

RESEARCH

Open Access



# Families of distributions arising from the quantile of generalized lambda distribution

Mahmoud Aldeni<sup>\*</sup> , Carl Lee and Felix Famoye

<sup>\*</sup> Correspondence:  
alden1ms@cmich.edu  
Department of Mathematics, Central  
Michigan University, Mount  
Pleasant, MI 48859, USA

## Abstract

In this paper, the class of  $T-R$  {generalized lambda} families of distributions based on the quantile of generalized lambda distribution has been proposed using the  $T-R\{Y\}$  framework. In the development of the  $T-R\{Y\}$  framework, the support of  $Y$  and  $T$  must be the same. It is typical that the random variable  $Y$  has one type of support and  $T$  is restricted to the same support. Taking  $Y$  to be a generalized lambda random variable leads to three different types of supports, thus, making the choice of the generator  $T$  to be much more broad and flexible. This is interesting and unique. By allowing  $T$  with different supports makes the  $T-R$ {generalized lambda} a desirable method for generating new versatile and broad families of generalized distributions for any given random variable  $R$ . Some general properties of these families of distributions are studied. Four members of the  $T-R$ {generalized lambda} families of distributions are derived. The shapes of these distributions can be symmetric, skewed to the left, skewed to the right, or bimodal. Two real life data sets are applied to illustrate the flexibility of the distributions.

**Keywords:**  $T-R\{Y\}$  framework, Quantile function, Shannon's entropy

## Introduction

Statistical distributions play an important role in theory and applications, which are used to fit model and describe real world phenomena. For this reason, statistical distributions and their properties are of great importance especially in the social sciences (such as economics, political science) and engineering disciplines such as computer science, as well as in the natural sciences (such as biology, chemistry, physics). Although a large number of distributions have been defined and studied over the years, seeking for more flexibility in fitting data remains a strong reason for researchers to develop and study new distributions.

In the last two decades, there has been a growing body of research concerned with developing new and more flexible univariate statistical distributions. For example, Eugene et al. (2002) introduced a new method to develop the beta-generated family of distributions. Using this methodology, a significant number of new families of distributions have been defined and studied. Examples of the beta-generated family of distributions include the beta-normal distribution introduced by Eugene et al. (2002), the beta-exponential distribution (Nadarajah and Kotz, 2006), the beta-Weibull distribution (Famoye, Lee and Olumolade, 2005), the beta-Pareto distribution (Akinsete, Famoye and Lee, 2008), and others.

An extension of the beta-generated family of distributions was proposed in Jones (2009) and Cordeiro and de Castro (2011) by replacing the beta distribution with the Kumaraswamy distribution (Kumaraswamy, 1980). Many statistical properties of Kumaraswamy-generated ( $Kw-G$ ) family have been studied in the literature. Examples of this family include  $Kw$ -Weibull distribution (Cordeiro et al., 2010),  $Kw$ -Pareto distribution (Pereira et al., 2012),  $Kw$ -Burr XII distribution (Paranaíba et al., 2013) and  $Kw$ -log-logistic distribution (de Santana et al., 2012).

In the beta- and  $Kw$ -generated families, the use of distributions with support between 0 and 1 was a limitation in generating different classes of distributions. A more general family, called the  $T-X(W)$  family, was introduced by Alzaatreh et al. (2013) to derive new families of distributions by using continuous random variable as a generator.

Let  $r(t)$  and  $R(t)$  be the probability density function (PDF) and the cumulative distribution function (CDF) of a random variable  $T \in [a, b]$ , for  $-\infty \leq a < b \leq \infty$ , and let  $F(x)$  be the CDF of a random variable  $X$  such that the link function  $W(\cdot): [0, 1] \rightarrow [a, b]$  is monotonic and absolutely continuous with  $W(0) \rightarrow a$  and  $W(1) \rightarrow b$ . If the interval  $[a, b]$  is open or half-open, we replace  $W(0) \rightarrow a$  and/or  $W(1) \rightarrow b$  with  $\lim_{\lambda \rightarrow 0^+} W(\lambda) = a$  and/or  $\lim_{\lambda \rightarrow 1^-} W(\lambda) = b$ . The CDF and the PDF of the  $T-X(W)$  family of distributions are defined, respectively, as

$$G(x) = \int_a^{W(F(x))} r(t) dt = R\{W(F(x))\} \text{ and } g(x) = \left\{ \frac{d}{dx} W(F(x)) \right\} r\{W(F(x))\}.$$

Based on this method, the use of different  $W(\cdot)$  functions generates a large number of distributions. For example, Alzaatreh et al. (2012) used  $W(F(x)) = -\log\{1 - F(x)\}$  to define and study the gamma-Pareto distribution. In a similar way, Al-Aqtash et al. (2015) used the logit of the CDF  $F(x)$ , which is defined as  $W(F(x)) = \log\{F(x)/(1 - F(x))\}$ , to generate the Gumbel-Weibull distribution.

Aljarrah et al. (2014) proposed quantile based approach to refine the  $T-X(W)$  family by replacing the function  $W(\cdot)$  with  $Q_Y(\cdot)$ , where  $Q_Y$  is the quantile function of a random variable  $Y$ . This family was first named as the  $T-X\{Y\}$  family. The methodology is called the  $T-R\{Y\}$  framework after the following unified notation given in Alzaatreh et al. (2014):

Let  $F_T(x) = P(T \leq x)$ ,  $F_R(x) = P(R \leq x)$ , and  $F_Y(x) = P(Y \leq x)$  be the CDFs of the random variables  $T$ ,  $R$ , and  $Y$ , respectively, with corresponding quantile functions  $Q_T(u)$ ,  $Q_R(u)$ , and  $Q_Y(u)$ , where  $Q_Z(u) = \inf\{z : F_Z(z) \geq u\}$ ,  $u \in (0, 1)$ . The PDFs (if they exist) will be denoted by  $f_T(x)$ ,  $f_R(x)$ , and  $f_Y(x)$ , respectively. The CDF of the random variable  $X$  is defined as

$$F_X(x) = \int_a^{Q_Y(F_R(x))} f_T(t) dt = F_T\{Q_Y(F_R(x))\}, \quad T, Y \in (a, b), \text{ for } -\infty \leq a < b \leq \infty, \quad (1.1)$$

and accordingly the corresponding PDF associated with (1.1) is

$$f_X(x) = f_R(x) \times f_T\{Q_Y(F_R(x))\} \times Q'_Y(F_R(x)) = f_R(x) \times \frac{f_T\{Q_Y(F_R(x))\}}{f_Y\{Q_Y(F_R(x))\}}. \quad (1.2)$$

If  $Y$  follows the standard uniform distribution and  $T$  follows the beta distribution (or Kumaraswamy distribution), then the  $T-R\{Y\}$  family reduces to beta-generated family (or  $Kw-G$  family). Different choices of the random variables  $T$  and  $Y$  lead to different

families of generalized  $R$ -distributions. Some research articles in the literature have proposed several generalizations of some  $R$ -distributions based on the  $T$ - $R\{Y\}$  framework. Examples include  $T$ -normal $\{Y\}$  by Alzaatreh et al. (2014) and  $T$ -Weibull $\{Y\}$  by Almheidat et al. (2015). In this paper, we use the quantile function of generalized lambda distribution (GLD) proposed by Ramberg and Schmeiser (1974) to develop new generalization of different  $R$  distributions, by using the  $T$ - $R\{Y\}$  framework. For a review of the recent development of generalized distributions, one may refer to Lee et al. (2013).

The rest of this paper is organized as follows. In Section 2, we briefly review the development of the GLDs, and define the  $T$ - $R\{\text{generalized lambda}\}$  ( $T$ - $R\{GL\}$ ) families of distributions based on the quantile function of GLD. Some general properties of the proposed families are investigated in Section 3. In Section 4, four members of the  $T$ - $R\{GL\}$  families of distributions are derived, and some of their properties are studied. In Section 5, we address parameter estimation and simulation for the uniform-exponential $\{\text{generalized lambda}\}$  distribution. In Section 6, we present two applications illustrating the usefulness of the uniform-exponential $\{\text{generalized lambda}\}$  distribution in fitting real data and compare the results with other existing distributions. Section 7 summarizes the main findings and concludes the article.

## The $T$ - $R\{\text{generalized lambda}\}$ families of distributions

### A brief review of generalized lambda distribution

The family of generalized lambda distributions (GLDs) is known for its high flexibility. It produces distributions with a wide range of various shapes, and provides good approximations to many of the commonly used distributions such as the uniform, normal, exponential, Weibull, and logistic. For these reasons, there is an extensive amount of literature that presented and discussed different techniques for estimating the parameters of the GLDs, as well as fitting its quantile regression model to empirical data.

Ramberg and Schmeiser (1974) proposed the four-parameter generalized lambda distribution (GLD), which is the most discussed member of the different GLDs. The GLD is defined in terms of the quantile function  $Q(u) = Q(u; \lambda_1, \lambda_2, \lambda_3, \lambda_4) = \lambda_1 + \frac{u^{\lambda_3} - (1-u)^{\lambda_4}}{\lambda_2}$ ,  $u \in (0, 1)$ . The parameters  $\lambda_1$  and  $\lambda_2$  are, respectively, the location and the scale parameters, whereas  $\lambda_3$  and  $\lambda_4$  are shape parameters and determine the skewness and kurtosis of the GLD. When  $\lambda_1 = 0$  and  $\lambda_2 = \lambda_3 = \lambda_4$ , we obtain the Tukey lambda distribution (Tukey, 1960). The GLD is asymmetric when  $\lambda_3 \neq \lambda_4$ , and has different shapes (unimodal, monotone, U-shaped, and S-shaped).

