Hacettepe Journal of Mathematics and Statistics
Volume 46 (4) (2017), 645 – 667

Generalized transmuted family of distributions: properties and applications

Morad Alizadeh* †, Faton Merovci‡ and G.G. Hamedani§

Abstract

We introduce and study general mathematical properties of a new generator of continuous distributions with two extra parameters called the *Generalized Transmuted Family of Distributions*. We investigate the shapes and present some special models. The new density function can be expressed as a linear combination of exponentiated densities in terms of the same baseline distribution. We obtain explicit expressions for the ordinary and incomplete moments and generating function, Bonferroni and Lorenz curves, asymptotic distribution of the extreme values, Shannon and Rényi entropies and order statistics, which hold for any baseline model. Further, we introduce a bivariate extension of the new family. We discuss the different methods of estimation of the model parameters and illustrate the potential application of the model via real data. A brief simulation for evaluating Maximum likelihood estimator is done. Finally certain characterziations of our model are presented.

Keywords: Transmuted distribution, Generated family, Maximum likelihood, Moment, Order statistic, Quantile function, Rényi entropy, Characterizations.

2000 AMS Classification: 60E05, 62E10, 62N05

Received : 29.12.2014 *Accepted* : 18.11.2015 *Doi* : 10.15672/HJMS.201610915478

*Department of Statistics, Faculty of Sciences, Persian Gulf University, Bushehr, 75169, Iran
Email: moradalizadeh78@gmail.com

†Corresponding Author.

‡Department of Mathematics, University of Prishtina , Kosovo
Email: fmerovci@yahoo.com

§Department of Mathematics, Statistics and Computer Science,Marquette University, USA
Email: gholamhoss.hamedani@marquette.edu

1. Introduction

In many practical situations, classical distributions do not provide adequate fit to real data. For example, if the data are asymmetric, the normal distribution will not be a good choice. So, several generators based on one or more parameters have been proposed to generate new distributions. Some well-known generators are Marshal-Olkin generated family (MO-G) [33], the beta-G by Eugene et al. [20] , Jones [31], Kumaraswamy-G (Kw-G for short) by Cordeiro and de Castro [16], McDonald-G (Mc-G) by Alexander et al. [1], gamma-G (type 1) by Zografos and Balakrishanan [57], gamma-G (type 2) by Ristić and Balakrishanan [47] , , gamma-G (type 3) by Torabi and Montazari [55], log-gamma-G by Amini et al. [7], logistic-G by Torabi and Montazari [56], exponentiated generalized-G by Cordeiro et al. [18], Transformed-Transformer (T-X) by Alzaatreh et al. [5], exponentiated (T-X) by Alzaghal et al. [6], Weibull-G by Bourguignon et al. [12], Exponentiated half logistic generated family by Cordeiro et al. [15], Kumaraswamy Odd log-logistic by Alizadeh et al. [3], Lomax Generator by Cordeiro et al. [19], a new Weibull-G by Tahir et al. [51], Logistic-X by Tahir et al. [52], Kumaraswamy Marshal-Olkin family by Alizadeh et al. [4], Beta Marshal-OLkin family by Alizadeh et al. [2], type I half-logistic family by Cordeiro et al. [14] and Odd Generalized Exponential family by Tahir et al. [53].

Let $p(t)$ be the probability density function (pdf) of a random variable $T \in [a,b]$ for $-\infty < a < b < \infty$ and let $W\left[G(x)\right]$ be a function of the cumulative distribution function (cdf) of a random variable X such that $W\left[G(x)\right]$ satisfies the following conditions:

$$(1.1) \qquad \begin{cases} (i) & W\left[G(x)\right] \in [a,b], \\ (ii) & W\left[G(x)\right] \text{is differentiable and monotonically non-decreasing, and} \\ (iii) & W\left[G(x)\right] \to a \text{ as } x \to -\infty \text{ and } W\left[G(x)\right] \to b \text{ as } x \to \infty. \end{cases}$$

Recently, Alzaatreh *et al.* [5] defined the *T-X family* of distributions by

$$(1.2) \qquad F(x) = \int_a^{W[G(x)]} p(t)\, dt,$$

where $W\left[G(x)\right]$ satisfies conditions (1.1). The pdf corresponding to (1.2) is given by

$$(1.3) \qquad f(x) = \left\{ \frac{d}{dx} W\left[G(x)\right] \right\} p\left\{ W\left[G(x)\right] \right\}.$$

For $W[G(x)] = [G(x)]^{\alpha}$ and $p(t) = 1 + \lambda - 2\,\lambda\, t \;\; , 0 < t < 1$, we define the cdf of the new *Generalized Transmuted family ("GT-G" for short)* of distributions by

$$\begin{aligned} F(x;\lambda,\alpha,\boldsymbol{\xi}) &= \int_0^{[G(x;\boldsymbol{\xi})]^{\alpha}} (1 + \lambda - 2\,\lambda\, t)\, dt \\ &= (1+\lambda)\,[G(x;\boldsymbol{\xi})]^{\alpha} - \lambda\,[G(x;\boldsymbol{\xi})]^{2\alpha} \\ (1.4) \qquad &= 1 - \left\{1 - \lambda\,[G(x;\boldsymbol{\xi})]^{\alpha}\right\}\left\{1 - [G(x;\boldsymbol{\xi})]^{\alpha}\right\} \quad , \alpha > 0 \;, |\lambda| \le 1, \end{aligned}$$

where $G(x;\boldsymbol{\xi})$ is the baseline cdf depending on a parameter vector $\boldsymbol{\xi}$ and $\alpha > 0$, $|\lambda| \le 1$ are two additional shape parameters. GT-G is a wider class of continuous distributions. It includes the *Transmuted Family* of Distributions and the Proportional Reversed Hazard Rate models. Some special models are given in Table 1.

This paper is organized as follows. In Section 2, we provide motivation for considering GT-G class. Three special cases of this family are defined in Section 3. In Section 4, the shape of the density and hazard rate functions are described analytically. Some useful expansions are derived in Section 5. In Section 6, we propose explicit expressions for the moments, incomplete moments, generating function and mean deviation. General expressions for the Rényi and Shannon entropies are presented in Section 7. General

Table 1. Some known special cases of the GT-G model.

α	λ	$G(x)$	Reduced distribution
1	-	$G(x)$	Transmuted G family of distributions [42]
-	0	$G(x)$	Proportioanl reversed hazard rate family [25]
1	0	$G(x)$	$G(x)$
1	-	exponential	Transmuted exponential distribution [42]
1	-	pareto	Transmuted pareto distribution [41]
1	-	Gumbel	Transmuted Gumbel distribution [10]
1	-	weibull	Transmuted Weibull distribution [11]
1	-	Lindley	Transmuted Lindley distribution [36]
-	-	exponentiated exponentoial	Transmuted exponentiated exponential distribution [34]
1	-	Lindley-geometric	Transmuted Lindley-geometric [39]
1	-	weibull-geometric	Transmuted Weibull-geometric [40]
1	-	Rayleigh	Transmuted Rayleigh distribution [38]
-	-	Generalized Rayleigh	Transmuted Generalized Rayleigh distribution [35]
1	-	extreme value	Transmuted extreme value distribution [10]
1	-	log-logistic	Transmuted log-logistic distribution [9]
1	-	inverse Rayleigh	Transmuted inverse Rayleigh [50]

results for order statistics are obtained in section 8. In Section 9, we introduce a bivariate extension of GT-G. Estimation procedures of the model parameters are performed in Section 10. Application to a real data set illustrating the performance of GT-G is given in Section 11. We give a simulation study in section 12. We present certain characterizations of GT-G family in section 13. The concluding remarks are given in section 14.

2. The new family

The pdf corresponding to (1.4) is given by

$$f(x;\lambda,\alpha,\boldsymbol{\xi}) = \alpha g(x,\boldsymbol{\xi})\left[G(x,\boldsymbol{\xi})\right]^{\alpha-1}\left\{1+\lambda-2\lambda\left[G(x,\boldsymbol{\xi})\right]^{\alpha}\right\}, \tag{2.1}$$

where $g(x;\boldsymbol{\xi})$ is the baseline pdf. Equation (2.1) will be most tractable when the cdf $G(x)$ and the pdf $g(x)$ have simple analytic expressions. Hereafter, a random variable X with density function (2.1) is denoted by $X \sim$ GT-G$(\alpha,\lambda,\boldsymbol{\xi})$. Further, we may omit the dependence on the vector $\boldsymbol{\xi}$ of the parameters and simply write $G(x) = G(x;\boldsymbol{\xi})$.

The hazard rate function (hrf) of X becomes

$$\begin{aligned} h(x;\lambda,\alpha,\boldsymbol{\xi}) &= \frac{\alpha g(x,\boldsymbol{\xi})\left[G(x,\boldsymbol{\xi})\right]^{\alpha-1}\left\{1+\lambda-2\lambda\left[G(x,\boldsymbol{\xi})\right]^{\alpha}\right\}}{1-(1+\lambda)\left[G(x,\boldsymbol{\xi})\right]^{\alpha}+\lambda\left[G(x,\boldsymbol{\xi})\right]^{2\alpha}} \\ &= \frac{\alpha\lambda\, g(x,\boldsymbol{\xi})\left[G(x,\boldsymbol{\xi})\right]^{\alpha-1}}{1-\lambda\left[G(x,\boldsymbol{\xi})\right]^{\alpha}} + \frac{\alpha\, g(x,\boldsymbol{\xi})\left[G(x,\boldsymbol{\xi})\right]^{\alpha-1}}{1-\left[G(x,\boldsymbol{\xi})\right]^{\alpha}}. \end{aligned} \tag{2.2}$$

To motivate the introduction of this new family, let Z_1, Z_2 be *i.i.d* random variables from $\left[G(x;\boldsymbol{\xi})\right]^{\alpha}$ and $Z_{1:2} = \min(Z_1, Z_2)$ $Z_{2:2} = \max(Z_1, Z_2)$. Let

$$V = \begin{cases} Z_{1:2}, & \text{with probability } \frac{1+\lambda}{2}; \\ Z_{2:2}, & \text{with probability } \frac{1-\lambda}{2}, \end{cases}$$

then $F_V(x;\alpha,\lambda,\boldsymbol{\xi}) = (1+\lambda)\left[G(x,\boldsymbol{\xi})\right]^{\alpha} - \lambda\left[G(x,\boldsymbol{\xi})\right]^{2\alpha}$,which is the proposed family. The GT-G family of distributions is easily simulated by inverting (1.4) as follows: if U has a uniform $U(0,1)$ distribution,then

$$X_U = G^{-1}\left\{\left[\frac{1+\lambda-\sqrt{(1+\lambda)^2-4U\lambda}}{2\lambda}\right];\boldsymbol{\xi}\right\} \quad for\ \lambda \neq 0 \tag{2.3}$$

has the density function (2.1).

