A new method for generating families of continuous distributions

Ayman Alzaatreh · Carl Lee · Felix Famoye

Abstract In this paper, a new method is proposed for generating families of continuous distributions. A random variable X, "the transformer", is used to transform another random variable T, "the transformed". The resulting family, the T-X family of distributions, has a connection with the hazard functions and each generated distribution is considered as a weighted hazard function of the random variable X. Many new distributions, which are members of the family, are presented. Several known continuous distributions are found to be special cases of the new distributions.

Keywords Hazard function · Pearson system · Beta-family · Generalized distribution · Entropy · Shannon entropy

1 Introduction

Statistical distributions are commonly applied to describe real world phenomena. Due to the usefulness of statistical distributions, their theory is widely studied and new distributions are developed. The interest in developing more flexible statistical distributions remains strong in statistics profession. Many generalized classes of distributions have been developed and applied to describe various phenomena. A common feature of these generalized distributions is that they have more parameters. Johnson et al. [19] stated that the use of four-parameter distributions should be sufficient for most practical purposes. According to these authors, at least three parameters are needed but they doubted any noticeable improvement arising from including a fifth or sixth parameter.

A. Alzaatreh
Department of Mathematics and Statistics, Austin Peay State University,
Clarksville, TN 37044, USA

C. Lee · F. Famoye (✉)
Department of Mathematics, Central Michigan University,
Mount Pleasant, MI 48859, USA
e-mail: felix.famoye@cmich.edu

The Pearson system of continuous distributions, as developed by Pearson [31], is a system for which every probability density function (p.d.f.) $f(x)$ satisfies a differential equation of the form

$$\frac{1}{f(x)}\frac{df(x)}{dx} = \frac{a+x}{b_0 + b_1 x + b_2 x^2},\tag{1.1}$$

where a, b_0, b_1, and b_2 are the parameters (see Johnson et al. [19, Chapter 12]). The shape of the function $f(x)$ depends on the parameters. The different shapes of the distribution were classified by Pearson into a number of types. The different types correspond to the different forms of solution to (1.1). The form of solution of (1.1) depends on the roots of the equation $b_0 + b_1 x + b_2 x^2 = 0$. An example is when $b_1 = b_2 = 0$, which led to the normal distribution and this is not assigned to a particular type. For a detailed discussion of the various types, see Chapter 12 of Johnson et al. [19]

Burr [5] presented a system of continuous distributions which can take on a wide variety of shapes. The system of distributions satisfy the differential equation

$$dF = F(1-F)g(x)dx,\tag{1.2}$$

where $0 \leq F \leq 1$ and $g(x)$ is a non-negative function over x. Burr [5] gave 12 solutions to the equation in (1.2) and these correspond to the choices of $g(x)$. See Fry [15] and Johnson et al. [19] for a list of Burr Types I–XII distributions.

Johnson [18] proposed a system for generating distributions using normalization transformation with the general form

$$Z = \gamma + \delta f\left(\frac{x-\xi}{\lambda}\right),\tag{1.3}$$

where $f(.)$ is the transformation function, Z is a standardized normal random variable, γ and δ are shape parameters, λ is a scale parameter and ξ is a location parameter. Without loss of generality, Johnson assumed that δ and λ are positive. He proposed three transformation functions and defined the lognormal family, the bounded system of distributions and unbounded system of distributions. These families of distributions cover many commonly used distributions such as normal, log-normal, gamma, beta, exponential distributions, and others. For more discussions, see Johnson et al. [19, p.33].

Tukey [37] proposed lambda distribution, which was generalized by Ramberg and Schmeiser [32,33] and Ramberg et al. [34] as the so-called generalized lambda distributions (GLD). This family of distributions is defined in terms of percentile function

$$Q(y) = Q(y; \lambda_1, \lambda_2, \lambda_3, \lambda_4) = \lambda_1 + \frac{y^{\lambda_3} - (1-y)^{\lambda_4}}{\lambda_2}, \quad \text{where } 0 \leq y \leq 1.\tag{1.4}$$

The parameters λ_1 and λ_2 are, respectively, location and scale parameters, while λ_3 and λ_4 determine the skewness and kurtosis. The corresponding p.d.f. is given by

$$f(x) = \frac{\lambda_2}{\lambda_3 y^{\lambda_3 - 1} + \lambda_4 (1-y)^{\lambda_4 - 1}}, \quad \text{with } x = Q(y).\tag{1.5}$$

The existence of a valid p.d.f. requires the condition that $\lambda_3 y^{\lambda_3 - 1} + \lambda_4(1-y)^{\lambda_4 - 1}$ has the same sign for all y in [0, 1] and that λ_2 takes the same sign. Freimer et al. [14] discussed the similarity and differences between the Pearson's system and the GLD. They pointed out that Pearson's family does not include logistic distribution, while GLD does not cover all

skewness and kurtosis values. An extended GLD was proposed by Karian and Dudewicz [23] that consists of both GLD and generalized beta distribution defined as

$$f(x) = \begin{cases} \frac{(x-\beta_1)^{\beta_3}(\beta_1+\beta_2-x)^{\beta_4}}{B(\beta_3+1,\,\beta_4+1)\beta_2^{(\beta_3+\beta_4+1)}}, & \text{for } \beta_1 \le x \le \beta_1+\beta_2 \\ 0, & \text{otherwise,} \end{cases} \quad (1.6)$$

where $B(.,.)$ is the complete beta function. For detailed discussion about the GLD and extended GLD, see Karian and Dudewicz [23].

Azzalini [4] introduced the skew normal family of distributions. Suppose X and Y are independent random variables, each with a p.d.f. that is symmetric about zero. For any λ,

$$0.5 = P(X - \lambda Y < 0) = \int_{-\infty}^{\infty} f_Y(y) F_X(\lambda y) dy. \quad (1.7)$$

Thus, $2 f_Y(y) F_X(\lambda y)$ is a probability density function. If X and Y are each standard normal, $N(0, 1)$, then the skew-normal family of distributions has the p.d.f.

$$2\varphi(x)\Phi(\lambda x), \quad (1.8)$$

where $\varphi(x)$ and $\Phi(x)$ are $N(0, 1)$ p.d.f. and cumulative distribution function respectively. The distribution in (1.8) is characterized by a single parameter λ. Location and scale parameters can be added to the distribution in (1.8) by using the translation $Y = \mu + \sigma X$. For skew normal distribution and other systems of continuous distributions, see Johnson et al. [19, Chapter 12].

Eugene et al. [10] used the beta distribution as a generator to develop the so-called family of beta-generated distributions. The cumulative distribution function (c.d.f.) of a beta-generated random variable X is defined as

$$G(x) = \int_0^{F(x)} b(t) dt, \quad (1.9)$$

where $b(t)$ is the p.d.f. of the beta random variable and $F(x)$ is the c.d.f. of any random variable. The p.d.f. corresponding to the beta-generated distribution in (1.9) is given by

$$g(x) = \frac{1}{B(\alpha, \beta)} f(x) F^{\alpha-1}(x)(1 - F(x))^{\beta-1}. \quad (1.10)$$

This family of distributions is a generalization of the distributions of order statistics for the random variable X with c.d.f. $F(x)$ as pointed out by Eugene et al. [10] and Jones [21]. Since the paper by Eugene et al. [10], many beta-generated distributions have been studied in the literature including the beta-Gumbel distribution by Nadarajah and Kotz [29], beta-exponential distribution by Nadarajah and Kotz [30], beta-Weibull distribution by Famoye et al. [12], beta-gamma by distribution Kong et al. [24], beta-Pareto by distribution Akinsete et al. [1], and others.

