Thailand Statistician
July 2022; 20(3): 489-507
http://statassoc.or.th
Contributed paper

The Beta Topp-Leone Generated Family of Distributions and Theirs Applications

Atchariya Watthanawisut, Winai Bodhisuwan* and Thidaporn Supapakorn
Department of Statistics, Faculty of Science, Kasetsart University, Bangkok, Thailand.
*Corresponding author; e-mail: fsciwnb@ku.ac.th

Received: 1 October 2020
Revised: 16 June 2021
Accepted: 6 July 2021

Abstract

This work aims to establish a new family of distributions, namely the beta Topp-Leone generated family of distributions. The proposed family of distributions is combined from two families: the beta generated family and the Topp-Leone generated family. Some statistical properties of the proposed family are derived, e.g., linear representation, ordinary moments, and moment generating function. Furthermore, a new modification of Weibull distribution, namely the beta Topp-Leone Weibull distribution, is studied. The beta Topp-Leone Weibull distribution has flexible hazard shapes. Some statistical properties of the proposed distribution are studied, e.g., transformation, quantile function, ordinary moments, and moment generating function. The distribution parameters are estimated with the methods of maximum likelihood estimation. The proposed distribution shows more appropriate than other candidate distributions for fitting with the complete and censored datasets based on the values of Akaike's information criterion, Bayesian information criterion, Akaike's information corrected criterion, and Hannon and Quinn's information criterion.

Keywords: Beta generated family, Topp-Leone generated family, T-X family, Weibull distribution, lifetime distribution, censored data.

1. Introduction

Several classical continuous distributions are widely used for modeling data in many areas such as biostatistics, business analytics, econometrics, environmental statistics, reliability engineering, and others. Recent developments focus on defining the new families of distributions that extend classical distributions and, meanwhile, offer considerable flexibility in modeling data. Hence, several families of distributions have been proposed by adding more (location, shape, or scale) parameters to generate new distributions in the statistical literary work recently.

Eugene et al. (2002) introduced the beta generated (BG) family of distributions and noted that the BG family provides excellent flexibility for modeling data. Jones (2004) studied some properties of BG family. Let $G(x;\boldsymbol{\xi})$ be a baseline cumulative distribution function (cdf), let $g(x;\boldsymbol{\xi}) = dG(x;\boldsymbol{\xi})/dx$ be a baseline probability distribution function (pdf) of a random variable X and $\boldsymbol{\xi}$ the $p \times 1$ vector of associated parameter. The BG family pdf is expressed as

$$f_{BG}(x;a,b,\boldsymbol{\xi}) = \frac{1}{B(a,b)} g(x;\boldsymbol{\xi}) G(x;\boldsymbol{\xi})^{a-1}[1-G(x;\boldsymbol{\xi})]^{b-1}, \quad a,b>0, \tag{1}$$

where $B(a,b) = \int_0^1 t^{a-1}(1-t)^{b-1}dt$ is the complete beta function. Thus, the BG family of distributions has its cdf as

$$F_{BG}(x; a, b, \boldsymbol{\xi}) = I_{G(x;\boldsymbol{\xi})}(a, b), \tag{2}$$

where the function $I_{G(x;\boldsymbol{\xi})}(a,b)$ denotes the incomplete beta ratio function defined by

$$I_y(a,b) = \frac{B_y(a,b)}{B(a,b)},$$

where $B_y(a,b) = \int_0^y t^{a-1}(1-t)^{b-1}dt$, $0 < t < 1$, is the incomplete beta function.

In addition, some well-known families are proposed, such as the Kumaraswamy generated (Cordeiro and de Castro, 2011), McDonald generated (Alexander et al., 2012), and Kummer beta generated (Pescim et al., 2012) families of distributions. Furthermore, Alzaatreh et al. (2013) developed a new method for generating family distributions, referred to as the T-X family of distributions. A good review of methodologies for generating continuous distributions is referred to the work of Lee et al. (2013).

Sangsanit and Bodhisuwan (2016) introduced the Topp-Leone generated (TLG) family of distributions, its pdf and cdf are, respectively,

$$f_{TLG}(x; c, \boldsymbol{\xi}) = 2cg(x;\boldsymbol{\xi})(1 - G(x;\boldsymbol{\xi}))[1 - (1 - G(x;\boldsymbol{\xi}))^2]^{c-1}, \quad c > 0, \tag{3}$$

and

$$F_{TLG}(x; c, \boldsymbol{\xi}) = [1 - (1 - G(x;\boldsymbol{\xi}))^2]^c, \tag{4}$$

Using this method, the Topp-Leone generalized exponential distribution was proposed by Sangsanit and Bodhisuwan (2016) with applying to maximum stress per cycle 31,000 psi and breaking stress of carbon fibers datasets.

The rest of paper is structured as follows. Section 2 is the definition of the beta Topp-Leone generated (BTLG) family. In Sections 3 and 4, linear representation of the proposed family and some of its statistical properties are obtained. In Section 5, we propose a new modification of Weibull distribution called the beta Topp-Leone Weibull (BTLW) distribution. In Section 6, some of its statistical properties are investigated. In Section 7, the distribution parameters are estimated by maximum likelihood estimation (MLE). In Section 8, the flexibility of the proposed distribution will be explored through two applications to real datasets. Finally, Section 9 is the conclusion.

2. The Beta Topp-Leone Generated Family of Distributions

Let $r(t)$ be the pdf of a random variable $T \in [p,q]$ for $-\infty < p < q < \infty$ and let $W[G(x)]$ be a function of the cdf of a random variable X satisfying the following conditions:

(i) $W[G(x)] \in [p,q]$,

(ii) $W[G(x)]$ is differentiable and monotonically non-decreasing,

(iii) $W[G(x)] \to p$ as $x \to -\infty$ and $W[G(x)] \to q$ as $x \to \infty$.

The cdf of the T-X family defined by Alzaatreh et al. (2013) is

$$F(x) = \int_p^{W[G(x)]} r(t)dt \tag{5}$$

where $W[G(x)]$ satisfies the above conditions. The pdf corresponding to Equation (5) is given by

$$f(x) = \left\{\frac{d}{dx}W[G(x)]\right\} r\{W[G(x)]\} \tag{6}$$

Setting $W[G(x)] = F_{TLG}(x; c, \boldsymbol{\xi})$ and $r(t)$ is the pdf beta distribution, we define the cdf of BTLG family by

$$F_{BTLG}(x; a, b, c, \boldsymbol{\xi}) = I_{[1-(1-G(x;\boldsymbol{\xi}))^2]^c}(a, b). \tag{7}$$

The pdf corresponding to Equation (7) is

$$\begin{aligned} f_{BTLG}(x; a, b, c, \boldsymbol{\xi}) =& \frac{2c}{B(a,b)} g(x;\boldsymbol{\xi})(1 - G(x;\boldsymbol{\xi}))[1 - (1 - G(x;\boldsymbol{\xi}))^2]^{ac-1} \\ & \times [1 - [1 - (1 - G(x;\boldsymbol{\xi}))^2]^c]^{b-1}, \end{aligned} \tag{8}$$

where $G(x; \boldsymbol{\xi})$ is the baseline with a parameter vector $\boldsymbol{\xi}$ and $a, b, c > 0$ are shape parameters. Hereafter, a random variable X with cdf in Equation (7) is denoted by $X \sim BTLG(a, b, c, \boldsymbol{\xi})$. Further, the parameters a, b, c and the vector of the baseline parameter $\boldsymbol{\xi}$ can be omitted. That is, we can write $G(x) = G(x; \boldsymbol{\xi})$ and $F(x) = F(x; a, b, c, \boldsymbol{\xi})$.

