
	3	0.5663	0.0406	-0.1825	2.3087
	6	0.6777	0.0279	-0.4212	2.6675
	9	0.7321	0.0225	-0.5230	2.9148

From this table, we observe that for a fixed parameter λ, as parameter a increases, the mean increases, whereas the variance decreases. Conversely, for a fixed parameter a, the mean and variance increase as parameter λ increases. Hence, we assert that the mean of the OpTL distribution is an increasing function of parameter a. The results further described the OpTL distribution as one possessing both symmetry and asymmetry properties.

3.6. Parameter estimation

In a basic statistical scenario, we have observations of a certain variable considered as random and whose distribution fits with the OpTL distribution. We aim to estimate the related unknown parameters λ and a based on them. The ML method is employed to this end. It can be described as follows: Let $x_1, \ldots, x_n$ represent n independent observations from a random variable X following the OpTL distribution. The related parameters λ and a are supposed to be unknown. For convenience in the presentation, we suppose that $\lambda \in (0,1)/\{1/2\}$, corresponding to the special CB distribution case. Then, based on the pdf indicated in equation (15), the corresponding likelihood function is given as

$$\begin{aligned} L_{\text{OpTL}}(\lambda, a; x_1, \ldots, x_n) &= \prod_{i=1}^{n} f_{\text{OpTL}}(x_i; \lambda, a) \\ &= c_\lambda^n 2^n a^n \prod_{i=1}^{n}(1-x_i)\left[\prod_{i=1}^{n} x_i(2-x_i)\right]^{a-1} \lambda^{\sum_{i=1}^{n}[x_i(2-x_i)]^a}(1-\lambda)^{n-\sum_{i=1}^{n}[x_i(2-x_i)]^a} \end{aligned}$$

The ML estimates (MLEs) of λ and a are obtained as

$$(\widehat{\lambda}, \widehat{a}) = \text{argmax}_{(\lambda,a)}\, L_{\text{OpTL}}(\lambda, a; x_1, \ldots, x_n)$$

We can eventually replace the likelihood function by its logarithmic version defined as

$$\begin{aligned} \ell_{\text{OpTL}}(\lambda, a; x_1, \ldots, x_n) &= \ln\left[L_{\text{OpTL}}(\lambda, a; x_1, \ldots, x_n)\right] \\ &= n\ln(c_\lambda) + n\ln(2) + n\ln(a) + \sum_{i=1}^{n}\ln(1-x_i) + (a-1)\sum_{i=1}^{n}[\ln(x_i) + \ln(2-x_i)] \\ &+ \ln(\lambda)\sum_{i=1}^{n}[x_i(2-x_i)]^a + \ln(1-\lambda)\left\{n - \sum_{i=1}^{n}[x_i(2-x_i)]^a\right\} \end{aligned}$$

Based on it, on the computational side, the MLEs can be determined by solving the following non-linear equations with respect to λ and a:

$$\frac{\partial \ell_{\text{OpTL}}(\lambda, a; x_1, \ldots, x_n)}{\partial \lambda} = 0, \quad \frac{\partial \ell_{\text{OpTL}}(\lambda, a; x_1, \ldots, x_n)}{\partial a} = 0$$

The ML method has numerous advantages, including the asymptotic unbiasedness and normality properties of the produced estimates. Details on the benefits of applying the ML method can be found in [5].

Based on $\widehat{\lambda}$ and $\widehat{a}$, the functions of the OpTL distribution can be estimated by the substitution approach. In particular, an estimate of $F_{\text{OpTL}}(x; \lambda, a)$ is given by $\widehat{F}(x) = F_{\text{OpTL}}(x; \widehat{\lambda}, \widehat{a})$ and an estimate of the pdf $f_{\text{OpTL}}(x; \lambda, a)$ is obtained as $\widehat{f}(x) = f_{\text{OpTL}}(x; \widehat{\lambda}, \widehat{a})$.