The corresponding PDF from the quantile function of the GLD is given by

$$f(x) = f(Q(u)) = \lambda_2 \left[ \lambda_3 u^{\lambda_3-1} + \lambda_4 (1-u)^{\lambda_4-1} \right]^{-1}, \text{ at } x = Q(u),$$

and accordingly the quantile density function is

$$q(u) = \lambda_2^{-1} \left[ \lambda_3 u^{\lambda_3-1} + \lambda_4 (1-u)^{\lambda_4-1} \right].$$

In order to have a valid distribution, the PDF of GLD must satisfy the following conditions:

- (i) For all  $x$  over the allowed domain,  $f(x) \geq 0$ , and (ii)  $\int_{Q(0)}^{Q(1)} f(x) dx = 1$ .

As originally indicated by Ramberg and Schmeiser (1974) and Karian and Dudewicz (2000), there are six regions of parameter values where GLD is valid, see Fig. 1. The conditions on the parameters and the support regions for the PDF of GLD are listed in Table 1.

**Definition of  $T$ - $R$ {generalized lambda} families of distributions**

In this sub-section we define the class of  $T$ - $R$ {GL} families of distributions based on the quantile function of GLD.

Let  $Y$  be a random variable that follows the GLD, then the definition in (1.1) gives the CDF (in general) of the random variable  $X$  in  $T$ - $R$ {GL} families of distributions as

$$F_X(x) = F_T(Q_Y(F_R(x))) = F_T\left(\lambda_1 + \frac{F_R^{\lambda_3}(x) - (1 - F_R(x))^{\lambda_4}}{\lambda_2}\right). \quad (2.1)$$

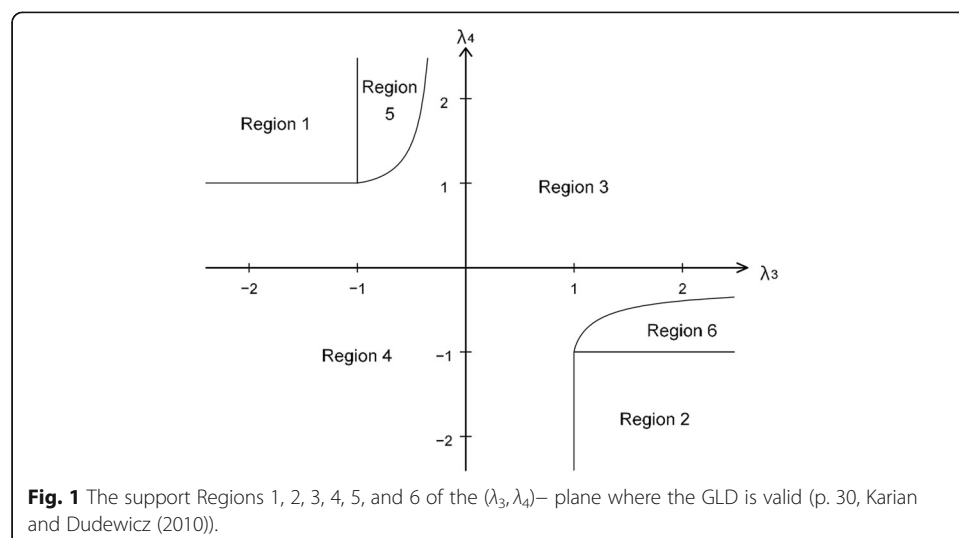
The corresponding PDF associated with (2.1) is

$$f_X(x) = f_R(x) \left( \frac{\lambda_3 F_R^{\lambda_3-1}(x) + \lambda_4 (1 - F_R(x))^{\lambda_4-1}}{\lambda_2} \right) f_T\left(\lambda_1 + \frac{F_R^{\lambda_3}(x) - (1 - F_R(x))^{\lambda_4}}{\lambda_2}\right). \quad (2.2)$$

The hazard function of the  $T$ - $R$ {GL} families of distributions can be obtained from the definition  $h_X(x) = f_X(x)/(1 - F_X(x))$ .

Based on the  $T$ - $R$ {Y} framework, the link function  $Q_Y: [0, 1] \rightarrow [a, b]$ , for  $-\infty \leq a < b \leq \infty$ , is absolutely continuous and monotonic with  $\lim_{u \rightarrow 0^+} Q_Y(u) = a$  and  $\lim_{u \rightarrow 1^-} Q_Y(u) = b$ , where  $[a, b]$  is the support of the random variable  $T$ . In other words, the choice of the random variable  $T$  is not arbitrary and it depends on the choice of the random variable  $Y$  in order to have a valid distribution. Figure 1 and Table 1 show that valid GLDs are defined in different domains. Accordingly, the valid PDF of  $T$ - $R$ {GL} in (2.2) is associated with different domains of the GLDs. The valid PDF's and the associated restrictions on the parameters are summarized as cases (i)-(vi) and given in Table 2.

There are good reasons to let the random variable  $Y$  in the  $T$ - $R$ {Y} framework be the quantile function of GLD. First, adding one or more shape parameters may allow the derived distribution to have different shapes as well as being flexible enough to fit a



**Table 1** The conditions on the parameters and the support regions of GLD (p. 39, Karian and Dudewicz (2010))

Region	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	$Q(0)$	$Q(1)$
1	all	$<0$	$<-1$	$>1$	$-\infty$	$\lambda_1 + (1/\lambda_2)$
5	all	$<0$	$\begin{cases} -1 < \lambda_3 < 0, \lambda_4 > 1 \\ \frac{(1-\lambda_3)^{1-\lambda_4}(\lambda_4-1)^{\lambda_4-1}}{(\lambda_4-\lambda_3)^{\lambda_4-\lambda_3}} < \frac{-\lambda_3}{\lambda_4} \end{cases}$		$-\infty$	$\lambda_1 + (1/\lambda_2)$
2	all	$<0$	$>1$	$<-1$	$\lambda_1 - (1/\lambda_2)$	$\infty$
6	all	$<0$	$\begin{cases} \lambda_3 > 1, -1 < \lambda_4 < 0 \\ \frac{(1-\lambda_3)^{1-\lambda_4}(\lambda_3-1)^{\lambda_4-1}}{(\lambda_3-\lambda_4)^{\lambda_4-\lambda_3}} < \frac{-\lambda_4}{\lambda_3} \end{cases}$		$\lambda_1 - (1/\lambda_2)$	$\infty$
3	all	$>0$	$>0$	$>0$	$\lambda_1 - (1/\lambda_2)$	$\lambda_1 + (1/\lambda_2)$
			$=0$	$>0$	$\lambda_1$	$\lambda_1 + (1/\lambda_2)$
			$>0$	$=0$	$\lambda_1 - (1/\lambda_2)$	$\lambda_1$
4	all	$<0$	$<0$	$<0$	$-\infty$	$\infty$
			$=0$	$<0$	$\lambda_1$	$\infty$
			$<0$	$=0$	$-\infty$	$\lambda_1$

wide variety of data sets. Second, the support of GLD covers the three types of intervals: bounded, semi-infinite and whole real line, which places no restrictions in the process of choosing the random variable  $T$  other than those with only bounded, semi-infinite, or whole real line support. In the development of the  $T-R\{Y\}$  framework so far, the random variable  $Y$  has one type of support. Taking  $Y$  to be a generalized lambda random variable leads to three different types of supports for the generator random variable  $T$ . By allowing one to apply different generators,  $T$ , with different supports makes the  $T-R\{\text{generalized lambda}\}$  a desirable method for generating new versatile and broad generalized families of distributions for any given random variable  $R$ . This unique and quite attractive property of the  $T-R\{GL\}$  family motivates us to study this family of distributions. Similar to existing distributions, the interpretations of parameters are often application dependent. We hope that researchers in different disciplines will apply this family of distributions in their respective disciplines with specific interpretations for the parameters of the  $T-R\{GL\}$  distributions.

### Some general properties of $T-R\{\text{generalized lambda}\}$ families of distributions

In this section, we highlight some of the general properties of the  $T-R\{GL\}$  families of distributions.

The following lemma shows the relationship between the random variable  $X$  that follows the  $T-R\{GL\}$  distributions and the random variable  $T$ .

