



Original Article

The Topp-Leone generator of distributions: properties and inferences

Yuwadee Sangsanit and Winai Bodhisuwan*

Department of Statistics, Faculty of Science,
Kasetsart University, Chatuchak, Bangkok, 10900 Thailand.

Received: 6 October 2015; Accepted: 27 January 2016

Abstract

In this paper, a new framework for generating lifetime distributions is introduced, called the Topp-Leone generated (TLG) family of distributions. The generator of the TLG family is the Topp-Leone distribution, which was proposed by Topp and Leone in 1955. Some various properties of the TLG distribution are discussed, e.g., survival function, hazard function, moments and generating function. In addition, the TLG family is capable of improving fitted results and tail behavior of existing distributions. We present the Topp-Leone generalized exponential (TLGE) distribution as an example of the TLG distribution. Some graphical representations related to the probability density function and hazard function of the TLGE distribution are provided. In application study, the goodness of fit test based on the TLGE, the generalized exponential (GE), and exponentiated generalized exponential (EGE) distributions are compared. The results emphasize that the TLGE distribution can be considered as a competitive distribution for the GE and EGE distributions.

Keywords: Topp-Leone family, parameter-adding method, beta-generated, probability weighted moments, generalized exponential

1. Introduction

Lifetime data plays an important role in a wide range of applications such as medicine, engineering, biological science, management, and public health. Statistical distributions are used to model the life of an item in order to study its important properties. Proper distribution may provide useful information that result in sound conclusions and decisions. When there is a need for more flexible distributions, many researchers are about to use the new one with more generalization. An excellent review of Lee *et al.* (2013) has provided thorough knowledge of several methods for generating families of continuous univariate distributions. According to their work, there are some general methods introduced prior to 1980, which were developed by the strategies based on differential equation, transformation and quantile function. In addition, they also put emphasis on the movement of

those methods proposed since 1980s, which changed the momentum by adding extra parameters or combining existing distributions.

The beta generated (BG) family of distributions belongs to a parameter-adding method (Lee *et al.*, 2013). Some existing distributions incorporated with the BG family will have two additional parameters, which are the parameters of the beta distribution. For an arbitrary distribution with a cumulative distribution function (cdf) $G(\cdot)$ and a probability density function (pdf) $g(\cdot)$, this method generates it by letting $X = G^{-1}(B)$ where B has the beta distribution with parameters a and b , $B \sim \text{Beta}(a, b)$ (Alexander *et al.*, 2012). According to the work of Eugene *et al.* (2002), a random variable X is considered as having the BG distribution, a cdf and a pdf of which can be characterized by

$$F_{BG}(x) = \frac{1}{B(a, b)} \int_0^{G(x)} t^{a-1} (1-t)^{b-1} dt, \quad a > 0, \quad b > 0$$

and

$$f_{BG}(x) = \frac{1}{B(a, b)} g(x) G(x)^{a-1} (1-G(x))^{b-1},$$

* Corresponding author.

Email address: fsciwnb@ku.ac.th

respectively, where $B(a, b) = \Gamma(a)\Gamma(b) / \Gamma(a + b)$ is the beta function, $\Gamma(\cdot)$ is the gamma function.

Based on the discussion of Jones (2004) and the work of Alzaatreh *et al.* (2013), the distribution with $F_{BG}(x)$ and $f_{BG}(x)$ has $G(\cdot)$ and the beta distribution, respectively, as its parent distribution and its generator. Moreover, the work of Eugene *et al.* (2002) motivated many researchers to raise various novel distributions by using other parent distributions. Some examples include the beta Gumbel distribution (Nadarajah and Kotz, 2004), the beta exponential distribution (Nadarajah and Kotz, 2006), the beta-Weibull distribution (Lee *et al.*, 2007), the beta-Pareto distribution (Akinsete *et al.*, 2008), the beta generalized exponential distribution (Barreto-Souza *et al.*, 2010), the beta generalized Weibull distribution (Singla *et al.*, 2012), the beta exponentiated Weibull distribution (Cordeiro *et al.*, 2013b), the beta generalized Rayleigh distribution (Cordeiro *et al.*, 2013a), and the beta exponentiated Weibull Poisson distribution (Insuk *et al.*, 2015).

Recently, applying new generators for continuous distributions became more interesting. This methodology can improve on the goodness of fit and determine tail properties. These features have been established by the results of many generators such as beta distribution (Eugene *et al.*, 2002; Jones, 2004), Kumaraswamy distribution (Jones, 2009), generalized beta distribution (Alexander *et al.*, 2012), McDonald distribution (Cordeiro *et al.*, 2014a), gamma distribution (Zografos and Balakrishnan, 2009), Kummer beta distribution (Pescim *et al.*, 2012), Transformed Transformer distribution (Alzaatreh *et al.*, 2013), Log gamma distribution (Amini *et al.*, 2014), Weibull distribution (Bourguignon *et al.*, 2014), Lomax distribution (Cordeiro *et al.*, 2014b), Lindley distribution (Bhati *et al.*, 2015), Kumaraswamy Marshal-Olkin distribution (Alizadeh *et al.*, 2015), and odd generalized exponential distribution (Tahir *et al.*, 2015). Although the generator seems to be a more complex function, some experts still prefer to use the more simplified generator. In other words, the generator with both closed form of cdf and a few number of parameters make deriving inference part easily accessible.

The Topp-Leone distribution (TL) is one of the continuous distributions that is attractive as a generator. This distribution was proposed by Topp and Leone (1955). It provides closed forms of the cdf and pdf. According to the number of parameters, the estimation part for the TL distribution is not complicated. However, the TL distribution had not received much attention until Nadarajah and Kotz (2003) discovered it. In addition, they studied some properties of the TL distribution and provided its moments, central moments and characteristic function. Furthermore, there were numerous authors who were interested in this distribution. For example, Ghitany *et al.* (2005) provided some reliability measures of the TL distribution, a discussion on kurtosis of the TL distribution was reported by Kotz and Seier (2007), Vicaria *et al.* (2008) introduced two-sided generalized Topp and Leone distributions, and Al-Zahrani (2012) derived goodness of fit tests for the TL distribution. Evidently, a number

of works reveal significant impact of the TL distribution on lifetime data analysis. Accordingly, in this paper, we introduce TL distribution as a generator for continuous distributions.

The aim of this paper is to propose the Topp-Leone generated (TLG) family of distributions. We also study its properties and derive associated inferences. The rest of this paper is outlined as follows. In the preliminaries section, we provide fundamental properties of the TL distribution. In Section 3, the TLG family of distributions is introduced together with main properties, corresponding order statistics, moments, and moment generating function. In Section 4, we discuss one of its special cases called the TLGE distribution and provide some properties. The parameter estimation based on the maximum likelihood method is derived in Section 5. Moreover, applications with real data are also demonstrated.

2. Preliminaries

In 1955, Topp and Leone constructed the distribution for empirical data with J-shaped histogram such as powered band tool failures, and automatic calculating machine failure. If a random variable T belongs to the TL distribution, it can have either finite ($0 < t < b$) or infinite ($0 < t < b < \infty$) support. In this paper, we focus primarily on the TL distribution with $b = 1$ to avoid any additional function for creating a new generated family of distributions (Zografos and Balakrishnan, 2009; Lee *et al.*, 2013; Alzaatreh *et al.*, 2013). In order to pave the way for the TLG family of distributions, some properties of the TL distribution are provided.

