Austrian Journal of Statistics
MMMMMM YYYY, Volume VV, 1–xx.
http://www.ajs.or.at/

The Complementary Generalized Transmuted Poisson-G Family of Distributions

M. Alizadeh Persian Gulf Univ.
H. M. Yousof Benha Univ.
A. Z. Afify Benha Univ.
G. M. Cordeiro Univ. Federal de Pernambuco
M. Mansoor Islamia Univ. of Bahawalpur

Abstract

We introduce a new class of continuous distributions called the complementary generalized transmuted Poisson-G family, which extends the transmuted class pioneered by Shaw and Buckley (2007). We provide some special models and derive general mathematical properties including quantile function, explicit expressions for the ordinary and incomplete moments, generating function, Rényi and Shannon entropies and order statistics. The estimation of the model parameters is performed by maximum likelihood. The flexibility of the new family is illustrated by means of two applications to real data sets.

Keywords: Entropy, Generating Function, Maximum Likelihood, Order Statistic, Poisson-G family, Transmuted Family..

1. Introduction

In the last few decades, there have been an increased interest among statisticians to define new generators of univariate distributions by adding one or more shape parameter(s) to a baseline distribution. The extra parameters can provide great flexibility for modelling data in several applied areas such as reliability, engineering, economics, biological studies, environmental and medical sciences. The extra parameters of a good generator can usually give lighter tails and heavier tails, accommodate unimodal, bimodal, symmetric, bimodal and right-skewed and bimodal and left-skewed density function, increase and decrease skewness and kurtosis and, more important, yield the four types of the hazard function (increasing, decreasing, bathtub and unimodal). There are several well-known generators such as the following ones: the Marshall-Olkin-G by Marshall and Olkin (1997), beta-G by Eugene et al. (2002), Kumaraswamy-G by Cordeiro and de Castro (2011), McDonald-G by Alexander et al. (2012), gamma-G by Zografos and Balakrishanan (2009), Kumaraswamy odd log-logistic-G by Alizadeh et al. (2015), beta odd log-logistic generalized by Cordeiro et al. (2015), transmuted exponentiated generalized-G by Yousof et al. (2015), generalized transmuted-G by Nofal et al. (2016), transmuted geometric-G by Afify et al. (2016a), Kumaraswamy transmuted-G by Afify et al. (2016b), beta transmuted-H by Afify et al. (2016c), Burr X-G by Yousof et al. (2016) and odd-Burr generalized-G by Alizadeh et al. (2016) families, among others.

Consider a baseline cumulative distribution function (cdf) $G(x;\xi)$ and probability density function (pdf) $g(x;\xi)$ depending on a parameter vector ξ, where $\xi = (\xi_k) = (\xi_1, \xi_2, \ldots)$. Thus, the cdf and pdf of the *transmuted-G* (TG) family of distributions are defined by

$$F\left(x;\lambda,\xi\right) = G\left(x;\xi\right)\left[1+\lambda-\lambda G\left(x;\xi\right)\right] \tag{1}$$

and

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

respectively. Henceforth, let $h_\delta(x) = \delta\, g(x)\, G(x)^{\delta-1}$ be the exponentiated-G (Exp-G) density with power parameter $\delta > 0$. Clearly, the TG density is a linear mixture of the baseline density and Exp-G density with power parameter two. If $\lambda = 0$ in (2), we obtain the baseline distribution. Further details were explored by Shaw and Buckley (2007).