**Table 2** The support of the random variable  $T$  corresponding to the cases (i)-(vi)

Case	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	$Q_Y(u)$	$F_Y(x)$	Support of $T$
(i)	free	$>0$	$>0$	$>0$	$\lambda_1 + \frac{u^{\lambda_3} - (1-u)^{\lambda_4}}{\lambda_2}$	Computed numerically. No closed form.	$[\lambda_1 - \lambda_2^{-1}, \lambda_1 + \lambda_2^{-1}]$
(ii)	1/2	2	$>0$	$>0$	$\frac{1+u^{\lambda_3} - (1-u)^{\lambda_4}}{2}$	Computed numerically. No closed form.	$[0, 1]$
(iii)	free	$<0$	$<0$	$<0$	$\lambda_1 + \frac{u^{\lambda_3} - (1-u)^{\lambda_4}}{\lambda_2}$	Computed numerically. No closed form.	$(-\infty, \infty)$
(iv)	free	$<0$	$=0$	$<0$	$\lambda_1 + \frac{1 - (1-u)^{\lambda_4}}{\lambda_2}$	$1 - (1 - \lambda_2(x - \lambda_1))^{\lambda_4}$	$[\lambda_1, \infty)$
(v)	$=0$	$<0$	$=0$	$<0$	$(1 - (1-u)^{\lambda_4})/\lambda_2$	$1 - (1 - \lambda_2 x)^{1/\lambda_4}$	$[0, \infty)$
(vi)	$=0$	$<0$	$<0$	$=0$	$(u^{\lambda_3} - 1)/\lambda_2$	$(1 + \lambda_2 x)^{1/\lambda_3}$	$(-\infty, 0]$

**Lemma 1 (Transformation):** Let  $T$  be any random variable with PDF  $f_T(x)$ .

- If  $T$  has the support  $[\lambda_1, \infty)$  as in case (iv) in Table 2, then the random variable  $X = Q_R\left(1 - [1 - \lambda_2(T - \lambda_1)]^{1/\lambda_4}\right)$  belongs to the  $T\text{-}R\{GL\}$  families of distributions.
- If  $T$  has the support  $[0, \infty)$  as in case (v) in Table 2, then the random variable  $X = Q_R\left(1 - [1 - \lambda_2 T]^{1/\lambda_4}\right)$  belongs to the  $T\text{-}R\{GL\}$  families of distributions.
- If  $T$  has the support  $(-\infty, 0]$  as in case (vi) in Table 2, then the random variable  $X = Q_R\left([1 + \lambda_2 T]^{1/\lambda_3}\right)$  belongs to the  $T\text{-}R\{GL\}$  families of distributions.

*Proof:* The proof follows directly from the definition of the  $T\text{-}R\{GL\}$  families of distributions in (2.1) and Table 2.

Note that in the first three cases (i)-(iii) in Table 2, the relationships between the random variables  $X$  and  $T$  can be evaluated numerically.

The relation  $F_X(x) = F_T(Q_Y(F_R(x)))$ , where  $T = Q_Y(F_R(X))$  implies  $X = Q_R(F_Y(T))$ , provides an important connection between the random variables  $X$  and  $T$ . For example, one can apply the transformation  $X = Q_R(F_Y(T))$  to generate random samples from  $X$  which has the CDF  $F_X(x)$  by first simulating the random variable  $T$  from the PDF  $f_T(t)$ . Moreover, the  $r^{\text{th}}$  moments (if they exist) of the  $T\text{-}R\{Y\}$  family of distributions can be obtained using  $E_X[X^r] = E_T[Q_R(F_Y(T))]^r$ .

The next lemma makes a connection between the quantile function for the random variable  $X$  which follows the  $T\text{-}R\{GL\}$  families of distributions and the quantile functions of the random variables  $T$  and  $R$ .

**Lemma 2 (Quantiles):** Let  $Q_T(u)$  and  $Q_R(u)$  be the quantile functions of the random variables  $T$  and  $R$ , respectively.

- If  $T$  has the support  $[\lambda_1, \infty)$  as in case (iv) in Table 2, then the quantile function of the random variable  $X$  which follows the  $T\text{-}R\{GL\}$  distributions is  $Q_X(u) = Q_R\left(1 - [1 - \lambda_2(Q_T(u) - \lambda_1)]^{1/\lambda_4}\right)$ .
- If  $T$  has the support  $[0, \infty)$  as in case (v) in Table 2, then the quantile function of the random variable  $X$  which follows the  $T\text{-}R\{GL\}$  distributions is  $Q_X(u) = Q_R\left(1 - [1 - \lambda_2 Q_T(u)]^{1/\lambda_4}\right)$ .
- If  $T$  has the support  $(-\infty, 0]$  as in case (vi) in Table 2, then the quantile function of the random variable  $X$  which follows the  $T\text{-}R\{GL\}$  distributions is  $Q_X(u) = Q_R\left([1 + \lambda_2 Q_T(u)]^{1/\lambda_3}\right)$ .

*Proof:* The results follow directly by solving  $F_X(Q_X(u)) = u$  for  $Q_X(u)$ , where  $F_X(\cdot)$  is the CDF of the random variable  $X$ .

In the literature, some of the quantile functions do not have closed form expressions. For instance, in the first three cases (i)-(iii) in Table 2 the random variable  $X$  has no closed form expression for its quantile function, and it has to be evaluated numerically.

An implicit formula for the mode(s) of the  $T\text{-}R\{GL\}$  families of distributions is presented in the following theorem.

**Theorem 1** *The mode(s) of the  $T\text{-}R\{GL\}$  families of distributions in (2.2) can be obtained from the roots of the equation*

$$\frac{f_R'(x)}{f_R^2(x)} = -Q_Y'(F_R(x)) \left[ \frac{Q_Y''(F_R(x))}{(Q_Y'(F_R(x)))^2} + \frac{f_T'(\lambda_1 + \lambda_2^{-1}[F_R^{\lambda_3}(x) - \bar{F}_R^{\lambda_4}(x)])}{f_T(\lambda_1 + \lambda_2^{-1}[F_R^{\lambda_3}(x) - \bar{F}_R^{\lambda_4}(x)])} \right], \quad (3.1)$$

where  $\bar{F}_R(x) = 1 - F_R(x)$  is the survival function of the random variable  $R$  with PDF  $f_R(x)$ ,  $Q_Y'(F_R(x)) = \lambda_2^{-1}[\lambda_3 F_R^{\lambda_3-1}(x) + \lambda_4 \bar{F}_R^{\lambda_4-1}(x)]$ , and  $Q_Y''(F_R(x)) = \lambda_2^{-1}[\lambda_3(\lambda_3-1)F_R^{\lambda_3-2}(x) - \lambda_4(\lambda_4-1)\bar{F}_R^{\lambda_4-2}(x)]$ .

*Proof:* One can show the result in (3.1) by setting the first derivative of  $f_X(x)$  in (2.2) to 0.

Note that the result in Theorem 1 does not necessarily guarantee that the mode is unique. It is possible that some members of the  $T-R\{GL\}$  families of distributions have more than one mode. For example, the uniform-exponential $\{GL\}$  distribution in section 4 is bimodal, depending on the values of its shape parameters.

Shannon (1948) defined the entropy of a random variable  $X$  as a measure of uncertainty variation by  $\eta_X = E_X[-\log(f_X(X))]$ . The next theorem defines the Shannon's entropy of the random variable  $X$  that follows the  $T-R\{GL\}$  families of distributions with PDF  $f_X(x)$  in terms of the Shannon's entropy of the random variable  $T$  with PDF  $f_T(x)$ .

**Theorem 2** *The Shannon's entropy for the  $T-R\{GL\}$  families of distributions is given by*

$$\eta_X = \eta_T + E_T[\log f_Y(T)] + E_T[\log q_R(F_Y(T))], \quad (3.2)$$

where  $\eta_T$  is the Shannon's entropy of the random variable  $T$  with PDF  $f_T(x)$  and  $q_R(\cdot)$  is the quantile density function of the random variable  $R$ .

*Proof:* By the definition of Shannon's entropy,

$$\begin{aligned} \eta_X = E_X[-\log(f_X(X))] &= E_X \left[ \log \left( \frac{\lambda_3 F_R^{\lambda_3-1}(X) + \lambda_4 (1-F_R(X))^{\lambda_4-1}}{\lambda_2} \right)^{-1} \right] \\ &+ E_X[-\log f_R(X)] + E_X \left[ -\log f_T \left( \lambda_1 + \frac{F_R^{\lambda_3}(X) - (1-F_R(X))^{\lambda_4}}{\lambda_2} \right) \right]. \end{aligned} \quad (3.3)$$

The random variable  $T = Q_Y(F_R(X)) = \left( \lambda_1 + \frac{F_R^{\lambda_3}(X) - (1-F_R(X))^{\lambda_4}}{\lambda_2} \right)$ , or equivalently,  $X = Q_R(F_Y(T))$ . This implies that

$$E_X \left[ -\log f_T \left( \lambda_1 + \frac{F_R^{\lambda_3}(X) - (1-F_R(X))^{\lambda_4}}{\lambda_2} \right) \right] = E_T[-\log f_T(T)] = \eta_T, \quad (3.4)$$

$$\begin{aligned} E_X \left[ \log \left( \frac{\lambda_3 F_R^{\lambda_3-1}(X) + \lambda_4 (1-F_R(X))^{\lambda_4-1}}{\lambda_2} \right)^{-1} \right] &= E_X[\log f_Y(Q_Y(F_R(X)))] \\ &= E_T[\log f_Y(T)], \end{aligned} \quad (3.5)$$

$$E_X[-\log f_R(X)] = E_T[-\log f_R(Q_R(F_Y(T)))] = E_T[\log q_R(F_Y(T))]. \quad (3.6)$$

Substituting (3.4) through (3.6) into (3.3) gives (3.2).