3.7. Simulation

Investigating the performance of the parameter estimates of any proposed probability model is very significant. This gives us an insight of how well an estimate predicts the true parameter values. In this subsection, we treat the parameters of the OpTL distribution arising from the Op family. Random samples of size $n = 25, 50, 100, 200$ and 500 are generated from OpTL distribution at three different sets of parameter values $(\lambda = 0.1, a = 0.2), (\lambda = 0.2, a = 0.6)$ and $(\lambda = 0.3, a = 0.8)$. A Monte Carlo simulation is conducted N times, with $N = 5000$. For $\zeta = a$ and $\zeta = \lambda$, the following quantities are computed:

1. Mean estimate defined by $\frac{1}{N}\sum_{i=1}^{N}\widehat{\zeta}_i$,
2. Bias indicated as $\frac{1}{N}\sum_{i=1}^{N}(\widehat{\zeta}_i - \zeta_i)$,
3. Root mean square error (RMSE) specified by $\sqrt{\frac{1}{N}\sum_{i=1}^{N}(\widehat{\zeta}_i - \zeta_i)^2}$.

Table 3 presents the mean estimates, bias and root mean square error of the parameter of the OpTL distribution.

Table 3. The mean estimate, bias and RMSE of the parameter of the OpTL distribution

Parameters		Mean estimate		Bias		RMSE	
	n	a	λ	a	λ	a	λ
	25	0.2142	0.1147	0.0142	0.0147	0.0512	0.1125
$a = 0.2$	50	0.2083	0.1136	0.0083	0.0136	0.0352	0.0856
$\lambda = 0.1$	100	0.2047	0.1043	0.0047	0.0043	0.0245	0.0575
	200	0.2033	0.1007	0.0033	0.0008	0.0166	0.0370
	500	0.2032	0.0971	0.0032	-0.0029	0.0107	0.0229
	25	0.6402	0.2162	0.0402	0.0163	0.1706	0.1773
$a = 0.6$	50	0.6202	0.2112	0.0202	0.0112	0.1166	0.1305
$\lambda = 0.2$	100	0.6145	0.2042	0.0145	0.0042	0.0886	0.1003
	200	0.6087	0.2009	0.0087	0.0009	0.0609	0.0692
	500	0.6051	0.1977	0.0051	-0.0023	0.0373	0.0420
	25	0.8248	0.3163	0.0248	0.0163	0.1719	0.1889
$a = 0.8$	50	0.8124	0.3103	0.0124	0.0103	0.1486	0.1565
$\lambda = 0.3$	100	0.8114	0.3056	0.0114	0.0056	0.1162	0.1200
	200	0.8023	0.3009	0.0023	0.0009	0.0877	0.0904
	500	0.8018	0.3000	0.0018	0.00002	0.0482	0.0494

Notice that in Table 3, the mean estimate of both parameters converges to the true parameter value as n increases. Likewise, as n increases, the bias and RMSE of both parameter estimates decreases. These properties are consistent with the properties of a good estimator.

3.8. Data fitting

Here, we attempt to show the relevance of the Op family in modeling real world data sets. More precisely, two data sets of different nature are considered for data fitting purposes and in all fairness, five existing bounded lifetime distributions are selected as the main competitors. These competitors are presented below.

1. Unit-Weibull (UW) distribution introduced in [20], and defined by the following pdf:

$$f_{UW}(x;\alpha,\beta) = \frac{1}{x}\alpha\beta\left(-\ln x\right)^{\beta-1}\mathrm{e}^{-\alpha(-\ln x)^{\beta}}, \quad x \in (0,1]$$

   and $f_{UW}(x;\alpha,\beta) = 0$ for $x \notin (0,1]$.
2. Unit-Gompertz (UG) distribution studied in [19], and defined by the following pdf:

$$f_{UG}(x;a,b) = abx^{-(a+1)}e^{-b(x^{-a}-1)}, \quad x \in [0,1]$$

   and $f_{UG}(x;a,b) = 0$ for $x \notin [0,1]$.
3. Log-weighted exponential (LWE) distribution developed in [3], and defined by the following pdf:

$$f_{LWE}(x;\alpha,\lambda) = \frac{\alpha+1}{\alpha}\lambda x^{\lambda-1}\left(1-x^{\alpha\lambda}\right), \quad x \in [0,1]$$

   and $f_{LWE}(x;\alpha,\lambda) = 0$ for $x \notin [0,1]$.
4. TL distribution reported in [26], and defined by the following pdf:

$$f_{TL}(x;a) = 2a\left(1-x\right)\left[x(2-x)\right]^{a-1}, \quad x \in [0,1]$$

   and $f_{TL}(x;a) = 0$ for $x \notin [0,1]$.
5. CB distribution reported in [29], and defined by the following pdf:

$$f_{CB}(x;\lambda) = c_{\lambda}\lambda^{x}(1-\lambda)^{1-x}, \quad x \in [0,1]$$

   where c_{λ} is defined in equation (3) and $f_{CB}(x;\lambda) = 0$ for $x \notin [0,1]$.