A random variable T is distributed as the TL with parameter α denoted by $T \sim TL(\alpha)$, with a cdf

$$F_{TL}(t) = t^\alpha (2 - t)^\alpha$$

where $0 < t < 1$ and $\alpha > 0$. The corresponding pdf is

$$f_{TL}(t) = 2\alpha t^{\alpha-1}(1-t)(2-t)^{\alpha-1}.$$

Other important characteristics of lifetime data analysis are the survival and hazard functions. Those functions of the TL distribution are respectively

$$s_{TL}(t) = 1 - t^\alpha (2 - t)^\alpha,$$

and

$$h_{TL}(t) = \frac{2\alpha t^{\alpha-1}(1-t)(2-t)^{\alpha-1}}{1 - t^\alpha (2 - t)^\alpha}.$$

Furthermore, Nadarajah and Kotz (2003) pointed out that the TL distribution provides bathtub shape of the hazard function when $0 < \alpha < 1$. In addition, if $\alpha \geq 1$ the TL distribution has a non-increasing hazard function. Finally, the quantile function of the TL distribution is

$$t = F_{TL}^{-1}(u) = 1 - \sqrt[1/\alpha]{1 - u}, \tag{1}$$

where u is distributed as the uniform on the interval $(0, 1)$.

3. The Topp-Leone generated family of distributions

Creating a new family of distribution requires two principal components, which are a generator and a parent distribution (Jones, 2004; Alzaatreh *et al.*, 2013). Indeed, the pdf of a generator is transformed into a new pdf through the cdf $G(\cdot)$ of a parent distribution. Alzaatreh *et al.* (2013) presented a method for generating a new family of distributions with the following definition for a random variable $T \in [a, b]$, $-\infty \leq a < b \leq \infty$, and any random variable X with cdf $G(x)$.

Let $W(G(x))$ be a function of $G(x)$ and satisfy the conditions as follows:

- a) $W(G(x)) \in [a, b]$
- b) $W(G(x))$ is differentiable and monotonically non-decreasing
- c) $W(G(x)) \rightarrow a$ as $x \rightarrow -\infty$ and $W(G(x)) \rightarrow b$ as $x \rightarrow \infty$.

Definition 1:

Let T be a random variable of a generator distribution with pdf $r(t)$ defined on $[a, b]$. Let X be a continuous random variable with cdf $G(x)$. Thus, the cdf of a new family of distributions is given by

$$F(x) = \int_0^{W(G(x))} r(t) dt.$$

The corresponding pdf is

$$f(x) = r\{W(G(x))\} \left\{ \frac{d}{dx} W(G(x)) \right\}.$$

Consequently, the cdf and pdf of the TLG distribution are obtained with the use of Definition 1. as follows.

Proposition 1:

If a random variable T is distributed as the TL and bounded on $[0, 1]$. Let X be a continuous random variable with cdf $G(x)$. The TLG distribution has cdf written by

$$F_{TLG}(x) = G(x)^\alpha (2 - G(x))^\alpha,$$

where $\alpha > 0$ is a shape parameter. The associated pdf is

$$f_{TLG}(x) = 2\alpha g(x)(1 - G(x))G(x)^{\alpha-1}(2 - G(x))^{\alpha-1},$$

where $g(x) = dG(x) / dx$.

Proof:

According to Definition 1, let $T \sim TL(\alpha)$, then

$$r(t) = 2\alpha t^{\alpha-1}(1-t)(2-t)^{\alpha-1},$$

where $t \in [a = 0, b = 1]$. Therefore, the conditions of $W(G(x))$ defined by Alzaatreh *et al.* (2013) become basic properties of cdf $G(x)$ of any random variable X . Suppose $W(G(x)) = G(x)$, then the cdf of the TLG distribution is

$$F_{TLG}(x) = G(x)^\alpha (2 - G(x))^\alpha,$$

By differentiating, we get the pdf

$$f_{TLG}(x) = 2\alpha g(x)(1 - G(x))G(x)^{\alpha-1}(2 - G(x))^{\alpha-1}.$$

In addition, the TL random variable with finite support has the same bounds as the cdf $G(x)$ of any other random variable. Therefore, the relation of a random variable X having the TLG distribution and a random variable T having the TL distribution is

$$X = G^{-1}(T) \text{ with } T \sim TL(\alpha)$$

As mentioned by Alzaatreh *et al.* (2013), when a random variable T is bounded on $[0, 1]$, a term $W(G(x))$ can be $G(x)$ or $G(x)^\alpha$, that belongs to the exponentiated family of distributions (AL-Hussaini and Ahsanullah, 2015). Moreover, the fact that several works deal with new modifications of the distribution makes it beneficial to alter $W(G(x))$ in Definition 1. Some interesting examples include $1 - e^{-aG(x)}$ by Gurvich *et al.* (1997), $G(x)/(a + (1 - a)G(x))$ by Marshall and Olkin (1997), and $(1 + a)G(x) - aG^2(x)$ by Shaw and Buckley (2007).

3.1 Main properties

If a random variable X is distributed as the TLG, the survival and hazard functions are respectively

$$s_{TLG}(x) = 1 - G(x)^\alpha (2 - G(x))^\alpha,$$

and

$$h_{TLG}(x) = \frac{2\alpha g(x)(1 - G(x))G(x)^{\alpha-1}(2 - G(x))^{\alpha-1}}{1 - G(x)^\alpha (2 - G(x))^\alpha}.$$

When $Q_G(\cdot)$ is the quantile function of a parent distribution, we can simulate the TLG random variate from

$$x = Q_G(1 - \sqrt{1 - u^{1/\alpha}}) \tag{2}$$

where $1 - \sqrt{1 - u^{1/\alpha}}$ is the quantile function of the TL distribution in Eq. (1).

3.2 A general expansion of the density function

Many well-known families of distributions can be written as infinite or finite weighted series of their parent distributions (Eugene *et al.*, 2002; Jones, 2009; Cordeiro *et al.*, 2014a). This also means that the properties and inferences can be obtained from the same measures of its parent distribution. Therefore, it is useful to derive expansion of the density function. The following series allow us to get an expansion for the density function

$$(2 - x)^\alpha = \sum_{j=0}^{\infty} \binom{\alpha}{j} (-1)^j 2^{\alpha-j} x^j$$

For a real non-integer α , we apply the binomial series to

$$(2 - G(x))^\alpha = \sum_{j=0}^{\infty} \binom{\alpha}{j} (-1)^j 2^{\alpha-j} G(x)^j \tag{3}$$

Therefore,

$$F_{TLG}(x) = \sum_{j=0}^{\infty} b_j(\alpha) G(x)^{\alpha+j}$$

where

$$b_j(\alpha) = \binom{\alpha}{j} (-1)^j 2^{\alpha-j} \tag{4}$$

As α is a real non-integer, $G(x)^{\alpha+j}$ can be written as

$$\begin{aligned} G(x)^{\alpha+j} &= (1 - (1 - G(x)))^{\alpha+j} = \sum_{k=0}^{\infty} (-1)^k \binom{\alpha+j}{k} (1 - G(x))^k \\ G(x)^{\alpha+j} &= \sum_{k=0}^{\infty} \sum_{r=0}^k (-1)^{k+r} \binom{\alpha+j}{k} \binom{k}{r} G(x)^r \\ &= \sum_{k=0}^{\infty} q_r(\alpha+j, k) G(x)^r \end{aligned}$$

where

$$q_r(\alpha+j, k) = \sum_{r=0}^k (-1)^{k+r} \binom{\alpha+j}{k} \binom{k}{r} \tag{5}$$