The result in (3.2) can be used to find the Shannon entropies of  $T-R\{GL\}$  for the cases (iv)-(vi) since  $F_Y(x)$  and  $f_Y(x)$  are in closed form. For example, consider case (v) in Table 2 and let the random variable  $T$  follow the Lomax distribution with PDF  $f_T(t) = (\alpha/\beta)[1$



$+(t/\beta)]^{-(\alpha+1)}$ ,  $t \geq 0$ ,  $\alpha, \beta > 0$ , and let the random variable  $R$  follow the standard exponential distribution with PDF  $f_R(x) = e^{-x}$  and quantile function  $Q_R(u) = -\log(1-u)$ . The Shannon entropy of the Lomax distribution is given by

$$\begin{aligned}\eta_T &= ((\alpha+1)/\alpha) - \log(\alpha/\beta). \\ E_T[\log f_Y(T)] &= E_T \left[ \log \left( (\lambda_2/\lambda_4)(1-\lambda_2 t)^{(1/\lambda_4)-1} \right) \right] \\ &= (\lambda_2\beta(\lambda_4-1)/\lambda_4) \Gamma(\alpha) {}_2\tilde{F}_1(1, 1; 1+\alpha; 1+\lambda_2\beta) + \log(\lambda_2/\lambda_4),\end{aligned}$$

where  $\Gamma(\alpha)$  is the gamma function and  ${}_2\tilde{F}_1(1, 1; 1+\alpha; 1+\lambda_2\beta)$  is the regularized hypergeometric function.

$$\begin{aligned}E_T[\log q_R(F_Y(T))] &= E_T \left[ -\log(1-\lambda_2 t)^{1/\lambda_4} \right] \\ &= (-(-\lambda_2\beta)^\alpha/\lambda_4) \Gamma(\alpha) {}_2\tilde{F}_1(\alpha, \alpha; 1+\alpha; 1+\lambda_2\beta).\end{aligned}$$

Therefore, the Shannon entropy of the Lomax-exponential{GL} is given by

$$\begin{aligned}\eta_X &= \left( \frac{\alpha+1}{\alpha} \right) - \log\left(\frac{\alpha}{\beta}\right) + \log\left(\frac{\lambda_2}{\lambda_4}\right) + \left( \frac{\lambda_2\beta(\lambda_4-1)}{\lambda_4} \right) \Gamma(\alpha) {}_2\tilde{F}_1(1, 1; 1+\alpha; 1+\lambda_2\beta) \\ &\quad - \left( \frac{(-\lambda_2\beta)^\alpha}{\lambda_4} \right) \Gamma(\alpha) {}_2\tilde{F}_1(\alpha, \alpha; 1+\alpha; 1+\lambda_2\beta).\end{aligned}$$

**Moments** In general, the non-central moments (if they exist) for the  $T-R\{GL\}$  family of distributions can be obtained by using  $E_X[X^n] = E_T[Q_R(F_Y(T))]^n = \int_T [Q_R(F_Y(t))]^n f_T(t) dt$ . However,  $F_Y(\cdot)$  in the first three cases in Table 2 have no closed form and one may use  $E_X[X^n] = \int_X x^n f_X(x) dx$  to find the moments. For the cases (iv)-(vi) in Table 2, the quantile function  $Q_X(u)$  is in closed form, so the  $n^{th}$  moment of the random variable  $X$  may be obtained from  $E_X[X^n] = \int_0^1 [Q_X(u)]^n du$ . The following Theorem 3 derives an approximation for computing the  $n^{th}$  moment of a member, Uniform-R{GL} of case (i).

**Theorem 3** Let  $T$  be a random variable that follows the uniform distribution with support as in case (i) in Table 2, and let  $X$  be a random variable having the Uniform-R{GL} PDF. The  $n^{th}$  moment of  $X$  can be expressed in terms of the quantile function  $Q_R(u)$  and is given by:

$$E_X(X^n) = \frac{1}{2} \left[ \lambda_3 \int_0^1 u^{\lambda_3-1} [Q_R(u)]^n du + \lambda_4 \sum_{k=0}^{\infty} (-1)^k \binom{\lambda_4-1}{k} \int_0^1 u^k [Q_R(u)]^n du \right]. \quad (3.7)$$

*Proof:*

$$E_X[X^n] = \int_{-\infty}^{\infty} x^n f_X(x) dx = \frac{1}{2} \int_{-\infty}^{\infty} x^n f_R(x) \left[ \lambda_3 F_R^{\lambda_3-1}(x) + \lambda_4 (1-F_R(x))^{\lambda_4-1} \right] dx. \quad (3.8)$$

Using the substitution  $u = F_R(x)$ , then equation (3.8) can be written as

$$E_X[X^n] = \frac{1}{2} \left[ \int_0^1 \lambda_3 u^{\lambda_3-1} [Q_R(u)]^n du + \int_0^1 \lambda_4 (1-u)^{\lambda_4-1} [Q_R(u)]^n du \right]. \quad (3.9)$$

For  $\lambda_4 > 0$ , a real non-integer, then by the generalized binomial theorem, we have



$$(1-u)^{\lambda_4-1} = \sum_{k=0}^{\infty} (-1)^k \binom{\lambda_4-1}{k} u^k, \quad (3.10)$$

and if  $\lambda_4 > 1$  and it is an integer, then the upper summation stops at  $\lambda_4 - 1$ . The series in (3.10) converges uniformly on  $(0, 1)$  since  $0 < u < 1$ . By the dominated convergence theorem of series, the second integral in (3.9) can be integrated term by term. This completes the proof.

**Remark 1:** Let  $r = \lambda_3 - 1$  and  $s = \lambda_4 - 1$ , the  $n^{\text{th}}$  moment of  $Uniform-R\{GL\}$  in (3.9) can be expressed as the probability weighted moments of random variable  $R$ ,  $E_X[X^n] = \frac{1}{2} \{\lambda_3 M_{n,r,0} + \lambda_4 M_{n,0,s}\}$ , where  $M_{n,r,s} = E_R[X^n F_R^r(X) \bar{F}_R^s(X)]$  is the probability weighted moments of the random variable  $R$  of order  $(n, r, s)$ .

**Remark 2:** If the random variable  $R$  has finite  $n^{\text{th}}$  moment (i.e.  $[Q_R(u)]^n$  is integrable), then by applying Cauchy-Schwarz inequality to (3.9), the  $n^{\text{th}}$  moment of  $X$  is bounded by the upper bound

$$\begin{aligned} E_X[X^n] &\leq \frac{\lambda_3}{2} \left\{ \int_0^1 [Q_R(u)]^{2n} du \int_0^1 u^{2\lambda_3-2} du \right\}^{1/2} \\ &\quad + \frac{\lambda_4}{2} \left\{ \int_0^1 [Q_R(u)]^{2n} du \int_0^1 (1-u)^{2\lambda_4-2} du \right\}^{1/2} \\ &= \frac{1}{2} [E_R(X^{2n})]^{1/2} \left[ \left\{ \frac{\lambda_3^2}{2\lambda_3-1} \right\}^{1/2} + \left\{ \frac{\lambda_4^2}{2\lambda_4-1} \right\}^{1/2} \right]. \end{aligned}$$

### Some examples of $T-R\{GL\}$ families of distributions with different $T$ and $R$ distributions

In this section different  $T$  and  $R$  distributions are used to generate various members of the  $T-R\{GL\}$  families of distributions. We present four new  $T-R\{GL\}$  distributions namely, uniform-exponential{generalized lambda}, normal-uniform{generalized lambda}, Pareto-Weibull{generalized lambda} and log-logistic-logistic{generalized lambda}.

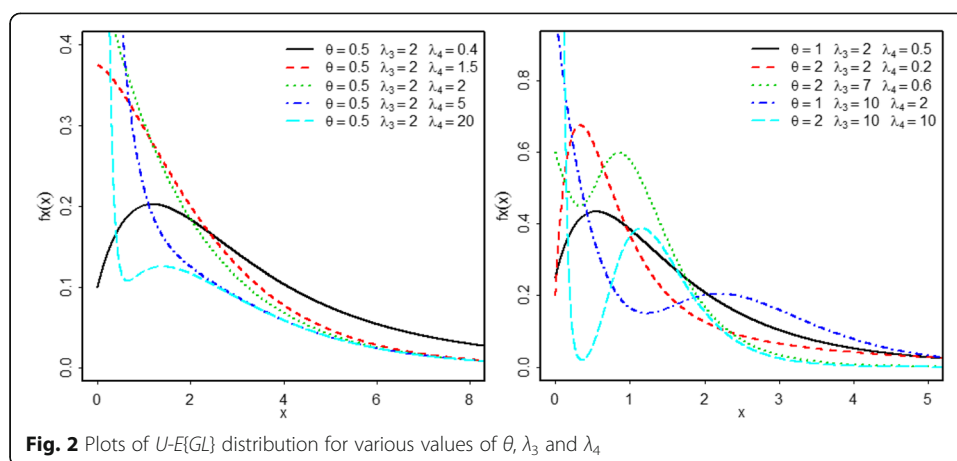
#### The uniform-exponential{generalized lambda} distribution

Consider case (i) in Table 2 and let  $T$  be a random variable that follows the uniform distribution. If a random variable  $R$  follows the exponential distribution with a rate parameter  $\theta > 0$  and CDF  $F_R(x) = 1 - e^{-\theta x}$ ,  $x \geq 0$ , then the CDF and PDF of the uniform-exponential{generalized lambda}(U-E{GL}) distribution are given by, respectively:

$$F_X(x) = \frac{1}{2} \left[ 1 + (1 - e^{-\theta x})^{\lambda_3} - (e^{-\theta x})^{\lambda_4} \right], \quad (4.1)$$

$$f_X(x) = \frac{1}{2} \theta e^{-\theta x} \left[ \lambda_3 (1 - e^{-\theta x})^{\lambda_3-1} + \lambda_4 (e^{-\theta x})^{\lambda_4-1} \right], x \geq 0, \theta, \lambda_3, \lambda_4 > 0. \quad (4.2)$$

The parameters  $\lambda_3$  and  $\lambda_4$  are shape parameters (see Figure 2). It is clear that the U-E{GL} distribution reduces to the exponential distribution when  $\lambda_3 = \lambda_4 = 1$  or  $\lambda_3 = \lambda_4 = 2$ . Various shapes of the U-E{GL} distribution for different values of the parameters  $\theta, \lambda_3, \lambda_4$  are provided in Fig. 2. These graphs indicate that the U-E{GL} distribution can be monotonically decreasing (reversed J-shape), right skewed, unimodal or bimodal.