The importance of the BTLG family is that there are contained several sub-family generated of distributions. The BTLG family reduces to the TLG family (Sangsanit and Bodhisuwan, 2016) when $a = 1$ and $b = 1$. If $c = 1$ it reduces to the beta transmuted generated (BTG) family (Afify et al., 2017) with transmuted parameter equals to 1. If $b = 1$, it gives as special case the exponentiated Topp-Leone generated (ETLG) family. The exponentiated transmuted generated (ETG) family (Merovci et al., 2017) with transmuted parameter equals to 1 is also a sub-family when $b = 1$ and $c = 1$. The transmuted generated (TG) family (Shaw and Buckley, 2009) with transmuted parameter equals to 1 is clearly a special case for $a = 1, b = 1$ and $c = 1$. In Table 1, the relationship between sub-family defined from the BTLG family is provided.

Table 1 Sub-families of the BTLG family of distributions

Families	Parameters				$F(x)$	References
	a	b	c	$\boldsymbol{\xi}$		
BTG	a	b	1	$\boldsymbol{\xi}$	$I_{[1-(1-G(x;\boldsymbol{\xi}))^2]}(a, b)$	Afify et al. (2017)
ETLG	a	1	c	$\boldsymbol{\xi}$	$[1 - (1 - G(x;\boldsymbol{\xi}))^2]^{ac}$	–
ETG	a	1	1	$\boldsymbol{\xi}$	$[2G(x;\boldsymbol{\xi}) - G(x;\boldsymbol{\xi})^2]^a$	Merovci et al. (2017)
TLG	1	1	c	$\boldsymbol{\xi}$	$[1 - (1 - G(x;\boldsymbol{\xi}))^2]^c$	Sangsanit and Bodhisuwan (2016)
TG	1	1	1	$\boldsymbol{\xi}$	$2G(x;\boldsymbol{\xi}) - G(x;\boldsymbol{\xi})^2$	Shaw and Buckley (2009)

The survival function and hazard function of the BTLG family are, respectively,

$$S(x; a, b, c, \boldsymbol{\xi}) = 1 - I_{[1-(1-G(x;\boldsymbol{\xi}))^2]^c}(a, b) = I_{1-[1-(1-G(x;\boldsymbol{\xi}))^2]^c}(b, a), \tag{9}$$

and

$$\begin{aligned} h(x; a, b, c, \boldsymbol{\xi}) =& \frac{2c}{B(a,b) I_{1-[1-(1-G(x;\boldsymbol{\xi}))^2]^c}(b, a)} g(x;\boldsymbol{\xi})(1 - G(x;\boldsymbol{\xi})) \\ & \times [1 - (1 - G(x;\boldsymbol{\xi}))^2]^{ac-1}[1 - [1 - (1 - G(x;\boldsymbol{\xi}))^2]^c]^{b-1}. \end{aligned} \tag{10}$$

The quantile function of the BTLG family can be derived by solving Equation (7) as

$$Q_{BTLG}(u; a, b, c, \boldsymbol{\xi}) = F^{-1}_{BTLG}(u; a, b, c, \boldsymbol{\xi}) = G^{-1}\{1 - [1 - (I^{-1}_u(a, b))^{1/c}]^{1/2}; \boldsymbol{\xi}\}, \tag{11}$$

where $I^{-1}_u(a, b)$ represents the inverse of the incomplete beta ratio function (Majumder and Bhattacharjee, 1973).

3. Linear Representation of the BTLG Family

Some useful expansions of Equation (7) and Equation (8) can be derived using the exponentiated generated (EG) family of distributions. For any baseline cdf $G(x)$, a random variable X is distributed as the EG family of distributions with power parameter $\theta > 0$, $X_\theta \sim EG(\theta)$ and the pdf and cdf of EG family are $f_\theta(x) = ag(x)G^{\theta-1}(x)$ and $F_\theta(x) = G^\theta(x)$ respectively.

Firstly, for non-integer real value $b > 0$, the term of $(1-t)^{b-1}$ under the integral is replaced by the power series, and is expressed as

$$\begin{aligned}\int_0^x t^{a-1}(1-t)^{b-1}dt &= \sum_{j=0}^{\infty}(-1)^j\binom{b-1}{j}\int_0^x t^{a+j-1}dt\\ &= \sum_{j=0}^{\infty}\frac{1}{a+j}(-1)^j\binom{b-1}{j}x^{a+j},\end{aligned}$$

where the binomial coefficient $\binom{b-1}{j} = \Gamma(b)/\Gamma(b-j)j!$ is defined for any real value of b and $\Gamma(\cdot)$ is gamma function. Consequently, we obtain

$$F_{BTLG}(x;a,b,c,\boldsymbol{\xi}) = \frac{1}{B(a,b)}\sum_{j=0}^{\infty}\frac{1}{a+j}(-1)^j\binom{b-1}{j}[1-(1-G(x;\boldsymbol{\xi}))^2]^{c(a+j)}.$$

Furthermore, by using binomial expansion, the cdf of BTLG family will be

$$F_{BTLG}(x;a,b,c,\boldsymbol{\xi}) = \sum_{j,k=0}^{\infty}\sum_{m=0}^{2k} w_{j,k,m}G^m(x;\boldsymbol{\xi}) = \sum_{j,k=0}^{\infty}\sum_{m=0}^{2k} w_{j,k,m}F_m(x;\boldsymbol{\xi}) \tag{12}$$

where,

$$w_{j,k,m} = \frac{(-1)^{j+k+m}}{B(a,b)(a+j)}\binom{b-1}{j}\binom{c(a+j)}{k}\binom{2k}{m}.$$

By differentiating Equation (12), we obtain

$$f_{BTLG}(x;a,b,c,\boldsymbol{\xi}) = \sum_{j,k=0}^{\infty}\sum_{m=0}^{2k} w_{j,k,m}mg(x;\boldsymbol{\xi})G(x;\boldsymbol{\xi})^{m-1} = \sum_{j,k=0}^{\infty}\sum_{m=0}^{2k} w_{j,k,m}f_m(x;\boldsymbol{\xi}). \tag{13}$$

If $b > 0$ is an integer, the index j in Equations (12) and (13) will run from 0 to $b-1$, and if both a and c are integers, then the index k will run from 0 to $c(a+j)$.

Equation (13) shows that the pdf of BTLG distribution can be expressed as a linear representation of the pdf of EG distribution. Thus, several statistical properties of the BTLG family can be derived from the EG family. Many members of EG family are studied over the past three decades or so, for instance, exponentiated Weibull (Mudholkar and Srivastava, 1993), exponentiated exponential (Gupta and Kundu, 1999) and exponentiated Gumbel (Nadarajah, 2006) distributions.

4. Statistical Properties of the BTLG Family

In this section, the ordinary moments and moment generating function (mgf) of the BTLG family are derived. The derived formulas will be processed in computer programming languages, such as Matlab, Maple, Mathematica, and R, which currently can handle with analytical expressions of enormous size and complexity. The infinite limit in the sums may be replaced by a large positive integer for most practical purposes.