If α is a real non-integer, the expansion of the cdf is

$$F_{TLG}(x) = \sum_{j,k=0}^{\infty} \sum_{r=0}^k \omega_{j,k,r} G(x)^r$$

where

$$\omega_{j,k,r} = \binom{\alpha}{j} \binom{\alpha+j}{k} \binom{k}{r} (-1)^{j+k+r} 2^{\alpha-j} \tag{6}$$

By differentiation of we have

$$f_{TLG}(x) = g(x) \sum_{j,k=0}^{\infty} \sum_{r=0}^k r \omega_{j,k,r} G(x)^{r-1} \tag{7}$$

On the other hand, when α is an integer, the index j in Eq. (3) stops at α . By the same manner established above, the expansion of the cdf and pdf with integer α can be written as

$$F_{TLG}(x) = \sum_{j=0}^{\alpha} b_j(\alpha) G(x)^{\alpha+j}$$

and

$$f_{TLG}(x) = g(x) \sum_{j=0}^{\alpha} (\alpha+j) b_j(\alpha) G(x)^{\alpha+j-1} \tag{8}$$

where the coefficient $b_j(\alpha)$ is defined in Eq. (4).

In fact, the TLG distribution can be considered as an infinite weighted sum of $G(x)^r$, which could be either the power of the parent distribution or one with a new expression of parameters. When α is an integer, the TLG distribution is a finite weighted sum of $G(x)^{\alpha+j}$. Moreover, the associated properties of the TLG distribution are mainly discussed based on Eqs. (7) and (8).

3.3 Order statistics

Let X_1, X_2, \dots, X_n be a random sample of size n from the TLG distribution. Then the pdf of the i th, $1 \leq i \leq n$, order statistic, $X_{i:n}$, can be obtained as follows

$$f_{i:n}(x) = \frac{f_{TLG}(x)}{B(i, n-i+1)} F_{TLG}(x)^{i-1} \{1 - F_{TLG}(x)\}^{n-i} \tag{9}$$

Using the series expansion for $\{1 - F_{TLG}(x)\}^{n-i}$, we get

$$f_{i:n}(x) = \frac{f_{TLG}(x)}{B(i, n-i+1)} \sum_{l=0}^{n-i} \binom{n-i}{l} (-1)^l F_{TLG}(x)^{i+l-1}$$

Referring to cdf of the TLG distribution in Proposition 1,

$F_{TLG}(x)^{i+j-1}$ becomes

$$F_{TLG}(x)^{i+j-1} = G(x)^{\alpha(i+j-1)} (2 - G(x))^{\alpha(i+j-1)}$$

Adjust the expansion of $(2 - G(x))^\alpha$ in Eq. (3) to be the expansion of $(2 - G(x))^{\alpha(i+j-1)}$, then

$$F_{TLG}(x)^{i+j-1} = \sum_{h=0}^{\infty} \sum_{m=0}^{\infty} b_h(\alpha(i+j-1)) q_v(\alpha(i+l-1), m) G(x)^{\alpha(i+j-1)+h} \tag{10}$$

where $b_h(\alpha(i+j-1))$ and $q_v(\alpha(i+l-1), m)$ are easily obtained from Eqs. (4) and (5), respectively.

Substituting $f_{TLG}(x)$ in Eq. (7) and $F_{TLG}(x)^{i+j-1}$ in Eq. (10) into $f_{i:n}(x)$ in Eq. (9), and changing indices then we have

$$f_{i:n}(x) = \frac{g(x)}{B(i, n-i+1)} \sum_{j=0}^{\infty} \sum_{l=0}^{n-i} \sum_{h,m=0}^{\infty} (-1)^l \binom{n-i}{l} r \omega_{j,k,r} \delta_{h,v} G(x)^{r+v-1} \tag{11}$$

where

$$\delta_{h,v} = \sum_{v=0}^m \binom{\alpha(i+l-1)}{h} \binom{\alpha(i+l-1)+h}{m} \binom{m}{v} (-1)^{h+m+v} 2^{\alpha(i+l-1)-h}$$

In the case of integer α ,

$$F_{TLG}(x)^{i+j-1} = \sum_{h=0}^{\alpha(i+j-1)} b_h(\alpha(i+j-1)) G(x)^{\alpha(i+j-1)} \tag{12}$$

Substituting $f_{TLG}(x)$ in Eq. (8) and $F_{TLG}(x)^{i+j-1}$ in Eq. (12) into Eq. (9), then we get

$$f_{in}(x) = \frac{g(x)}{B(i, n-i+1)} \sum_{j=0}^{\alpha} \sum_{l=0}^{n-i} \sum_{h=0}^{\alpha(i+l-1)} (-1)^l \binom{n-i}{l} (\alpha+j) \zeta_{j,h} G(x)^{\alpha(i+l)+j-1} \tag{13}$$

where

$$\zeta_j = \sum_{l=0}^{n-i} \sum_{h=0}^{\alpha(i+l-1)} \binom{n-i}{l} (-1)^l (\alpha+j) b_j(\alpha) b_h(\alpha(i+j-1)).$$

3.4 Moments

The moment of the TLG distribution can be computed by probability weighted moments order (s, r) of the parent distribution. Let a random variable Y follow the parent distribution $G(Y)$, then (s, r) th probability weighted moment (PWM) of Y (Greenwood *et al.*, 1979) is

$$\tau_{s,r} = E\{Y^s G(Y)^r\} = \int_{-\infty}^{\infty} y^s G(y)^r g(x) dx$$

or

$$\tau_{s,r} = \int_0^1 Q_G(u)^s u^r dx$$

Indeed, the moment of the TLG distribution from Eq. (7) can be obtained by

$$E(X^s) = \sum_{j,k=0}^{\infty} \sum_{r=0}^k \binom{\alpha}{j} \binom{\alpha+j}{k} \binom{k}{r} r (-1)^{j+k+r} 2^{\alpha-j} \tau_{s,r-1},$$

and for integer α in Eq. (8)

$$E(X^s) = \sum_{i=1}^{\alpha} \binom{\alpha}{i} (-1)^i 2^{\alpha-i} (\alpha+i) \tau_{s,\alpha+i-1}.$$

Moreover, when the TLG distribution is a mixture of the parent distribution with a new expression of parameter, then the ordinary, central, inverse and factorial moments of the TLG distribution can be derived directly from the corresponding expression of the parent distributions.

3.5 Moment generating function

We show two representations for moment generating function (mgf) $M(t)$ of the TLG distribution. Firstly, it requires the following series expansion

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

Thus, the mgf can be expressed in terms of the associated s th moment $\mu'_s = E(X^s)$

$$M(t) = \sum_{s=0}^{\infty} \frac{\mu'_s t^s}{s!}$$

Secondly, the quantile function of the parent distribution is utilizable to get $M(t)$.