3. Special GT-G distributions

3.1. The GT-Normal(GT-N) distribution. The GT-N pdf is obtained from (2.1) by taking the normal $N(\mu,\sigma)$ as the parent distribution, where $\xi = (\mu,\sigma)$, so that

$$f_{GT-N}(x;\lambda,\alpha,\mu,\sigma) = \alpha\,\phi\left(\frac{x-\mu}{\sigma}\right)\left[\Phi(\frac{x-\mu}{\sigma})\right]^{\alpha-1}\left\{1+\lambda-2\lambda\left[\Phi(\frac{x-\mu}{\sigma})\right]^{\alpha}\right\}, \tag{3.1}$$

where $x \in \mathbb{R}, \mu \in \mathbb{R}$ is location parameter, σ is a scale parameter and $\phi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of the standard normal distribution, respectively. The standard GT-N is the one with $\mu = 0$ and $\sigma = 1$,which reduces to power normal distribution [32] by taking $\lambda = 0$.

3.2. The GT-Exponential(GT-E) distribution. The GT-E pdf is obtained from (2.1) by taking the cdf $F(x;\gamma) = 1 - exp(-\gamma x)$ as the parent distribution, where $\xi = \gamma$ so that

$$f_{GT-E}(x;\lambda,\alpha,\gamma) = \alpha\gamma\exp(-\gamma x)(1-\exp(-\gamma x))^{\alpha-1}\left[1+\lambda-2\lambda(1-\exp(-\gamma x))^{\alpha}\right]. \tag{3.2}$$

For $\lambda = 0$ we obtain generalized exponential distribution [26] . Figures 1 and 2 illustrate some of the possible shapes of the pdf and cdf of GT-E distribution for selected values of the parameters λ, α and γ, respectively.

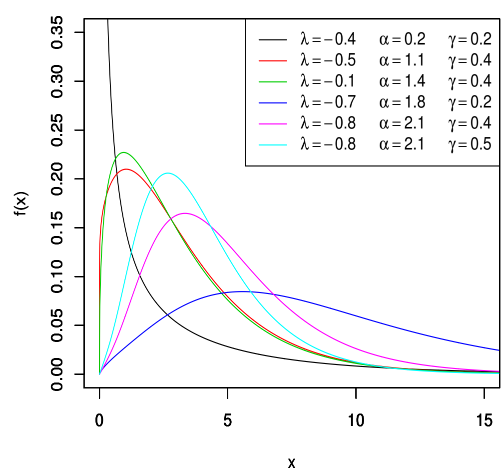


Figure 1. The pdf's of various GTE distributions.

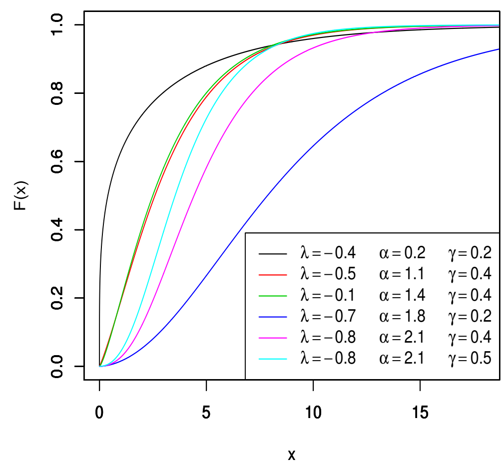


Figure 2. The cdf's of various GTE distributions.

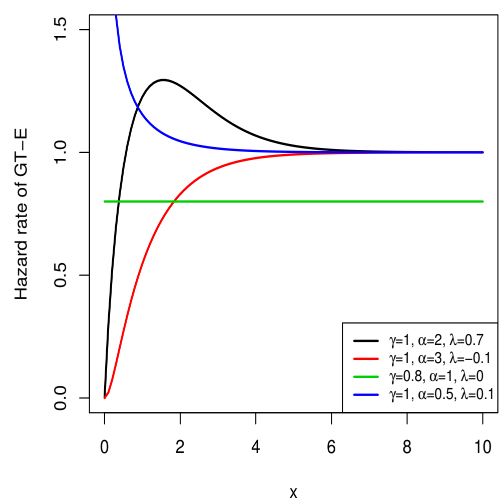


Figure 3. S The hrf's of various GTE distributions.

3.3. The GT-Weibull(GT-W) distribution. The pdf and cdf of Weibull distribution are $f(x;\eta,\sigma) = \frac{\eta}{\sigma}\left(\frac{x}{\sigma}\right)^{\eta-1}\exp\left(-\left(\frac{x}{\sigma}\right)^{\eta}\right), x > 0, \eta > 0, \sigma > 0$ and $F(x;\eta,\sigma) =$

$1-\exp\{-\left(\frac{x}{\sigma}\right)^{\eta}\}, x>0$. In this case, the generalized transmuted Weibull (GT-W) distribution has density

$$\begin{aligned} f_{GT-W}(x;\lambda,\alpha,\eta,\sigma) &= \alpha\frac{\eta}{\sigma}\left(\frac{x}{\sigma}\right)^{\eta-1}\exp\left(-\left(\frac{x}{\sigma}\right)^{\eta}\right)\left(1-\exp\{-\left(\frac{x}{\sigma}\right)^{\eta}\}\right)^{\alpha-1} \\ &\times \left[1+\lambda-2\lambda\left(1-\exp\{-\left(\frac{x}{\sigma}\right)^{\eta}\}\right)^{\alpha}\right]. \end{aligned} \tag{3.3}$$

For $\lambda=0$ we obtain exponentiated Weibull distribution [43] and [44].

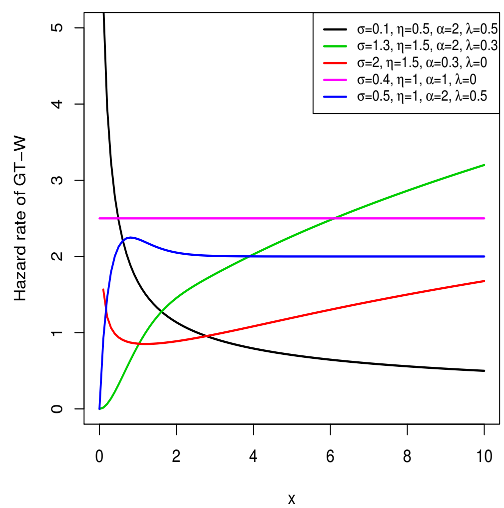


Figure 4. S The hrf's of various GTW distributions.

4. Shapes

The shapes of the density and hazard rate functions can be described analytically. The critical points of the GT-G density function are the roots of the equation:

$$\frac{d\log[f(x)]}{dx}=\frac{g'(x)}{g(x)}+(\alpha-1)\frac{g(x)}{G(x)}-\frac{2\alpha\lambda g(x)\,[G(x)]^{\alpha-1}}{1+\lambda-2\lambda\,[G(x)]^{\alpha}}. \tag{4.1}$$

There may be more than one root to (4.1). If $x=x_0$ is a root of (4.1), then it corresponds to a local maximum, a local minimum or a point of inflection depending on whether $S(x_0)<0$, $S(x_0)>0$ or $S(x_0)=0$, where $S(x)=\frac{d^2\log[f(x)]}{dx^2}$.

The critical points of $r(x)$ are the roots of the equation:

$$\begin{aligned} \frac{d\log(h(x))}{dx}=\frac{g'(x)}{g(x)} &+ \frac{(\alpha-1)g(x)}{G(x)}-\frac{2\alpha\lambda g(x)\,[G(x)]^{\alpha-1}}{1+\lambda-2\lambda\,[G(x)]^{\alpha}} \\ &+ \frac{\alpha g(x,\boldsymbol{\xi})\,[G(x)]^{\alpha-1}\,[1+\lambda-2\lambda\,[G(x)]^{\alpha}]}{1-(1+\lambda)\,[G(x)]^{\alpha}+\lambda\,[G(x)]^{2\alpha}}. \end{aligned} \tag{4.2}$$

There may be more than one root to (4.2). If $x = x_0$ is a root of (4.2) then it corresponds to a local maximum, a local minimum or a point of inflection depending on whether $\varsigma(x_0) < 0$, $\varsigma(x_0) > 0$ or $\varsigma(x_0) = 0$, where $\varsigma(x) = \frac{d^2 \log[h(x)]}{dx^2}$.