Let $p(t)$ be the 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 cdf of a random variable X such that the following conditions hold:

$$\begin{cases} (i) & W\left[G(x)\right] \in [a,b], \\ (ii) & W\left[G(x)\right] \text{ is differentiable and monotonically non-decreasing, and} \\ (iii) & W\left[G(x)\right] \rightarrow a \text{ as } x \rightarrow -\infty \text{ and } W\left[G(x)\right] \rightarrow b \text{ as } x \rightarrow \infty. \end{cases} \tag{3}$$

Recently, Alzaatreh et al. (2013) defined the cdf of the *T-X family* of distributions by

$$F(x) = \int_a^{W[G(x)]} p(t)\,dt, \tag{4}$$

where $W[G(x)]$ satisfies the conditions (3). The pdf corresponding to 4 is given by

$$f(x) = \left\{\frac{d}{dx}\,W\left[G(x)\right]\right\}\,p\left(W\left[G(x)\right]\right). \tag{5}$$

Based on the complementary power series distribution (Flores el al., 2013), the cdf and pdf of the *transmuted complementary Poisson* (TCP) distribution are given by

$$F(x) = \frac{\mathrm{e}^{\theta\left(1-\mathrm{e}^{-\beta\,x}\right)}-1}{\mathrm{e}^{\theta}-1}\left[1+\lambda\,\frac{\mathrm{e}^{\theta}-\mathrm{e}^{\theta\left(1-\mathrm{e}^{-\beta\,x}\right)}}{\mathrm{e}^{\theta}-1}\right]$$

and

$$f(x) = \frac{\theta\beta\,\mathrm{e}^{-\beta\,x}\mathrm{e}^{\theta\left(1-\mathrm{e}^{-\beta\,x}\right)}}{\mathrm{e}^{\theta}-1}\left[1-\lambda+2\lambda\,\frac{\mathrm{e}^{\theta}-\mathrm{e}^{\theta\left(1-\mathrm{e}^{-\beta\,x}\right)}}{\mathrm{e}^{\theta}-1}\right],$$

respectively, where $\theta > 0$, $\beta > 0$ and $|\lambda| \leq 1$.

Let $W[G(x)] = -\log[1-G(x;\boldsymbol{\xi})]$ and $p(t)$ be the pdf of the TCP model with unity scale. We define the cdf of the new *complementary generalized transmuted Poisson-G* (CGTP-G) family by

$$\begin{aligned} F(x) &= 1-\int_0^{-\log[G(x;\boldsymbol{\xi})]}\frac{\theta\,\mathrm{e}^{-\mathrm{t}}\mathrm{e}^{\theta\left(1-\mathrm{e}^{-\mathrm{t}}\right)}}{\mathrm{e}^{\theta}-1}\left[1-\lambda+2\lambda\,\frac{\mathrm{e}^{\theta}-\mathrm{e}^{\theta\left(1-\mathrm{e}^{-\mathrm{t}}\right)}}{\mathrm{e}^{\theta}-1}\right]dt \\ &= \frac{\mathrm{e}^{\theta}-\mathrm{e}^{\theta\overline{\mathrm{G}}(\mathrm{x};\boldsymbol{\xi})}}{\mathrm{e}^{\theta}-1}\left[1+\lambda\frac{\mathrm{e}^{\theta\overline{\mathrm{G}}(\mathrm{x};\boldsymbol{\xi})}-1}{\mathrm{e}^{\theta}-1}\right], \end{aligned} \tag{6}$$

where $G(x;\boldsymbol{\xi})$ is the baseline cdf depending on a parameter vector $\boldsymbol{\xi}$, $\overline{G}(x;\boldsymbol{\xi}) = 1-G(x;\boldsymbol{\xi})$, $\theta > 0$ and $|\lambda| \leq 1$ are two additional shape parameters. The CGTP-G family is a wider class of continuous distributions. It includes the TG family when $\theta \rightarrow 0$. The main advantage of the new family relies on the fact that practitioners will have a quite flexible two-parameter generator to fit real data from several fields. We provide a comprehensive account of some of its mathematical properties.

The rest of the paper is organized as follows. In Section 2, we define the CGTP-G family. In Section 3, we present two special models and plots of their pdfs and hazard rate functions (hrfs). We give a very useful linear representation for the family density function in Section 4. In Section 5, we derive some of its general mathematical properties including asymptotics, ordinary and incomplete moments, quantile and generating functions, residual life and reversed residual life functions and entropies. In Section 6, we investigate the order statistics and their moments. Maximum likelihood estimation of the model parameters is addressed in Section 7. Simulation results to assess the performance of the maximum likelihood estimation method are reported in Section 8. In Section 9, we provide two applications to real data to illustrate the flexibility of the new family. Finally, we offer some concluding remarks in Section 10.