Referring to the expansions of density function in Eqs. (7) and (8), their are

$$M(t) = \sum_{j,k=0}^{\infty} \sum_{r=0}^k r \omega_{j,k,r} \int_0^1 e^{tQ_G(u)} Q_G(u) u^{r-1} dx,$$

and

$$M(t) = \sum_{j=0}^{\alpha} (\alpha+j) b_j(\alpha) \int_0^1 e^{tQ_G(u)} Q_G(u) u^{\alpha+j-1} dx,$$

respectively.

4. The Topp-Leone Generalized Exponential Distribution

The results obtained in Section 3 can be applied to an example of the TLG family of distributions. In this section, we introduce an example called the Topp-Leone generalized exponential (TLGE) distribution. Let $G(x)$ be the generalized exponential (GE) distribution (Gupta and Kundu, 1999) with a cdf and a pdf as follows

$$G(x) = (1 - e^{-\lambda x})^\beta, \tag{14}$$

and

$$g(x) = \beta \lambda e^{-\lambda x} (1 - e^{-\lambda x})^{\beta-1}. \tag{15}$$

where $\beta > 0$ is a shape parameter and $\lambda > 0$ is a scale parameter.

Moreover, the quantile function of the GE distribution is

$$Q_G(u) = \frac{1}{\lambda} \log(1 - u^{1/\beta}). \tag{16}$$

Consequently, a random variable X of the TLGE distribution, $X \sim \text{TLGE}(\alpha, \lambda, \beta)$, has the cdf and pdf respectively

$$F_{\text{TLGE}}(x) = (1 - e^{-\lambda x})^{\beta\alpha} (2 - (1 - e^{-\lambda x})^\beta)^\alpha,$$

and

$$f_{\text{TLGE}}(x) = 2\alpha\beta\lambda e^{-\lambda x} (1 - (1 - e^{-\lambda x})^\beta) (1 - e^{-\lambda x})^{\beta\alpha-1} (2 - (1 - e^{-\lambda x})^\beta)^{\alpha-1}.$$

The behavior of the density function can be separated into two cases as shown in Figure 1. One is a decreasing function when $\alpha < 1$ and $\beta < 1$, and while the other is unimodal and right tailed for $\alpha > 1$ and $\beta > 1$.

The survival and hazard functions of the TLGE distribution are

$$s(x) = 1 - (1 - e^{-\lambda x})^{\beta\alpha} (2 - (1 - e^{-\lambda x})^\beta)^\alpha,$$

and

$$h(x) = \frac{2\alpha\beta\lambda e^{-\lambda x} (1 - (1 - e^{-\lambda x})^\beta) (1 - e^{-\lambda x})^{\beta\alpha-1} (2 - (1 - e^{-\lambda x})^\beta)^{\alpha-1}}{1 - (1 - e^{-\lambda x})^{\beta\alpha} (2 - (1 - e^{-\lambda x})^\beta)^\alpha}.$$

Some graphical representations of the hazard function are also illustrated in Figure 1. As λ is a scale parameter, the shape of the hazard function only depend on α and β . When $\alpha < 1$ and $\beta < 1$, the TLGE has decreasing hazard function, otherwise it has constant or increasing hazard function.

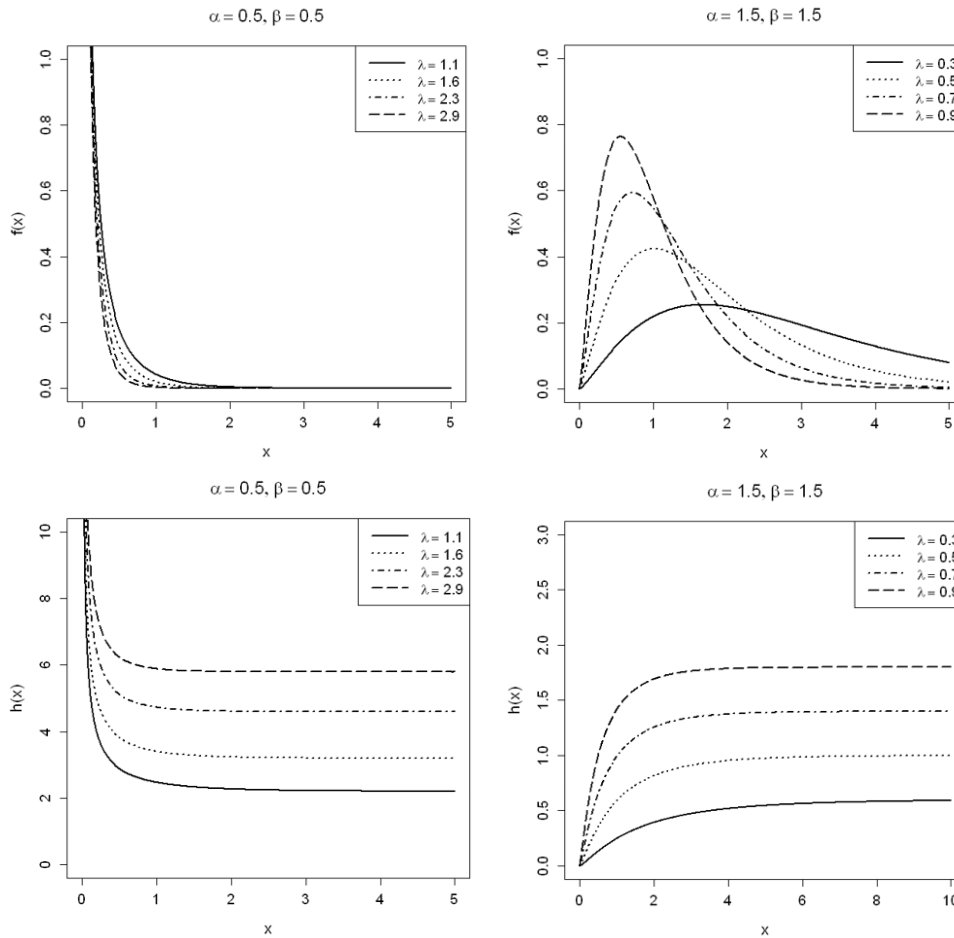


Figure 1. Plots of the TLGE density function and hazard function

The quantile function is obtained by substituting Eq. (16) for in Eq. (2)

$$x = -\frac{1}{\lambda} \log(1 - (1 - \sqrt{1 - u^{1/\alpha}})^{1/\beta}).$$

We expand the density function of the TLGE distribution to demonstrate its genesis in relation to the GE distribution. Following density expansions in Eqs. (6) and (7), in the case of real non-integer α , the cdf and pdf of the TLGE distribution can be written as

$$F_{TLGE}(x) = \sum_{j=0}^{\infty} \sum_{r=0}^k \omega_{j,k,r} G_{r\beta,\lambda}(x)$$

and

$$f_{TLGE}(x) = \sum_{j=0}^{\infty} \sum_{r=0}^k \omega_{j,k,r} g_{r\beta,\lambda}(x) \tag{18}$$

where $G_{r\beta,\lambda}(x)$ and $g_{r\beta,\lambda}(x)$ are, respectively, the cdf and pdf of the GE distribution with a shape parameter $r\beta$ and scale parameter λ , defined as $GE(r\beta, \lambda)$.