5. Useful expansions

First, using generalized binomial expansion, we have

$$\begin{aligned} [G(x)]^{\alpha} &= \{1 - [1 - G(x)]\}^{\alpha} = \sum_{i=0}^{\infty} (-1)^i \binom{\alpha}{i} [1 - G(x)]^i \\ &= \sum_{i=0}^{\infty} \sum_{k=0}^{i} (-1)^{i+k} \binom{\alpha}{i} \binom{i}{k} [G(x)]^k . \end{aligned} \tag{5.1}$$

Changing the summation over k, i we have

$$[G(x)]^{\alpha} = \sum_{k=0}^{\infty} \sum_{i=k}^{\infty} (-1)^{i+k} \binom{\alpha}{i} \binom{i}{k} [G(x)]^k = \sum_{k=0}^{\infty} s_k(\alpha) [G(x)]^k , \tag{5.2}$$

where

$$s_k(\alpha) = \sum_{i=k}^{\infty} (-1)^{i+k} \binom{\alpha}{i} \binom{i}{k}. \tag{5.3}$$

Using generalized binomial expansion, we can show that the cdf (1.4) of X has the expansion

$$F(x) = \sum_{k=0}^{\infty} c_k [G(x)]^k = \sum_{k=0}^{\infty} c_k H_k(x), \tag{5.4}$$

where

$$c_k = (1 + \lambda) s_k(\alpha) - \lambda s_k(2\alpha), \tag{5.5}$$

and $H_a(x) = [G(x)]^a$ and $h_a(x) = a\, g(x) [G(x)]^{a-1}$ denote the exponentiated-G ("exp-G" for short) cumulative distribution and density functions . Some structural properties of the exp-G distribution are studied by Mudholkar [43], Gupta and Kundu [26] and Nadarajah and Kotz [46], among others.

The density function of X can be expressed as an infinite linear combination of exp-G density functions

$$f(x; \alpha, \lambda, \boldsymbol{\xi}) = \sum_{k=0}^{\infty} c_{k+1} h_{k+1}(x). \tag{5.6}$$

Thus, some mathematical properties of the new model can be derived from those of the exp-G distribution. For example, the ordinary and incomplete moments and moment generating function (mgf) of X can be obtained from those of the exp-G distribution.

The formulae derived throughout the paper can be easily handled in most symbolic computation software plataforms such as Maple, Mathematica and Matlab. These plataforms have currently the ability to deal with analytic expressions of formidable size and complexity. Established explicit expressions to calculate statistical measures can be more efficient than computing them directly by numerical integration. The infinity limit in these sums can be substituted by a large positive integer such as 20 or 30 for most practical purposes.

6. Some measures

6.1. Moments. Let Y_k be a random variable having an exp-G distribution with power parameter $k+1$, i.e., with density $h_{k+1}(x)$. A first formula for the nth moment of $X \sim$GT-G follows from (5.6) as

$$E(X^n) = \sum_{k=0}^{\infty} c_{k+1}\, E(Y_k^n). \tag{6.1}$$

Expressions for moments of several exp-G distributions are given by Nadarajah and Kotz [45], which can be used to obtained $E(X^n)$.

A second formula for $E(X^n)$ which can be written from (6.1) in terms of the G quantile function as

$$E(X^n) = \sum_{k=0}^{\infty} (k+1)\, c_{k+1}\, \tau(n,k), \tag{6.2}$$

where $\tau(n,k) = \int_{-\infty}^{\infty} x^n\, [G(x)]^k\, g(x)dx = \int_0^1 [Q_G(u)]^n\, u^a du$. Cordeiro and Ndarajah [17] obtained $\tau(n,k)$ for some well known distribution such as Normal, Beta, Gamma and Weibull distributions, which can be used to find moments of GT-G.

For empirical purposes, the shape of many distributions can be usefully described by what we call the incomplete moments. These types of moments play an important role in measuring inequality, for example, income quantiles and Lorenz and Bonferroni curves, which depend upon the incomplete moments of a distribution. The nth incomplete moment of X is calculated as

$$m_n(y) = E(X^n | X < y) = \sum_{k=0}^{\infty} (k+1)\, c_{k+1} \int_0^{G(y)} [Q_G(u)]^n\, u^k du. \tag{6.3}$$

The last integral can be computed for most G distributions.

6.2. Moment generating function. Let $M_X(t) = E(\mathrm{e}^{\mathrm{t\,X}})$ be mgf of $X \sim$GT-G, then the first formula from (5.6) is

$$M_X(t) = \sum_{k=0}^{\infty} c_{k+1}\, M_k(t), \tag{6.4}$$

where $M_k(t)$ is the mgf of Y_k. Hence, $M_X(t)$ can be determined from the exp-G mgf.

A second formula for $M_X(t)$ can be derived from (5.6) as

$$M_X(t) = \sum_{i=0}^{\infty} (k+1)\, c_{k+1}\, \rho(t,k), \tag{6.5}$$

where $\rho(t,k) = \int_{-\infty}^{\infty} \mathrm{e}^{\mathrm{t\,x}}\, [\mathrm{G(x)}]^{\mathrm{k}}\, \mathrm{g(x)dx} = \int_0^1 \exp[\mathrm{t\,Q_G(u)}]\, \mathrm{u^k du}$.

We can obtain the mgfs of several distributions directly from equation (6.5).

6.3. Mean deviation. The mean deviation about the mean ($\delta_1 = E(|X - \mu_1'|)$) and about the median ($\delta_2 = E(|X - M|)$) of X can be expressed as

$$\delta_1(X) = 2\mu_1'\, F\left(\mu_1'\right) - 2m_1\left(\mu_1'\right) \quad \text{and} \quad \delta_2(X) = \mu_1' - 2m_1(M), \tag{6.6}$$

respectively, where $\mu_1' = E(X)$, $M = Median(X)$, the median, defined by $M = Q(0.5)$, $F(\mu_1')$ is easily calculated from the cdf (1.4) and $m_1(z) = \int_{-\infty}^{z} x f(x)dx$ is the first incomplete moment obtained from (6.3) with $n = 1$.

Now, we provide two alternative ways to compute δ_1 and δ_2. A general equation for $m_1(z)$ can be derived from (5.6) as

$$m_1(z) = \sum_{k=0}^{\infty} c_{k+1}\, J_k(z), \tag{6.7}$$

where $J_k(z) = \int_{-\infty}^{z} x\, h_{k+1}(x)dx$ is the basic quantity to compute the mean deviation for the exp-G distribution. Hence, the mean deviation in (6.6) depends only on the mean deviations of the exp-G distribution. So, alternative representations for δ_1 and δ_2 are

$$\delta_1(X) = 2\mu_1' F\left(\mu_1'\right) - 2\sum_{k=0}^{\infty} c_{k+1}\, J_k(\mu_1') \qquad \text{and} \qquad \delta_2(X) = \mu_1' - 2\sum_{k=0}^{\infty} c_{k+1}\, J_k(M).$$

A simple application of $J_k(z)$ refers to the the GT-G distribution discussed in Section 3.3. The exponentiated Weibull with parameter $k+1$ has pdf (for $x>0$) given by

$$h_{k+1}(x) = \frac{(k+1)\eta}{\sigma^\eta}\, x^{\eta-1}\, \exp\left[-(\frac{x}{\sigma})^\eta\right]\left\{1-\exp\left[-(\frac{x}{\sigma})^\eta\right]\right\}^k,$$

and then

$$\begin{aligned} J_k(z) &= \frac{(k+1)\eta}{\sigma^\eta}\int_0^z x^\eta\, \exp\left[-(\frac{x}{\sigma})^\eta\right]\left\{1-\exp\left[-(\frac{x}{\sigma})^\eta\right]\right\}^k dx \\ &= \frac{(k+1)\eta}{\sigma^\eta}\sum_{r=0}^{k}(-1)^r\binom{k}{r}\int_0^z x^\eta\, \exp\left[-(r+1)(\frac{x}{\sigma})^\eta\right]. \end{aligned}$$

The last integral is just the incomplete gamma function and mean deviation for the GT-G distribution can be determined from

$$m_1(z) = \sum_{k=0}^{\infty}\sum_{r=0}^{k} \frac{(k+1)b_{k+1}\,(-1)^r\binom{k}{r}}{(r+1)^{1+\eta^{-1}}\,\sigma^{2\eta+1}}\,\gamma(1+\eta^{-1},(r+1)(\frac{z}{\sigma})^\eta).$$

A second general formula for $m_1(z)$ can be derived by setting $u = G(x)$ in (5.6)

$$m_1(z) = \sum_{k=0}^{\infty}(k+1)\, c_{k+1}\, T_k(z), \tag{6.8}$$

where $T_k(z) = \int_0^{G(z)} Q_G(u)\, u^k du$ is a simple integral defined via the baseline quantile function.

Remarks 1. These equations can be used to obtain Bonferroni and Lorenz curves defined for a given probability π by

$$B(\pi) = \frac{T(q)}{\pi\mu_1'} \qquad \text{and} \qquad L(\pi) = \frac{T(q)}{\mu_1'},$$

respectively, where $\mu_1' = E(X)$ and $q = Q(\pi)$ is the quantile function of X at π.

7. Entropies

An entropy is a measure of variation or uncertainty of a random variable X. Two popular entropy measures are the Rényi and Shannon entropies [48] , [49]. The Rényi entropy of a random variable with pdf $f(x)$ is defined as

$$I_R(\gamma) = \frac{1}{1-\gamma}\log\left(\int_0^\infty f^\gamma(x)dx\right),$$

for $\gamma > 0$ and $\gamma \neq 1$. The Shannon entropy of a random variable X is defined by $E\{-\log[f(X)]\}$. It is the special case of the Rényi entropy when $\gamma \uparrow 1$. Direct calculation yields

$$\begin{aligned} \mathrm{E}\{-\log[f(X)]\} &= -\log(\alpha) - \mathrm{E}\{\log[g(X;\boldsymbol{\xi})]\} + (1-\alpha)\,\mathrm{E}\{\log[G(X;\boldsymbol{\xi})]\} \\ &- \mathrm{E}\{\log[1+\lambda-2\lambda G(X;\boldsymbol{\xi})^{\alpha}]\}. \end{aligned}$$

First we define and compute

$$A(a_1,a_2;\lambda,\alpha) = \int_0^1 x^{a_1}(1-\frac{2\lambda}{1+\lambda}x^{\alpha})^{a_2}dx. \tag{7.1}$$

Using generalized binomial expansion and then after some algebraic manipulations, we obtain

$$A(a_1,a_2;\lambda,\alpha) = \sum_{i=0}^{\infty}\frac{(-1)^i\binom{a_2}{i}}{a_1+\alpha\, i+1}\,(\frac{2\lambda}{1+\lambda})^i.$$