2. The CGTP-G family

The pdf corresponding to (6) is given by

$$f(x)=\frac{\theta g(x;\xi)\mathrm{e}^{\theta\overline{\mathrm{G}}(\mathrm{x};\xi)}}{\mathrm{e}^{\theta}-1}\left[1-\lambda\frac{\left(\mathrm{e}^{\theta}+1\right)-2\mathrm{e}^{\theta\overline{\mathrm{G}}(\mathrm{x};\xi)}}{\mathrm{e}^{\theta}-1}\right]. \tag{7}$$

We denote by X ∼CGTP-$G(\lambda,\theta,\xi)$ a random variable having density function (7). The reliability function (rf), hrf and cumulative hazard rate function (chrf) of X are, respectively, given by

$$R(x)=1-\frac{\mathrm{e}^{\theta}-\mathrm{e}^{\theta\overline{\mathrm{G}}(\mathrm{x};\xi)}}{\mathrm{e}^{\theta}-1}\left[1+\lambda\frac{\mathrm{e}^{\theta\overline{\mathrm{G}}(\mathrm{x};\xi)}-1}{\mathrm{e}^{\theta}-1}\right], \tag{8}$$

$$\tau(x)=\frac{\theta g(x;\xi)\mathrm{e}^{\theta\overline{\mathrm{G}}(\mathrm{x};\xi)}\left[1-\lambda\frac{\left(\mathrm{e}^{\theta}+1\right)-2\mathrm{e}^{\theta\overline{\mathrm{G}}(\mathrm{x};\xi)}}{\mathrm{e}^{\theta}-1}\right]}{\left[\mathrm{e}^{\theta\overline{\mathrm{G}}(\mathrm{x};\xi)}-1\right]\left[1+\lambda\frac{\mathrm{e}^{\theta\overline{\mathrm{G}}(\mathrm{x};\xi)}-1}{\mathrm{e}^{\theta}-1}\right]}$$

and

$$H(x)=-\log\left\{1-\frac{\mathrm{e}^{\theta}-\mathrm{e}^{\theta\overline{\mathrm{G}}(\mathrm{x};\xi)}}{\mathrm{e}^{\theta}-1}\left[1+\lambda\frac{\mathrm{e}^{\theta\overline{\mathrm{G}}(\mathrm{x};\xi)}-1}{\mathrm{e}^{\theta}-1}\right]\right\}.$$

A simple motivation for this family follows if we consider independent identically distributed (iid) random variables $Z_1,\ldots,Z_N$ having common cdf $G(x)$ and let N be a random variable with probability mass function (pmf)

$$P(N=n)=\frac{1}{(\mathrm{e}^{\theta}-1)}\frac{\theta^n}{n!},\quad n=1,2,\ldots,\ \theta>0.$$

Next, we define $M_N=\min(Z_1,\ldots,Z_N)$. The cdf of M_N reduces to

$$\Pi(x)=\sum_{n=1}^{\infty}P(M_N\le x|N=n)P(N=n)=\frac{\mathrm{e}^{\theta}-\mathrm{e}^{\theta\overline{\mathrm{G}}(\mathrm{x};\xi)}}{\mathrm{e}^{\theta}-1}.$$

Further, let Y_1 and Y_2 be iid random variables with cdf $\Pi(x)$ and define $Y_{1:2}=\min(Y_1,Y_2)$, $Y_{2:2}=\max(Y_1,Y_2)$ and

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

Then,

$$F_V(x;\theta,\lambda,\boldsymbol{\xi})=\frac{\mathrm{e}^{\theta}-\mathrm{e}^{\theta\overline{\mathrm{G}}(\mathrm{x};\xi)}}{\mathrm{e}^{\theta}-1}\left[1+\lambda\frac{\mathrm{e}^{\theta\overline{\mathrm{G}}(\mathrm{x};\xi)}-1}{\mathrm{e}^{\theta}-1}\right],$$

which is identical to the cdf $F(x)=1-R(x)$, where $R(x)$ is given by (8).