On the other hand, if α is an integer, the index i stops at b .

$$F_{TLGE}(x) = \sum_{j=0}^{\alpha} b_j(\alpha) G_{\beta(\alpha+j),\lambda}(x)$$

and

$$f_{TLGE}(x) = \sum_{j=0}^{\alpha} b_j(\alpha) g_{\beta(\alpha+j),\lambda}(x) \tag{19}$$

where $G_{\beta(\alpha+j),\lambda}(x)$ and $g_{\beta(\alpha+j),\lambda}(x)$ are the cdf and pdf of the GE distribution with a shape parameter $\beta(\alpha + j)$ and scale parameter λ , denoted by $GE(\beta(\alpha + j), \lambda)$, respectively. In summary, the TLGE distribution is infinite or finite mixture of the GE distribution in relation to the value of α .

We can calculate $M(t)$ of the TLGE distribution with the use of the same measure of $GE(r\beta, \lambda)$ and $GE(\beta(\alpha + j), \lambda)$. Without loss of generality, we provide an example based only on real non-integer α . As a result, $M(t)$ of the TLGE distribution in Eq. (18) is

$$M(t) = \sum_{j=0}^{\infty} \sum_{r=0}^k \omega_{j,k,r} \frac{\Gamma(r\beta + 1)\Gamma(1 - t / \lambda)}{\Gamma(r\beta - t / \lambda + 1)},$$

Therefore,

$$E(X^n) = \sum_{j,k=0}^{\infty} \sum_{r=0}^k \omega_{j,k,r} \left(\frac{\partial}{\partial t} \right)^n \frac{\Gamma(r\beta + 1)\Gamma(1 - t / \lambda)}{\Gamma(r\beta - t / \lambda + 1)} \Bigg|_{t=0}$$

The first and second moments are

$$E(X) = \sum_{j,k=0}^{\infty} \sum_{r=0}^k \omega_{j,k,r} \frac{\Gamma(1)}{\lambda} (\psi^{(0)}(r\beta + 1) - \psi^{(0)}(1))$$

$$E(X^2) = \sum_{j,k=0}^{\infty} \sum_{r=0}^k \omega_{j,k,r} \frac{\Gamma(1)}{\lambda^2} (\psi^{(1)}(1) - \psi^{(1)}(r\beta + 1) + (\psi^{(0)}(1) - \psi^{(0)}(r\beta + 1))^2)$$

The mean and variance of the TLGE distribution are $E(X)$ and $Var(X) = E(X^2) - E^2(X)$ respectively.

5. Maximum Likelihood Estimation

In this section, we focus on estimation of unknown parameter vector, $\theta = (\alpha, \lambda, \beta)^T$, based on the maximum likelihood estimation. Suppose x_1, x_2, \dots, x_n are an observed sample from the TLGE distribution.

Consequently, the likelihood function can be expressed as

$$L(\theta) = (2\alpha\beta\lambda)^n \prod_{i=1}^n e^{-\lambda x_i} (1 - v_i^\beta) v_i^{\beta\alpha - 1} (2 - v_i^\beta)^{\alpha - 1},$$

where $v_i = 1 - e^{-\lambda x_i}$.

The log-likelihood function is

$$\begin{aligned} \ell(\theta) = & n \log(2) + n \log(\alpha) + n \log(\beta) \\ & + n \log(\lambda) - \lambda \sum_{i=1}^n x_i + \sum_{i=1}^n \log(1 - v_i^\beta) \\ & + (\beta\alpha - 1) \sum_{i=1}^n \log(v_i) + (\alpha - 1) \sum_{i=1}^n \log(2 - v_i^\beta). \end{aligned} \tag{20}$$

By differentiating $\ell(\theta)$ with respect to α , λ , and β , respectively, the components of the unit score vector are

$$\begin{aligned} U_\alpha(\theta) &= \frac{n}{\alpha} + \beta \sum_{i=1}^n \log(v_i) + \sum_{i=1}^n \log(2 - v_i^\beta), \\ U_\lambda(\theta) &= \frac{n}{\lambda} - \sum_{i=1}^n x_i - \sum_{i=1}^n \left(\frac{\beta x_i e^{-\lambda x_i} v_i^{\beta-1}}{1 - v_i^\beta} \right) \\ &+ (\beta\alpha - 1) \sum_{i=1}^n \left(\frac{x_i e^{-\lambda x_i}}{v_i} \right) \\ &- \beta(\alpha - 1) \sum_{i=1}^n \left(\frac{x_i e^{-\lambda x_i} v_i^{\beta-1}}{2 - v_i^\beta} \right), \\ U_\beta(\theta) &= \frac{n}{\beta} - \sum_{i=1}^n \frac{v_i^\beta \log(v_i)}{1 - v_i^\beta} + \alpha \sum_{i=1}^n \log(v_i) \\ &- (\alpha - 1) \sum_{i=1}^n \left(\frac{v_i^\beta \log(v_i)}{2 - v_i^\beta} \right). \end{aligned} \tag{21}$$

Furthermore, a system of non-linear equations is derived by setting $U_\alpha(\theta) = U_\lambda(\theta) = U_\beta(\theta) = \mathbf{0}$. Then, we obtain the maximum likelihood estimate $\hat{\theta} = (\hat{\alpha}, \hat{\lambda}, \hat{\beta})^T$ of $\theta = (\alpha, \lambda, \beta)^T$ by solving the system of non-linear equations numerically through the Newton-Raphson procedure. For interval estimation and hypothesis tests on the model parameters, we require the 3×3 unit observed information matrix $J(\theta)$. Its elements are given in the Appendix.

6. Applications

In this section, we model real datasets with the TLGE, the GE, and the exponentiated generalized exponential (EGE) distributions (Cordeiro *et al.*, 2013c). In addition, the GE distribution can be considered as special case of the TLGE and EGE distributions. To verify which distribution fits better to real datasets, the Kolmogorov-Smirnov test (KS test) will be employed. Other criteria including the Akaike information criterion (AIC), Bayesian information criterion (BIC), and consistent Akaike information criterion (CAIC) are provided for the purpose of model selection. Furthermore, the aforementioned values of the TLGE and the GE distributions are obtained by using the function `goodness.fit` in `AdequacyModel` package (Diniz Marinho *et al.*, 2013) of R language (R Core Team, 2014). In case of the EGE distribution, we computed those values by function `meg` in `Newdistns` package (Nadarajah and Rocha, 2015).

The first dataset consists of 101 observations with maximum stress per cycle 31,000 psi (Birnbau and Saunders, 1969), and the second dataset is breaking stress of carbon fibers provided in `AdequacyModel` package, shown in Tables 1 and 2, respectively.

In Figure 2, the TTT-Transform plots are demonstrated that the datasets have increasing hazard function. Furthermore, all of the competitive distributions, which have increasing shaped of hazard function, would be appropriate for analyzing these lifetime data.

Tables 3 and 4 show MLEs with the corresponding standard errors in parentheses, KS test, AIC, BIC, and CAIC for maximum stress data and carbon data.

For both of the datasets, the p-value of KS-test under the TLGE distribution is greater than any other distributions. Additionally, AIC, BIC, and CAIC values of the TLGE distribution are the smallest ones of the candidate models. Figures 3-4 display the closeness of the distributions to raw datasets, which also suggest the use of the TLGE distribution. For these reasons, the TLGE distribution is more appropriate for fitting these datasets than the GE and EGE distributions.