Recently, Jones [22] and Cordeiro and de Castro [6] extended the beta-generated family of distributions by replacing the beta distribution in (1.9) with the Kumaraswamy distribution, $b(t) = \alpha\beta x^{\alpha-1}(1 - x^\alpha)^{\beta-1}$, $x \in (0, 1)$, Kumaraswamy [25]. The p.d.f. of the Kumaraswamy generalized distributions (KW-G) is given by

$$g(x) = \alpha\beta f(x) F^{\alpha-1}(x)(1 - F^\alpha(x))^{\beta-1}. \quad (1.11)$$

Several generalized distributions from (1.11) have been studied in the literature including the Kumaraswamy Weibull distribution by Cordeiro et al. [7], the Kumaraswamy generalized gamma distribution by de Castro et al. [9], and the Kumaraswamy generalized half-normal distribution by Cordeiro et al. [8].

Ferreira and Steel [13] introduced a method to generate skewed distributions through inverse probability integral transformations. According to Ferreira and Steel [13], a distribution G is said to be a skewed version of the symmetric distribution F, generated by the skewing mechanism P, if its p.d.f. is of the form

$$g(y|F, P) = f(y)p(F(y)). \tag{1.12}$$

Note that the p.d.f. (1.12) is a weighted function of $f(.)$ with the weight $p(F(.))$. The skewed normal family in (1.8) is a special case of this family. By relaxing the assumption $F(.)$ being symmetric, the beta-generated family (1.10) is a special case of (1.12).

This article presents yet another technique to generate families of continuous probability distributions. The article is organized as follows: Sect. 2 presents a new technique for generating families of continuous distributions. Section 3 gives examples of classes of generalized families developed using the technique in Sect. 2. The paper ends with a summary and conclusion in Sect. 4.

2 Method for generating families of continuous probability distributions

The beta-generated family of distributions in (1.10) and the KW-G family of distributions in (1.11) are generated by using distributions with support between 0 and 1 as the generator. The beta random variable and the KW random variable lie between 0 and 1, so is the c.d.f. $F(x)$ of any other random variable. The limitation of using a generator with support lying between 0 and 1 raises an interesting question: 'Can we use other distributions with different support as the generator to derive different classes of distributions?' This section will address this question and introduce a new technique to derive families of distributions by using any p.d.f. as a generator.

Let $r(t)$ be the p.d.f. of a random variable $T \in [a, b]$, for $-\infty \le a < b \le \infty$. Let $W(F(x))$ be a function of the c.d.f. $F(x)$ of any random variable X so that $W(F(x))$ satisfies the following conditions:

$$\left.\begin{array}{l} W(F(x)) \in [a, b] \\ W(F(x)) \text{ is differentiable and monotonically non-decreasing} \\ W(F(x)) \to a \text{ as } x \to -\infty \text{ and } W(F(x)) \to b \text{ as } x \to \infty. \end{array}\right\} \tag{2.1}$$

A method for generating new families of distribution is presented in the following definition.

Definition Let X be a random variable with p.d.f. $f(x)$ and c.d.f. $F(x)$. Let T be a continuous random variable with p.d.f. $r(t)$ defined on $[a, b]$. The c.d.f. of a new family of distributions is defined as

$$G(x) = \int_{a}^{W(F(x))} r(t)dt, \tag{2.2}$$

where $W(F(x))$ satisfies the conditions in (2.1). The c.d.f. $G(x)$ in (2.2) can be written as $G(x) = R\{W(F(x))\}$, where $R(t)$ is the c.d.f. of the random variable T. The corresponding p.d.f. associated with (2.2) is

$$g(x) = \left\{\frac{d}{dx}W(F(x))\right\} r\{W(F(x))\}. \tag{2.3}$$

Note that:

- The c.d.f. in (2.2) is a composite function of $(R \bullet W \bullet F)(x)$.
- The p.d.f. $r(t)$ in (2.2) is "transformed" into a new c.d.f. $G(x)$ through the function, $W(F(x))$, which acts as a "transformer". Hence, we shall refer to the distribution $g(x)$ in (2.3) as transformed from random variable T through the transformer random variable X and call it "Transformed-Transformer" or "T-X" distribution.
- The random variable X may be discrete and in such a case, $G(x)$ is the c.d.f. of a family of discrete distributions.
- The distribution (1.12) introduced by Ferreira and Steel [13] is a special case of (2.3) by defining $W(F(x)) = F(x)$ and $r(.)$ plays the same role as the weight function.

Different $W(F(x))$ will give a new family of distributions. The definition of $W(F(x))$ depends on the support of the random variable T. The following are some examples of $W(.)$.

1. When the support of T is bounded: Without loss of generality, we assume the support of T is [0, 1]. Distributions for such T include uniform (0, 1), beta, Kumaraswamy and other types of generalized beta distributions. $W(F(x))$ can be defined as $F(x)$ or $F^\alpha(x)$. This is the beta-generated family of distributions which have been well studied during the recent decade.
2. When the support of T is $[a, \infty), a \geq 0$: Without loss of generality, we assume $a = 0$. $W(F(x))$ can be defined as $-\log(1 - F(x))$, $F(x)/(1 - F(x))$, $-\log(1 - F^\alpha(x))$, and $F^\alpha(x)/(1 - F^\alpha(x))$, where $\alpha > 0$.
3. When the support of T is $(-\infty, \infty)$: $W(F(x))$ can be defined as $\log[-\log(1 - F(x))]$, $\log[F(x)/(1 - F(x))]$, $\log[-\log(1 - F^\alpha(x))]$, and $\log[F^\alpha(x)/(1 - F^\alpha(x))]$.

By using the $W(F(x)) = -\log(1 - F(x))$ in the second example, the $G(x)$ in (2.2) is a c.d.f. of the new family of distributions which is given by

$$G(x) = \int_0^{-\log(1-F(x))} r(t)dt = R\{-\log(1 - F(x))\}, \tag{2.4}$$

where $R(t)$ is the c.d.f. of the random variable T. The corresponding p.d.f. associated with (2.4) is

$$g(x) = \frac{f(x)}{1 - F(x)}r(-\log(1 - F(x))) = h(x)r(-\log(1 - F(x))), \tag{2.5}$$

where $h(x)$ is the hazard function for the random variable X with the c.d.f. $F(x)$.

The corresponding families of distributions generated from the other $W(.)$ functions mentioned in examples 2 and 3 are given in Table 1.

In the remainder of this article, we will focus on the case when T has the support $[0, \infty)$ and $W(F(x)) = -\log(1 - F(x))$. For simplicity, we will use the name T-X family of distributions for the new family of distributions in (2.5).

Some remarks on the family of distributions defined in (2.5):

(a) The p.d.f. in (2.5) can be written as $g(x) = h(x)r(H(x))$ and the corresponding c.d.f. is $G(x) = R(-\log(1 - F(x))) = R(H(x))$, where $h(x)$ and $H(x)$ are hazard and cumulative hazard functions of the random variable X with c.d.f. $F(x)$. Hence, this family of distributions can be considered as a family of distributions arising from a weighted hazard function.