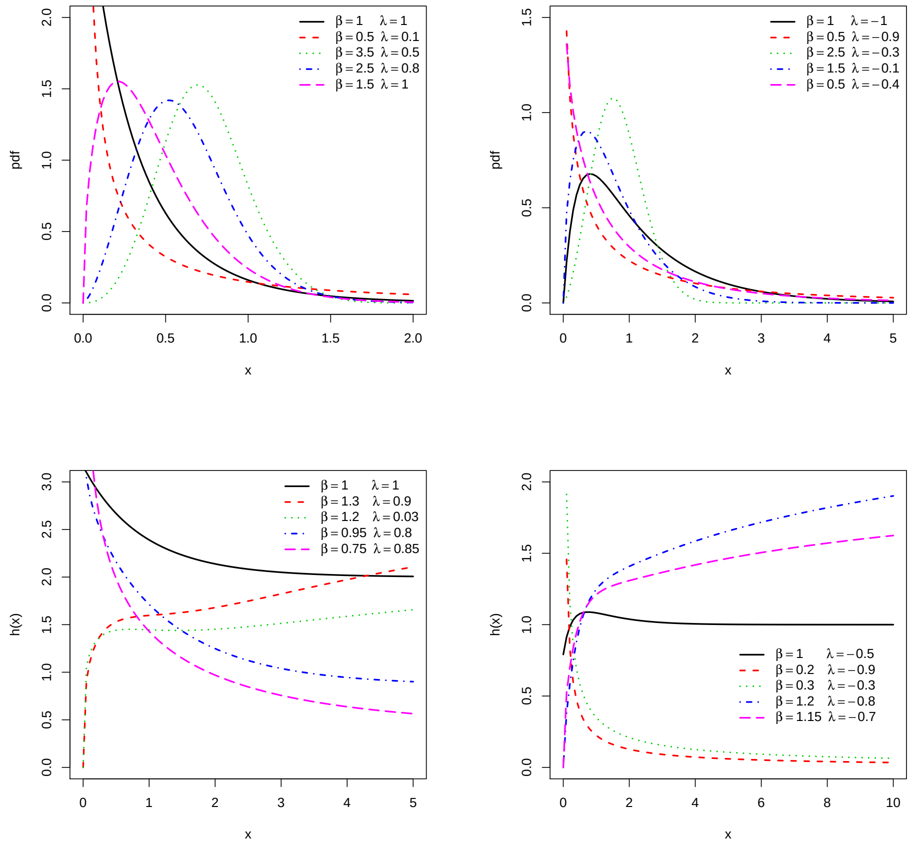

Figure 1: Plots of the CGTPW density and hazard rate.

3. Special models

The pdf (7) will be most tractable when $G(x;\xi)$ and $g(x;\xi)$ have simple analytic expressions. In this section, we provide two special models of the CGTP-G family corresponding to the baseline Weibull (W) and Lindley (Li) distributions, which generalize some well-known distributions in the literature.

3.1. The CGTPW model

Consider the cdf and pdf (for $x>0$) $G(x)=1-\exp(-x^{\beta})$ and $g(x)=\beta x^{\beta-1}\exp(-x^{\beta})$, respectively, of the Weibull distribution with shape parameter β and scale one. Then, the pdf and cdf of the CGTPW model are given by

$$f(x)=\frac{\theta\beta x^{\beta-1}\,\mathrm{e}^{-x^{\beta}}\,\mathrm{e}^{\theta\exp\left(-x^{\beta}\right)}}{\mathrm{e}^{\theta}-1}\left\{1-\lambda\frac{\left(\mathrm{e}^{\theta}+1\right)-2\mathrm{e}^{\theta\exp\left(-x^{\beta}\right)}}{\mathrm{e}^{\theta}-1}\right\}$$

and

$$F(x)=\frac{\mathrm{e}^{\theta}-\mathrm{e}^{\theta\exp\left(-x^{\beta}\right)}}{\mathrm{e}^{\theta}-1}\left\{1+\lambda\frac{\mathrm{e}^{\theta\exp\left(-x^{\beta}\right)}-1}{\mathrm{e}^{\theta}-1}\right\},$$

respectively. Some plots of the density and hrf of the CGTPW model for selected parameter values are displayed in Figure 1.