7. Conclusions

In this paper, we propose the TLG family of distributions. The method for generating the TLG distribution is presented in Section 3. Some important properties of the the TLG distribution are discussed. This TLG family of distributions has closed forms of cdf, pdf, survival function and

Table 1. The maximum stress per cycle

| | | | | | | | | | | | |
|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 70 | 90 | 96 | 97 | 99 | 100 | 103 | 104 | 104 | 105 | 107 | 108 |
| 108 | 108 | 109 | 109 | 112 | 112 | 113 | 114 | 114 | 114 | 116 | 119 |
| 120 | 120 | 120 | 121 | 121 | 123 | 124 | 124 | 124 | 124 | 124 | 128 |
| 128 | 129 | 129 | 130 | 130 | 130 | 131 | 131 | 131 | 131 | 131 | 132 |
| 132 | 132 | 133 | 134 | 134 | 134 | 134 | 136 | 136 | 137 | 138 | 138 |
| 138 | 139 | 139 | 141 | 141 | 142 | 142 | 142 | 142 | 142 | 142 | 144 |
| 144 | 145 | 146 | 148 | 148 | 149 | 151 | 151 | 152 | 155 | 156 | 157 |
| 157 | 157 | 157 | 158 | 159 | 162 | 163 | 163 | 164 | 166 | 166 | 168 |
| 170 | 174 | 201 | 212 | | | | | | | | |

Table 2. Breaking stress of carbon fibers

| | | | | | | | | | | | |
|------|------|------|------|------|------|------|------|------|------|------|------|
| 3.70 | 2.74 | 2.73 | 2.50 | 3.60 | 3.11 | 3.27 | 2.87 | 1.47 | 3.11 | 4.42 | 2.41 |
| 3.19 | 3.22 | 1.69 | 3.28 | 3.09 | 1.87 | 3.15 | 4.90 | 3.75 | 2.43 | 2.95 | 2.97 |
| 3.39 | 2.96 | 2.53 | 2.67 | 2.93 | 3.22 | 3.39 | 2.81 | 4.20 | 3.33 | 2.55 | 3.31 |
| 3.31 | 2.85 | 2.56 | 3.56 | 3.15 | 2.35 | 2.55 | 2.59 | 2.38 | 2.81 | 2.77 | 2.17 |
| 2.83 | 1.92 | 1.41 | 3.68 | 2.97 | 1.36 | 0.98 | 2.76 | 4.91 | 3.68 | 1.84 | 1.59 |
| 3.19 | 1.57 | 0.81 | 5.56 | 1.73 | 1.59 | 2.00 | 1.22 | 1.12 | 1.71 | 2.17 | 1.17 |
| 5.08 | 2.48 | 1.18 | 3.51 | 2.17 | 1.69 | 1.25 | 4.38 | 1.84 | 0.39 | 3.68 | 2.48 |
| 0.85 | 1.61 | 2.79 | 4.70 | 2.03 | 1.80 | 1.57 | 1.08 | 2.03 | 1.61 | 2.12 | 1.89 |
| 2.88 | 2.82 | 2.05 | 3.65 | | | | | | | | |

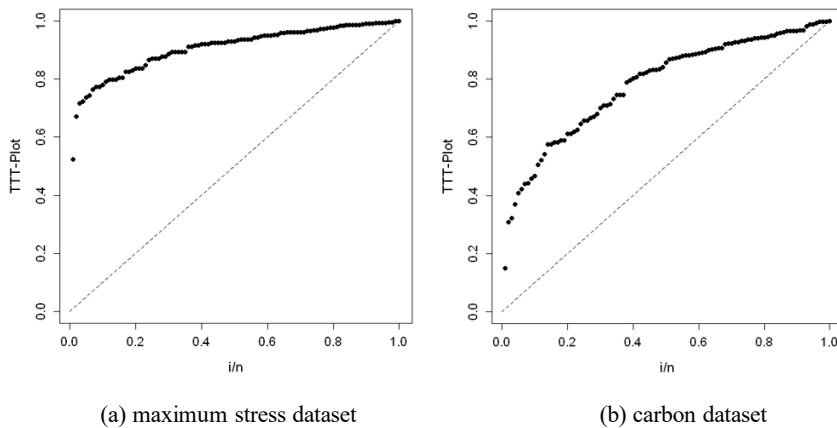


Figure 2. TTT-Transform plots for both of the datasets

Table 3. Summary of fitting, goodness-of-fit testing results , AIC , BIC, and CAIC values for maximum stress dataset

| | MLEs | KS test (p-value) | AIC | BIC | CAIC |
|------|--|-------------------|----------|----------|----------|
| TLGE | $\hat{\alpha} = 6.8607(5.5692)$ $\hat{\lambda} = 0.0275(0.0023)$ $\hat{\beta} = 13.4886(9.1228)$ | 0.0997(0.2735) | 917.8901 | 925.7056 | 918.1401 |
| GE | $\hat{\lambda} = 0.0460(0.0032)$ $\hat{\beta} = 281.8876(106.2584)$ | 0.1068(0.2041) | 921.2854 | 926.4958 | 921.4091 |
| EGE | $\hat{\alpha} = 0.2391(0.0185)$ $\hat{b} = 105.6216(27.4080)$ $\hat{\lambda} = 0.1593(0.0080)$ | 0.1229(0.0971) | 929.6721 | 937.4876 | 929.9221 |

Table 4. Summary of fitting, goodness-of-fit testing results , AIC , BIC, and CAIC values for carbon dataset

| | MLEs | KS test (p-value) | AIC | BIC | CAIC |
|------|---|-------------------|----------|----------|----------|
| TLGE | $\hat{\alpha} = 0.2921(0.2903)$ $\hat{\lambda} = 0.8101(0.1296)$ $\hat{\beta} = 21.0731(23.0517)$ | 0.0919(0.3669) | 292.0925 | 299.908 | 292.3425 |
| GE | $\hat{\lambda} = 1.0131(0.0875)$ $\hat{\beta} = 7.7883(1.4962)$ | 0.1077(0.1962) | 296.3646 | 301.5749 | 296.4883 |
| EGE | $\hat{\alpha} = 0.8479(3.8201)$ $\hat{b} = 7.7953(1.4983)$ $\hat{\lambda} = 1.1953(5.3850)$ | 0.1076(0.1965) | 298.3646 | 306.1801 | 298.6146 |

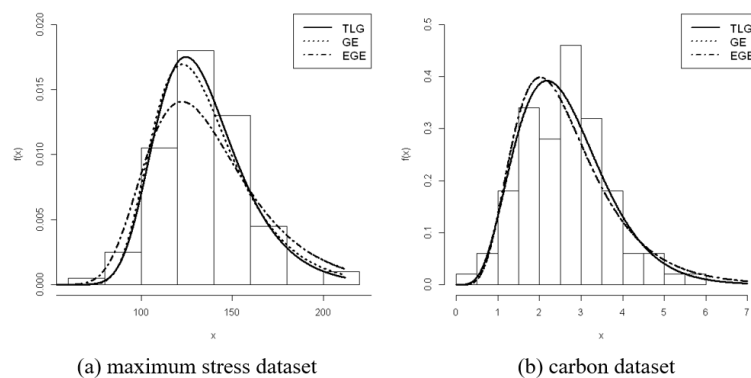


Figure 3. Empirical and fitted distributions of the TLGE, GE, and GEG distributions

hazard function. The moments can be derived from PWM’s moments of the parent distribution. The special case of the TLG family of distributions is presented, which is called the TLGE distribution. Then, we apply the general properties to the TLGE distribution. Parameter estimation and its observed Fisher information matrix of the TLGE distribution are provided. In Section 6, the applications of the TLGE distribution are demonstrated, and then we compare the fitted results with its parent distribution and EGE distribution. According to the values of KS test, AIC, BIC, and CAIC, the TLGE distribution can be considered a competitive distribution for the GE and EGE distributions.