Data set 1

The first data set constitutes of trade share data reported in [4]. It contains the following values:

0.140501976,	0.156622976,	0.157703221,	0.160405084,	0.160815045,	0.22145839,	0.299405932,
0.31307286,	0.324612707,	0.324745566,	0.329479247,	0.330021679,	0.337879002,	0.339706242,
0.352317631,	0.358856708,	0.393250912,	0.41760394,	0.425837249,	0.43557933,	0.442142904,
0.444374621,	0.450546652,	0.4557693,	0.46834656,	0.473254889,	0.484600782,	0.488949597,
0.509590268,	0.517664552,	0.527773321,	0.534684658,	0.543337107,	0.544243515,	0.550812602,
0.552722335,	0.56064254,	0.56074965,	0.567130983,	0.575274825,	0.582814276,	0.603035331,
0.605031252,	0.613616884,	0.626079738,	0.639484167,	0.646913528,	0.651203632,	0.681555152,
0.699432909,	0.704819918,	0.729232311,	0.742971599,	0.745497823,	0.779847085,	0.798375845,
0.814710021,	0.822956383,	0.830238342,	0.834204197,	.979355395.		

Data set 2

The second data set reported in [24] is concerned with ordered failure of components. The data are presented as follows:

0.0009,	0.004,	0.0142,	0.0221,	0.0261,	0.0418,	0.0473,	0.0834,	0.1091,	0.1252,
0.1404,	0.1498,	0.175,	0.2031,	0.2099,	0.2168,	0.2918,	0.3465,	0.4035,	0.6143.

Table 4 shows the descriptive statistics of the two data sets.

Table 4. Descriptive statistics for data sets 1 and 2

Data set	Minimum	Mean	Variance	Skewness	Kurtosis	Maximum
1	0.1405	0.5142	0.0375	0.0059	2.4696	0.9794
2	0.0009	0.1613	0.0248	1.2317	4.0743	0.6143

Information drawn from Table 4 reveals that data sets 1 and 2 are, respectively, approximately symmetric and right-skewed in nature. More information about them can be obtained by inspecting the box plot and total time on test (TTT) plot of the two data sets as shown in Figures 3 and 4, respectively.