3.2. The CGTPLi model

The Lindley distribution with shape parameter $\alpha>0$ has pdf and cdf (for $x>0$) given

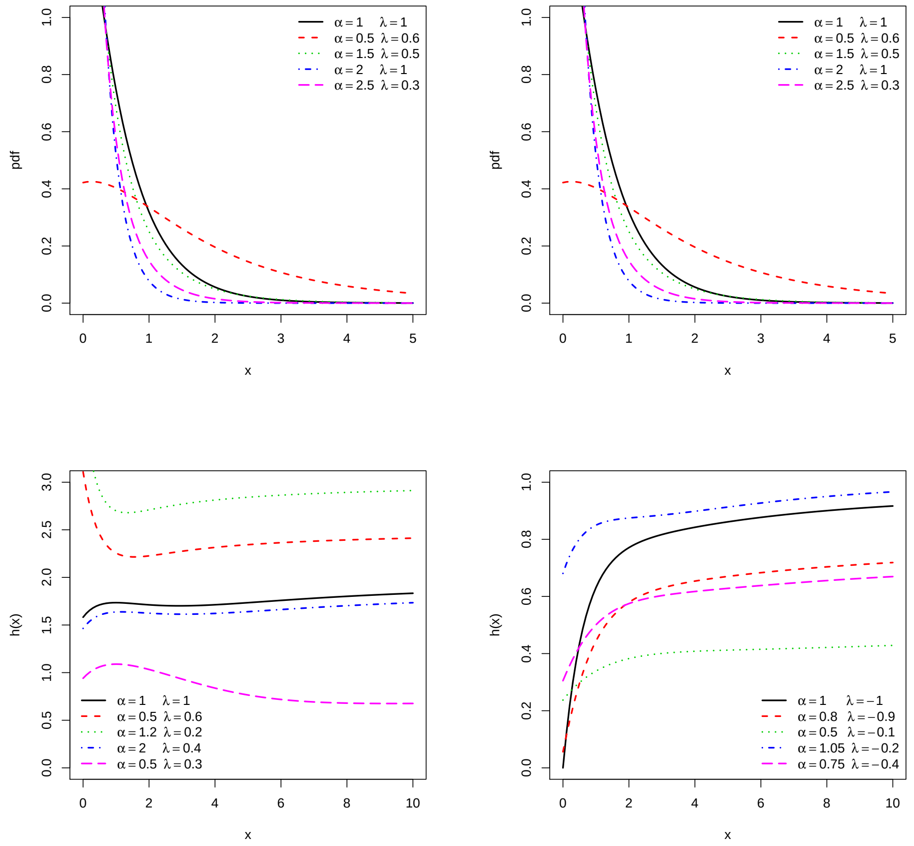

Figure 2: Plots of the CGTPLi density and hazard rate.

by $g(x) = \frac{\alpha^2}{1+\alpha}(1+x)\exp(-\alpha x)$ and $G(x) = 1 - \frac{1+\alpha+\alpha x}{1+\alpha}\exp(-\alpha x)$, respectively. Then, the pdf and cdf of the CGTPLi model are given by

$$f(x) = \frac{\theta(1+x)\exp\left(\frac{1+\alpha+\alpha x}{1+\alpha}\theta \mathrm{e}^{-\alpha x}\right)}{\alpha^{-2}(1+\alpha)(\mathrm{e}^{\theta}-1)\,\mathrm{e}^{\alpha x}}\left\{1-\lambda\frac{\left(\mathrm{e}^{\theta}+1\right)-2\exp\left(\frac{1+\alpha+\alpha x}{1+\alpha}\theta \mathrm{e}^{-\alpha x}\right)}{\mathrm{e}^{\theta}-1}\right\}$$