7.1. Proposition. *Let X be a random variable with pdf* (2.1). *Then,*

$$\mathrm{E}\{\log[G(X)]\} = \frac{\alpha}{1+\lambda}\,\frac{\partial}{\partial t}A(\alpha+t-1,1;\lambda,\alpha)\big|_{t=0},$$

$$\mathrm{E}\{\log[1+\lambda-2\lambda\,[G(X)]^{\alpha}]\} = \frac{\alpha}{1+\lambda}\,\frac{\partial}{\partial t}\,\frac{1}{(1+\lambda)^t}\,A(\alpha-1,t+1;\lambda,\alpha)\big|_{t=0}.$$

The simplest formula for the entropy of X is given by

$$\begin{aligned} \mathrm{E}\{-\log[f(X)]\} &= -\log(\alpha) - \mathrm{E}\{\log[g(X;\boldsymbol{\xi})]\} \\ &+ (1-\alpha)\frac{\alpha}{1+\lambda}\,\frac{\partial}{\partial t}A(\alpha+t-1,1;\lambda,\alpha)\big|_{t=0} \\ &- \frac{\alpha}{1+\lambda}\,\frac{\partial}{\partial t}\,\frac{1}{(1+\lambda)^t}\,A(\alpha-1,t+1;\lambda,\alpha)\big|_{t=0}. \end{aligned}$$

After some algebraic manipulations, we obtain an alternative expression for $I_R(\gamma)$

$$I_R(\gamma) = \frac{\gamma}{1-\gamma}\log(\frac{\alpha}{1+\lambda}) + \frac{1}{1-\gamma}\log\left\{\sum_{i=0}^{\infty} w_i^*\,\mathrm{E}_{Y_i}[g^{\gamma-1}[G^{-1}(Y)]]\right\}, \tag{7.2}$$

where $Y_i \sim Beta(\gamma(\alpha-1)+\alpha i+1,1)$ and

$$w_i^* = \frac{(-1)^i\binom{\gamma}{i}}{\gamma(\alpha-1)+\alpha i+1}\,(\frac{2\lambda}{1+\lambda})^i.$$

8. Order statistics

Order statistics have been employed in many areas of statistical theory and practice. Suppose $X_1,\ldots,X_n$ is a random sample from the GT-G distribution. We can write the pdf of the ith order statistic, say $X_{i:n}$, as

$$f_{i:n}(x) = K\,f(x)\,F^{i-1}(x)\,\{1-F(x)\}^{n-i} = K\sum_{j=0}^{n-i}(-1)^j\binom{n-i}{j}f(x)\,F(x)^{j+i-1},$$

where $K = n!/[(i-1)!\,(n-i)!]$.

Following similar algebraic manipulations , we can write the density function of $X_{i:n}$ as

$$f_{i:n}(x) = \sum_{r,k=0}^{\infty} m_{r,k}\, h_{r+k+1}(x), \tag{8.1}$$

where $h_{r+k+1}(x)$ is the exp-G density function with power parameter $r+k+1$,

$$m_{r,k} = \frac{n!\,(r+1)\,(i-1)!\,c_{r+1}}{(r+k+1)} \sum_{j=0}^{n-i} \frac{(-1)^j\, f_{j+i-1,k}}{(n-i-j)!\,j!},$$

and c_k is defined in equation (5.5). Here, the quantities $f_{j+i-1,k}$ are obtained recursively, $f_{j+i-1,0} = c_0^{j+i-1}$ and (for $k \geq 1$)

$$f_{j+i-1,k} = (k\,c_0)^{-1} \sum_{m=1}^{k} [m(j+i) - k]\, c_m\, f_{j+i-1,k-m}.$$

The last equation is the main result of this section. It reveals that the pdf of the GT-G order statistics is a linear combination of exp-G density functions. So, several mathematical quantities of the GT-G order statistics such as ordinary, incomplete and factorial moments, mgf, mean deviation and several others can be obtained from those of the exp-G distribution.

9. Bivariate extention

In this section we introduce a bivariate version of proposed model. A joint cdf is given by

$$F_{X,Y}(x,y) = (1+\lambda)\,[G(x,y;\boldsymbol{\xi})]^{\alpha} - \lambda\,[G(x,y;\boldsymbol{\xi})]^{2\alpha}, \tag{9.1}$$

where $G(x,y;\boldsymbol{\xi})$ is a bivariate continuous distribution with mariginal cdf's $G_1(x;\boldsymbol{\xi})$ and $G_2(y;\boldsymbol{\xi})$. We call this distribution by *Bivariate Generalized Termuted G* (BGT-G) distribution. The marginal cdf's are given by

$$F_X(x) = (1+\lambda)\,[G_1(x;\boldsymbol{\xi})]^{\alpha} - \lambda\,[G_1(x;\boldsymbol{\xi})]^{2\alpha} \quad \text{and}$$
$$F_Y(y) = (1+\lambda)\,[G_2(y;\boldsymbol{\xi})]^{\alpha} - \lambda\,[G_2(y;\boldsymbol{\xi})]^{2\alpha}.$$

The joint pdf of (X,Y) is easily obtained from $f_{X,Y}(x,y) = \dfrac{\partial^2 F_{X,Y}(x,y)}{\partial x \partial y}$

$$f_{X,Y}(x,y) \quad = \quad \alpha A(x,y;\alpha,\lambda\boldsymbol{\xi})\,[G(x,y;\boldsymbol{\xi})]^{\alpha-1}\left\{1+\lambda-2\lambda\,[G(x,y;\boldsymbol{\xi})]^{\alpha}\right\},$$

where

$$\begin{aligned} A(x,y;\lambda,\alpha,\boldsymbol{\xi}) \quad &= \quad g(x,y;\boldsymbol{\xi}) + \frac{\partial G(x,y,\boldsymbol{\xi})}{\partial x}\frac{\partial G(x,y,\boldsymbol{\xi})}{\partial y} \times \frac{\alpha-1}{G(x,y,\boldsymbol{\xi})} \\ &- \quad \frac{\partial G(x,y,\boldsymbol{\xi})}{\partial x}\frac{\partial G(x,y,\boldsymbol{\xi})}{\partial y} \times \frac{2\alpha\lambda\, G(x,y,\boldsymbol{\xi})^{\alpha-1}}{1+\lambda-2\lambda\, G(x,y,\boldsymbol{\xi})^{\alpha}} \end{aligned}$$
.

The marginal pdf's are

$$f_X(x) = \alpha g_1(x,\boldsymbol{\xi})\,[G_1(x;\boldsymbol{\xi})]^{\alpha-1}\left\{1+\lambda-2\lambda\,[G_1(x;\boldsymbol{\xi})]^{\alpha}\right\},$$

and

$$f_Y(y) = \alpha g_2(y,\boldsymbol{\xi})\,[G_2(y;\boldsymbol{\xi})]^{\alpha-1}\left\{1+\lambda-2\lambda\,[G_2(y;\boldsymbol{\xi})]^{\alpha}\right\}.$$

The conditional cdf's are

$$F_{X|Y}(x|y) \quad = \quad \frac{(1+\lambda)\,[G(x,y;\boldsymbol{\xi})]^{\alpha} - \lambda\,[G(x,y;\boldsymbol{\xi})]^{2\alpha}}{(1+\lambda)\,[G_2(y;\boldsymbol{\xi})]^{\alpha} - \lambda\,[G_2(y;\boldsymbol{\xi})]^{2\alpha}},$$

and

$$F_{Y|X}(y|x) = \frac{(1+\lambda)\,[G(x,y;\boldsymbol{\xi})]^{\alpha} - \lambda\,[G(x,y;\boldsymbol{\xi})]^{2\alpha}}{(1+\lambda)\,[G_1(x;\boldsymbol{\xi})]^{\alpha} - \lambda\,[G_1(x;\boldsymbol{\xi})]^{2\alpha}}.$$

The conditional density functions are

$$f_{X|Y}(x|y) = \frac{A(x,y;\alpha,\lambda\boldsymbol{\xi})\,[G(x,y;\boldsymbol{\xi})]^{\alpha-1}\,[1+\lambda-2\lambda\,[G(x,y;\boldsymbol{\xi})]^{\alpha}]}{g_2(y,\boldsymbol{\xi})\,[G_2(y;\boldsymbol{\xi})]^{\alpha-1}\,[1+\lambda-2\lambda\,[G_2(y;\boldsymbol{\xi})]^{\alpha}]},$$

and

$$f_{Y|X}(y|x) = \frac{A(x,y;\alpha,\lambda\boldsymbol{\xi})\,[G(x,y;\boldsymbol{\xi})]^{\alpha-1}\,[1+\lambda-2\lambda\,[G(x,y;\boldsymbol{\xi})]^{\alpha}]}{g_1(x,\boldsymbol{\xi})\,[G_1(x;\boldsymbol{\xi})]^{\alpha-1}\,[1+\lambda-2\lambda\,[G_1(x;\boldsymbol{\xi})]^{\alpha}]}.$$

10. Estimation procedures

10.1. Maximum likelihood estimates. Here, we obtain the maximum likelihood estimates (MLEs) of the parameters of the GT-G distribution from complete samples only. Let $x_1,\ldots,x_n$ be observed values from the GT-G distribution with parameters α,λ and $\boldsymbol{\xi}$. Let $\Theta=(\alpha,\lambda,\boldsymbol{\xi})^{\top}$ be the $r\times 1$ parameter vector. The total log-likelihood function for Θ is given by

$$\begin{aligned}\ell_n &= \ell_n(\Theta) = n\,\log(\alpha) + \sum_{i=1}^{n}\log[g(x_i;\boldsymbol{\xi})] + (\alpha-1)\sum_{i=1}^{n}\log[G(x_i;\boldsymbol{\xi})] \\ &+ \sum_{i=1}^{n}\log[1+\lambda-2\lambda\,[G(x_i;\boldsymbol{\xi})]^{\alpha}]\,. \end{aligned} \tag{10.1}$$