4.1. Ordinary moments

The three formulas for the rth ordinary moment $\mu_r' = E(X^r)$ of X are derived. The first formula for ordinary moments can be straightforwardly obtained from Equation (13) as

$$\mu_r^{'} = \sum_{j,k=0}^{\infty} \sum_{m=0}^{2k} w_{j,k,m} E(X_m^r). \tag{14}$$

The second formula is simply obtained as in term of the probability weighted moments $\tau_{r,s} = E(Y^r G(Y)^s)$, where Y have the baseline distribution $G(x)$, for $r, s = 0, 1, \ldots$. The term of $G(x)^\theta$ for real non-integer $\theta > 0$ can be expressed as

$$G(x)^\theta = \sum_{s=0}^{\infty} t_s(\theta) G(x)^s, \tag{15}$$

where

$$t_s(\theta) = \sum_{p=s}^{\infty} (-1)^{s+p} \binom{\theta}{p} \binom{p}{s}.$$

By substituting Equation (15) in Equation (13) yields

$$\mu_r^{'} = E(X^r) = \sum_{s=0}^{\infty} \pi_s \tau_{r,s}, \tag{16}$$

where

$$\pi_s = \sum_{j,k=0}^{\infty} \sum_{m=0}^{2k} m w_{j,k,m} t_s(m-1).$$

The third formula for μ_r' can be derived from Equation (16) in terms of the baseline quantile function $Q_G(u) = G^{-1}(u)$

$$\tau_{r,s} = \int_0^{\infty} y^r G(y)^s g(y) \mathrm{d}y = \int_0^1 Q(u)^r u^s \mathrm{d}u. \tag{17}$$

In addition, the central moments (μ_r) and cumulants (κ_r) of X are obtained by employing the ordinary moments as

$$\mu_r = \sum_{k=0}^{r} \binom{r}{k} (-1)^k \mu_1'^k \mu_{r-k}' \quad \text{and} \quad \kappa_r = \mu_r' - \sum_{k=1}^{r-1} \binom{r-1}{k-1} \kappa_k \mu_{r-k}',$$

respectively, where $\kappa_1 = \mu_1'$. The skewness and kurtosis of X can be derived from the second, third and fourth cumulants with their relationships.

4.2. Moment generating function

The three formulas for the mgf of X, say $M_X(t) = E(e^{tX})$, are provided. First, it requires the following series expansion:

$$e^{tx} = \sum_{r=0}^{\infty} \frac{(tx)^r}{r!}.$$

Hence, the first for mgf of BTLG family can be expressed in terms of the rth ordinary moment from Equation (14) as

$$M_X(t) = \sum_{i=0}^{\infty} \frac{\mu_i^{'} t^i}{i!}. \tag{18}$$

The second formula for the mgf of BTLG family can be written from Equation (13)

$$M_X(t) = \sum_{j,k=0}^{\infty} \sum_{m=0}^{2k} w_{j,k,m} M_{X_m}(t), \tag{19}$$

where $M_{X_m}(t)$ is the mgf of X_m. Therefore, the mgf of BTLG can be determined from the mgf of EG such as those determined by Nadarajah and Kotz (2006).

Lastly, the third formula for the mgf of BTLG family can be expressed from Equation (13) as

$$M_X(t) = \sum_{j,k=0}^{\infty} \sum_{m=0}^{2k} m w_{j,k,l,m} \rho(t, m-1), \tag{20}$$

where

$$\rho(t, \theta) = \int_{-\infty}^{\infty} e^{tx} g(x) G(x)^{\theta} \mathrm{d}x = \int_0^1 e^{tQ(u)} u^{\theta} \mathrm{d}u.$$

5. The Beta Topp-Leone Weibull Distribution

The Weibull distribution (Weibull, 1951) with exponential and Rayleigh distributions are special cases which is one of the most commonly used distributions for monotone hazards modeling. Nevertheless, it does not provide an appropriate fit for modeling lifetime data with non-monotone hazards such as bathtub-shaped and unimodal hazards.

For the reasons outlined above, the primary purpose of the modification or extension of the Weibull distribution is to describe and fit the data sets with non-monotonic hazard, such as, the bathtub-shaped and unimodal hazards. Some modifications of Weibull distribution are exponentiated Weibull (EW) (Mudholkar and Srivastava, 1993), modified Weibull (Lai et al., 2003), beta Weibull (BW) (Lee et al., 2007), generalized modified Weibull (Carrasco et al., 2008), Kumaraswamy Weibull (KW) (Cordeiro et al., 2010), beta modified Weibull (Silva, 2010), generalized modified Weibull power series (Bagheri et al., 2016) and Topp-Leone Weibull (TLW) (Aryal et al., 2017) distributions. Furthermore, the Almalki and Nadarajah (2014) provide review of the literature on modifications of Weibull distribution.

The two-parameter Weibull distribution is specified by its cdf

$$G_W(x; \alpha, \lambda) = 1 - e^{-\alpha x^{\lambda}}, \quad x > 0 \tag{21}$$

where $\alpha > 0$ and $\lambda > 0$ are the scale and shape parameters, respectively. The corresponding pdf is

$$g_W(x; \alpha, \lambda) = \alpha \lambda x^{\lambda - 1} e^{-\alpha x^{\lambda}}, \quad x > 0. \tag{22}$$

The corresponding hazard function is

$$h_W(x; \alpha, \lambda) = \alpha \lambda x^{\lambda - 1}, \quad x > 0, \tag{23}$$

which can be increasing, decreasing or constant depending on $\lambda > 1$, $\lambda < 1$ or $\lambda = 1$, respectively.

In this work, a new modification of Weibull distribution called the BTLW distribution is introduced by using the BTLG family. Substituting Equation (21) and Equation (22) in Equation (7) and Equation (8), the cdf and pdf of BTLW distribution $(a, b, c, \alpha, \lambda > 0)$ are obtained as

$$F_{BTLW}(x; a, b, c, \alpha, \lambda) = I_{(1-e^{-2\alpha x^{\lambda}})^c}(a, b), \quad x > 0, \tag{24}$$

and

$$\begin{aligned} f_{BTLW}(x; a, b, c, \alpha, \lambda) =& \frac{2c\alpha\lambda}{B(a, b)} x^{\lambda-1} e^{-2\alpha x^{\lambda}} (1 - e^{-2\alpha x^{\lambda}})^{ac-1} \\ & \times [1 - (1 - e^{-2\alpha x^{\lambda}})^c]^{b-1} \quad x > 0, \end{aligned} \tag{25}$$

respectively. A random variable X is distributed according to the cdf Equation (24) is denoted by $X \sim BTLW(a, b, c, \alpha, \lambda)$.

Some pdf plots of BTLW distribution with specified parameter values a, b, c, α and λ are illustrated as in Figure 1.

The survival function and hazard function of the BTLW distribution are, respectively,

$$S_{BTLW}(x; a, b, c, \alpha, \lambda) = 1 - I_{(1-e^{-2\alpha x^{\lambda}})^c}(a, b) = I_{1-(1-e^{-2\alpha x^{\lambda}})^c}(b, a), \quad x > 0, \tag{26}$$

and

$$\begin{aligned} h_{BTLW}(x; a, b, c, \alpha, \lambda) =& \frac{2c\alpha\lambda}{B(a, b) I_{1-(1-e^{-2\alpha x^{\lambda}})^c}(b, a)} x^{\lambda-1} e^{-2\alpha x^{\lambda}} (1 - e^{-2\alpha x^{\lambda}})^{ac-1} \\ & \times [1 - (1 - e^{-2\alpha x^{\lambda}})^c]^{b-1}, \quad x > 0. \end{aligned} \tag{27}$$

Some BTLW hazard plots for different values of a, b, c, α and λ are shown in Figure 2. The BTLW hazard can be monotonically increasing, monotonically decreasing, bathtub-shaped, upside-down bathtub-shaped and modified bathtub-shaped depending on the values of its parameters.