### The normal-uniform{generalized lambda} distribution

Consider case (iii) in Table 2 and let  $T$  be a standard normal random variable with PDF  $f_T(x) = \phi(x) = \frac{1}{\sqrt{2\pi}} \exp(-x^2/2)$  and CDF  $F_T(x) = \Phi(x)$ . Let  $R$  be a random variable that follows the uniform distribution with parameters  $a, b$  and CDF  $F_R(x) = \frac{x-a}{b-a}$ ,  $x \in [a, b]$ ,  $-\infty < a < b < \infty$ . If we set the location parameter  $\lambda_1 = 0$  in case (iii) in Table 2, then the CDF and PDF of the normal-uniform{generalized lambda} ( $N-U\{GL\}$ ) distribution are given by, respectively:

$$F_X(x) = \Phi \left[ \frac{1}{\lambda_2} \left( \left( \frac{x-a}{b-a} \right)^{\lambda_3} - \left( 1 - \frac{x-a}{b-a} \right)^{\lambda_4} \right) \right], \quad (4.3)$$

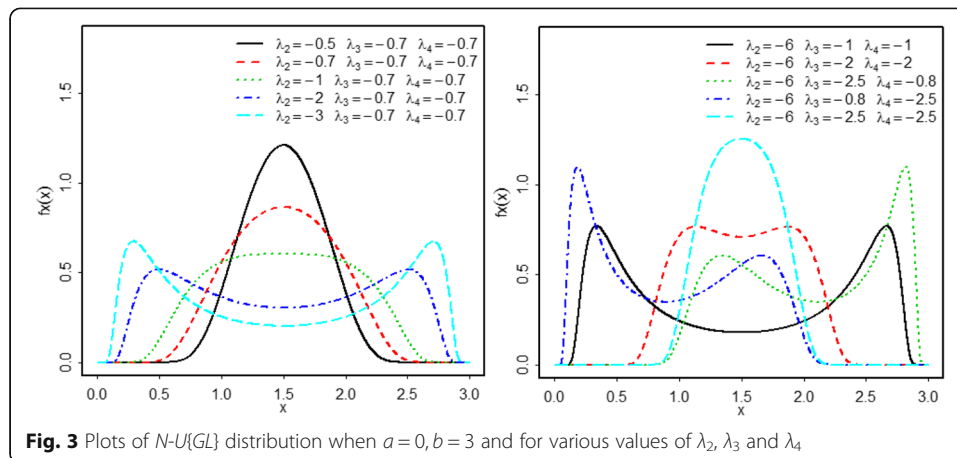
$$f_X(x) = \frac{1}{\lambda_2(b-a)\sqrt{2\pi}} \left( \lambda_3 \left( \frac{x-a}{b-a} \right)^{\lambda_3-1} + \lambda_4 \left( 1 - \frac{x-a}{b-a} \right)^{\lambda_4-1} \right) \times \exp \left[ -\frac{1}{2\lambda_2^2} \left( \left( \frac{x-a}{b-a} \right)^{\lambda_3} - \left( 1 - \frac{x-a}{b-a} \right)^{\lambda_4} \right)^2 \right], x \in [a, b], -\infty < a < b < \infty, \lambda_2, \lambda_3, \lambda_4 < 0. \quad (4.4)$$

The parameters  $\lambda_2$ ,  $\lambda_3$ , and  $\lambda_4$  are shape parameters (see Fig. 3). Plots of  $N-U\{GL\}$  distribution when  $a = 0, b = 3$  and various values of  $\lambda_2, \lambda_3$  and  $\lambda_4$  are given in Fig. 3. The plots show that the  $N-U\{GL\}$  distribution can be symmetric, left skewed or right skewed and it can be either unimodal or bimodal.

### The Pareto-Weibull{generalized lambda} distribution:

Consider case (iv) in Table 2 and let the random variable  $T$  follow the Pareto distribution with CDF  $F_T(x) = 1 - (\lambda_1/x)^s$ ,  $x \geq \lambda_1, \lambda_1 > 0, s > 0$ , and take the random variable  $R$  to be the Weibull distribution with CDF  $F_R(x) = 1 - e^{-(x/\gamma)^c}$ ,  $x \geq 0, \gamma, c > 0$ . If we replace  $\lambda_1\lambda_2$  by  $-\beta, \beta > 0$  and let  $\lambda_4 = -1$ , then the CDF and PDF of the Pareto-Weibull{generalized lambda} ( $P-W\{GL\}$ ) distribution are respectively given by

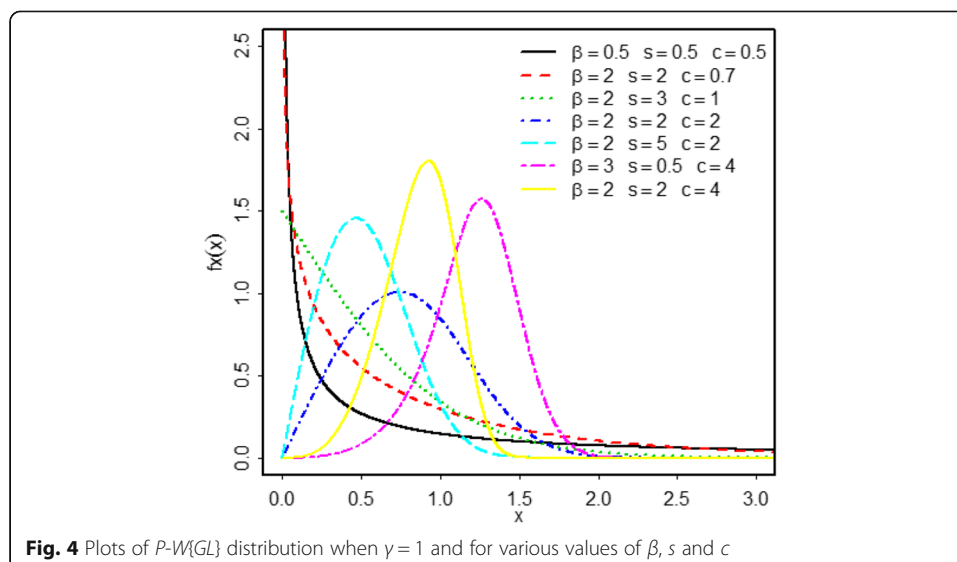
$$F_X(x) = 1 - \left( \beta / \left[ \beta - 1 + e^{(x/\gamma)^c} \right] \right)^s, \quad (4.5)$$



$$f_X(x) = sc\beta^s \gamma^{-1} (x/\gamma)^{c-1} e^{(x/\gamma)^c} \left( \beta - 1 + e^{(x/\gamma)^c} \right)^{-s-1}, x \geq 0, s, \beta, c, \gamma > 0. \quad (4.6)$$

When  $s = \beta = 1$ , the  $P-W\{GL\}$  distribution reduces to the Weibull distribution, which has Rayleigh and exponential distributions as special cases. It is worth mentioning that by setting  $c = 1$  in (4.6), the  $P-W\{GL\}$  distribution reduces to the Pareto-exponential{generalized lambda} distribution, which is called in the literature the gamma/Gompertz distribution. Thus, the  $P-W\{GL\}$  distribution is a generalization of the gamma/Gompertz distribution, which was derived using a different approach (Bemmaor and Glady, 2012).

Figure 4 illustrates some possible shapes of the density function in (4.6) when  $\gamma = 1$  and for selected parameter values. The graphs in Fig. 4 indicate that the  $P-W\{GL\}$  distribution can be skewed to the left, skewed to the right or monotonically decreasing (reversed J-shape). Different combinations of the values of the parameters were tried but in all cases the graph of  $P-W\{GL\}$  distribution appears to be unimodal.