Table 1 Probability density functions of some T-X families based on different $W(.)$ functions

Support of T	$W(F(x))$	$g(x)$
$[0, \infty)$	$\frac{F(x)}{1-F(x)}$	$\frac{f(x)}{(1-F(x))^2} \, r\left\{\frac{F(x)}{1-F(x)}\right\}$
$[0, \infty)$	$-\log(1 - F^\alpha(x))$	$\frac{\alpha f(x) F^{\alpha-1}(x)}{1-F^\alpha(x)} r\{-\log(1 - F^\alpha(x))\}$
$[0, \infty)$	$\frac{F^\alpha(x)}{1-F^\alpha(x)}$	$\frac{\alpha f(x) F^{\alpha-1}(x)}{(1-F^\alpha(x))^2} \, r\left\{\frac{F^\alpha(x)}{1-F^\alpha(x)}\right\}$
$(-\infty, \infty)$	$\log(-\log(1 - F(x)))$	$\frac{f(x) \, r\{\log(-\log(1-F(x)))\}}{(F(x)-1)\log(1-F(x))}$
$(-\infty, \infty)$	$\log\left(\frac{F(x)}{1-F(x)}\right)$	$\frac{f(x)}{F(x)(1-F(x))} \, r\left\{\log\left(\frac{F(x)}{1-F(x)}\right)\right\}$
$(-\infty, \infty)$	$\log(-\log(1 - F^\alpha(x)))$	$\frac{\alpha f(x) F^{\alpha-1}(x)}{(F^\alpha(x)-1)\log(1-F^\alpha(x))} r\{\log(-\log(1 - F^\alpha(x)))\}$
$(-\infty, \infty)$	$\log\left(\frac{F^\alpha(x)}{1-F^\alpha(x)}\right)$	$\frac{\alpha f(x)}{F(x)(1-F^\alpha(x))} \, r\left\{\log\left(\frac{F^\alpha(x)}{1-F^\alpha(x)}\right)\right\}$

(b) The fact that $G(x) = R(-\log(1 - F(x)))$ gives the relationship between random variables X and T: $X = F^{-1}(1 - e^{-T})$. This provides an easy way for simulating random variable X by first simulating random variable T from p.d.f. $r(t)$ and computing $X = F^{-1}(1 - e^{-T})$, which has the c.d.f. $G(x)$. Thus, $E(X)$ can be obtained using $E(X) = E\{F^{-1}(1 - e^{-T})\}$.

The quantile function, $Q(\lambda)$, $0 < \lambda < 1$, for the T-X family of distributions can be computed by using the formula

$$Q(\lambda) = F^{-1}\{1 - e^{-R^{-1}(\lambda)}\}. \tag{2.6}$$

The Shannon [35] entropy of a random variable X is a measure of variation of uncertainty. Shannon entropy is defined as $E\{-\log(g(X))\}$. Theorem 1 shows the connection between the Shannon entropy of the new family of distributions, $g(x)$, and the Shannon entropy of the generator, $r(t)$.

Theorem 1 *If a random variable X follows the family of distributions* $g(x) = \frac{f(x)}{1-F(x)} r(-\log (1 - F(x)))$ *in* (2.5)*, then the Shannon entropy of X,* η_X*, is given by*

$$\eta_X = -E\{\log f(F^{-1}(1 - e^{-T}))\} - \mu_T + \eta_T, \tag{2.7}$$

where μ_T *and* η_T *are the mean and the Shannon entropy for the random variable T with p.d.f.* $r(t)$.

Proof By definition,

$$\eta_X = E(-\log[g(X)])$$
$$= -E(\log f(X)) + E(\log(1 - F(X))) + E(-\log r\{-\log(1 - F(X))\}).$$

From (2.4), the random variable $T = -\log(1 - F(X))$ has the p.d.f. $r(t)$ which implies the following: $E(\log f(X)) = E\{\log f(F^{-1}(1-e^{-T}))\}$, $E(\log(1-F(X))) = -E(T) = -\mu_T$, and $E(-\log r\{-\log(1 - F(X))\}) = E(-\log r(t)) = \eta_T$.

Hence, $\eta_X = -E\{\log f(F^{-1}(1 - e^{-T}))\} - \mu_T + \eta_T$, which is the result in (2.7).   □

Skewness and kurtosis of a parametric distribution are often measured by $\alpha_3 = \mu_3/\sigma^3$ and $\alpha_4 = \mu_4/\sigma^4$, respectively. When the third or fourth moment does not exist, for example,

Cauchy, Lévy and Pareto distributions, α_3 and α_4 cannot be computed. For the T-X family, one may encounter some difficulty in computing the third and fourth moments. Alternative measures for skewness and kurtosis, based on quantile functions, are sometimes more appropriate for such distributions. The measure of skewness S defined by Galton [16] and the measure of kurtosis K defined by Moors [27] are based on quantile functions and they are defined as

$$S = \frac{Q(6/8) - 2Q(4/8) + Q(2/8)}{Q(6/8) - Q(2/8)}, \tag{2.8}$$

$$K = \frac{Q(7/8) - Q(5/8) + Q(3/8) - Q(1/8)}{Q(6/8) - Q(2/8)}. \tag{2.9}$$

Skewness measures the degree of the long tail (towards left or right side). Kurtosis is a measure of the degree of tail heaviness. When the distribution is symmetric, $S = 0$ and when the distribution is right (or left) skewed, $S > 0$ (or < 0). As K increases, the tail of the distribution becomes heavier. For the T-X family, Galton's skewness and Moors' kurtosis can be computed by using the quantile function in (2.6) and the appropriate T and X distributions.

3 Some families of T-X distributions with different T distributions

The T-X family of distributions can be further classified into two sub-families: One sub-family has the same X distribution but different T distributions and the other sub-family has the same T distribution but different X distributions. For example, by letting T be a Weibull random variable, we generate a sub-family of Weibull-X distributions. By letting X be a Weibull random variable, we generate a sub-family of T-Weibull distributions. In this section, we consider the sub-family with different T distributions. Table 2 gives several such sub-families with the same X and different T random variables.

In each of the following sub-sections, we discuss the properties of the gamma-X family, beta-exponential-X family, and Weibull-X family.

3.1 Gamma-X family

If a random variable T follows the gamma distribution with parameters α and β, then $r(t) = (\Gamma(\alpha)\beta^\alpha)^{-1} t^{\alpha-1} e^{-t/\beta}$, $t > 0$. From (2.5), the p.d.f. of gamma-X family is defined as

$$g(x) = \frac{1}{\Gamma(\alpha)\beta^\alpha} f(x)(-\log(1 - F(x)))^{\alpha-1}(1 - F(x))^{\frac{1}{\beta}-1}. \tag{3.1}$$

By using (2.4) and expressing the c.d.f. of the gamma distribution in terms of the incomplete gamma function, $R(t) = (1/\Gamma(\alpha))\gamma(\alpha, t/\beta)$, where $\gamma(\alpha, t) = \int_0^t u^{\alpha-1} e^{-u} du$, the c.d.f. of the gamma-X family in (3.1) is $G(x) = \gamma\{\alpha, -\log(1 - F(x))\}/\Gamma(\alpha)$. We will refer to distributions of the form $(F(x))^c$ and $(1 - F(x))^c$ as, respectively, $\mathrm{Exp}(F)$ and $\mathrm{Exp}(1 - F)$ family of distributions.