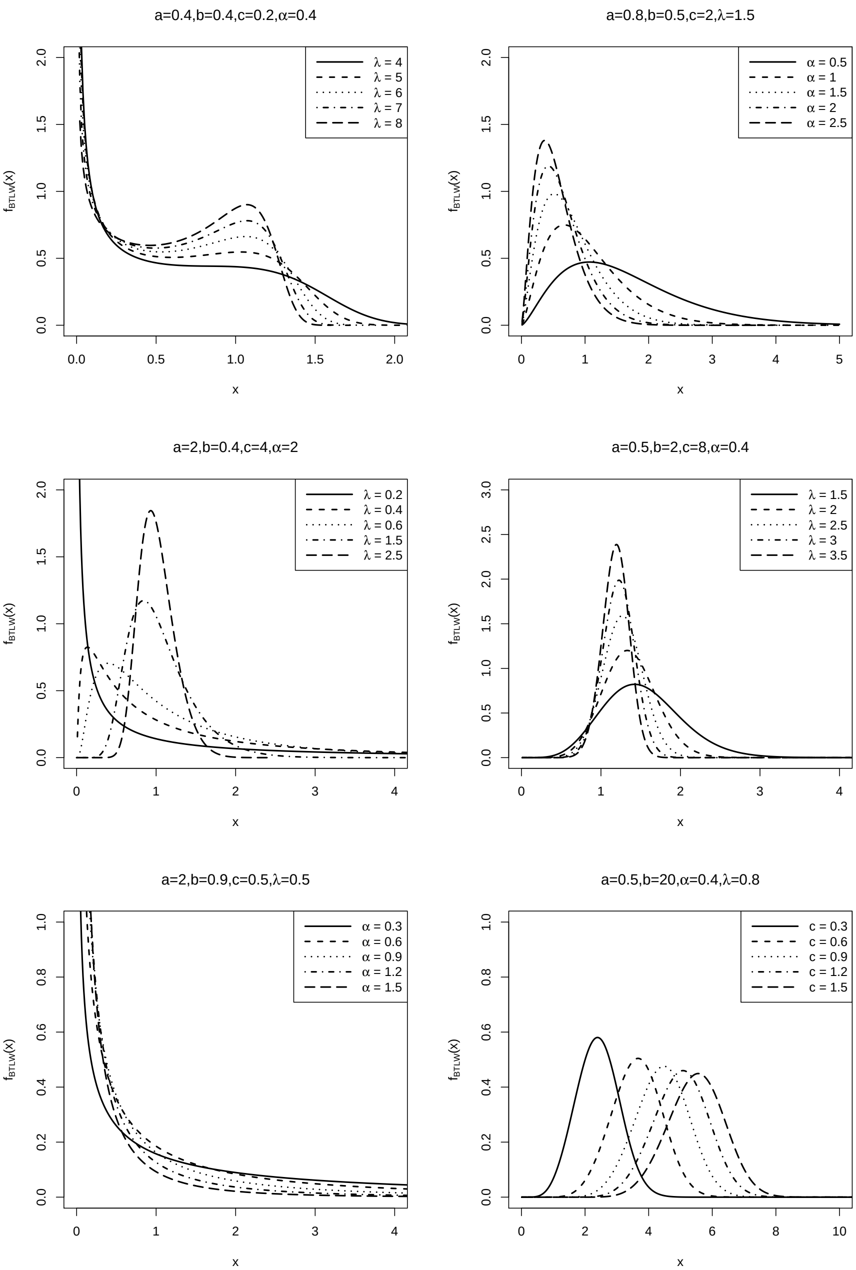

Figure 1 Some pdf plots of the BTLW distribution with specified parameter values

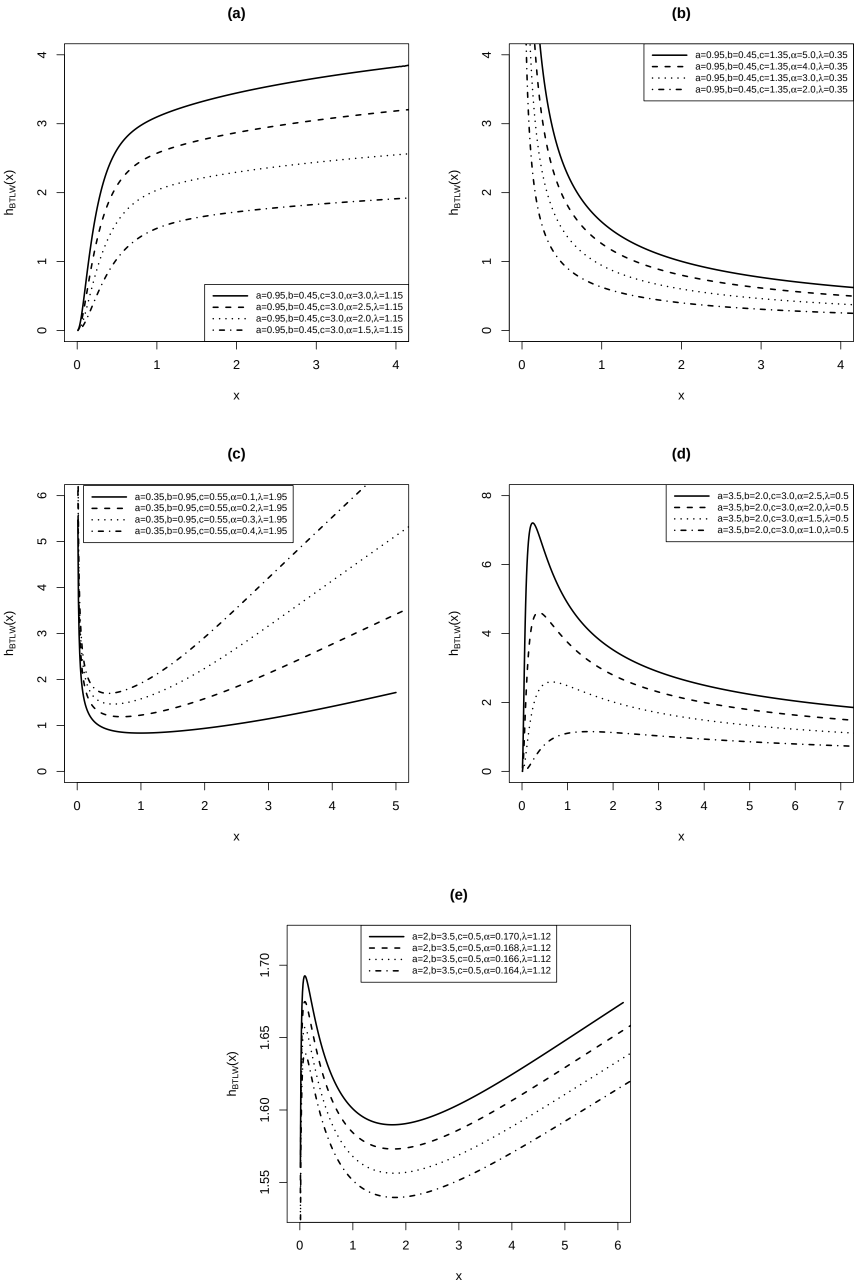

Figure 2 Some BTLW hazard shapes (a) increasing, (b) decreasing, (c) bathtub-shaped, (d) upside-down bathtub-shaped and (e) modified bathtub-shaped

6. Statistical Properties of the BTLW Distribution

In this section, some statistical properties of the BTLW distribution, including transformation, quantile function, ordinary moments and moment generating function, are provided.