**The log-logistic-logistic{generalized lambda} distribution:**

Consider case (v) in Table 2 and let  $T$  be a random variable that follows the standard log-logistic distribution with CDF  $F_T(x) = \frac{x}{1+x}$ ,  $x \geq 0$ . Take the random variable  $R$  to be the logistic distribution with CDF  $F_R(x) = \frac{1}{2} + \frac{1}{2} \tanh\left(\frac{x-\mu}{2\sigma}\right) = \frac{e^{(x-\mu)/\sigma}}{1+e^{(x-\mu)/\sigma}}$ ,  $-\infty < x < \infty$ , where the location parameter  $\mu \in \mathbb{R}$  and the scale parameter  $\sigma > 0$ . If we replace  $\lambda_2$  by  $-c$ ,  $c > 0$  and  $\lambda_4$  by  $-w$ ,  $w > 0$ , then the CDF and PDF of the log-logistic-logistic{generalized lambda} ( $LL-L\{GL\}$ ) distribution are respectively given by

$$F_X(x) = \frac{1-2^w(1-\tanh[(x-\mu)/2\sigma])^{-w}}{1-c-2^w(1-\tanh[(x-\mu)/2\sigma])^{-w}}, \quad (4.7)$$

$$f_X(x) = \frac{cw(2-2\tanh[(x-\mu)/2\sigma])^{w-1} \operatorname{sech}^2[(x-\mu)/2\sigma]}{\sigma\{2^w + (c-1)(1-\tanh[(x-\mu)/2\sigma])^w\}^2},$$

$$-\infty < x < \infty, \mu \in \mathbb{R}, \sigma, c, w > 0. \quad (4.8)$$

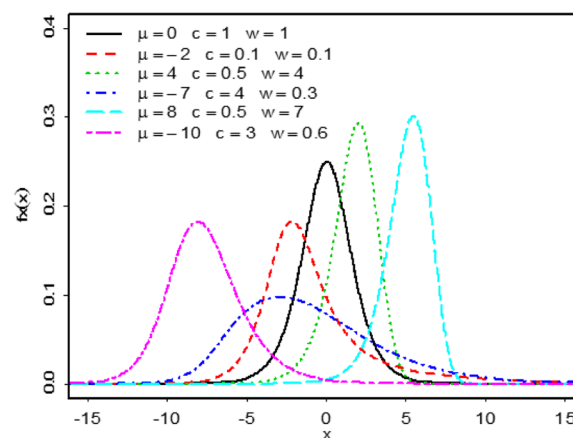
Note that the  $LL-L\{GL\}$  distribution reduces to the logistic distribution when  $c = w = 1$ . In Fig. 5, various graphs of the density in (4.8) when  $\sigma = 1$  and for various values of the parameters  $\mu$ ,  $c$  and  $w$  are provided. The plots show that the  $LL-L\{GL\}$  distribution can be symmetric, left skewed or right skewed. The graph of  $LL-L\{GL\}$  distribution appears to be unimodal from trying many different combinations of the parameter values.

**Parameter estimation and simulation for  $U-E\{GL\}$  distribution**

In this section, we use the method of maximum likelihood to address the parameter estimation and conduct a simulation to examine the performance of this method. Let  $x_1, x_2, \dots, x_n$  be a random sample of size  $n$  from a  $U-E\{GL\}$  distribution defined in equation (4.2), then the log-likelihood function is given by

$$\ell(\theta, \lambda_3, \lambda_4) = n \log(\theta/2) - \theta \sum_{i=1}^n x_i + \sum_{i=1}^n \log[\lambda_3(1-e^{-\theta x_i})^{\lambda_3-1} + \lambda_4(e^{-\theta x_i})^{\lambda_4-1}]. \quad (5.1)$$

To measure the performance of the MLEs, we conduct a simulation study to evaluate the MLEs in terms of the bias (actual – estimate) and standard deviation of the parameter estimates for different parameter combinations and sample sizes.



**Fig. 5** Plots of  $LL-L\{GL\}$  distribution when  $\sigma = 1$  and for various values of  $\mu$ ,  $w$  and  $c$

The  $U-E\{GL\}$  is a generalization of the exponential distribution. It reduces to the exponential distribution with mean  $1/\theta$  when  $\lambda_3 = \lambda_4 = 1$  or  $\lambda_3 = \lambda_4 = 2$ . In the simulation study, we take the initial estimates of parameters  $\lambda_3$  and  $\lambda_4$  to be 1 and the initial estimate of parameter  $\theta$  to be the MLE of  $\theta$  by taking the simulated data to have an exponential distribution. We obtain a random sample  $x_1, x_2, \dots, x_n$  of size  $n$  from a  $U-E\{GL\}$  distribution by first generating a random sample  $t_1, t_2, \dots, t_n$  from standard uniform distribution and then transforming it to  $U-E\{GL\}$  using the relationship  $X = Q_R(F_Y(T)) = -(1/\theta) \log(1 - F_Y(T))$ , where  $F_Y(T)$  is computed numerically in SAS for different parameter combinations of  $\lambda_3$  and  $\lambda_4$ .

In this simulation, five sample sizes are considered ( $n = 50, 100, 250, 500, 1000$ ). The NLMIXED procedure in SAS is used to maximize the log-likelihood function in Equation (5.1). We consider different parameter combinations to cover different shapes of the distribution, including monotonically decreasing, right skewed, unimodal or bimodal. The parameter combinations considered are  $(\lambda_3, \lambda_4, \theta) = \{(0.8, 0.6, 0.5), (1, 2, 2), (2, 0.8, 1), (3, 0.5, 3), (4, 0.7, 2)\}$ . The MLEs of the parameters  $\lambda_3, \lambda_4$  and  $\theta$  are computed and the process is repeated 1000 times for each sample size and each parameter

**Table 3** Average bias (standard deviation) for the MLEs

Actual values				Average bias			Mode(s)
$\lambda_3$	$\lambda_4$	$\theta$	$n$	$\lambda_3$	$\lambda_4$	$\theta$	
0.8	0.6	0.5	50	-0.0120 (0.1765)	0.0332 (0.1666)	-0.0405 (0.1156)	Reversed J-shape with one mode at $x = 0$
			100	-0.0415 (0.1498)	0.0382 (0.1565)	-0.0495 (0.1092)	
			250	-0.0357 (0.1100)	0.0294 (0.1348)	-0.0385 (0.0944)	
			500	-0.0240 (0.0828)	0.0127 (0.1216)	-0.0238 (0.0807)	
			1000	-0.0130 (0.0591)	-0.0000 (0.1081)	-0.0121 (0.0658)	
1	2	2	50	-0.0289 (0.2693)	0.1108 (0.5991)	-0.1246 (0.4180)	Unimodal at $x = 0$
			100	-0.0336 (0.2405)	0.0757 (0.5425)	-0.1135 (0.3728)	
			250	-0.0345 (0.2089)	0.0496 (0.4811)	-0.0942 (0.2870)	
			500	-0.0284 (0.1839)	0.0436 (0.4072)	-0.0713 (0.2122)	
			1000	-0.0469 (0.1769)	-0.0014 (0.3396)	-0.0549 (0.1534)	
2	0.8	1	50	0.1337 (0.5625)	0.0237 (0.2040)	-0.0222 (0.2059)	Right-skewed with one mode at $x > 0$
			100	0.0073 (0.5240)	0.0424 (0.1978)	-0.0504 (0.1845)	
			250	-0.1297 (0.4199)	0.0381 (0.1895)	-0.0684 (0.1531)	
			500	-0.1440 (0.3393)	0.0170 (0.1692)	-0.0506 (0.1224)	
			1000	-0.1121 (0.2746)	-0.0019 (0.1484)	-0.0281 (0.0925)	
3	0.5	3	50	-0.0571 (0.7754)	-0.0030 (0.1311)	-0.0908 (0.5682)	Right-skewed with one mode at $x > 0$
			100	-0.0851 (0.7237)	-0.0018 (0.1190)	-0.0870 (0.5159)	
			250	-0.1058 (0.6366)	-0.0054 (0.0990)	-0.0529 (0.4076)	
			500	-0.1073 (0.5306)	-0.0088 (0.0836)	-0.0309 (0.3365)	
			1000	-0.0796 (0.4066)	-0.0034 (0.0632)	-0.0256 (0.2526)	
4	0.7	2	50	0.0921 (1.0819)	0.0284 (0.1753)	-0.0651 (0.3435)	Bimodal with two modes at $x = 0$ and $x > 0$
			100	-0.0580 (1.0216)	0.0110 (0.1722)	-0.0609 (0.3031)	
			250	-0.1211 (0.8777)	-0.0007 (0.1428)	-0.0393 (0.2348)	
			500	-0.1794 (0.7393)	-0.0030 (0.1183)	-0.0360 (0.1914)	
			1000	-0.1301 (0.5861)	0.0010 (0.0868)	-0.0256 (0.1454)	

combination. The average bias and standard deviation of the MLEs are computed and the results are presented in Table 3.

The simulation results show that the maximum likelihood estimation method performs quite well in estimating the  $U-E\{GL\}$  distribution parameters. It is observed that the standard deviations of the MLEs decrease as the sample size increases and the average biases of the MLEs are somewhat small and seem to be reasonable. As the sample size increases, it is also noticed that the average biases do not show a clear decreasing or increasing pattern. In addition, it appears that the MLEs of  $\theta$  tend to be overestimated. In conclusion, the simulation results suggest that the maximum likelihood estimation method is appropriate and it can be used to estimate the parameters of the  $U-E\{GL\}$  distribution.

### Applications

In order to illustrate the flexibility of the members of  $T-R\{GL\}$  families of distributions in fitting real data, we present some applications of the  $U-E\{GL\}$  distribution using two different real data sets. We use the method of maximum likelihood to estimate the parameters of the fitted distribution. The fits of the  $U-E\{GL\}$  distribution are compared to other distributions based on the log-likelihood value, the Kolmogorov-Smirnov (K-S) statistic, the  $p$ -value of (K-S) statistic and the Akaike information criterion (AIC).