Lemma 1 *The Shannon entropy of the gamma-X family of distributions is given by $\eta_X = -E\{\log f(F^{-1}(1 - e^{-T}))\} + \alpha(1 - \beta) + \log \beta + \log \Gamma(\alpha) + (1 - \alpha)\psi(\alpha)$, where ψ is the digamma function.*

Proof It follows from Theorem 1 by using $\mu_T = \alpha\beta$ and the Shannon entropy for the gamma distribution, which is given by Song [36] as $\eta_T = \alpha + \log \beta + \log \Gamma(\alpha) + (1 - \alpha)\psi(\alpha)$. □

Table 2 Families of generalized distributions derived from different T distributions

Name	The density $r(t)$	The density of the family $g(x)$
Exponential	$\theta e^{-\theta t}$	$\theta f(x)(1-F(x))^{\theta-1}$
Beta-exponential	$\frac{\lambda e^{-\lambda \beta x}(1-e^{-\lambda x})^{\alpha-1}}{B(\alpha,\beta)}$	$\frac{\lambda f(x)}{B(\alpha,\beta)}(1-F(x))^{\lambda\beta-1}\{1-(1-F(x))^{\lambda}\}^{\alpha-1}$
Exponentiated-exponential	$\frac{\alpha\lambda(1-e^{-\lambda x})^{\alpha-1}}{e^{\lambda x}}$	$\alpha\lambda f(x)\{1-(1-F(x))^{\lambda}\}^{\alpha-1}(1-F(x))^{\lambda-1}$
Gamma	$\frac{1}{\Gamma(\alpha)\beta^{\alpha}}t^{\alpha-1}e^{-t/\beta}$	$\frac{f(x)}{\Gamma(\alpha)\beta^{\alpha}}(-\log(1-F(x)))^{\alpha-1}(1-F(x))^{\frac{1}{\beta}-1}$
Half normal	$\frac{1}{\sigma}(\frac{2}{\pi})^{1/2}e^{-t^2/2\sigma^2}$	$\frac{1}{\sigma}(\frac{2}{\pi})^{1/2}\frac{f(x)}{1-F(x)}\exp(-\{\log(1-F(x))\}^2/2\sigma^2)$
Levy	$(\frac{c}{2\pi})^{1/2}\frac{e^{-c/2t}}{t^{3/2}}$	$(\frac{c}{2\pi})^{1/2}\frac{f(x)}{1-F(x)}\frac{\exp(-c/2\log(1-F(x)))}{\{-\log(1-F(x))\}^{3/2}}$
Log logistic	$\frac{\beta(t/\alpha)^{\beta-1}}{\alpha\{1+(t/\alpha)^{\beta}\}^2}$	$\frac{\beta}{\alpha^{\beta}}\frac{f(x)}{1-F(x)}\{-\log(1-F(x))\}^{\beta-1}$ $\times\{1+(-\log(1-F(x))/\alpha)^{\beta}\}^{-2}$
Rayleigh	$\frac{t}{\sigma^2}e^{-t^2/2\sigma^2}$	$\frac{-f(x)\log(1-F(x))}{\sigma^2(1-F(x))}\exp(-\{\log(1-F(x))\}^2/2\sigma^2)$
Type-2 Gumbel	$\alpha\beta t^{-\alpha-1}e^{-\beta t^{-\alpha}}$	$\frac{\alpha\beta f(x)}{1-F(x)}\{-\log(1-F(x))\}^{-\alpha-1}$ $\times\exp(-\beta\{-\log(1-F(x))\}^{-\alpha})$
Lomax	$\frac{\lambda k}{(1+\lambda t)^{k+1}}$	$\frac{f(x)}{1-F(x)}\frac{\lambda k}{\{1-\lambda\log(1-F(x))\}^{k+1}}$
Inverted beta	$\frac{t^{\beta-1}(1+t)^{-\beta-\gamma}}{B(\beta,\gamma)}$	$\frac{f(x)}{B(\beta,\gamma)(1-F(x))}\{-\log(1-F(x))\}^{\beta-1}$ $\times\{1-\log(1-F(x))\}^{-\beta-\gamma}$
Inverse Gaussian	$\sqrt{\frac{\lambda}{2\pi t^3}}e^{-\frac{\lambda}{2\mu^2 t}(t-\mu)^2}$	$\frac{f(x)}{1-F(x)}\left(\frac{\lambda}{2\pi\{-\log(1-F(x))\}^3}\right)^{1/2}$ $\times\exp\left(-\frac{\lambda\{-\log(1-F(x))-\mu\}^2}{2\mu^2\{-\log(1-F(x))\}}\right)$
Weibull	$\frac{c}{\gamma}(\frac{t}{\gamma})^{c-1}e^{-(t/\gamma)^c}$	$\frac{c}{\gamma}\frac{f(x)}{1-F(x)}\{-(1/\gamma)\log(1-F(x))\}^{c-1}$ $\times\exp(-\{-(1/\gamma)\log(1-F(x))\}^c)$

When $\alpha = 1$, the gamma-X family in (3.1) reduces to $\text{Exp}(1 - F)$ distributions. When $\alpha = n$ and $\beta = 1$, the gamma-X family is the density function of the nth upper record value arising from a sequence $\{X_i\}$ of identically independent random variables with the p.d.f. $f(x)$ and c.d.f. $F(x)$ (see Johnson et al. [19, p. 99]). The generalized gamma distribution defined by Amoroso [3] is a member of gamma-X family where X is the Weibull random variable. If $f(x)$ is the p.d.f. of the Weibull distribution, then (3.1) becomes

$$g(x) = \frac{c}{\gamma^{\alpha c}\Gamma(\alpha)\beta^{\alpha}}x^{\alpha c-1}e^{-\frac{1}{\beta}(x/\gamma)^c}, \quad x > 0; \ \alpha, \ \gamma, \ \beta, \ c > 0. \tag{3.2}$$

Setting $\delta = \beta\gamma^c$ in Eq. (3.2), the distribution reduces to the generalized gamma distribution in Amoroso [3]. When $c = \gamma = 1$, (3.2) reduces to the gamma distribution.

If $f(x)$ is the p.d.f. of the Pareto distribution, then from (3.1) we get

$$g(x) = \frac{k^{\alpha}}{x\Gamma(\alpha)\beta^{\alpha}}\left(\frac{\theta}{x}\right)^{k/\beta}\left(\log\left(\frac{x}{\theta}\right)\right)^{\alpha-1}, \quad x > \theta.$$
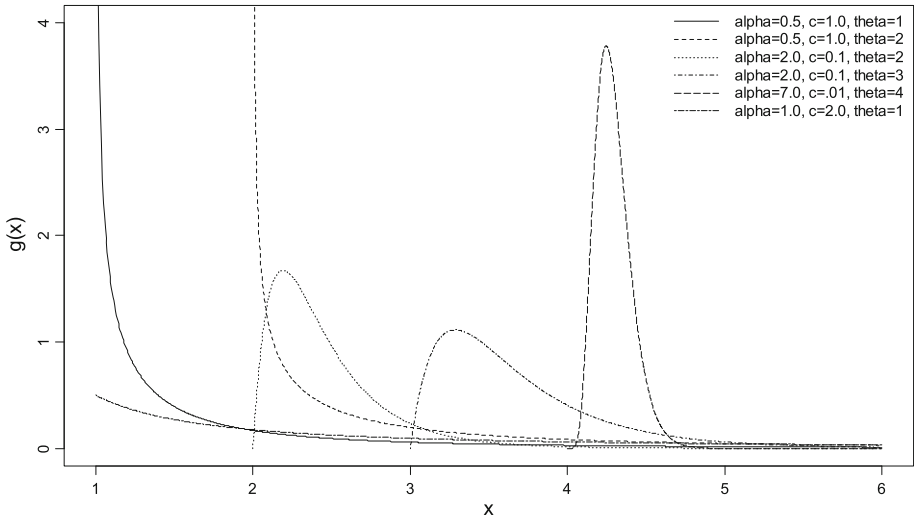
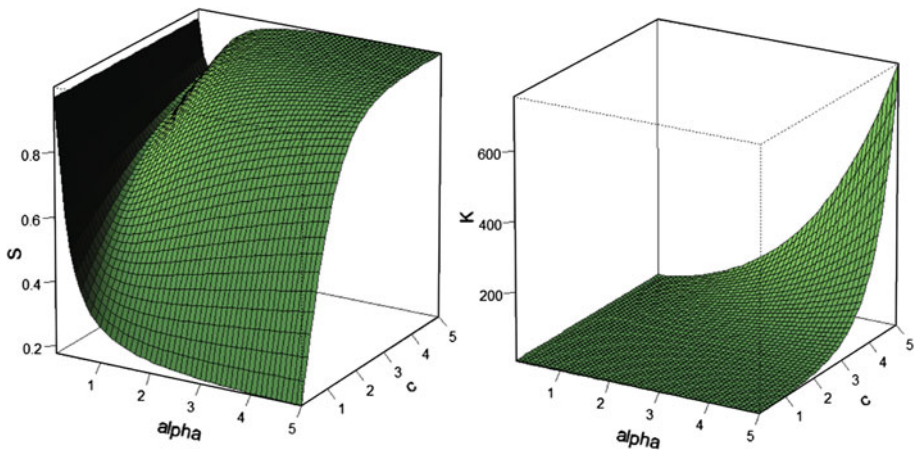**Fig. 1** Graphs of the gamma-Pareto distribution for various parameter values