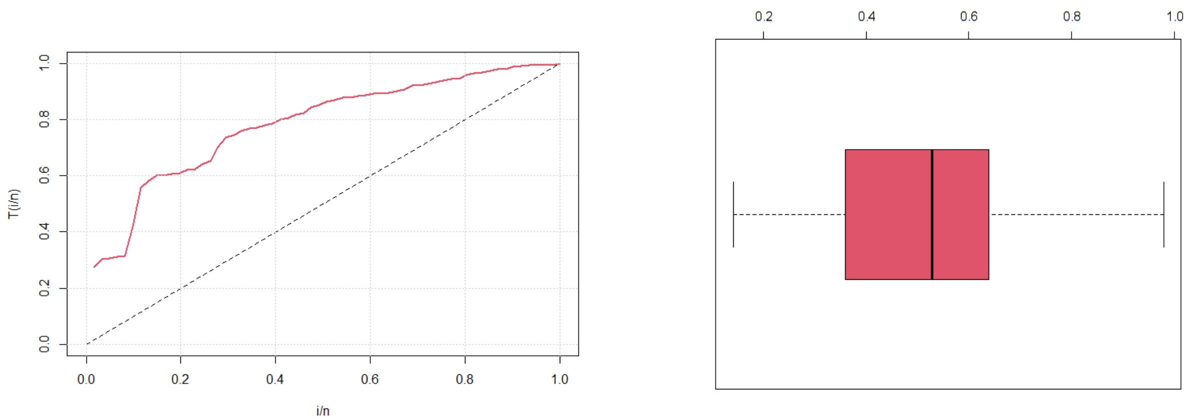

Figure 3. TTT plot (left) and box plot (right) for data set 1

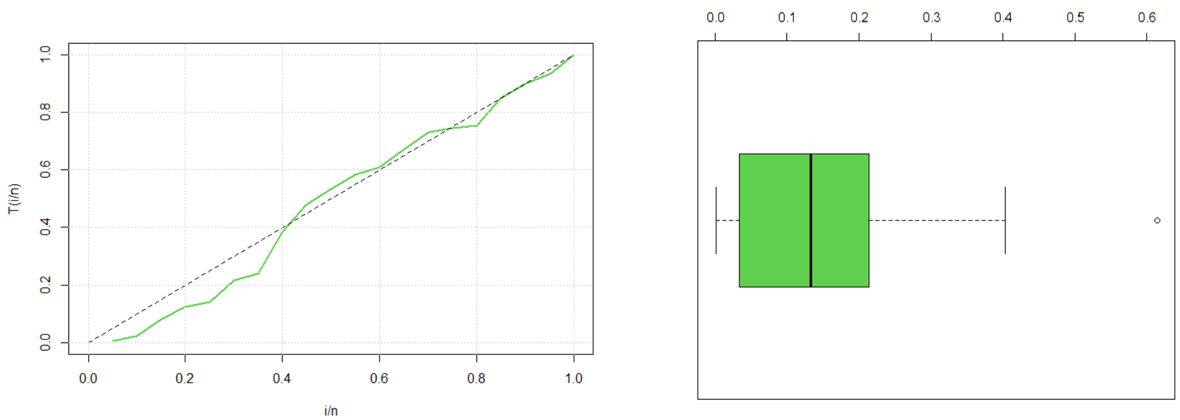

Figure 4. TTT plot (left) and box plot (right) for data set 2

In Figure 3, the TTT line is concave, implying that data set 1 exhibits an increasing failure rate property, while the box plot indicates that the data set is approximately symmetric with no outlier. Also, we observe that the TTT line in Figure 4 is first convex and then concave, suggesting that the data set exhibits an inverted bathtub failure rate property. Whereas, the box plot shows that the data set is right-skewed with the presence of an outlier in the data.

For model comparison and investigation of an appropriate model suitable for fitting the two data sets, we adopt the computation of the maximized log-likelihood (LL), Akaike information criterion (AIC), corrected Akaike information criterion (AICc), Kolmogorov–Smirnov (K-S), and Cramér–Von-Mises (W^*) test statistics with their corresponding p-value. The smaller the value of the AIC and AICc, and the higher the value of the p-values, the more preferable the corresponding model is to fit the data under consideration. Tables 5 and 6 present the summary results for the data sets.

Table 5. Summary results for data set 1

Distribution	Parameter estimates	LL	AIC	AICc	K–S (p-value)	W^* (p-value)
OpTL	$a = 3.3726$ $\lambda = 0.2938$	14.4208	-24.8417	-24.6348	0.0575 (0.9809)	0.0416 (0.9259)
UW	$\alpha = 1.3395$ $\beta = 1.7345$	14.2436	-24.4872	-24.2803	0.0681 (0.9210)	0.0617 (0.8049)
UG	$a = 0.6161$ $b = 1.0921$	10.8758	-17.7518	-17.5449	0.1098 (0.4235)	0.2076 (0.2535)
LWE	α =-2.9$\times 10^{-5}$ $\lambda = 2.6578$	13.0829	-22.1659	-21.9540	0.1025 (0.5108)	0.1356 (0.4376)
CB	$\lambda = 0.5424$	0.0734	1.8532	1.9210	0.1834 (0.0287)	0.6877 (0.0134)
TL	$a = 2.7391$	13.3202	-24.6404	-24.5726	0.0859 (0.7258)	0.0878 (0.6495)

Table 6. Summary results for data set 2

Distribution	Parameter estimates	LL	AIC	AICc	K–S (p-value)	W^* (p-value)
OpTL	$a = 0.8349$ $\lambda = 0.0873$	16.8009	-29.6017	-28.8958	0.1158 (0.9234)	0.0382 (0.9462)
UW	$\alpha = 0.1598$ $\beta = 1.7269$	16.4575	-28.9150	-28.2091	0.1318 (0.8335)	0.0531 (0.8625)
UG	$a = 0.7741$ $b = 0.2782$	14.7625	-25.5251	-24.8192	0.1494 (0.7093)	0.0996 (0.5911)
LWE	α =0.0003 $\lambda = 0.7807$	16.4329	-28.8659	-28.1600	0.1351 (0.812)	0.0521 (0.8689)
CB	$\lambda = 0.0022$	15.5370	-29.074	-28.8518	0.1221 (0.8713)	0.0410 (0.9325)
TL	$a = 0.5112$	15.6167	-29.2337	-29.0115	0.1848 (0.4481)	0.1114 (0.5360)

3.9. Discussion of the results

An appropriate model suitable for fitting any real data set is traceable to the one having the maximized log-likelihood value and the least value judging from AIC, AICc, K-S, and W^*.