6.1. Transformation

If the random variable B follows beta distribution with parameters a and b, denoted by $Beta(a,b)$, then the random variable

$$X = \left[\frac{\log(1-B^{1/c})}{-2\alpha}\right]^{1/\lambda} \tag{28}$$

follows a BTLW distribution with parameters a, b, c, α and λ. A random variable $X \sim BTLW(a,b,c,\alpha,\lambda)$ can be generated by utilizing the transformation in Equation (28).

6.2. Quantile function

The quantile function of the BTLW distribution is

$$Q_{BTLW}(u;a,b,c,\alpha,\lambda) = \left\{\frac{\log[1-(I_u^{-1}(a,b))^{1/c}]}{-2\alpha}\right\}^{1/\lambda}, \quad 0 < u < 1, \tag{29}$$

where $I_u^{-1}(a,b)$ represents the inverse of the incomplete beta ratio function (Majumder and Bhattacharjee, 1973).

6.3. Ordinary moments

Nadarajah (2006) show that the rth ordinary moment of the EW distribution for any $r > -\lambda$ are

$$E[X_\theta^r] = \theta\alpha^{-r/\lambda}\Gamma\left(\frac{r}{\lambda}+1\right)\sum_{i=0}^{\infty}\frac{(1-\theta)_i}{i!(i+1)^{(r+\lambda)/\lambda}}, \tag{30}$$

where X_θ is the EW random variable with parameters α, λ and power parameter θ, $\Gamma(p) = \int_0^\infty t^{p-1}e^{-t}\mathrm{d}t$ is the gamma function and $(1-\theta)_i = (-1)^i\Gamma(\theta)/\Gamma(\theta-i)$.

From Equations (14) and (30), the rth ordinary moment of BTLW distribution can be expressed as

$$\mu_r^{'} = \alpha^{-r/\lambda}\Gamma\left(\frac{r}{\lambda}+1\right)\sum_{j,k,n=0}^{\infty}\sum_{m=0}^{2k} m w_{j,k,m}\frac{(1-m)_n}{n!(n+1)^{(r+\lambda)/\lambda}}, \tag{31}$$

for any $r > -\lambda$.

6.4. Moment generating function

The mgf of BTLW distribution can be derived using Equations (18) and (31) as

$$M_X(t) = \sum_{i,j,k,n=0}^{\infty}\frac{\alpha^{-i/\lambda}t^i}{i!}\Gamma\left(\frac{i}{\lambda}+1\right)\sum_{m=0}^{2k} m w_{j,k,n}\frac{(1-m)_n}{n!(n+1)^{(i+\lambda)/\lambda}}, \tag{32}$$

for any $i > -\lambda$.

7. Maximum Likelihood Estimation

Let $\boldsymbol{x} = (x_1, \ldots, x_n)$ be a random sample of size n from the BTLW distribution and let F and C be the sets of individuals for which x_i is the failure or censoring time, respectively. The likelihood function for the vector of parameters $\Theta = (a, b, c, \alpha, \lambda)^T$ of the BTLW distribtuion can be written as

$$\begin{aligned}L(\Theta; \boldsymbol{x}) &= \prod_{i\in F} f_{BTLW}(x_i; \Theta) \prod_{i\in C} S_{BTLW}(x_i; \Theta)\\ &= \prod_{i\in F}\left[\frac{2c\alpha\lambda}{B(a,b)} x_i^{\lambda-1} e^{-2\alpha x_i^\lambda}(1-e^{-2\alpha x_i^\lambda})^{ac-1}[1-(1-e^{-2\alpha x_i^\lambda})^c]^{b-1}\right]\\ &\quad\times \prod_{i\in C}\left[I_{1-(1-e^{-2\alpha x_i^\lambda})^c}(b,a)\right].\end{aligned}$$

The log-likelihood (LL) function is

$$\begin{aligned}\ell(\Theta; \boldsymbol{x}) &= \log L(\Theta; \boldsymbol{x})\\ &= \sum_{i\in F} \log f_{BTLW}(x_i; \Theta) + \sum_{i\in C} \log S_{BTLW}(x_i; \Theta)\\ &= -n\log B(a,b) + n\log(2) + n\log(c) + n\log(\alpha) + n\log(c)\\ &\quad + (\lambda-1)\sum_{i\in F}\log(x_i) - 2\alpha\sum_{i\in F} x_i^\lambda + (ac-1)\sum_{i\in F}\log(1-v_i^2)\\ &\quad + (b-1)\sum_{i\in F}\log[1-(1-v_i^2)^c] + \sum_{i\in C}\log I_{1-(1-v_i^2)^c}(b,a), \end{aligned} \tag{33}$$

where $v_i = e^{-\alpha x_i^\lambda}$ is a transformed observation.

The maximum likelihood estimate $\hat{\Theta} = (\hat{a}, \hat{b}, \hat{c}, \hat{\alpha}, \hat{\lambda})^T$ of the vector of unknown parameters in Equation (33) can be obtained by the score function

$$U(\Theta) = \frac{\partial \ell(\Theta; \boldsymbol{x})}{\partial \Theta} = 0. \tag{34}$$

The asymptotic distribution of $\sqrt{n}(\hat{\Theta} - \Theta)$ is multivariate normal $N_5(\mathbf{0}, [I(\Theta)]^{-1})$ where $I(\Theta)$ is the expected Fisher information matrix which is given by

$$I(\Theta) = E\left[-\frac{\partial^2 \ell(\Theta; \boldsymbol{x})}{\partial\Theta\partial\Theta^T}\right]. \tag{35}$$

However, the expected Fisher information matrix $I(\Theta)$ is not available unless the censoring process is fully specified. Therefore, the asymptotic covariance $Var(\Theta) = [I(\Theta)]^{-1}$ can be approximated by using the observed Fisher information matrix, which is defined as

$$J(\Theta) = -\frac{\partial^2 \ell(\Theta; \boldsymbol{x})}{\partial\Theta\partial\Theta^T} = -\begin{bmatrix} J_{aa} & J_{ab} & J_{ac} & J_{a\alpha} & J_{a\lambda}\\ \cdot & J_{bb} & J_{bc} & J_{b\alpha} & J_{b\lambda}\\ \cdot & \cdot & J_{cc} & J_{c\alpha} & J_{c\lambda}\\ \cdot & \cdot & \cdot & J_{\alpha\alpha} & J_{\alpha\lambda}\\ \cdot & \cdot & \cdot & \cdot & J_{\lambda\lambda}\end{bmatrix}. \tag{36}$$

The score function and the observed Fisher information matrix corresponding to Equations (34) and (36) are too complicated to be presented in close-form expressions. Therefore, the LL function can be maximized to obtained $\hat{\Theta}$ by a procedure of Newton-Raphson iteration utilizing the `optimr` package (Nash, 2016) in the `R` programming language (R Core Team, 2020). Furthermore, the second partial derivatives for the observed Fisher information matrix can be numerically computed by using `numDeriv` package (Gilbert and Varadhan, 2019).