These non-linear equations can be routinely solved using Newton's method or fixed point iteration techniques. The subroutines to solve non-linear optimization problem are available in R ([54]) software namely *optim*, *nlm* and *bbmle* etc. We used *nlm* package for optimizing (10.1) . The log-likelihood function can be maximized either directly by using the SAS (PROC NLMIXED) or the Ox program (sub-routine MaxBFGS) or by solving the nonlinear likelihood equations obtained by differentiating (10.1). The components of the score function $U_n(\Theta)=(\partial\ell_n/\partial\alpha,\partial\ell_n/\partial\lambda,\partial\ell_n/\partial\boldsymbol{\xi})^{\top}$ are

$$\frac{\partial\ell_n}{\partial\lambda} = \frac{n}{\alpha} + \sum_{i=1}^{n}\log[G(x_i;\boldsymbol{\xi})] - 2\lambda\sum_{i=1}^{n}\frac{[G(x_i;\boldsymbol{\xi})]^{\alpha}\log[G(x_i;\boldsymbol{\xi})]}{1+\lambda-2\lambda\,[G(x_i;\boldsymbol{\xi})])^{\alpha}},$$

$$\frac{\partial\ell_n}{\partial\lambda} = \sum_{i=1}^{n}\frac{1-2\,[G(x_i;\boldsymbol{\xi})]^{\alpha}}{1+\lambda-2\lambda\,[G(x_i;\boldsymbol{\xi})]^{\alpha}},$$

and

$$\frac{\partial\ell_n}{\partial\boldsymbol{\xi}} = \sum_{i=1}^{n}\frac{g'_{\boldsymbol{\xi}}(x_i,\boldsymbol{\xi})}{g(x_i,\boldsymbol{\xi})} + (\alpha-1)\sum_{i=1}^{n}\frac{G'_{\boldsymbol{\xi}}(x_i,\boldsymbol{\xi})}{G(x_i,\boldsymbol{\xi})} - 2\alpha\lambda\sum_{i=1}^{n}\frac{G'_{\boldsymbol{\xi}}(x_i,\boldsymbol{\xi})\,[G(x_i;\boldsymbol{\xi})]^{\alpha-1}}{1+\lambda-2\lambda\,[G(x_i;\boldsymbol{\xi})]^{\alpha}},$$

where $h'_{\boldsymbol{\xi}}(\cdot)$ means the derivative of the function h with respect to ξ.

We can compute the maximized unrestricted and restricted log-likelihood functions to construct the likelihood ratio (LR) test statistic for testing on some transmuted Rayleigh sub-models. For example, we can use the LR test statistic to check whether a transmuted Rayleigh distribution for a given data set is statistically *superior* to the Rayleigh distribution. In any case, hypothesis tests of the type $H_0:\theta=\theta_0$ versus $H_0:\theta\neq\theta_0$ can be performed using a LR test. In this case, the LR test statistic for testing H_0 versus H_1 is $\omega=2(\ell(\hat{\theta};x)-\ell(\hat{\theta}_0;x))$, where $\hat{\theta}$ and $\hat{\theta}_0$ are the MLEs under H_1 and H_0, respectively.

The statistic ω is asymptotically (as $n \to \infty$) distributed as χ^2_k, where k is the length of the parameter vector θ of interest. The LR test rejects H_0 if $\omega > \chi^2_{k;\gamma}$, where $\chi^2_{k;\gamma}$ denotes the upper $100\gamma\%$ quantile of the χ^2_k distribution.

10.2. Maximum product spacing estimates. The maximum product spacing (MPS) method has been proposed by [13]. This method is based on an idea that the differences (spacings) of the consecutive points should be identically distributed. The geometric mean of the differences is given as

$$(10.2)\quad GM = \sqrt[n+1]{\prod_{i=1}^{n+1} D_i},$$

where the difference D_i is defined as

$$(10.3)\quad D_i = \int_{x_{(i-1)}}^{x_{(i)}} f\left(x, \lambda, \alpha, \boldsymbol{\xi}\right) dx; \quad i = 1, 2, \ldots, n+1.$$

Note that $F(x_{(0)}, \lambda, \alpha, \xi) = 0$ and $F(x_{(n+1)}, \lambda, \alpha, \boldsymbol{\xi}) = 1$. The MPS estimators $\hat{\alpha}_{PS}, \hat{\beta}_{PS}$ and $\hat{\boldsymbol{\xi}}_{PS}$ of λ, α and $\boldsymbol{\xi}$ are obtained by maximizing the geometric mean (GM) of the differences. Substituting pdf of GT-G in (10.3) and taking logarithm of the above expression, we have

$$(10.4)\quad LogGM = \frac{1}{n+1} \sum_{i=1}^{n+1} \log\left[F(x_{(i)}, \lambda, \alpha, \boldsymbol{\xi}) - F(x_{(i-1)}, \lambda, \alpha, \boldsymbol{\xi})\right].$$

The MPS estimators $\hat{\lambda}_{PS}, \hat{\alpha}_{PS}$ and $\hat{\boldsymbol{\xi}}_{PS}$ of λ, α and $\boldsymbol{\xi}$ can be obtained as the simultaneous solution of the following non-linear equations:

$$\frac{\partial LogGM}{\partial \lambda} = \frac{1}{n+1} \sum_{i=1}^{n+1} \left[\frac{F'_{\lambda}(x_{(i)}, \lambda, \alpha, \boldsymbol{\xi}) - F'_{\lambda}(x_{(i-1)}, \lambda, \alpha, \boldsymbol{\xi})}{F(x_{(i)}, \lambda, \alpha, \boldsymbol{\xi}) - F(x_{(i-1)}, \lambda, \alpha, \boldsymbol{\xi})}\right] = 0$$

$$\frac{\partial LogGM}{\partial \alpha} = \frac{1}{n+1} \sum_{i=1}^{n+1} \left[\frac{F'_{\alpha}(x_{(i)}, \lambda, \alpha, \boldsymbol{\xi}) - F'_{\alpha}(x_{(i-1)}, \lambda, \alpha, \boldsymbol{\xi})}{F(x_{(i)}, \lambda, \alpha, \boldsymbol{\xi}) - F(x_{(i-1)}, \lambda, \alpha, \boldsymbol{\xi})}\right] = 0$$

$$\frac{\partial LogGM}{\partial \boldsymbol{\xi}} = \frac{1}{n+1} \sum_{i=1}^{n+1} \left[\frac{F'_{\boldsymbol{\xi}}(x_{(i)}, \lambda, \alpha, \boldsymbol{\xi}) - F'_{\boldsymbol{\xi}}(x_{(i-1)}, \lambda, \alpha, \boldsymbol{\xi})}{F(x_{(i)}, \lambda, \alpha, \boldsymbol{\xi}) - F(x_{(i-1)}, \lambda, \alpha, \boldsymbol{\xi})}\right] = 0,$$

where

$$F'_{\lambda}(x_i; \boldsymbol{\xi}) = G(x_i; \boldsymbol{\xi})^{\alpha}\left[1 - G(x_i; \boldsymbol{\xi})^{\alpha}\right]$$

$$F'_{\alpha}(x_i; \boldsymbol{\xi}) = G(x_i; \boldsymbol{\xi})^{\alpha}\left[1 + \lambda - 2\lambda\, G(x_i; \boldsymbol{\xi})^{\alpha}\right] \log\left[G(x_i; \boldsymbol{\xi})\right]$$

$$(10.5)\quad F'_{\boldsymbol{\xi}}(x_i; \boldsymbol{\xi}) = \alpha\, G'_{\boldsymbol{\xi}}(x_i; \boldsymbol{\xi})\, G(x_i; \boldsymbol{\xi})^{\alpha-1}\left[1 + \lambda - 2\lambda\, G(x_i; \boldsymbol{\xi})^{\alpha}\right].$$

10.3. Least square estimates. Let $x_{(1)}, x_{(2)}, \ldots, x_{(n)}$ be the ordered sample of size n from GT-G distribution. Then, the expectation of the empirical cumulative distribution function is defined as

$$(10.6)\quad E\left[F\left(X_{(i)}\right)\right] = \frac{i}{n+1}; i = 1, 2, \ldots, n.$$

The least square estimates (LSEs) $\hat{\lambda}_{LS}, \hat{\alpha}_{LS}$ and $\hat{\boldsymbol{\xi}}_{LS}$ of λ, α and $\boldsymbol{\xi}$ are obtained by minimizing

$$Z\left(\lambda, \alpha, \boldsymbol{\xi}\right) = \sum_{i=1}^{n} \left(F\left(x_{(i)}, \lambda, \alpha, \boldsymbol{\xi}\right) - \frac{i}{n+1}\right)^2.$$

Therefore, $\hat{\lambda}_{LS}, \hat{\alpha}_{LS}$ and $\hat{\boldsymbol{\xi}}_{LS}$ of λ, α and $\boldsymbol{\xi}$ can be obtained as the solution of the following system of equations:

$$\frac{\partial Z(\lambda,\alpha,\boldsymbol{\xi})}{\partial\lambda} = \sum_{i=1}^{n} F_{\lambda}^{'}(x_{(i)},\lambda,\alpha,\boldsymbol{\xi})\left(F\left(x_{(i)},\lambda,\alpha,\boldsymbol{\xi}\right) - \frac{i}{n+1}\right) = 0$$

$$\frac{\partial Z(\lambda,\alpha,\boldsymbol{\xi})}{\partial\alpha} = \sum_{i=1}^{n} F_{\alpha}^{'}(x_{(i)},\lambda,\alpha,\boldsymbol{\xi})\left(F\left(x_{(i)},\lambda,\alpha,\boldsymbol{\xi}\right) - \frac{i}{n+1}\right) = 0$$

$$\frac{\partial Z(\lambda,\alpha,\boldsymbol{\xi})}{\partial\boldsymbol{\xi}} = \sum_{i=1}^{n} F_{\boldsymbol{\xi}}^{'}(x_{(i)},\lambda,\alpha,\boldsymbol{\xi})\left(F\left(x_{(i)},\alpha,\theta\right) - \frac{i}{n+1}\right) = 0.$$

11. Applications

Now we use a real data set to show that the GT-E can be a better model than the beta-exponential([45]), Kumaraswamy-exponential distributions and exponential distribtuion.