A careful study of Tables 5 and 6 reveals that the OpTL distribution has satisfied the aforementioned criteria across the two data sets (except the AICc in Table 6 which was in favor of the TL distribution) and performed better than the competitor distributions, thus being declared the most appropriate model for fitting the two data sets. The superiority of the OpTL distribution is also established judging from the p-value of the test statistics in Tables 5 and 6, since the OpTL distribution has the highest p-value in both tables. It is appealing to reckon that only the CB distribution with (p-value < 0.05), failed to model the first data set adequately, hence the usefulness of the proposed Op family of distributions.

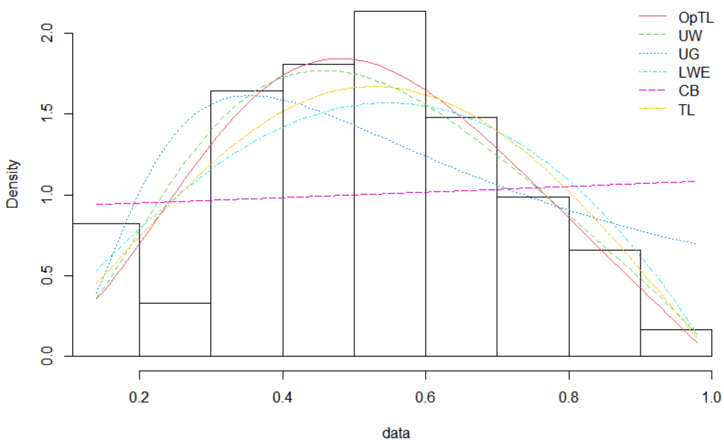

Figure 5. Estimated pdf fit of the distributions for data set 1

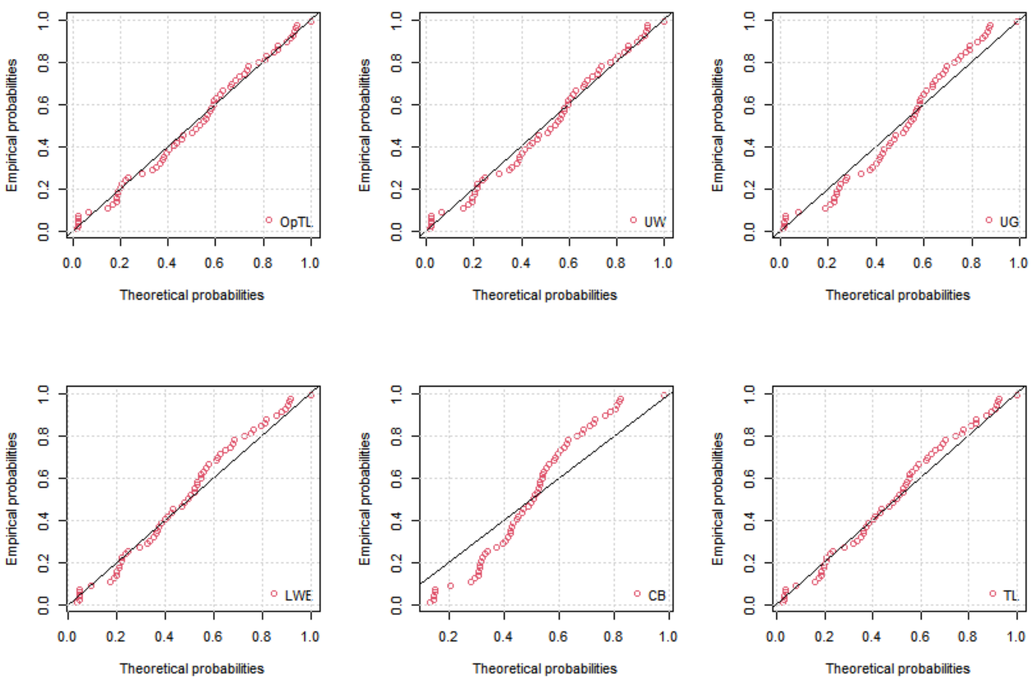

Figure 6. P-P plot of the distributions for data set 1

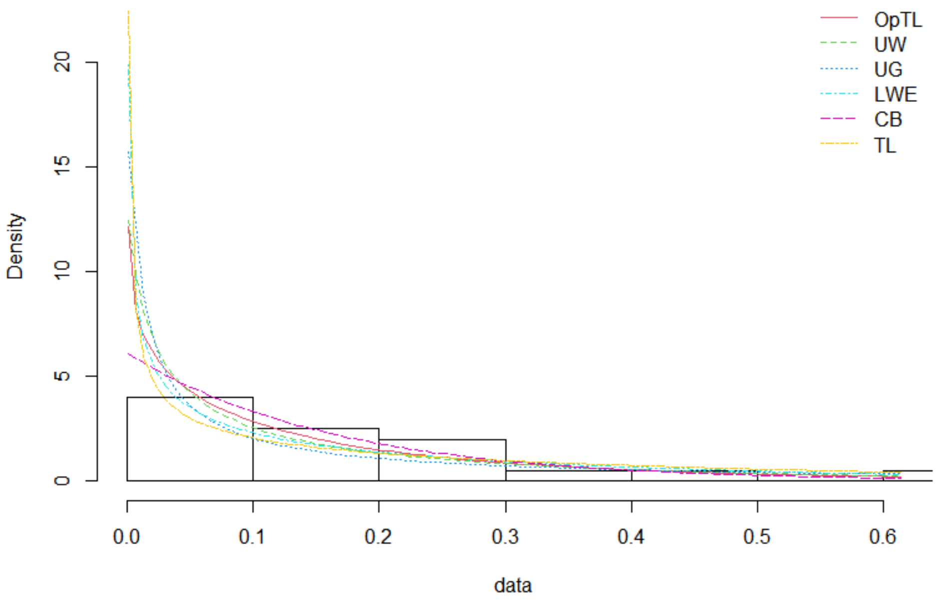

Figure 7. Estimated pdf fit of the distributions for data set 2

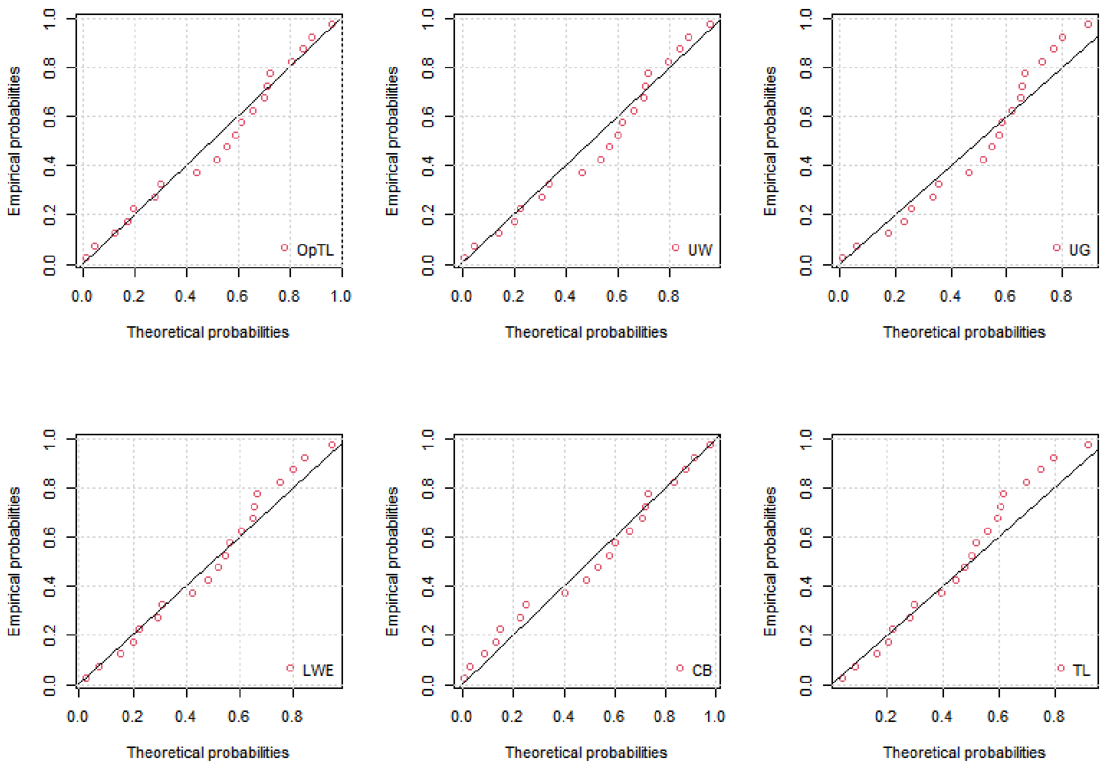

Figure 8. P-P plot of the distributions for data set 2

Alternatives to numerical computation for goodness-of-fit are graphical plots such as the estimated pdf fit and Probability-Probability (P-P) plots. We present them for the data sets in Figures 5–8.