#### *Remission times of bladder cancer patients:*

The data in Table 4 represents the remission times (in months) of a random sample of 128 bladder cancer patients. This data was previously used by Zea et al. (2012) to compare the fits of the five-parameter beta exponentiated Pareto (BEP) distribution and other sub-models such as the beta-Pareto (BP) distribution. The data is also recently studied and analyzed by Almheidat et al. (2015) to show the flexibility of the four-parameter Cauchy-Weibull{logistic} (C-W{L}) distribution in fitting real data. The data is unimodal and is highly skewed to the right (skewness = 3.286 and kurtosis = 18.483).

We apply the  $U-E\{GL\}$  distribution to fit the same data. The MLEs (with corresponding standard errors) of the parameters, the log-likelihood, the AIC, the K-S statistic and the  $p$ -value of (K-S) statistic for the  $U-E\{GL\}$  distribution and the other fitted distributions are provided in Table 5. The results in Table 5 for the C-W{L} distribution are taken from Almheidat et al.

**Table 4** Remission times (in months) of bladder cancer patients

0.080	0.200	0.400	0.500	0.510	0.810	0.900	1.050	1.190	1.260	1.350	1.400
1.460	1.760	2.020	2.020	2.070	2.090	2.230	2.260	2.460	2.540	2.620	2.640
2.690	2.690	2.750	2.830	2.870	3.020	3.250	3.310	3.360	3.360	3.480	3.520
3.570	3.640	3.700	3.820	3.880	4.180	4.230	4.260	4.330	4.340	4.400	4.500
4.510	4.870	4.980	5.060	5.090	5.170	5.320	5.320	5.340	5.410	5.410	5.490
5.620	5.710	5.850	6.250	6.540	6.760	6.930	6.940	6.970	7.090	7.260	7.280
7.320	7.390	7.590	7.620	7.630	7.660	7.870	7.930	8.260	8.370	8.530	8.650
8.660	9.020	9.220	9.470	9.740	10.06	10.34	10.66	10.75	11.25	11.64	11.79
11.98	12.02	12.03	12.07	12.63	13.11	13.29	13.80	14.24	14.76	14.77	14.83
15.96	16.62	17.12	17.14	17.36	18.10	19.13	20.28	21.73	22.69	23.63	25.74
25.82	26.31	32.15	34.26	36.66	43.01	46.12	79.05				

**Table 5** MLEs for remission times of bladder cancer patient's data (standard errors in parentheses)

Distribution	<sup>a</sup> BP	<sup>a</sup> BEP	<sup>b</sup> Four-parameter C-W{L}	U-E{GL}
Parameter estimates	$a = 4.805$ (0.055) $b = 100.502$ (0.251) $k = 0.011$ (0.001) $\beta = 0.080$	$a = 0.348$ (0.097) $b = 159831$ (183.7501) $k = 0.051$ (0.019) $\beta = 0.080$ $\alpha = 8.612$ (2.093)	$\alpha = -2.3040$ (1.0937) $\beta = 2.0205$ (0.4585) $k = 3.0673$ (0.7319) $\lambda = 12.663$ (2.6326)	$\theta = 0.2757$ (0.0665) $\lambda_3 = 2.5904$ (0.9285) $\lambda_4 = 0.2894$ (0.0858)
Log-likelihood	-480.446	-432.41	-416.0965	-409.45
AIC	968.893	874.819	840.2	824.9
K-S statistic ( <i>p</i> -value)	0.217 (1.105E-5)	0.142 (0.0121)	0.06672 (0.6189)	0.02876 (0.9999)

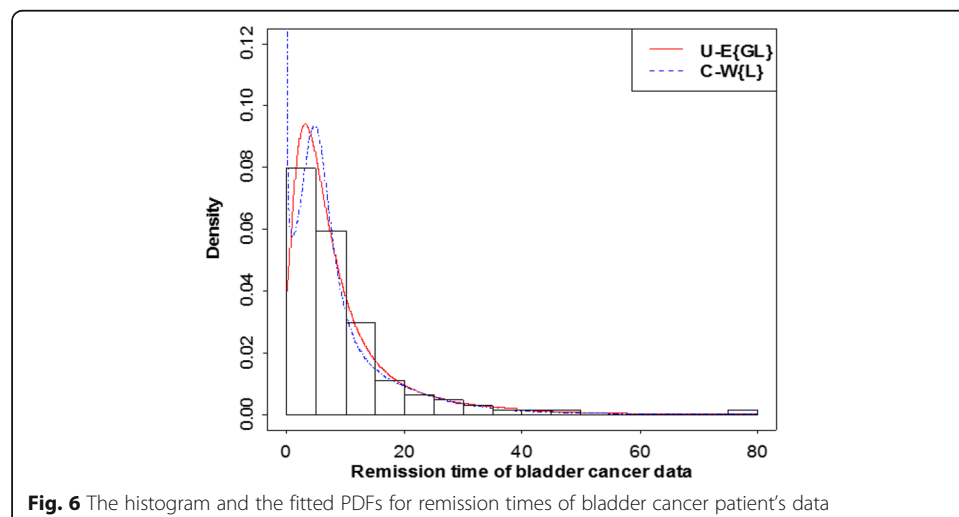
<sup>a</sup> MLEs, log likelihood, K-S (*p*-value), and AIC are from Zea et al. (2012).<sup>b</sup> MLEs, log likelihood, K-S (*p*-value), and AIC are from Almheidat et al. (2015).

(2015) whereas the results for the BP and BEP distributions are obtained from Zea et al. (2012). The other results in Table 5 are obtained by using the SAS (PROC NLMIXED) software.

The values in Table 5 indicate that the three-parameter  $U-E\{GL\}$  distribution outperforms the other three distributions and provide the best fit (based on the log-likelihood, the AIC, the K-S statistic and the *p*-value of K-S statistic) to the remission times of bladder cancer patient's data. This application suggests that the  $U-E\{GL\}$  distribution can fit very well highly right skewed data with long tail. Figure 6 displays the histogram of the data and the density functions of the fitted distributions that provide adequate fits to the data.

#### The Wheaton river data

In Table 6, the data set with  $n = 72$  is on exceedances (in  $m^3/s$ ) of flood peaks of the Wheaton River, Yukon Territory, Canada, for the years 1958 to 1984, rounded to one decimal place. This data was considered by Akinsete et al. (2008) to illustrate the application of the beta-Pareto (BP) distribution. The data is also used by Alshawarbeh et al. (2013) and fitted to the beta-Cauchy (BC) distribution. Recently, Al-Aqtash et al.





**Table 6** Exceedances of the Wheaton River data.

1.7	2.2	14.4	1.1	0.4	20.6	5.3	0.7	1.9	13.0	12.0	9.3
1.4	18.7	8.5	25.5	11.6	14.1	22.1	1.1	2.5	14.4	1.7	37.6
0.6	2.2	39.0	0.3	15.0	11.0	7.3	22.9	1.7	0.1	1.1	0.6
9.0	1.7	7.0	20.1	0.4	2.8	14.1	9.9	10.4	10.7	30.0	3.6
5.6	30.8	13.3	4.2	25.5	3.4	11.9	21.5	27.6	36.4	2.7	64.0
1.5	2.5	27.4	1.0	27.1	20.2	16.8	5.3	9.7	27.5	2.5	27.0

(2015) used the data in an application of the Gumbel-Weibull (GW) distribution. The data is skewed to the right (skewness = 1.5 and kurtosis = 3.19).

To show the applicability of the  $U-E\{GL\}$  distribution, the distribution is applied to fit the data set and the results are compared with the BP distribution, the BC distribution and GW distribution. The maximum likelihood estimates, the log-likelihood value, the AIC, the K-S test statistic, and the  $p$ -value of the K-S statistics for the fitted distributions are presented in Table 7. The MLEs, the values of the K-S statistic and its corresponding  $p$ -value for the BP distribution, the BC distribution, and GW distribution in Table 7 are obtained from Al-Aqtash et al. (2015).

Based on the  $p$ -value of K-S statistic, the results in Table 7 show that the  $U-E\{GL\}$  distribution and GW distribution are superior to the other two distributions. Among the fitted distributions, it seems that the  $U-E\{GL\}$  provides the best fit to the data with smallest AIC, K-S statistic, and largest log-likelihood value. Figure 7 provides the histogram of the data and the density functions of the  $U-E\{GL\}$  and GW distributions.

From sub-sections 6.1 and 6.2, we observe that the  $U-E\{GL\}$  distribution seems to be very competitive to other distributions in fitting highly right skewed data with long tail.