We consider a data set of the life of fatigue fracture of Kevlar 373/epoxy that are subject to constant pressure at the 90% stress level until all had failed, so we have complete data with the exact times of failure. For the previous studies with the data sets see Andrews & Herzberg [8]. These data are:
0.0251, 0.0886, 0.0891, 0.2501, 0.3113, 0.3451, 0.4763, 0.5650, 0.5671, 0.6566, 0.6748, 0.6751, 0.6753, 0.7696, 0.8375, 0.8391, 0.8425, 0.8645, 0.8851, 0.9113, 0.9120, 0.9836, 1.0483, 1.0596, 1.0773, 1.1733, 1.2570, 1.2766, 1.2985, 1.3211, 1.3503, 1.3551, 1.4595, 1.4880, 1.5728, 1.5733, 1.7083, 1.7263, 1.7460, 1.7630, 1.7746, 1.8275, 1.8375, 1.8503, 1.8808, 1.8878, 1.8881, 1.9316, 1.9558, 2.0048, 2.0408, 2.0903, 2.1093, 2.1330, 2.2100, 2.2460, 2.2878, 2.3203, 2.3470, 2.3513, 2.4951, 2.5260, 2.9911, 3.0256, 3.2678, 3.4045, 3.4846, 3.7433, 3.7455, 3.9143, 4.8073, 5.4005, 5.4435, 5.5295, 6.5541, 9.0960.

Table 2. Estimated parameters of the GT-E, BE and KwE distributions for data set.

Model	ML Estimate	Standard Error	$-\ell(\cdot;x)$	LSE	PS Estimator
GT-Exponential	$\hat{\lambda} = -0.733$	0.274	121.3219	-0.636	-0.760
	$\hat{\alpha} = 1.197$	0.344		1.631	1.038
	$\hat{\gamma} = 0.769$	0.101		0.907	0.704
Beta	$\hat{a} = 1.679$	0.374	122.227	2.235	1.520
Exponential	$\hat{b} = 1.508$	6.760		1.558	1.082
	$\hat{\lambda} = 0.484$	1.981		0.586	0.598
Kumaraswamy	$\hat{a} = 1.556$	0.401	122.0942	1.987	1.426
Exponential	$\hat{b} = 2.448$	6.065		2.228	2.243
	$\hat{\lambda} = 0.328$	0.691		0.453	0.316
Exponential	$\hat{\lambda} = 0.510$	0.058	127.114	0.981	0.926
GT-Weibull	$\hat{\lambda} = -0.656$	0.340	195.133	-0.635	-0.751
	$\hat{\alpha} = 1.729$	1.065		1.644	1.269
	$\hat{\eta} = 0.864$	0.201		0.996	0.905
	$\hat{\sigma} = 1.397$	0.700		1.095	1.193

The variance covariance matrix $I(\hat{\varphi})^{-1}$ of the MLEs under the GT-E distribution is computed as

$$\begin{pmatrix} 0.075472979 & 0.07213909 & 0.005236484 \\ 0.072139088 & 0.11889578 & 0.021543094 \\ 0.005236484 & 0.02154309 & 0.010248121 \end{pmatrix}.$$

Thus, the variances of the MLE of λ, α and γ are $var(\hat{\lambda}) = 0.075472979, var(\hat{\alpha}) = 0.11889578$ and $var(\hat{\gamma}) = 0.010248121$.

Therefore, 95% confidence intervals for λ, α and γ are $[-1, -0.195], [0.521, 1.873]$ and $[0.571, 0.968]$ respectively.

The LR test statistic to test the hypotheses $H_0 : a = b = 1$ versus $H_1 : a \neq 1 \vee b \neq 1$ for data set is $\omega = 11.584 > 5.991 = \chi^2_{2;0.05}$, so we reject the null hypothesis.

Table 3. Criteria for comparison.

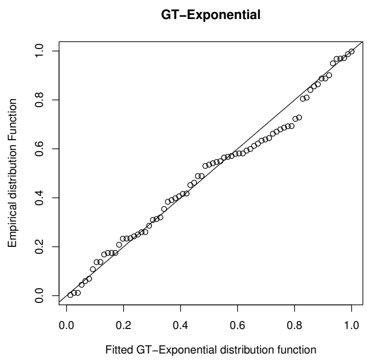
Model	K-S	-2ℓ	AIC	CAIC	BIC
GT-E	0.0956	242.643	248.643	249.143	255.636
Beta-E	0.0962	244.455	250.455	250.621	257.447
Kw-E	0.0988	244.188	250.188	250.521	257.180
Exponential	0.512	254.228	248.643	249.143	258.559
GT-Weibull	0.267	390.266	398.266	398.766	408.687

Further, we also applied the Statistical tools for model comparison such as Kolmogorov-Smirnov (K-S) statistics, Akaike information criterion (AIC) and Bayesian information criterion (BIC) to choose the best possible model for the data set among the competitive models. The statistical tools used are described as follows:

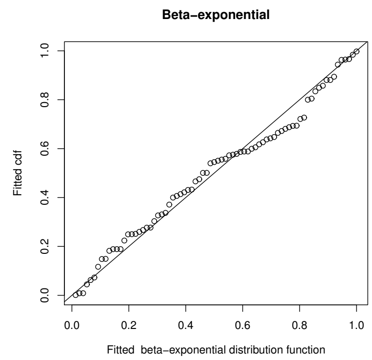
- K-S distance $D_n = \sup_x |F(x) - F_n(x)|$, where, $F_n(x)$ is the empirical distribution function,
- $AIC = -2 \log \ell \left(\underset{\sim}{x}, \alpha, \lambda, \boldsymbol{\xi} \right) + 2p$,
- $BIC = -2 \log \ell \left(\underset{\sim}{x}, \alpha, \lambda, \boldsymbol{\xi} \right) + p \log (n)$,

  where, p is the number of parameters which are to be estimated from the data.

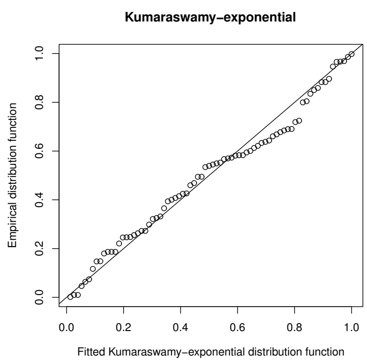
The selection criterion is that the lowest AIC and BIC correspond to the best fit model. Thus, the generalized transmuted exponential distribution provides the best fit for the data set as it shows the lowest AIC and BIC than other considered models, see Table 3. The P-P plots, fitted distribution functions and density functions of the considered models are plotted in Figures 5a, 5b, 5c, 5d, 5e, 6a, 6b for the data set.



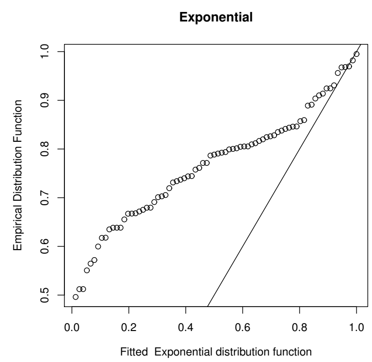
(a) The P-P plots of the GTE distribution for the real data set.



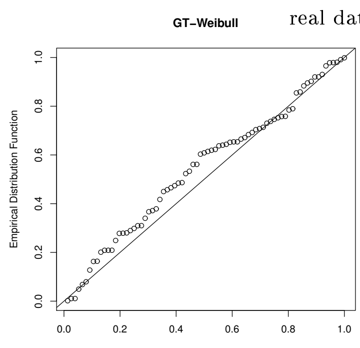
(b) The P-P plots of the BE distribution for the real data set.



(c) The P-P plots of the KE distribution for the real data set.

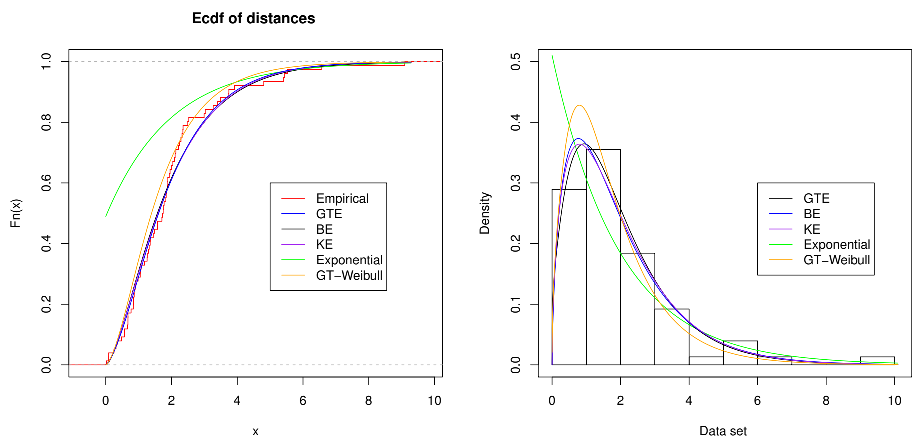


(d) The P-P plots of the Exponential distribution for the real data set.



(e) The P-P plots of the Exponential distribution for the real data set.

Figure 5



(a) Fitted cdfs plots of the considered distribution for the real data set.

(b) Fitted pdfs plots of the considered distribution for the real data set.

Figure 6

12. Simulation algorithm and study

12.1. Inverse CDF method. In this subsection we provide an algorithm to generate a random sample of size n from the $GT-E$ $(GT-E(\lambda,\alpha,\theta))$ distribution for the given values of its parameters and sample size n . The simulation process consists of the following steps:

Step 1. Set n, and $\Theta=(\lambda,\alpha,\gamma)$.
Step 2. Set initial value x_0 for the random starting.
Step 3. Set $j=1$.
Step 4. Generate $U\sim Uniform\,(0,1)$.
Step 5. Update x_0 by using the Newton's formula such as
$$x^\star=x_0-\left(\frac{G_\Theta(x)-U}{f_\Theta(x)}\right)\Big|_{x=x_0}$$
Step 6. If $|\,x_0-x^\star\,|\leq\epsilon$, (very small, $\epsilon>0$ tolerance limit). Then, $x^\star$ will be the desired sample from $F(x)$.
Step 7. If $|\,x_0-x^\star\,|>\epsilon$, then, set $x_0=x^\star$ and go to step 5.
Step 8. Repeat steps 4-7, for $j=1,2,\ldots,n$ and obtained $x_1,x_2,\ldots,x_n$.