Fig. 2 Galton's skewness(S) and Moors' kurtosis(K) for the gamma-Pareto distribution

On setting $\beta/k = c$, we get

$$g(x) = \frac{1}{x\Gamma(\alpha)c^\alpha}\left(\frac{\theta}{x}\right)^{1/c}\left(\log\left(\frac{x}{\theta}\right)\right)^{\alpha-1}, \quad x > \theta. \tag{3.3}$$

Based on our naming convention, the distribution in (3.3) will be called gamma-Pareto distribution. When $\alpha = 1$, (3.3) reduces to the Pareto distribution and hence the gamma-Pareto distribution can be considered as a generalization of the Pareto distribution. Figure 1 shows graphs of the gamma-Pareto density for different parameter values including the special cases. The figure shows that the shape parameter α adds extra flexibility to the distribution by changing the shape of the density function from reversed J-shape to concave down shape for certain parameter values.

The c.d.f. of the gamma-Pareto distribution in (3.3) is $G(x) = \gamma\{\alpha, c^{-1}\log(x/\theta)\}/\Gamma(\alpha)$, and hence the quantile function of the gamma-Pareto distribution is the solution of equation

$G(x) = p, \; 0 \le p \le 1$. To investigate the effect of the two shape parameters α and c on the gamma-Pareto density function, Eqs. (2.8) and (2.9) are used to obtain Galton's skewness and Moors' kurtosis. Figure 2 displays the Galton's skewness and Moors' kurtosis for the gamma-Pareto distribution in terms of the parameters α and c when $\theta = 1$.

From Fig. 2 and the corresponding data values (not included to save space), the Galton's skewness is always positive which indicates that the gamma-Pareto distribution is right skewed. For fixed $c \ge 1$, the Galton's skewness is an increasing function of α. For fixed $c < 1$, the Galton's skewness is a decreasing function of α and for fixed α, the Galton's skewness is an increasing function of c. The Moors' kurtosis is an increasing function of α and c.

3.2 Beta-exponential-X family

If a random variable T follows the beta-exponential distribution in Nadarajah and Kotz [30], then $r(t) = \lambda(B(\alpha, \beta))^{-1} e^{-\lambda \beta t} (1 - e^{-\lambda t})^{\alpha - 1}$. From (2.5), the p.d.f. of the beta-exponential-X family is defined as

$$g(x) = \frac{\lambda}{B(\alpha, \beta)} f(x)(1 - F(x))^{\lambda \beta - 1} \{1 - (1 - F(x))^\lambda\}^{\alpha - 1}. \tag{3.4}$$

The c.d.f. of (3.4) can be expressed in terms of the incomplete beta function $I_x(a, b)$. The c.d.f. of the beta-exponential-X family is $G(x) = 1 - I_{(1-F(x))^\lambda}(\lambda(\beta - 1) + 1, \alpha)$.

Lemma 2 *The Shannon entropy of the beta-exponential-X family of distributions is given by*

$$\eta_X = -E\{\log f(F^{-1}(1 - e^{-T}))\} + \log(\lambda^{-1} B(\alpha, \beta)) + (\alpha + \beta - 1)\psi(\alpha + \beta)$$
$$-(\alpha - 1)\psi(\alpha) - \beta\psi(\beta) - [\psi(\alpha + \beta) - \psi(\beta)]/\lambda.$$

Proof It follows from Theorem 1 by using the mean $\mu_T = [\psi(\alpha + \beta) - \psi(\beta)]/\lambda$ and the Shannon entropy $\eta_T = \log(\lambda^{-1} B(\alpha, \beta)) + (\alpha + \beta - 1)\psi(\alpha + \beta) - (\alpha - 1)\psi(\alpha) - \beta\psi(\beta)$ for the beta-exponential distribution, which are given by Nadarajah and Kotz [30]. □

Special cases of beta-exponential-X family:

(1) The beta-generated family in (1.10) is a special case of (3.4) when $\lambda = 1$. Hence, the family of distributions in (3.4) can be used to generate all the distributions belonging to the beta-generated family.
(2) When $\alpha = 1$, the beta-exponential-X family reduces to the Exp$(1 - F)$ distributions. When $\beta = 1$ and $\lambda = 1$, the beta-exponential-X reduces to the Exp(F) distributions.
(3) When $\beta = 1$, (3.4) reduces to the exponentiated-exponential-X family with p.d.f.

$$g(x) = \alpha\lambda f(x)\{1 - (1 - F(x))^\lambda\}^{\alpha - 1} (1 - F(x))^{\lambda - 1}. \tag{3.5}$$

The c.d.f of (3.5) can be written as $G(x) = \{1 - (1 - F(x))^\lambda\}^\alpha$.

By using $D(x) = 1 - F(x)$ in (3.5), the exponentiated-exponential-X family reduces to the KW-G family.

If X is the uniform random variable, then from (3.4) the beta-exponential-uniform is defined as

$$g(x) = \frac{\lambda}{B(\alpha, \beta)} \frac{1}{b-a} \left(\frac{b-x}{b-a} \right)^{\lambda\beta-1} \left\{ 1 - \left(\frac{b-x}{b-a} \right)^{\lambda} \right\}^{\alpha-1}, \quad a < x < b. \quad (3.6)$$

If we use the transformation $y = 1 - x$ in (3.6) then the distribution reduces to the (i) generalized beta distribution of the first kind (McDonald [26]), when $b = 1$, (ii) beta distribution when $a = 0$ and $b = \lambda = 1$, and (iii) Kumaraswamy's [25] double bounded distribution when $a = 0$ and $b = \beta = 1$.

The exponentiated-Weibull distribution defined by Mudholkar et al. [28] is a member of exponentiated-exponential-X family in (3.5) when X is the Weibull random variable. If $f(x)$ is the p.d.f. of the Weibull distribution, then (3.5) reduces to

$$g(x) = \frac{c\lambda\alpha}{\gamma} \left(\frac{x}{\gamma} \right)^{c-1} (1 - e^{-\lambda(x/\gamma)^c})^{\alpha-1} e^{-\lambda(x/\gamma)^c}, \quad x > 0; \ c, \gamma, \alpha, \lambda > 0. \quad (3.7)$$

Writing $\delta = \lambda\gamma^c$, (3.7) reduces to the exponentiated-Weibull distribution given by Mudholkar et al. [28]. When $\gamma = c = 1$, (3.7) reduces to the exponentiated-exponential distribution defined by Gupta and Kundu [17]. When $\lambda = c = 1$, (3.7) reduces to the Weibull distribution. When $\lambda = \gamma = c = 1$, (3.7) reduces to the exponential distribution.