8. Simulation Study

A Monte Carlo simulation study is conducted to investigate the performance of the maximum likelihood estimates based on bias and root mean square error (RMSE). We consider sample sizes $n = 15, 25, 50, 100, 250, 500$ and the different values of the parameters a, b, c, α and λ of the BTLW distribution: 1. $a = 0.5, b = 2, c = 8, \alpha = 0.4$ and $\lambda = 4$ and 2. $a = 2, b = 0.4, c = 4, \alpha = 2$ and $\lambda = 0.2$. The experiment is repeated 2000 times. An algorithm for generating a BTLW random variable X with parameters a, b, c, α and λ:

(i) Generate $B \sim Beta(a, b)$.

(ii) Set $X = \left[\dfrac{\log(1 - B^{1/c})}{-2\alpha}\right]^{1/\lambda}$.

Table 2 gives the average parameter estimates, average bias, and average RMSE of the maximum likelihood estimates. The results show that the maximum likelihood estimates are the asymptotically unbiased and consistent, i.e., the bias and RMSE decrease when the sample size increases.

Table 2 The average parameter estimates, average bias, and average RMSE

Sample size	Parameters	BTLW(0.5,2,8,0.4,4)			BTLW(2,0.4,4,2,0.2)		
		Parameter estimates	Bias	RMSE	Parameter estimates	Bias	RMSE
15	a	1.650	1.150	3.420	5.297	3.297	13.867
	b	8.416	6.416	27.268	0.471	0.071	1.153
	c	15.655	7.655	22.550	7.733	3.733	11.928
	α	0.723	0.323	0.878	3.359	1.359	2.345
	λ	7.086	3.086	5.471	0.337	0.137	0.273
25	a	1.261	0.761	2.694	4.595	2.595	10.004
	b	6.027	4.027	15.823	0.487	0.087	1.100
	c	14.057	6.057	19.218	6.514	2.514	7.757
	α	0.650	0.250	0.744	3.004	1.004	1.864
	λ	5.927	1.927	3.799	0.277	0.077	0.166
50	a	0.855	0.355	1.336	3.640	1.640	5.555
	b	4.257	2.257	8.679	0.513	0.113	0.640
	c	11.644	3.644	11.351	5.990	1.990	6.856
	α	0.530	0.130	0.467	2.696	0.696	1.451
	λ	4.926	0.926	2.309	0.234	0.034	0.099
100	a	0.717	0.217	0.918	3.175	1.175	3.738
	b	3.324	1.324	4.983	0.516	0.116	0.581
	c	10.922	2.922	9.800	5.080	1.080	4.225
	α	0.481	0.081	0.295	2.469	0.469	1.096
	λ	4.509	0.509	1.747	0.218	0.018	0.071
250	a	0.572	0.072	0.356	2.737	0.737	2.478
	b	2.636	0.636	2.408	0.499	0.099	0.460
	c	9.946	1.946	7.284	4.603	0.603	2.408
	α	0.431	0.031	0.162	2.282	0.282	0.767
	λ	4.303	0.303	1.381	0.206	0.006	0.052
500	a	0.545	0.045	0.296	2.446	0.446	1.494
	b	2.298	0.298	1.630	0.491	0.091	0.362
	c	9.415	1.415	5.628	4.424	0.424	2.029
	α	0.421	0.021	0.128	2.152	0.152	0.588
	λ	4.287	0.287	1.234	0.200	<0.001	0.042

9. Applications

In this section, we compare the fitted results of the BTLW, BW, KW, TLW, EW and Weibull distributions with two datasets to demonstrate the flexibility and applicability of the proposed model among the other lifetime parametric distributions. In order to evaluate whether the model is appropriate, the many statistical tools are considered.

9.1. Complete Data: Aarset Data

We consider the bathtub-shaped hazard data from (Aarset, 1987). The Aarset data consist of times to failure (in weeks) of 50 industrial devices put on life test at time 0.

Firstly, the Weibull distribution parameters are estimated by MLE to use as initial values for the numerically computed. Table 3 shows the maximum likelihood estimates (the standard errors are given in parentheses) of model parameters and Table 4 gives the values of Akaike's information criterion (AIC) by (Akaike, 1974), Bayesian information criterion (BIC) by Schwarz (1978), Akaike's information corrected criterion (AICC) by Hurvich and Tsai (1989) and Hannon and Quinn's information criterion (HQIC) by Hannan and Quinn (1979). These are defined as, respectively,

$$AIC = -2\ell(\hat{\Theta};\boldsymbol{x}) + 2k,$$
$$BIC = -2\ell(\hat{\Theta};\boldsymbol{x}) + k\log(n),$$
$$AICC = -2\ell(\hat{\Theta};\boldsymbol{x}) + 2k\left(\frac{n}{n-k-1}\right),$$
$$HQIC = -2\ell(\hat{\Theta};\boldsymbol{x}) + 2k\log(\log(n)),$$

where $\ell(\hat{\Theta};\boldsymbol{x})$ is the LL of maximum likelihood estimate $\hat{\Theta}$, k is the number of parameters and n is the sample size.

The lower the values of these statistics indicate a better fit to the data. Since these findings suggest that the BTLW distribution has the lowest AIC, BIC, AICC and HQIC values, it follows that the BTLW distribution could be a suitable model for the fitting of the data. In order to identify the type of empirical behavior of the hazard function, the total time on test (TTT) plot (Aarset, 1987) is used. Figure 3(a) shows that TTT plot for the data is initially a convex shape and then a concave shape. It points out that the data have a bathtub-shaped hazard. Also the estimated hazard function plot of the BTLW distribution in Figure 3(b) is bathtub-shaped. The histogram and the estimated pdf plots of the Aarset data are illustrated in Figure 4(a). In Figure 4(b), the empirical cdf and the estimated cdf plots for the Aarset data are shown. Furthermore, the goodness-of-fit plots for BTLW distribution that consist of Q-Q and P-P plots are presented in Figure 5(a) and Figure 5(b), respectively. The conclusion of these plots indicates that BTLW distribution provides a better fit for the Aarset data.

Table 3 Maximum likelihood estimates of the model parameters for the Aarset data

Distributions	Parameter estimates				
	a	b	c	α	λ
BTLW	0.0818 (0.0154)	0.0761 (0.0161)	0.9582 (0.0001)	4.347e-11 (<0.0001)	5.8936 (0.0002)
BW	0.1331 (0.0304)	0.0703 (0.0239)	-	1.63e-05 (<0.0001)	3.2032 (0.0839)
KW	0.0706 (0.0235)	0.2370 (0.0585)	-	1.607e-08 (<0.0001)	4.4762 (0.0652)
TLW	-	-	0.1459 (0.0211)	1.885e-10 (<0.0001)	4.8065 (0.0819)
EW	-	-	0.4670 (0.2260)	0.0010 (0.0030)	1.5940 (0.6250)
Weibull	-	-	-	0.027 (0.0139)	0.949 (0.1196)

Table 4 The measures of -LL, AIC, BIC, AICC and HQIC for the Aarset data

Distributions	-LL	AIC	BIC	AICC	HQIC
BTLW	**218.940**	**447.870**	**457.430**	**449.240**	**451.510**
BW	223.470	454.950	462.600	455.840	457.860
KW	222.480	452.960	460.610	453.850	455.880
TLW	228.980	463.950	469.690	464.480	466.140
EW	236.290	478.580	484.320	479.100	480.770
Weibull	241.000	486.000	486.000	486.260	487.460