12.2. Simulated data. This subsection deals with the comparisons of the proposed estimators in terms of their mean square error on the basis of simulated sample from pdf of GT-E with varying sample sizes. For this purpose, we take $\lambda=0.5\,,\alpha=2\,,\gamma=0.4$ arbitrarily and $n=10,20,\ldots,100$. All the algorithms are coded in R. We calculate MLEs, estimators of $\lambda\,,\alpha$ and γ based on each generated sample. This process is repeated 1000 of times, and average estimates and corresponding mean as reported in Table 4.

From Table 4, it can be clearly observed that as sample size increases the mean square error decreases, it proves the consistency of the estimators.

Table 4. Estimates and mean square errors (in 2nd row of each cell) of the proposed estimators with varying sample size

	λ	γ	α
10	0.5842	0.6901	2.0532
	1.3252	1.2395	0.0873
20	0.5693	0.59866	2.0302
	0.3873	0.8933	0.0422
30	0.5573	0.5086	2.0181
	0.2412	0.4582	0.0269
40	0.5215	0.4952	2.0146
	0.1678	0.3119	0.0202
50	0.5095	0.4734	2.0123
	0.1345	0.2282	0.0169
60	0.5084	0.4554	2.0106
	0.0996	0.2079	0.0137
70	0.5071	0.4443	2.0101
	0.0856	0.1575	0.0118
80	0.5049	0.4258	2.0100
	0.0694	0.1310	0.0100
90	0.5024	0.4204	2.0104
	0.0618	0.0981	0.0090
100	0.5024	0.4153	2.0087
	0.0504	0.0884	0.0083

13. Characterizations of GT-G distribution

In designing a stochastic model for a particular modeling problem, an investigator will be vitally interested to know if their model fits the requirements of a specific underlying probability distribution. To this end, the investigator will rely on the characterizations of the selected distribution. Generally speaking, the problem of characterizing a distribution is an important problem in various fields and has recently attracted the attention of many researchers. Consequently, various characterization results have been reported in the literature. These characterizations have been established in many different directions. In this Section, we present characterizations of GT-G distribution. These characterizations are based on a simple relationship between two truncated moments. We like to mention here the works of Glänzel [22], [23] and [21] , Glänzel and Hamedani [24] and Hamedani [28], [29]. in this direction. Our characterization results presented here will employ an interesting result due to Glänzel [22] (Theorem 1 below). The advantage of our main characterization given here is that, *cdf* F need not have a closed form and is given in terms of an integral whose integrand depends on the solution of a first order differential equation, which can serve as a bridge between probability and differential equation.

Theorem 1. Let $(\Omega, \Sigma, \mathbf{P})$ be a given probability space and let $I = [a, b]$ be an interval for some $a < b$ $(a = -\infty\ ,\ b = \infty$ might as well be allowed$)$. Let $X : \Omega \to I$ be a continuous random variable with the distribution function F and let q_1 and q_2 be two real functions defined on I such that

$$\mathbf{E}\left[q_1(X)\ \mid X \geq x\right] = \mathbf{E}\left[q_2(X)\ \mid X \geq x\right]\eta(x), \qquad x \in I\ ,$$

is defined with some real function η . Assume that $q_1, q_2 \in C^1(I)$, $\eta \in C^2(I)$ and F is twice continuously differentiable and strictly monotone function on the set I . Finally, assume that the equation $\eta q_2 = q_1$ has no real solution in the interior of I . Then F is uniquely determined by the functions q_1, q_2 and η , particularly

$$F(x) = \int_a^x C \left| \frac{\eta'(u)}{\eta(u) q_2(u) - q_1(u)} \right| \exp(-s(u)) \ du \ ,$$

where the function s is a solution of the differential equation $s' = \frac{\eta' \ q_2}{\eta \ q_2 \ - \ q_1}$ and C is a constant, chosen to make $\int_I dF = 1$.

We like to mention that this kind of characterization based on the ratio of truncated moments is stable in the sense of weak convergence, in particular, let us assume that there is a sequence $\{X_n\}$ of random variables with distribution functions $\{F_n\}$ such that the functions $q_{1,n}$, $q_{2,n}$ and η_n $(n \in \mathbb{N})$ satisfy the conditions of Theorem 1 and let $q_{1,n} \to q_1$, $q_{2,n} \to q_2$ for some continuously differentiable real functions q_1 and q_2 . Let, finally, X be a random variable with distribution F . Under the condition that $q_{1,n}(X)$ and $q_{2,n}(X)$ are uniformly integrable and the family $\{F_n\}$ is relatively compact, the sequence X_n converges to X in distribution if and only if η_n converges to η , where

$$\eta(x) = \frac{E[q_1(X) \mid X \geq x]}{E[q_2(X) \mid X \geq x]}.$$

This stability theorem makes sure that the convergence of distribution functions is reflected by corresponding convergence of the functions q_1 , q_1 and η , respectively. It guarantees, for instance, the 'convergence' of characterization of the Wald distribution to that of the Lévy-Smirnov distribution if $\alpha \to \infty$, as was pointed out in Glänzel and Hamedani [24].

A further consequence of the stability property of Theorem 1 is the application of this theorem to special tasks in statistical practice such as the estimation of the parameters of discrete distributions. For such purpose, the functions q_1 , q_2 and, specially, η should be as simple as possible. Since the function triplet is not uniquely determined it is often possible to choose η as a linear function. Therefore, it is worth analyzing some special cases which helps to find new characterizations reflecting the relationship between individual continuous univariate distributions and appropriate in other areas of statistics.

Remarks 2. (a) In Theorem 1, the interval I need not be closed since the condition is only on the interior of I. (b) Clearly, Theorem 1 can be stated in terms of two functions q_1 and η by taking $q_2(x) \equiv 1$, which will reduce the condition given in Theorem 1 to $E[q_1(X) \mid X \geq x] = \eta(x)$. However, adding an extra function will give a lot more flexibility, as far as its application is concerned.

Proposition 4. Let $X : \Omega \to \mathbb{R}$ be a continuous random variable and let $q_2(x) = [1 + \lambda - 2\lambda (G(x;\xi))^{\alpha}]^{-1}$ and $q_1(x) = q_2(x) G(x;\xi)$ for $x \in \mathbb{R}$. The *pdf* of X is (5) if and only if the function η defined in Theorem 1 has the form

$$\eta(x) = \frac{\alpha}{\alpha+1}\left(G(x;\xi)\right), \qquad x \in \mathbb{R}.$$

Proof. Let X have *pdf* (5) , then

$$(1 - F(x))\ \mathbf{E}\left[q_2(X)\ \mid X \geq x\right] = \left(G(x;\xi)\right)^{\alpha}, \qquad x \in \mathbb{R},$$

and

$$(1 - F(x))\ \mathbf{E}\left[q_1(X)\ \mid X \geq x\right] = \frac{\alpha}{\alpha+1}\left(G(x;\xi)\right)^{\alpha+1}, \qquad x \in \mathbb{R}$$

and finally

$$\eta(x)\, q_2(x) - q_1(x) = -\frac{1}{\alpha+1} q_2(x)\, G(x;\xi) < 0\ , \qquad x \in \mathbb{R}\ .$$

Conversely, if η is given as above, then

$$s'(x) = \frac{\eta'(x)\ q_2(x)}{\eta(x)\ q_2(x) - q_1(x)} = -\frac{\alpha g(x;\xi)}{G(x;\xi)}\ , \qquad x \in \mathbb{R}\ ,$$

and hence

$$s(x) = -\log(\left(G(x;\xi)\right)^{\alpha})\ , \qquad x \in \mathbb{R}.$$

Now, in view of Theorem 1, X has *pdf* (5) .

Corollary 1. Let $X : \Omega \to \mathbb{R}$ be a continuous random variable and let $q_2(x)$ be as in Proposition 1. The *pdf* of X is (5) if and only if there exist functions q_1 and η defined in Theorem 1 satisfying the differential equation

$$\frac{\eta'(x)\, q_2(x)}{\eta(x)\, q_2(x) - q_1(x)} = -\frac{\alpha g(x;\xi)}{G(x;\xi)}\ , \qquad x \in \mathbb{R}.$$

Remarks 3. (a) The general solution of the differential equation in Corollary 1 is

$$\eta(x) = \left(G(x;\xi)\right)^{-\alpha}\left[\int \alpha q_1(x)\, g(x;\xi)\left(G(x;\xi)\right)^{\alpha-1}\left(q_2(x)\right)^{-1} dx + D\right],$$

for $x \in \mathbb{R}$, where D is a constant. One set of appropriate functions is given in Proposition 1 with $D = 0$.

(b) Clearly there are other triplets of functions (q_1, q_2, η) satisfying the conditions of Theorem 1. We presented one such triplet in Proposition 4.

14. Concluding remarks

We introduce and study a new class of distributions called the Generalized Transumted-G (GT-G) family, which include as special cases, Transmuted-G and Proprtional Reversed Hazard Rate family. For each baseline distribution G, we define the corresponding GT-G distribution with two additional shape parameters using simple formulas to extend widely-known models such as the normal, exponential and Weibull distributions in order to provide more flexibility. Some characteristics of the new family, such as ordinary moments, generating function and mean deviations, have tractable mathematical properties. We estimate the parameters using maximum likelihood. An application to real data demonstrate the importance of the new family. Finally, certain characterizations of GT-G distributions are presented.