The type I generalized logistic distribution given by Johnson et al. [20, p. 140], is a special case of exponentiated-exponential-logistic distribution. If $f(x)$ is the p.d.f. of the standard logistic distribution then (3.5) reduces to

$$g(x) = \frac{\alpha\lambda e^{-\lambda x}}{(1 + e^{-x})^{\lambda+1}} \left(1 - \frac{e^{-\lambda x}}{(1 + e^{-x})^{\lambda}} \right)^{\alpha-1}, \quad -\infty < x < \infty, \alpha, \lambda > 0. \quad (3.8)$$

When $\lambda = 1$, the exponentiated-exponential-logistic distribution in (3.8) reduces to type I generalized logistic distribution. When $\alpha = \lambda = 1$, (3.8) reduces to standard logistic distribution.

Figure 3 shows graphs of the exponentiated-exponential-logistic density functions for different parameter values including the special cases.

The c.d.f. of the exponentiated-exponential-logistic distribution in equation (3.8) is $G(x) = (1 - (1 + e^x)^{-\lambda})^{\alpha}$, and hence the quantile function of the exponentiated-exponential-logistic distribution can be written as

$$Q(p) = \log((1 - p^{1/\alpha})^{-1/\lambda} - 1), \quad 0 \le p \le 1. \quad (3.9)$$

By using (3.9), (2.8) and (2.9), one can obtain the Galton's skewness and the Moors' kurtosis for the exponentiated-exponential-logistic distribution. Figure 4 displays the Galton's skewness and Moors' kurtosis for the exponentiated-exponential-logistic distribution in terms of the parameters α and λ.

From Fig. 4 and the corresponding data values (not included in order to save space), the exponentiated-exponential-logistic distribution can be left skewed, right skewed, and symmetric. For fixed $\lambda > 1$, the Galton's skewness is an increasing function of α, and for fixed α, the Galton's skewness is a decreasing function of λ. For fixed α, the Moors' kurtosis is a decreasing function of λ when $\lambda > 1$, and for fixed λ, the Moors' kurtosis is a decreasing function of α when $\alpha > 1$.
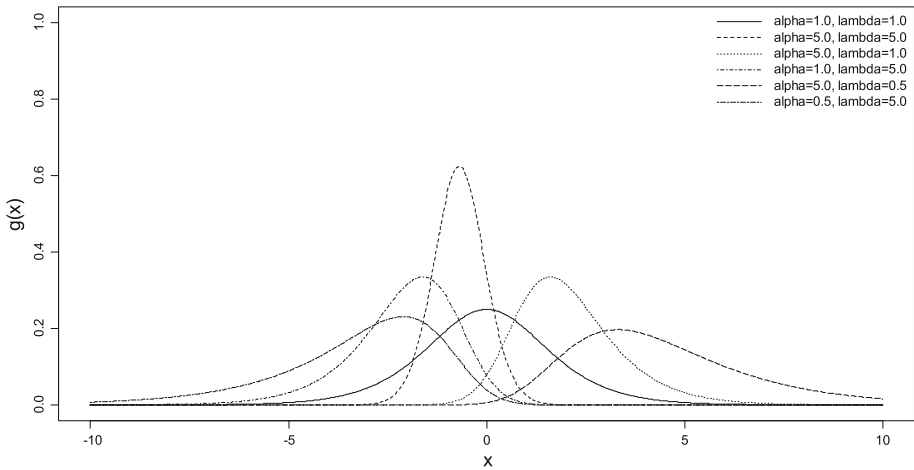
Fig. 3 Graphs of the exponentiated-exponential-logistic distribution for various parameter values
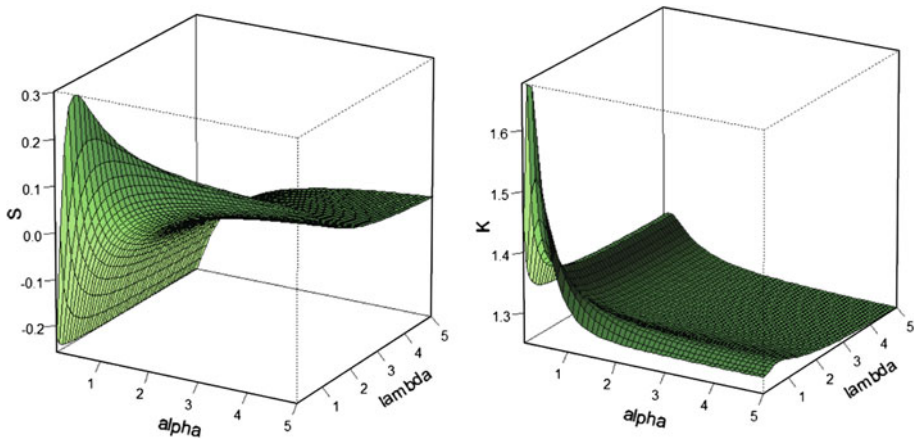

Fig. 4 Galton's skewness(S) and Moors' kurtosis(K) for the exponentiated-exponential-logistic distribution

3.3 Weibull-X family

If a random variable T follows the Weibull distribution with parameters c and γ, then $r(t) = (c/\beta)(t/\beta)^{c-1}e^{-(t/\beta)^c}$, $t \geq 0$. From (2.5) the Weibull-X family is given by

$$g(x) = \frac{c}{\beta} \frac{f(x)}{1 - F(x)} \left[\frac{-\log(1 - F(x))}{\beta} \right]^{c-1} \exp\left\{ -\left(\frac{-\log(1 - F(x))}{\beta} \right)^c \right\}. \quad (3.10)$$

The c.d.f. of the Weibull distribution is $R(t) = 1 - e^{-(t/\beta)^c}$ and hence from (2.4) the c.d.f. of the Weibull-X family is

$$G(x) = 1 - \exp\{-[-\log(1 - F(x))/\beta]^c\}. \quad (3.11)$$

Lemma 3 *The Shannon entropy of the Weibull-X family of distributions is given by*

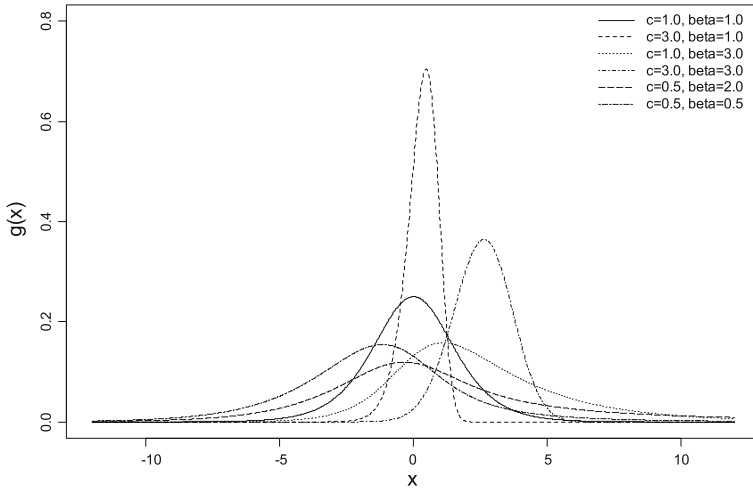$$\eta_X = -E\{\log f(F^{-1}(1 - e^{-T}))\} - \beta\,\Gamma(1 + 1/c) + \gamma(1 - 1/c) - \log(c/\beta) + 1,$$

Fig. 5 Graphs of the Weibull-logistic distribution for various parameter values

where γ is the Euler's constant.