A clear observation of all the plots validates the superiority of the OpTL distribution over the competitor distributions in fitting the two data sets, as its fit matches closer to the empirical pdf of each data set than the rest of the distributions. It is also obvious from the plots obtained from data sets 1 and 2 that the CB distribution performed poorly in fitting the data sets.

4. Conclusion

In this paper, we contributed to the development of the CB distribution by using its cdf as a generator of distributions, creating the Op family. We thus transposed the functionalities of the CB distribution to enhance the modeling capabilities of a given absolutely continuous univariate baseline distribution. In the first part, we provided all the necessary theory to apprehend the Op family as it is best understood, including the quantiles, expansion of important functions, and moments. In the second part, we exemplify it by considering the TL distribution; the OpTL distribution was thus created. It can be presented as a new two-parameter distribution with support of $[0, 1]$. By using mathematical arguments, graphics, and tables, we show that it has more modeling capacities than the TL distribution. Then, its quantile and moment properties were retained. We then use the OpTL distribution from a statistical perspective. The maximum likelihood method is applied. Two different data sets were considered, and it was revealed that the OpTL distribution could fit them better than the UW, UG, and LWE distributions. It also outperformed the TL and CB distributions.

The paper thus sets the basis for more in the direction of the use of the CB distribution in probability and statistics. New distributions are elaborated and can be used in various data-fitting scenarios in all areas of applied sciences. The perspectives of the Op family are numerous, including the developments of regression models, cluster models, and multivariate models.

Acknowledgement

The authors express their gratitude to two anonymous reviewers for their insightful critiques and recommendations on an earlier version of this paper.

Funding

The authors declare that there are no funding for the publication of this paper.

References

[1] Akata, I. U., Opone, F. C., and Osagiede, F. E. U. The Kumaraswamy unit-Gompertz distribution and its application to lifetime datasets. *Earthline Journal of Mathematical Sciences 11*, 1 (2023), 1–22.

[2] Al-Shomrani, A., Arif, O., Shawky, A., Hanif, S., and Shahbaz, M. Q. Topp–Leone family of distributions: Some properties and application. *Pakistan Journal of Statistics and Operation Research 12* (2016), 443–451.

[3] Altun, E. The log-weighted exponential regression model: alternative to the beta regression model. *Communications in Statistics - Theory and Methods 50*, 10 (2021), 2306–2321.

[4] Bantan, R. A. R., Jamal, F., Chesneau, C., and Elgarhy, M. Theory and applications of the unit Gamma/Gompertz distribution. *Mathematics 9*, 16 (2021), 1850.

[5] Casella, G., and Berger, R. L. *Statistical Inference*. Duxbury Resource Center, 2001.

[6] Chesneau, C. A note on an extreme left skewed unit distribution: Theory, modelling and data fitting. *Open Statistics 2*, 1 (2021), 1–23.

[7] Chesneau, C., and Opone, F. The power continuous Bernoulli distributions: Theory and applications. *Reliability: Theory & Applications 17*, 4(71) (2022), 232–248.

[8] Chesneau, C., Opone, F. C., and Ubaka, N. O. Theory and applications of the transmuted continuous Bernoulli distribution. *Earthline Journal of Mathematical Sciences 10*, 2 (2022), 385–407.

[9] Van Dorp, J. R., and Kotz, S. Modeling income distributions using elevated distributions on a bounded domain. In *Distribution Models Theory* (Singapure, 2006), R. Herrerías Pleguezuelo, J. Callejón Céspedes and J. M. Herrerías Velasco, Eds., World Scientific Publishing Co Pte Ltd., pp. 1–25.

[10] Eugene, N., Lee, C., and Famoye, F. Beta-normal distribution and its applications. *Communications in Statistics - Theory and Methods 31*, 4 (2002), 497–512.

[11] Genc, A. İ. Moments of order statistics of Topp–Leone distribution. *Statistical Papers 53* (2012), 117–131.