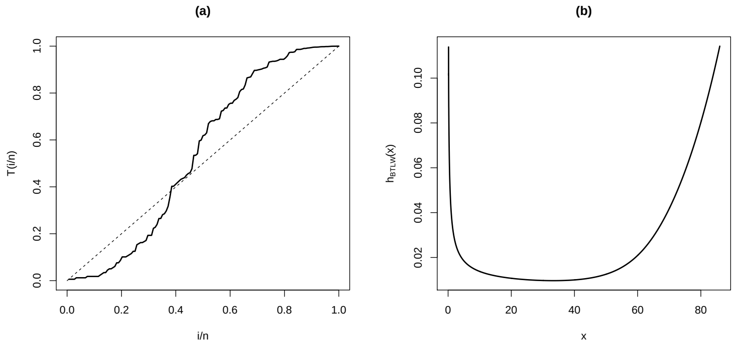

Figure 3 (a) TTT plot on the Aarset data (b) Estimated hazard function plot for the Aarset data

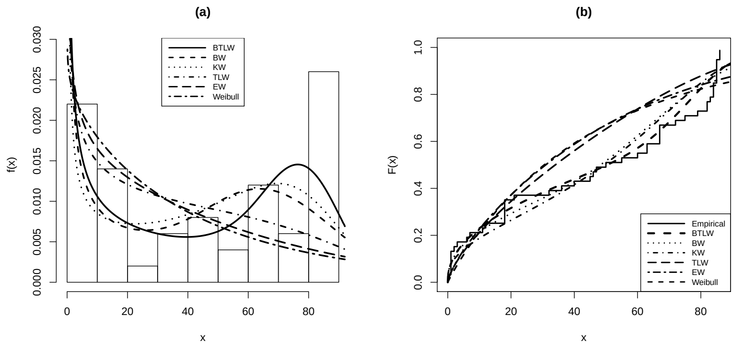

Figure 4 (a) The histogram and the estimated pdf plots of the Aarset data (b) The empirical cdf and the estimated cdf plots for the Aarset data

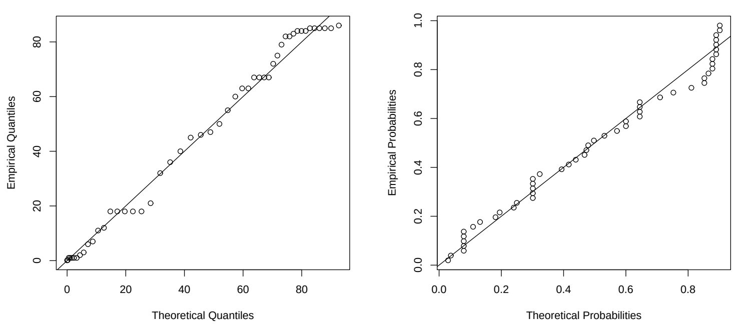

Figure 5 Goodness-of-fit plots for BTLW distribution fitted to Aarset data (a) Q-Q plot (b) P-P plot

9.2. Censored Data: Serum-Reversal Data

The disappearance of HIV antibodies in the patient blood that previously showed HIV-antibody-positive on serological testing is called the serum-reversal process. In the work by Silva (2004), the data pertains to the serum-reversal time (in days) of a random sample of 148 children born from mothers infected with HIV who have not received HIV treatment before pregnancy at the university hospital of the Ribeirão Preto School of Medicine, Brazil from 1986 to 2001. Also the dataset contains 88 right-censoring observations which constitute 59.46% of all data.

Table 5 and Table 6 give the maximum likelihood estimates of the model parameters (the standard errors are given in parentheses) and the measures of AIC, BIC, AICC and HQIC, respectively. Based on the lowest AIC, BIC, AICC and HQIC values, the BTLW distribution could be selected as the best model among the other distributions. Figure 6(a) shows that the TTT plot for the dataset is concave shape. It indicates that the data have a increasing hazard. Also the estimated hazard function plot of BTLW distribution in Figure 6(b) is a increasing hazard. Furthermore, the Kaplan-Meier curve (empirical survival function plot) (Kaplan and Meier, 1958) and the estimated survival function plots are provided in Figure 7. The conclusion of these plots indicates that BTLW distribution provides a better fit for the serum-reversal data.

Table 5 Maximum likelihood estimates of the model parameters for the serum-reversal data

Distributions	Parameter estimates				
	a	b	c	α	λ
BTLW	1.3661 (0.3218)	0.0832 (0.0092)	0.1470 (0.0002)	4.065e-12 (<0.0001)	4.8550 (0.0131)
BW	0.3749 (0.0770)	0.0513 (0.0069)	-	9.245e-10 (<0.0001)	4.1296 (0.0034)
KW	0.2824 (0.0008)	0.0545 (0.0074)	-	1.815e-08 (<0.0001)	3.6070 (0.0098)
TLW	-	-	0.5209 (0.0637)	1.726e-12 (<0.0001)	4.5044 (0.0361)
EW	-	-	0.6374 (0.0801)	7.843e-11 (<0.0001)	3.9943 (0.0330)
Weibull	-	-	-	1.797e-08 (<0.0001)	3.1130 (0.0218)

Table 6 The measures of -LL, AIC, BIC, AICC and HQIC for the serum-reversal data

Distributions	-LL	AIC	BIC	AICC	HQIC
BTLW	**389.240**	**788.480**	**799.280**	**789.520**	**792.740**
BW	392.420	792.830	801.470	793.510	796.230
KW	392.920	793.850	802.480	794.530	797.250
TLW	397.860	801.730	808.200	802.130	804.280
EW	398.790	803.590	810.050	803.980	806.130
Weibull	401.990	807.990	812.300	808.180	809.690

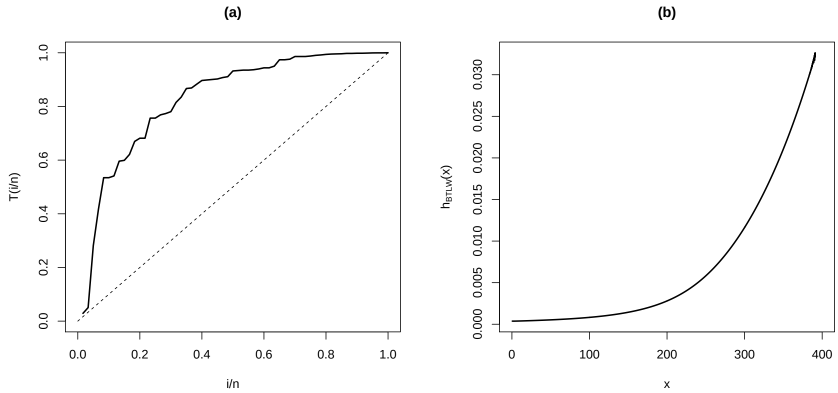

Figure 6 (a) TTT plot on the serum-reversal data. (b) Estimated hazard function plot for the serum-reversal data