Acknowledgements

It is our pleasure to thank the Department of Statistics, Faculty of Science, Kasetsart University, Graduate School Kasetsart University and the Research Professional Development Project under the Science Achievement Scholarship of Thailand (SAST).

References

Akinsete, A., Famoye, F., and Lee, C. 2008. The beta-Pareto distribution. *Statistics*. 42 (6), 547-563.

AL-Hussaini, E. K., and Ahsanullah, M., 2015. Exponentiated Distribution. *Atlantis Studies in Probability and Statistics*. Atlantis Press, Paris, France.

Al-Zahrani, B. 2012. Goodness-of-fit for the Topp-Leone distribution with unknown parameters. *Applied Mathematical Sciences*. 6(128).

Alexander, C., Cordeiro, G. M., Ortega, E. M., and Sarabia, J. M. 2012. Generalized beta-generated distributions. *Computational Statistics and Data Analysis*. 56, 1880-1897.

Alizadeh, M., Tahir, M. H., Cordeiro, G. M., Mansoor, M., Zubair, M., and Hamedani, G. G. 2015. The Kumaraswamy Marshal-Olkin family of distributions. *Journal of Statistical Distributions and Applications*. <http://dx.doi.org/10.1016/j.joems.2014.12.002>.

Alzaatreh, A., Lee, C., and Famoye, F. 2013. A new method for generating families of continuous distributions. *METRON*. 71, 63-79.

Amini, M., MirMostafae, S. M. T. K., and Ahmadi, J. 2014. Log-gamma-generated families of distributions. *Statistics*. 48(4), 913-932.

Barreto-Souza, W., Santos, A. H. S., and Cordeiro, G. M. 2010. The beta generalized exponential distribution. *Journal of Statistical Computation and Simulation*. 80(2), 159-172.

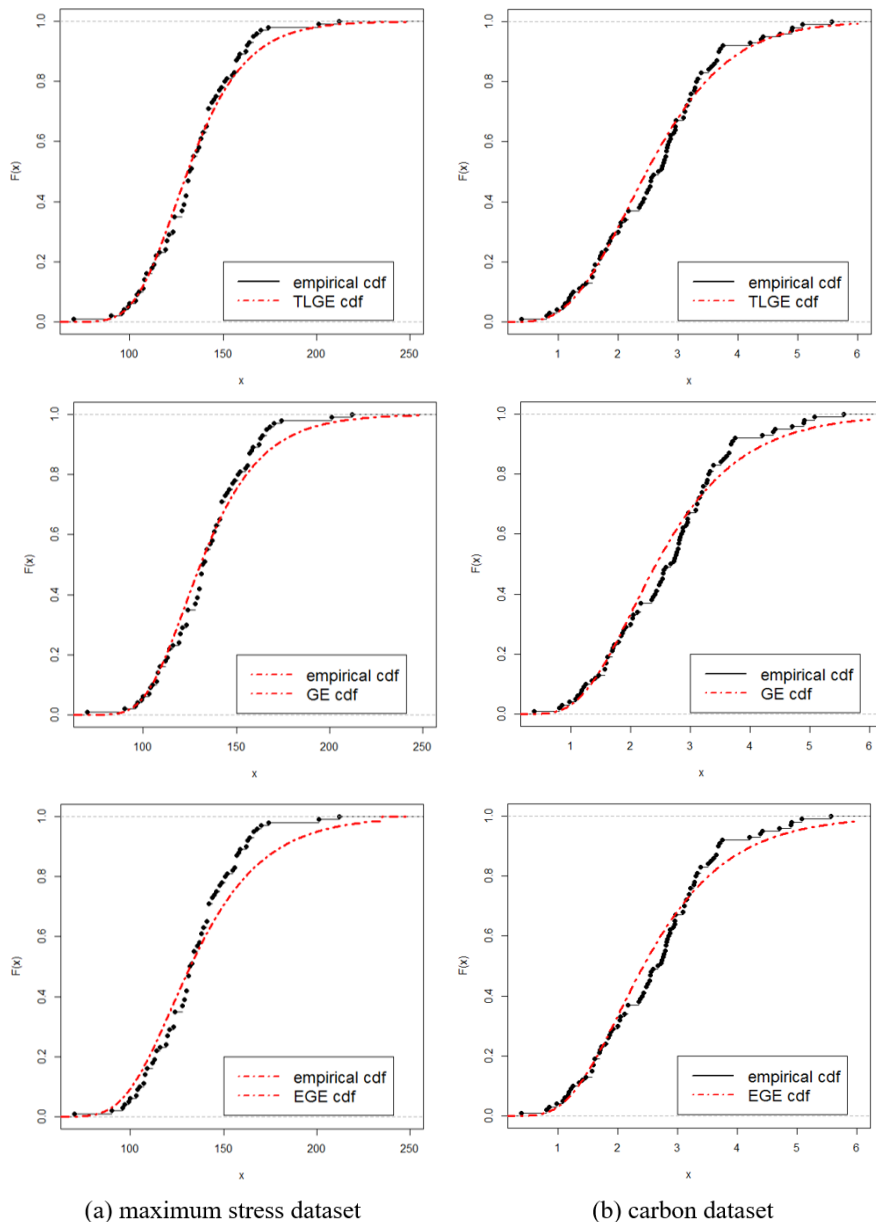


Figure 4. Empirical cdf and theoretical cdf of the TLGE, GE, and GEG distributions

- Bhati, D., Malik, M. A., and Vaman, H. J. 2015. Lindley-exponential distribution: properties and applications. *METRON*. 1–23. <http://dx.doi.org/10.1007/s40300-015-0060-9>.
- Birnbaum, Z. W. and Saunders, S. C. 1969. Estimation for a family of life distributions with applications to fatigue. *Journal of Applied Probability*. 6, 328-347.
- Bourguignon, M., Silva, R. B., and Cordeiro, G. M. 2014. The Weibull-g family of probability distributions. *Journal of Data Science*. 12, 53-68.
- Cordeiro, G. M., Cristino, C. T., Hashimoto, E. M., and Ortega, E. M. M. 2013a. The beta generalized Rayleigh distribution with applications to lifetime data. *Stat Papers*. 54, 133-161.
- Cordeiro, G. M., Gomes, A. E., da Silva, C. Q., and Ortega, E. M. M., 2013b. The beta exponentiated Weibull distribution. *Journal of Statistical Computation and Simulation*. 83 (1), 114-138.
- Cordeiro, G. M., Hashimoto, E. M., and Ortega, E. M. 2014a. The McDonald Weibull model. *A Journal of Theoretical and Applied Statistics*. 48(2), 256-278.
- Cordeiro, G. M., Ortega, E. M. M., and da Cunha, D. C. C. 2013c. The exponentiated generalized class of distributions. *Journal of Data Science*. 11, 1-27.
- Cordeiro, G. M., Ortega, E. M. M., and Pescim, B. V. P. R. R. 2014b. The Lomax generator of distributions: Properties, minification process and regression model. *Applied Mathematics and Computation*. 247, 465-486.