[12] Ghitany, M. E., Kotz, S., and Xie, M. On some reliability measures and their stochastic orderings for the Topp-Leone distribution. *Journal of Applied Statistics 32*, 7 (2005), 715–722.

[13] Gomaa, R. S., Magar, A. M., Alsadat, N., Almetwally, E. M., and Tolba, A. H. The unit alpha-power Kum-modified size-biased Lehmann type II distribution: Theory, simulation, and applications. *Symmetry 15*, 6 (2023), 1283.

[14] Gordon-Rodriguez, E., Loaiza-Ganem, G., and Cunningham, J. P. The continuous categorical: a novel simplex-valued exponential family. In *Proceedings of the 37th International Conference on Machine Learning, 13-18 July 2020, PMLR* (2020), H. Daumé III and A. Singh, Eds., PMLR, pp. 3637–3647.

[15] Hussain, M. A., Tahir, M. H., and Cordeiro, G. M. A new Kumaraswamy generalized family of distributions: Properties and applications. *Mathematica Slovaca 70*, 6 (2020), 1491–1510.

[16] Korkmaz, M. Ç., and Chesneau, C. On the unit Burr-XII distribution with the quantile regression modeling and applications. *Computational and Applied Mathematics 40* (2021), 29.

[17] Korkmaz, M. Ç., Leiva, V., and Martin-Barreiro, C. The continuous Bernoulli distribution: Mathematical characterization, fractile regression, computational simulations, and applications. *Fractal and Fractional 7*, 5 (2023), 386.

[18] Loaiza-Ganem, G., and Cunningham, J. P. The continuous Bernoulli: fixing a pervasive error in variational autoencoders. *Advances in Neural Information Processing Systems 32* (2019), 13266–13276.

[19] Mazucheli, J., Menezes, A. F., and Dey, S. Unit-Gompertz distribution with applications. *Statistica 79*, 1 (2019), 25-–43.

[20] Mazucheli, J., Menezes, A. F. B., Fernandes, L. B., de Oliveira, R. P., and Ghitany, M. E. The unit-Weibull distribution as an alternative to the Kumaraswamy distribution for the modeling of quantiles conditional on covariates. *Journal of Applied Statistics 47*, 6 (2020), 954–974.

[21] Muse, A. H., Tolba, A. H., Fayad, E., Ali, O. A. A., Nagy, M., and Yusuf, M. Modelling the COVID-19 mortality rate with a new versatile modification of the log-logistic distribution. *Computational Intelligence and Neuroscience 2021* (2021), 640794.

[22] Nadarajah, S., and Kotz, S. Moments of some J-shaped distributions. *Journal of Applied Statistics 30*, 3 (2003), 311–317.

[23] Nasiru, S., Abubakari, A. G., and Chesneau, C. New lifetime distribution for modeling data on the unit interval: Properties, applications and quantile regression. *Mathematical and Computational Applications 27*, 6 (2022), 105.

[24] Nigm, A. M., Al-Hussaini, E. K., and Jaheen, Z. F. Bayesian one-sample prediction of future observations under Pareto distribution. *Statistics 37*, 6 (2003), 527–536.

[25] Opone, F. C., Akata, I. U., and Altun, E. The Marshall-Olkin extended unit-Gompertz distribution: its properties, regression model and applications. *Statistica 82*, 2 (2022), 97–118.

[26] Opone, F. C., Ekhosuehi, N., and Omosigho, S. E. Topp-Leone power Lindley distribution(Tlpld): its properties and application. *Sankhya A 84*, 2 (2022), 597–608.

[27] Opone, F. C, and Iwerumor, B. A new Marshall-Olkin extended family of distributions with bounded support. *Gazi University Journal of Science 34*, 3 (2021), 899–914.

[28] Ramadan, A. T., Tolba, A. H., and El-Desouky, B. S. A unit half-logistic geometric distribution and its application in insurance. *Axioms 11*, 12 (2022), 676.

[29] Wang, K.-S., and Lee, M.-Y. (22 November 2020) *Continuous Bernoulli Distribution: Simulator and Test Statistic* (accesessed on 23 August 2021).

[30] Zhou, M., Yang, D. W., Wang, Y., and Nadarajah, S. Some j-shaped distributions: sums, products and ratios. In *RAMS '06. Annual Reliability and Maintainability Symposium* (Newport Beach, CA, USA, 2006), IEEE, pp. 175–181.