References

[1] Alexander, C., Cordeiro, G. M., Ortega, E. M., & Sarabia, J. M. (2012). Generalized beta-generated distributions. *Computational Statistics & Data Analysis*, **56**(6), 1880-1897.

[2] Alizadeh, M., Cordeiro, G. M., de Brito, E., & Demétrio, C. G. B. (2015). The beta Marshall-Olkin family of distributions. *Journal of Statistical Distributions and Applications*, **2**(1), 1-18.

[3] Alizadeh, M., Emadi, M., Doostparast, M., Cordeiro, G. M., Ortega, E. M., & Pescim, R. R. (2015). A new family of distributions: the Kumaraswamy odd log-logistic, properties and applications. *Hacettepe Journal of Mathematics and Statistics* (to appear).

[4] Alizadeh, M., Tahir, M. H., Cordeiro, G. M., Mansoor, M., Zubair, M., & Hamedani, G. G. (2015). The Kumaraswamy Marshal-Olkin family of distributions. *Journal of the Egyptian Mathematical Society.*

[5] Alzaatreh, A., Lee, C., & Famoye, F. (2013). A new method for generating families of continuous distributions. *Metron*, **71**(1), 63-79.

[6] Alzaghal, A., Famoye, F., & Lee, C. (2013). Exponentiated T-X Family of Distributions with Some Applications. *International Journal of Statistics and Probability*, **2**(3), p31.

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

[8] Andrews, D. F., Herzberg, A. M. (2012). Data: a collection of problems from many fields for the student and research worker. *Springer Science Business Media.*

[9] Aryal, G. R. (2013). Transmuted log-logistic distribution. *Journal of Statistics Applications & Probability*, **2**(1), 11-20.

[10] Aryal, G. R., & Tsokos, C. P. (2009). On the transmuted extreme value distribution with application. *Nonlinear Analysis: Theory, Methods & Applications*, **71**(12), e1401-e1407.

[11] Aryal, G. R., & Tsokos, C. P. (2011). Transmuted Weibull Distribution: A Generalization of theWeibull Probability Distribution. *European Journal of Pure and Applied Mathematics*, **4**(2), 89-102.

[12] Bourguignon, M., Silva, R. B., & Cordeiro, G. M. (2014). The Weibull–G family of probability distributions. *Journal of Data Science*, **12**(1), 53-68.

[13] Cheng, R. C. H., & Amin, N. A. K. (1983). Estimating parameters in continuous univariate distributions with a shifted origin. *Journal of the Royal Statistical Society. Series B (Methodological)*, 394-403.

[14] Cordeiro, G. M., Alizadeh, M., & Diniz Marinho, P. R. (2015). The type I half-logistic family of distributions. *Journal of Statistical Computation and Simulation*, (ahead-of-print), 1-22.

[15] Cordeiro, G. M., Alizadeh, M., & Ortega, E. M. (2014). The exponentiated half-logistic family of distributions: Properties and applications. *Journal of Probability and Statistics*, 2014.

[16] Cordeiro, G. M., & de Castro, M. (2011). A new family of generalized distributions. *Journal of Statistical Computation and Simulation*, **81**(7), 883-898.

[17] Cordeiro, G. M., & Nadarajah, S. (2011). Closed-form expressions for moments of a class of beta generalized distributions.*Brazilian journal of probability and statistics*, **25**(1), 14-33.

[18] Cordeiro, G. M., Ortega, E. M., & da Cunha, D. C. (2013). The exponentiated generalized class of distributions. *Journal of Data Science*, **11**(1), 1-27.

[19] Cordeiro, G. M., Ortega, E. M., Popović, B. V., & Pescim, R. R. (2014). The Lomax generator of distributions: Properties, minification process and regression model. *Applied Mathematics and Computation*, **247**, 465-486.

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

[21] Glänzel, W., Teles, A., Schubert, A. (1984). Characterization by truncated moments and its application to Pearson-type distributions. *Zeitschrift fur Wahrscheinlichkeitstheorie und verwandte Gebiete*, **66**(2), 173-183. Chicago

[22] Glänzel, W. (1987). A characterization theorem based on truncated moments and its application to some distribution families. In Mathematical statistics and probability theory (pp. 75-84). *Springer Netherlands.*

[23] Glänzel, W. (1990). Some consequences of a characterization theorem based on truncated moments. *Statistics*, **21**(4), 613-618.

[24] Glänzel, W., Hamedani, G. G. (2001). Characterization of univariate continuous distributions. *Studia Scientiarum Mathematicarum Hungarica*, **37**(1-2), 83-118

[25] Gupta, R. C., & Gupta, R. D. (2007). Proportional reversed hazard rate model and its applications. *Journal of Statistical Planning and Inference*, **137**(11), 3525-3536.

[26] Gupta, R. D., & Kundu, D. (1999). Theory & methods: Generalized exponential distributions. *Australian & New Zealand Journal of Statistics*, **41**(2), 173-188.

[27] Gupta, R. D., & Kundu, D. (2001). Generalized exponential distribution: different method of estimations. *Journal of Statistical Computation and Simulation*, **69**(4), 315-337.

[28] Hamedani, G. G. (2002). Characterizations of univariate continuous distributions. II. *Studia Scientiarum Mathematicarum Hungarica*, **39**(3-4), 407-424.

[29] Hamedani, G. G. (2006). Characterizations of univariate continuous distributions, III. *Studia Scientiarum Mathematicarum Hungarica*, **43**(3), 361-385.

[30] Hamedani, G. (2009). Characterizations of continuous univariate distributions based on the truncated moments of functions of order statistics. *Studia Scientiarum Mathematicarum Hungarica*, **47**(4), 462-484.

[31] Jones, M. C. (2004). Families of distributions arising from distributions of order statistics. *Test*, **13**(1), 1-43.

[32] Kundu, D., & Gupta, R. D. (2013). Power-normal distribution. *Statistics*, **47**(1), 110-125.

[33] Marshall, A. W., & Olkin, I. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. *Biometrika*, **84**(3), 641-652.

[34] Merovci, F. (2013). Transmuted exponentiated exponential distribution. *Mathematical Sciences And Applications E-Notes*, **1**(2):112–122.

[35] Merovci, F. (2014). Transmuted generalized Rayleigh distribution. *Journal of Statistics Applications and Probability*, **3**(1), 9-20.

[36] Merovci, F. (2013). Transmuted lindley distribution. *International Journal of Open Problems in Computer Science Mathematics*, **6**.

[37] Merovci, F., & Elbatal, I. (2013). Transmuted Lindley-geometric distribution and its applications. arXiv preprint arXiv:1309.3774.

[38] Merovci, F. (2013). Transmuted Rayleigh distribution. *Austrian Journal of Statistics*, **42**(1), 21-31.

[39] Merovci, F., & Elbatal, I. (2013). Transmuted Lindley-geometric distribution and its applications. arXiv preprint arXiv:1309.3774.

[40] Merovci, F., & Elbatal, I. (2014). Transmuted Weibull-geometric distribution and its applications. *School of Mathematics Northwest University Xi'an, Shaanxi, PR China*, **10**(1), 68-82.

[41] Merovci, F., & Pukab, L. (2014). Transmuted Pareto Distribution. *ProbStat*, **7**(1):1–11.

[42] Mirhossaini, S. M., & Dolati, A. (2008). On a New Generalization of the Exponential Distribution. *Journal of Mathematical Extension.*

[43] Mudholkar, G. S., & Hutson, A. D. (1996). The exponentiated Weibull family: some properties and a flood data application. *Communications in Statistics–Theory and Methods*, **25**(12), 3059-3083.

[44] Mudholkar, G. S., Srivastava, D. K., & Kollia, G. D. (1996). A generalization of the Weibull distribution with application to the analysis of survival data. *Journal of the American Statistical Association*, **91**(436), 1575-1583.

[45] Nadarajah, S., & Kotz, S. (2006). The beta exponential distribution. *Reliability engineering & system safety*, **91**(6), 689-697.

[46] Nadarajah, S., & Kotz, S. (2006). The exponentiated type distributions. *Acta Applicandae Mathematica*, **92**(2), 97-111.

[47] Ristić, M. M., & Balakrishnan, N. (2012). The gamma-exponentiated exponential distribution. *Journal of Statistical Computation and Simulation*, **82**(8), 1191-1206.

[48] Rényi, A. L. F. R. P. E. D. (1961). On measures of entropy and information. *In Fourth Berkeley symposium on mathematical statistics and probability* (Vol. **1**, pp. 547-561).

[49] Shannon, C. E. (2001). A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, **5**(1), 3-55.

[50] Sharma, V. K., Singh, S. K., & Singh, U. (2014). A new upside-down bathtub shaped hazard rate model for survival data analysis. *Applied Mathematics and Computation*, **239**, 242-253.

[51] Tahir, M. H., Zubair, M., Mansoor, M., Cordeiro, G. M., Alizadeh, M., Hamedani, G. G. (2015). A new Weibull-G family of distributions. *Hacet. J. Math. Stat.* forthcoming.

[52] Tahir, M. H., Cordeiro, G. M., Alzaatreh, A. Y. M. A. N., Mansoor, M., & Zubair, M. (2015). The Logistic-X family of distributions and its applications. *Commun. Stat. Theory Methods* (2015a). forthcoming.

[53] Tahir, M. H., Cordeiro, G. M., Alizadeh, M., Mansoor, M., Zubair, M., Hamedani, G. G. (2015). The odd generalized exponential family of distributions with applications. *Journal of Statistical Distributions and Applications*, **2**(1), 1-28.

[54] Team, R. C. (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2012.

[55] Torabi, H., & Hedesh, N. M. (2012). The gamma-uniform distribution and its applications. *Kybernetika*, **48**(1), 16-30.

[56] Zografos, K., & Balakrishnan, N. (2009). On families of beta-and generalized gamma-generated distributions and associated inference. *Statistical Methodology*, **6**(4), 344-362.

[57] Zografos, K., Balakrishnan, N. (2009). On families of beta and generalized gamma generated distributions and associated inference. *Statistical Methodology,* **6**(4), 344-362.