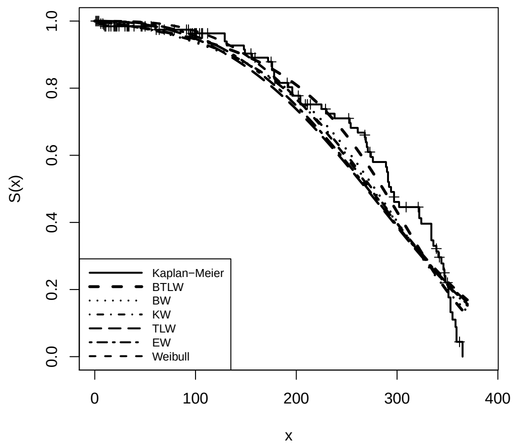

Figure 7 Kaplan-Meier curve (empirical survival function plot) and estimated survival function plots for the serum-reversal data

10. Conclusion

A new family of distributions called BTLG family is introduced by combining two families: the BG family and TLG family. Some of statistical properties for the proposed family consist of linear representation, ordinary moments and moment generating function are derived. In this work, a special case of the BTLG family namely the BTLW distribution is studied and some of its statistical properties are investigated containing transformation, quantile function, ordinary moments and moment generating function. The shapes of hazard function are monotonically increasing, monotonically decreasing, bathtub-shaped, upside-down bathtub-shaped and modified bathtub-shaped. The parameters of BTLW distribtuion for complete data and censored data are estimated by utilizing MLE. Two applications involving the Aarset and serum-reversal datasets has been analyzed. The BTLW distribution outperform than the others distribution by considering the information criteria and goodness-of-fit plots and it is an interesting alternative distribution for lifetime data in a variety of fields.

Acknowledgements

The authors would like to thank to Department of Statistics, Faculty of Science, and the graduate school of Kasetsart University. The first author would like to thank to Science Achievement Scholarship of Thailand (SAST) for financial supporting.

References

Aarset MV. How to identify a bathtub hazard rate. IEEE Trans Reliab. 1987; 36(1): 106-108.

Afify A, Yousof H, Nadarajah S. The beta transmuted-H family for lifetime data. Stat Interface. 2017; 10(3): 505-520.

Akaike H. A new look at the statistical model identification. IEEE Trans Automat Contr. 1974; 19(6): 716-723.

Alexander C, Cordeiro GM, Ortega EM, Sarabia JM. Generalized beta-generated distributions. Compu Stat Data Anal. 2012; 56(6): 1880-1897.

Almalki SJ, Nadarajah S. Modifications of the Weibull distribution: a review. Reliab Eng Syst Safe. 2014; 124: 32-55.

Alzaatreh A, Lee C, Famoye F. A new method for generating families of continuous distributions. Metron. 2013; 71(1): 63-79.

Aryal GR, Ortega EM, Hamedani G, Yousof HM. The Topp-Leone generated Weibull distribution: regression model, characterizations and applications. Int J Stat Probab. 2017; 6(1): 126-141.

Aryal GR, Tsokos CP. Transmuted Weibull distribution: A generalization of the Weibull probability distribution. Eur J Pure Appl Math. 2011; 4(2): 89-102.

Bagheri S, Samani EB, Ganjali M. The generalized modified Weibull power series distribution: theory and applications. Comput Stat Data Anal. 2016; 94: 136-160.

Carrasco JM, Ortega EM, Cordeiro GM. (2008). A generalized modified Weibull distribution for lifetime modeling. Comput Stat Data Anal. 2008; 53(2): 450-462.

Cordeiro GM, de Castro M. A new family of generalized distributions. J Stat Comput Simul. 2011; 81(7): 883-898.

Cordeiro GM, Ortega EM, Nadarajah S. The Kumaraswamy Weibull distribution with application to failure data. J Franklin Inst. 2010; 347(8): 1399-1429.

Eugene N, Lee C, Famoye F. Beta-normal distribution and its applications. Commun Stat Theo Meth. 2002; 31(4): 497512.

Gilbert P, Varadhan R. numDeriv: accurate numerical derivatives. 2019 [cited 2020 Sep 3]. Available from: https://CRAN.R-project.org/package=numDeriv.

Gupta RD, Kundu D. Theory & methods: generalized exponential distributions. Aust N Z J Stat. 1999; 41(2): 173-188.

Hannan EJ, Quinn BG. The determination of the order of an autoregression. J R Stat Soc Ser B Stat Methodol. 1979; 41(2): 190-195.

Hurvich CM, Tsai CL. Regression and time series model selection in small samples. Biometrika. 1989; 76(2): 297-307.

Jones M. Families of distributions arising from distributions of order statistics. Test. 2004; 13(1): 1-43.

Kaplan EL, Meier P. Nonparametric estimation from incomplete observations. J Am Stat Assoc. 1958; 53(282): 457-481.

Lai C, Xie M, Murthy D. A modified Weibull distribution. IEEE Trans Reliab. 2003; 52(1): 33-37.

Lee C, Famoye F, Alzaatreh AY. (2013). Methods for generating families of univariate continuous distributions in the recent decades. Wiley Interdiscip Rev Comput Stat. 2013; 5(3): 219-238.

Lee C, Famoye F, Olumolade O. Beta-Weibull distribution: some properties and applications to censored Data. J Mod Appl Stat Meth. 2007; 6(1): 173-186.

Majumder K, Bhattacharjee G. Algorithm as 64: inverse of the incomplete beta function ratio. J R Stat Soc Ser C Appl Stat, 1973; 22(3). 411-414.

Merovci F, Alizadeh M, Yousof HM, Hamedani GG. The exponentiated transmuted-g family of distributions: theory and applications. Commun Stat Theo Meth. 2017; 46(21): 10800-10822.

Mudholkar GS, Srivastava DK. Exponentiated Weibull family for analyzing bathtub failure-rate data. IEEE Trans Reliab. 1993; 42(2): 299-302.

Nadarajah S. The exponentiated Gumbel distribution with climate application. Environmetrics. 2006; 17(1): 13-23.

Nadarajah S, Kotz S. The exponentiated type distributions. Acta Appl Math. 2006; 92(2): 97-111.

Nash JC. optimr: A replacement and extension of the optim Function. 2016 [cited 2020 Sep 3]. Available from: https://CRAN.R-project.org/package=optimr.

Pescim RR, Cordeiro GM. Demétrio CG, Ortega EM, Nadarajah S. The new class of Kummer beta generalized distributions. SORT - Stat Oper Res T. 2012; 36(2): 153-180.

R Core Team. R: A language and environment for statistical computing. R foundation for statistical computing. Vienna: Austria; 2020 [cited 2020 Sep 3]. Available from: https://www.R-project.org.

Sangsanit Y, Bodhisuwan W. The Topp-Leone generator of distributions: properties and inferences. Songklanakarin J Sci Technol. 2016; 38(5): 537-548.

Schwarz G. Estimating the dimension of a model. Ann Stat. 1978; 6(2): 461-464.

Shaw WT, Buckley IRC. The alchemy of probability distributions: beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. 2009 [cited 2020 Sep 3]. Available from: https://arxiv.org/abs/0901.0434.

Silva ANFd. Estudo evolutivo das criancas expostas ao hiv e notificadas pelo núcleo de vigilância epidemiológica do hcfmrp-usp. M.Sc. [thesis], University of São Paulo; 2004.

Silva GO, Ortega EM, Cordeiro GM. The beta modified Weibull distribution. Lifetime Data Anal. 2010; 16(3): 409430.

Weibull W. A statistical distribution function of wide applicability. J Appl Mech. 1951; 18: 293-297.