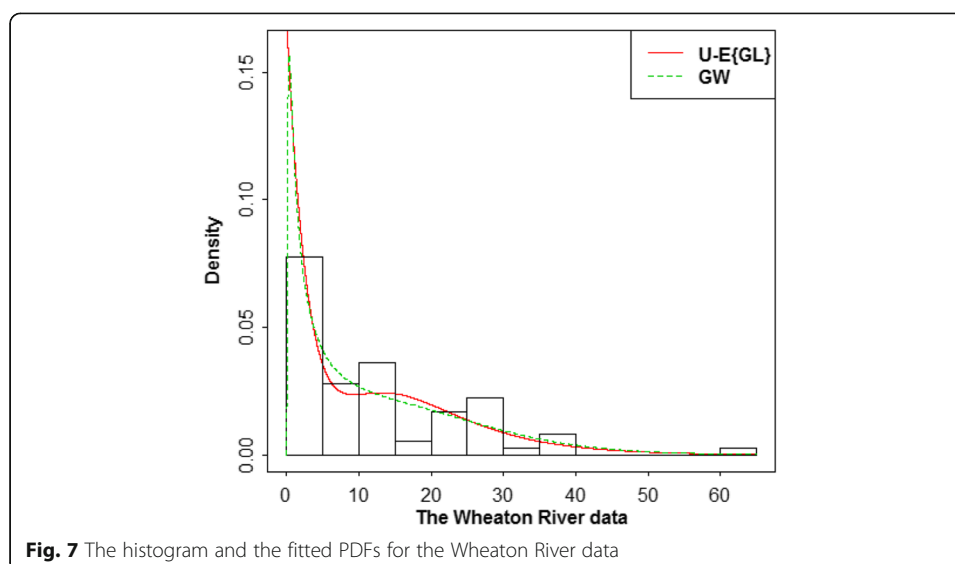
## Summary

In this article, the class of  $T-R\{\text{generalized } \lambda\}$  families of distributions based on the quantile of generalized lambda distribution is introduced using the  $T-R\{Y\}$  framework. One of the advantages for letting the random variable  $Y$  in the  $T-R\{Y\}$  framework to be the quantile function of GLD is that the generalized lambda random variable leads to three different types of support as shown in sub-section 2.2. For this reason, different families of the  $T-R\{\text{generalized } \lambda\}$  distributions can be derived based on the choices of the random variables  $T$  and  $R$ . Some general properties of  $T-R\{\text{generalized } \lambda\}$  families of distributions are studied.

**Table 7** Parameter estimates for Wheaton river data (standard errors in parentheses)

Distribution	<sup>a</sup> BP	<sup>a</sup> BC	<sup>a</sup> GW	U-E{GL}
Parameter estimates	$a = 7.6954$ $b = 85.75$ $\theta = 0.1$ $k = 0.0208$	$a = 317.0256$ $(312.5864)$ $b = 1.4584$ $(0.4899)$ $\theta = -0.0482$ $(1.2301)$ $\lambda = 0.09617$ $(0.0688)$	$\mu = -0.6548$ $(1.1214)$ $\sigma = 3.3672$ $(0.7295)$ $a = 1.4848$ $(0.3665)$ $\lambda = 8.0323$ $(2.8206)$	$\theta = 0.1134$ $(0.0201)$ $\lambda_3 = 5.3192$ $(2.2026)$ $\lambda_4 = 3.0133$ $(1.0588)$
Log-likelihood	- 272.1280	- 260.4813	- 247.8373	- 247.7
AIC	552.256	528.952	503.7	501.4
K-S statistic ( $p$ -value)	0.1625 (0.0446)	0.1219 (0.2350)	0.0662 (0.9101)	0.0531 (0.9873)

<sup>a</sup> MLEs, K-S statistic ( $p$ -value) are from Al-Aqtash et al. (2015).



**Fig. 7** The histogram and the fitted PDFs for the Wheaton River data

Four new generalized  $R$  distributions in the  $T-R\{\text{generalized } \lambda\}$  families of distributions are defined, namely, the uniform-exponential{generalized  $\lambda$ }, the normal-uniform{generalized  $\lambda$ }, the Pareto-Weibull{generalized  $\lambda$ } and the log-logistic-logistic{generalized  $\lambda$ }. As mentioned in sub-section 4.3, the Pareto-Weibull{generalized  $\lambda$ } distribution has the gamma/Gompertz distribution and other distributions as special cases.

The uniform-exponential{generalized  $\lambda$ } distribution is applied to fit two real data sets. The results show that the uniform-exponential{generalized  $\lambda$ } distribution has the ability to fit right skewed data with long tail.

#### Acknowledgments

We are grateful for many constructive comments and suggestions from the handling editor and the reviewer. These comments and suggestions have greatly improved the presentation of the paper.

#### Funding

The third author (Felix Famoye) gratefully acknowledges the financial support received from the U.S. Department of State, Bureau of Education and Cultural Affairs under the Fulbright Grant # PS00230565.

#### Authors' contributions

The authors, viz MA, CL and FF with the consultation of each other carried out this work and drafted the manuscript together. All authors read and approved the final manuscript. The authors confirmed that the content of the manuscript has not been published, or submitted for publication elsewhere.

#### Competing interests

The authors declare that they have no competing interests.

#### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 7 June 2017 Accepted: 27 October 2017

Published online: 22 November 2017

#### References

- Akinsete, A., Famoye, F., Lee, C.: The beta-Pareto distribution. *Statistics*. **42**(6), 547–563 (2008)
- Al-Aqtash, R., Famoye, F., Lee, C.: On generating a new family of distributions using the logit function. *Journal of Probability and Statistical Science*. **13**(1), 135–152 (2015)
- Aljarrah, M.A., Lee, C., Famoye, F.: On generating  $T-X$  family of distributions using quantile functions. *Journal of Statistical Distributions and Applications*. **1**, 1–17 (2014)
- Almheidat, M., Famoye, F., Lee, C.: Some generalized families of Weibull distribution: Properties and applications. *International Journal of Statistics and Probability*. **4**, 18–35 (2015)

- Alshawarbeh, E., Famoye, F., Lee, C.: Beta-Cauchy distribution: some properties and applications. *Journal of Statistical Theory and Applications*. **12**(4), 378–391 (2013)
- Alzaatreh, A., Famoye, F., Lee, C.: Gamma-Pareto distribution and its applications. *Journal of Modern Applied Statistical Methods*. **11**(1), 78–94 (2012)
- Alzaatreh, A., Lee, C., Famoye, F.: A new method for generating families of continuous distributions. *Metron*. **71**(1), 63–79 (2013)
- Alzaatreh, A., Lee, C., Famoye, F.: T-normal family of distributions: a new approach to generalize the normal distribution. *Journal of Statistical Distributions and Applications*. **1**, 1–16 (2014)
- Bemmaor, A.C., Glady, N.: Modeling purchasing behavior with sudden “Death”: A flexible customer lifetime model. *Management Science*. **58**(5), 1012–1021 (2012)
- Cordeiro, G.M., de Castro, M.: A new family of generalized distributions. *Journal of Statistical Computation and Simulation*. **81**(7), 883–898 (2011)
- Cordeiro, G.M., Ortega, E.M.M., Nadarajah, S.: The Kumaraswamy Weibull distribution with application to failure data. *Journal of the Franklin Institute*. **347**, 1399–1429 (2010)
- de Santana, T.V.F., Ortega, E.M., Cordeiro, G.M., Silva, G.O.: The Kumaraswamy-log-logistic distribution. *Journal of Statistical Theory and Applications*. **3**, 265–291 (2012)
- Eugene, N., Lee, C., Famoye, F.: Beta-normal distribution and its applications. *Communications in Statistics-Theory and Methods*. **31**(4), 497–512 (2002)
- Famoye, F., Lee, C., Olumolade, O.: The beta-Weibull distribution. *Journal of Statistical Theory and Applications*. **4**(2), 121–136 (2005)
- Jones, M.C.: Kumaraswamy's distribution: A beta-type distribution with tractability advantages. *Statistical Methodology*. **6**, 70–81 (2009)
- Karian, Z.A., Dudewicz, E.J.: *Fitting statistical distributions: the generalized lambda distribution and generalized bootstrap methods*. Chapman and Hall/CRC Press, Boca Raton, FL (2000)
- Karian, Z.A., Dudewicz, E.J.: *Handbook of fitting statistical distributions with R*. Chapman and Hall/CRC Press, Boca Raton, FL (2010)
- Kumaraswamy, P.: A generalized probability density function for double-bounded random processes. *Journal of Hydrology*. **46**, 79–88 (1980)
- Lee, C., Famoye, F., Alzaatreh, A.: Methods for generating families of univariate continuous distributions in the recent decades. *WIREs. Computational Statistics*. **5**(3), 219–238 (2013)
- Nadarajah, S., Kotz, S.: The beta exponential distribution. *Reliability Engineering and System Safety*. **91**(6), 689–697 (2006)
- Paranaíba, P.F., Ortega, E.M., Cordeiro, G.M., Pascoa, M.A.D.: The Kumaraswamy Burr XII distribution: theory and practice. *Journal of Statistical Computation and Simulation*. **83**(11), 2117–2143 (2013)
- Pereira, M.B., Silva, R.B., Zea, L.M., Cordeiro, G.M.: The Kumaraswamy Pareto distribution. *Journal of Statistical Theory and Applications*. **12**(2), 129–144 (2012)
- Ramberg, J.S., Schmeiser, B.W.: An approximate method for generating asymmetric random variables. *Communications of the ACM*. **17**(2), 78–82 (1974)
- Shannon, C.E.: A mathematical theory of communication. *Bell System Technical Journal*. **27**, 379–432 (1948)
- Tukey, J.W.: The practical relationship between the common transformations of percentages of counts and of amounts. *Technical Report 36. Princeton University, Statistical Techniques Research Group* (1960)
- Zea, L.M., Silva, R.B., Bourguignon, M., Santos, A.M., Cordeiro, G.M.: The beta exponentiated Pareto distribution with application to bladder cancer susceptibility. *International Journal of Statistics and Probability*. **1**(2), 8–21 (2012)

**Submit your manuscript to a SpringerOpen<sup>®</sup> journal and benefit from:**

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)