Proof It follows from Theorem 1 by using the mean $\mu_T = \beta \, \Gamma(1 + 1/c)$ and the Shannon entropy $\eta_T = \gamma(1 - 1/c) - \log(c/\beta) + 1$ for the Weibull distribution, which is given by Song [36]. □

When $c = 1$, the Weibull-X family reduces to the $\text{Exp}(1 - F(x))$ distributions. The type II generalized logistic distribution is a special case of Weibull-logistic distribution. If $F(x)$ is the c.d.f. of the standard logistic distribution then (3.11) reduces to

$$G(x) = 1 - \exp\{-[\log(1 + e^x)/\beta]^c\}, \quad -\infty < x < \infty. \tag{3.12}$$

When $c = 1$, the distribution in (3.12) reduces to type II generalized logistic distribution.

Figure 5 shows graphs of the Weibull-logistic density functions for different parameter values including the special case.

From (3.12), the quantile function of the Weibull-logistic distribution can be written as

$$Q(p) = \log\{\exp[\beta(-\log(1 - p))^{1/c}] - 1\}, \quad 0 \le p \le 1. \tag{3.13}$$

Equations (3.13), (2.8) and (2.9) can be used to obtain Galton's skewness and Moors' kurtosis. Figure 6 displays the Galton's skewness and Moors' kurtosis for the Weibull-logistic distribution in terms of parameters β and c.

Figure 6 and the corresponding data values (not included to save space) indicate that the Weibull-logistic distribution can be left skewed, right skewed, and symmetric. For fixed β, the Galton's skewness is a decreasing function of c, and for fixed c, the Galton's skewness is an increasing function of β. For fixed c, the Moors' kurtosis is an increasing function of β when $c \le 1$ and a decreasing function of β when $c > 1$.

4 Summary and conclusion

A method to generate new families of distributions is introduced. This technique defines new family of distributions using the composite function $(R \bullet W \bullet F)(x)$ with R and F being the c.d.f.s of the random variables T and X, respectively. The $W(.)$ function is defined to link the
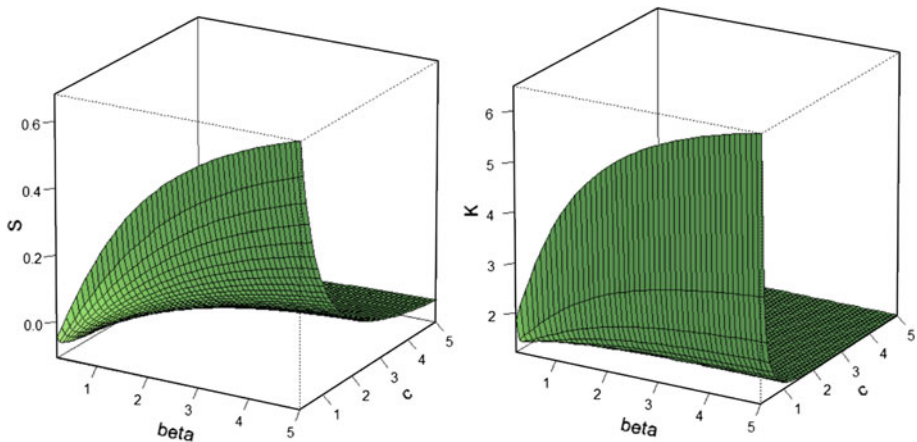
Fig. 6 Galton's skewness(S) and Moors' kurtosis(K) for the Weibull-logistic distribution

support of T to the range of X. This technique generates a large number of new distributions as well as existing distributions as special cases. Table 1 contains several different variants of T-X families using different $W(.)$ functions.

This article focuses on $W(F(x)) = -\log(1 - F(x))$, where the support of T is $[0, \infty)$. Some properties of this T-X family are studied. Besides using functions of moments for measuring skewness and kurtosis, we suggest Galton's measure of skewness and Moors' measure of kurtosis. Three sub-families of T-X family, namely gamma-X family, beta-exponential-X family and Weibull-X family are discussed. These sub-families demonstrate that the T-X family consists of many sub-families of distributions. Within each sub-family, one can define many new distributions as well as relate its members to many existing distributions.

Table 2 summarizes various sub-families based on different T distributions with the same X distribution. New distributions discussed include gamma-Pareto, exponentiated-exponential-logistic and Weibull-logistic distributions. In general, it is difficult to see how the shapes of the T and X distributions will affect the T-X distribution. We believe that a relationship may exist for some specific T and X distributions. For the gamma distribution, α is a shape parameter while β is a scale parameter. For the Pareto distribution, θ is a scale parameter and k is a shape parameter. After forming the gamma-Pareto distribution, θ remains a scale parameter, $\beta/k = c$ becomes a shape parameter. The study of the properties, parameter estimation and applications of these new distributions are currently under investigation. For example, Alzaatreh et al. [2] defined and studied the gamma-Pareto distribution, a member of the gamma-X family. Three real data sets were used to illustrate the applications of the gamma-Pareto distribution. The illustration showed that the gamma-Pareto distribution is a good model to fit data sets with various kinds of shapes.

Figure 7 provides a tree-relationship of the T-X family addressed in this article. As Fig. 7 shows, the T-X family consists of many sub-families, in which new distributions can be defined and various existing distributions are special cases.

The variants of T-X families in Table 1 will define many potential new distributions that deserve further study. Some of these variants are currently under investigation. Future research for the T-X family may include (i) the investigation of general properties of distributions generated using different $W(.)$ functions, (ii) defining and investigating the properties of specific new distributions, (iii) studying new methods for estimating the parameters in addition to the well-known moments and maximum likelihood (ML) methods, and (iv) applying
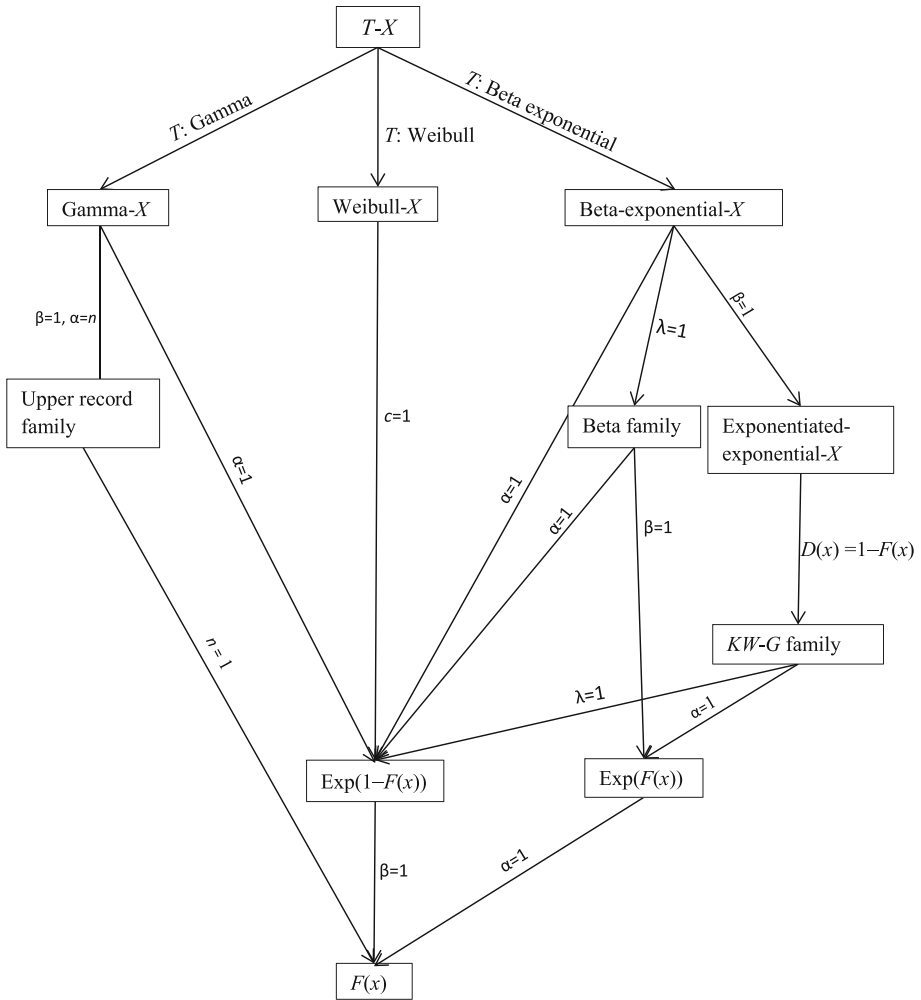
Fig. 7 Sub-families of T-X family of distributions

these new distributions to fit different types of data sets. Based on our experience, the ML method may be challenging for more than three parameters. A better estimation method will be needed for distributions with four or more parameters.