- Diniz Marinho, P. R., Bourguignon, M., and Barros Dias, C. R. 2013. AdequacyModel: Adequacy of probabilistic models and generation of pseudo-random numbers. R package version 1.0.8. <http://CRAN.R-project.org/package=AdequacyModel>
- Eugene, N., Lee, C., and Famoye, F. 2002. Beta-normal distribution and its applications. *Communications in Statistics - Theory and Methods*. 31(4), 497-512.
- Ghitany, M. E., Kotz, S., and Xie, M. 2005. On some reliability measures and their stochastic orderings for the Topp-Leone distribution. *Journal of Applied Statistics*. 32(7), 715-722.
- Greenwood, J. A., Landwehr, J., Matalas, N., and Wallis, J. 1979. Probability weighted moments definition and relation to parameters of several distributions expressible in inverse form. *Water Resources Research*. 15, 1049-1054.
- Gupta, R. D. and Kundu, D. 1999. Generalized exponential distributions. *Australian and New Zealand Journal of Statistics*. 41(2), 173-188.
- Gurvich, M. R., Dibenedetto, A. T., and Rande, S. V. 1997. A new statistical distribution for characterizing the random strength of brittle materials. *Journal of Materials Science*. 32, 2559-2564.
- Insuk, T., Bodhisuwan, W., and Jaroengertakun, U. 2015. A new mixed beta distribution and structural properties with applications. *Songklanakar J. Sci. Technol.* 37(1), 97-108.
- Jones, M. C. 2004. Families of distributions arising from distributions of order statistics. *TEST*. 13 (1), 1-43.
- Jones, M. C. 2009. Kumaraswamy's distribution: A beta-type distribution with some tractability advantages. *Statistical Methodology*. 6, 70-81.
- Kotz, S. and Seier, E. 2007. Kurtosis of the Topp-Leone distributions. *Interstat*, 2007 (July).
- Lee, C., Famoye, F., and Alzaatreh, A. Y. 2013. Methods for generating families of univariate continuous distributions in the recent decades. *WIREs Computational Statistics*. 5, 219-238.
- Lee, C., Famoye, F., and Olumolade, O. 2007. Beta-Weibull distribution: some properties and applications to censored data. *Journal of Modern Applied Statistical Methods*. 6 (1), 173-186.
- Marshall, A. W. and Olkin, I. 1997. A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika*. 84, 641-652.
- Nadarajah, S. and Kotz, S. 2003. Moments of some J-shaped distributions. *Journal of Applied Statistics*. 30(3), 311-317.
- Nadarajah, S. and Kotz, S. 2004. The beta gumbel distribution. *Mathematical Problems in Engineering*. 2004(4), 323-332.
- Nadarajah, S. and Rocha, R. 2015. Newdistns: Computes Pdf, Cdf, Quantile and Random Numbers, Measures of Inference for 19 General Families of Distributions. R package version 2.0. Available from: <https://cran.r-project.org/web/packages/Newdistns/index.html>. [December 17, 2014].
- Nadarajah, S. and Kotz, S. 2006. The beta exponential distribution. *Reliability Engineering and System Safety*. 91, 689-697.
- Pescim, R. R., Cordeiro, G. M., Demetrio, C. G. B., Ortega, E., and Nadarajah, S. 2012. The new class of Kummer beta generalized distributions. *Statistics and Operations Research Transactions – SORT*. 36, 153-180.
- R Core Team. 2014. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>
- Shaw, W. and Buckley, I. 2007. The alchemy of probability distributions: beyond gram-charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. *Research Report*.
- Singla, N., Jain, K., and Sharma, S. K. 2012. The beta generalized Weibull distribution: properties and applications. *Reliability Engineering and System Safety*. 102, 5-15.
- Tahir, M. H., Cordeiro, G. M., Alizadeh, M., Mansoor, M., Zubair, M., and Hamedani, G. G. 2015. The Kumaraswamy Marshal-Olkin family of distributions. *Journal of Statistical Distributions and Applications*. 2 (1), <http://dx.doi.org/10.1186/s40488-014-0024-2>.
- Topp, C. W. and Leone, F. C. 1955. A family of J-shaped frequency functions. *Journal of the American Statistical Association*. 50(269), 209-219.
- Vicaria, D., Dorpb, J. R. V., and Kotzb, S. 2008. Two-sided generalized Topp and Leone (TS-GTL) distributions. *Journal of Applied Statistics*. 35(10), 1115-1129.
- Zografos, K. and Balakrishnan, N. 2009. On families of beta- and generalized gamma-generated distributions and associated inference. *Statistical Methodology*. 6, 344-362.

Appendix

$$J_{\beta\beta} = -\frac{n}{\beta^2} - \sum_{i=1}^n \left(\frac{\log^2(v_i)}{v_i^\beta} \left(\frac{1}{(v_i^{-\beta} - 1)^2} - \frac{2(\alpha - 1)}{(2v_i^{-\beta} - 1)^2} \right) \right)$$

$$J_{\beta\lambda} = \sum_{i=1}^n \left\{ \frac{x_i e^{-\lambda x_i}}{v_i} \left(\frac{1}{v_i^{-\beta} - 1} + \frac{\beta v_i^{-\beta} \log(v_i)}{(v_i^{-\beta} - 1)^2} \right) \right\} \\ - (\alpha - 1) \sum_{i=1}^n \left\{ \frac{x_i e^{-\lambda x_i}}{v_i} \left(\frac{1}{2v_i^{-\beta} - 1} + \frac{2\beta v_i^{-\beta} \log(v_i)}{(2v_i^{-\beta} - 1)^2} \right) \right\} + \alpha \sum_{i=1}^n \left(\frac{x_i e^{-\lambda x_i}}{v_i} \right)$$

$$J_{\lambda\lambda} = -\frac{n}{\lambda^2} + \sum_{i=1}^n \left(\frac{\beta x_i e^{-2\lambda x_i} v_i^{-\beta-2}}{(v_i^{-\beta} - 1)^2} - \frac{x_i e^{-\lambda x_i} v_i^{-2}}{v_i^{-\beta} - 1} \right) + (\alpha - 1) \sum_{i=1}^n \left(\frac{2\beta x_i e^{-2\lambda x_i} v_i^{-\beta-2}}{(2v_i^{-\beta} - 1)^2} - \frac{x_i e^{-\lambda x_i} v_i^{-2}}{2v_i^{-\beta} - 1} \right) \\ - (\beta\alpha - 1) \sum_{i=1}^n \left(\frac{x_i^2 e^{-\lambda x_i}}{v_i^2} \right)$$

$$J_{\alpha\alpha} = -\frac{n}{\alpha^2}$$

$$J_{\alpha\beta} = \sum_{i=1}^n \log(v_i) - \sum_{i=1}^n \left(\frac{\log(v_i)}{2v_i^{-\beta} - 1} \right)$$

$$J_{\alpha\lambda} = \sum_{i=1}^n \left(\frac{\beta x_i e^{-\lambda x_i}}{v_i} \right) - \sum_{i=1}^n \left(\frac{\beta x_i e^{-\lambda x_i} (v_i)^{\beta-1}}{2 - v_i^\beta} \right)$$