During the recent decade, many new distributions developed in the literature seem to focus on more general and flexible distributions. Using the technique that generates the T-X family, one can develop new distributions that may be very general and flexible or for fitting specific types of data distributions such as highly left-tailed (right-tailed, thin-tailed, or heavy-tailed) distribution as well as bimodal distributions. There are only a few existing distributions that are known to be capable of fitting bimodal shapes. One of such distributions that have been successfully applied to fit real world data sets is the beta-normal distribution by Eugene et al. [10] and Famoye et al. [11]. Our limited investigation in the T-X family suggests that there are new distributions that can fit not only unimodal and bimodal, but also multimodal distributions.

This article focuses on the case when both T and X are continuous random variables. This technique can be extended to develop discrete T-X family of distributions where T is continuous and X is discrete. Different considerations for the $W(.)$ functions will be needed.

References

1. Akinsete, A., Famoye, F., Lee, C.: The beta-Pareto distribution. Statistics **42**, 547–563 (2008)
2. Alzaatreh, A., Famoye, F., Lee, C.: Gamma-Pareto distribution and its applications. J. Mod. Appl. Stat. Methods **11**(1), 78–94 (2012)
3. Amoroso, L.: Ricerche intorno alla curva dei redditi. Annali de Mathematica **2**, 123–159 (1925)
4. Azzalini, A.: A class of distributions which includes the normal ones. Scand. J. Stat. **12**, 171–178 (1985)
5. Burr, I.W.: Cumulative frequency functions. Ann. Math. Stat. **13**, 215–232 (1942)
6. Cordeiro, G.M., de Castro, M.: A new family of generalized distributions. J. Stat. Comput. Simul. **81**(7), 883–898 (2011)
7. Cordeiro, G.M., Ortega, E.M.M., Nadarajah, S.: The Kumaraswamy Weibull distribution with application to failure data. J. Frankl. Inst. **347**, 1399–1429 (2010)
8. Cordeiro, G.M., Pescim, R.R., Ortega, E.M.M.: The Kumaraswamy generalized half-normal distribution for skewed positive data. J. Data Sci. **10**, 195–224 (2012)
9. de Castro, M.A.R., Ortega, E.M.M., Cordeiro, G.M.: The Kumaraswamy generalized gamma distribution with application in survival analysis. Stat. Methodol. **8**(5), 411–433 (2011)
10. Eugene, N., Lee, C., Famoye, F.: The beta-normal distribution and its applications. Commun. Stat. Theory Methods **31**(4), 497–512 (2002)
11. Famoye, F., Lee, C., Eugene, N.: Beta-normal distribution: bimodality properties and applications. J. Mod. Appl. Stat. Methods **3**(1), 85–103 (2004)
12. Famoye, F., Lee, C., Olumolade, O.: The beta-Weibull distribution. J. Stat. Theory Appl. **4**(2), 121–136 (2005)
13. Ferreira, J.T.A.S., Steel, M.F.J.: A constructive representation of univariate skewed distributions. J. Am. Stat. Assoc. **101**(474), 823–829 (2006)
14. Freimer, M., Kollia, G., Mudholkar, G.S., Lin, C.T.: A study of the generalized Tukey lambda family. Commun. Stat. Theory Methods **17**, 3547–3567 (1988)
15. Fry, T.R.L.: Univariate and multivariate Burr distributions: a survey. Pak. J. Stat. Ser. A **9**, 1–24 (1993)
16. Galton, F.: Enquiries into Human Faculty and its Development. Macmillan & Company, London (1883)
17. Gupta, R.D., Kundu, D.: Exponentiated-exponential family: an alternative to gamma and Weibull distributions. Biom. J. **43**, 117–130 (2001)
18. Johnson, N.L.: Systems of frequency curves generated by methods of translation. Biometrika **36**, 149–176 (1949)
19. Johnson, N.L., Kotz, S., Balakrishnan, N.: Continuous Univariate Distributions, vol. 1, 2nd edn. Wiley, New York (1994)
20. Johnson, N.L., Kotz, S., Balakrishnan, N.: Continuous Univariate Distributions, vol. 2, 2nd edn. Wiley, New York (1995)
21. Jones, M.C.: Families of distributions arising from distributions of order statistics. Test **13**(1), 1–43 (2004)
22. Jones, M.C.: Kumaraswamys distribution: a beta-type distribution with tractability advantages. Stat. Methodol. **6**, 70–81 (2009)
23. Karian, Z.A., Dudewicz, E.: Fitting Statistical Distributions—The Generalized Lambda Distribution and Generalized Bootstrap Methods. Chapman & Hall/CRC Press, Boca Raton (2000)
24. Kong, L., Lee, C., Sepanski, J.H.: On the properties of beta-gamma distribution. J. Mod. Appl. Stat. Methods **6**(1), 187–211 (2007)
25. Kumaraswamy, P.: A generalized probability density functions for double-bounded random processes. J. Hydrol. **46**, 79–88 (1980)
26. McDonald, J.B.: Some generalized functions for the size distribution of income. Econometrica **52**, 647–663 (1984)
27. Moors, J.J.: A quantile alternative for kurtosis. Statistician **37**, 25–32 (1988)
28. Mudholkar, G.S., Srivastava, D.K., Freimer, M.: The exponentiated Weibull family: a reanalysis of the bus-motor-failure data. Technometrics **37**(4), 436–445 (1995)

29. Nadarajah, S., Kotz, S.: The beta Gumbel distribution. Math. Probl. Eng. **4**, 323–332 (2004)
30. Nadarajah, S., Kotz, S.: The beta exponential distribution. Reliab. Eng. Syst. Saf. **91**(6), 689–697 (2005)
31. Pearson, K.: Contributions to the mathematical theory of evolution. II. Skew variation in homogeneous material. Philos. Trans. Royal Soc. Lond. A **186**, 343–414 (1895)
32. Ramberg, J.S., Schmeiser, B.W.: An approximate method for generating symmetric random variables. Commun. Assoc. Comput. Mach. **15**, 987–990 (1972)
33. Ramberg, J.S., Schmeiser, B.W.: An approximate method for generating asymmetric random variables. Commun. Assoc. Comput. Mach. **17**, 78–82 (1974)
34. Ramberg, J.S., Tadikamalla, P.R., Dudewicz, E.J., Mykytka, E.F.: A probability distribution and its uses in fitting data. Technometrics **21**, 201–214 (1979)
35. Shannon, C.E.: A mathematical theory of communication. Bell Syst. Tech. J. **27**, 379–432 (1948)
36. Song, K.-S.: Rényi information, log likelihood and an intrinsic distribution measure. J. Stat. Plan. Inference **93**, 51–69 (2001)
37. Tukey, J.W.: The practical relationship between the common transformations of percentages of counts and amounts. Technical Report 36, Statistical Techniques Research Group, Princeton University, Princeton, NJ (1960)