and

$$F(x) = \frac{\mathrm{e}^{\theta}-\exp\left(\frac{1+\alpha+\alpha x}{1+\alpha}\theta \mathrm{e}^{-\alpha x}\right)}{\mathrm{e}^{\theta}-1}\left\{1+\lambda\frac{\exp\left(\frac{1+\alpha+\alpha x}{1+\alpha}\theta \mathrm{e}^{-\alpha x}\right)-1}{\mathrm{e}^{\theta}-1}\right\},$$

respectively. Some plots of the pdf and hrf of the CGTPLi model are displayed in Figure 2 for some parameter values.

4. Linear representation

We provide a useful representation for (7) using the concept of exponentiated distributions. The pdf (7) can be expressed as

$$f(x) = \frac{\theta(1-\lambda)\,g(x)\mathrm{e}^{\theta\overline{\mathrm{G}}(x)}}{\mathrm{e}^{\theta}-1} + \frac{2\lambda\theta g(x)\mathrm{e}^{2\theta\overline{\mathrm{G}}(x)}}{\left(\mathrm{e}^{\theta}-1\right)^2}.$$

Expanding the quantities $\exp\left[\theta\overline{G}(x)\right]$ and $\exp\left[2\theta\overline{G}(x)\right]$ in power series, we can write

$$f(x) = g(x)\sum_{i=0}^{\infty}\left[\frac{(1-\lambda)\,\theta^{i+1}}{\left(\mathrm{e}^{\theta}-1\right)i!} + \frac{\lambda(2\theta)^{i+1}}{\left(\mathrm{e}^{\theta}-1\right)^2 i!}\right][1-G(x)]^i.$$

Applying the binomial expansion to $[1 - G(x)]^i$, we have

$$f(x) = g(x) \sum_{i=0}^{\infty} \sum_{k=0}^{i} (-1)^k \left[\frac{(1-\lambda)\,\theta^{i+1}}{(e^{\theta}-1)\; i!} + \frac{\lambda\,(2\theta)^{i+1}}{(e^{\theta}-1)^2\; i!} \right] \binom{i}{k} G(x)^i .$$

By changing the sums over the indices k and i, we obtain

$$f(x) = g(x) \sum_{k=0}^{\infty} \sum_{i=k}^{i} (-1)^k \left[\frac{(1-\lambda)\,\theta^{i+1}}{(e^{\theta}-1)\; i!} + \frac{\lambda\,(2\theta)^{i+1}}{(e^{\theta}-1)^2\; i!} \right] \binom{i}{k} G(x)^i .$$

Then, the pdf of the CGTP-G family reduces to

$$f(x) = \sum_{k=0}^{\infty} b_k\, h_{k+1}(x), \tag{9}$$

where $h_{k+1}(x) = (k+1)\, g(x)\, G(x)^k$ is the Exp-G pdf of a random variable Y_{k+1} with power parameter $k+1$ and

$$b_k = \sum_{i=k}^{\infty} \frac{(-1)^k}{(k+1)} \binom{i}{k} \left[\frac{\lambda\,(2\theta)^{i+1}}{i!\,(e^{\theta}-1)^2} + \frac{(1-\lambda)\theta^{i+1}}{i!\,(e^{\theta}-1)} \right].$$

Equation (9) reveals that the CGTP-G density function is a linear combination of Exp-G densities. Thus, some mathematical properties of the new family can be derived from those properties of the Exp-G class.

By integrating (9), we obtain the same linear representation for the cdf of X

$$F(x) = \sum_{k=0}^{\infty} b_k\, H_{k+1}(x),$$

where $H_{k+1}(x)$ is the cdf of the Exp-G family with power parameter $k+1$.

5. Mathematical properties

The formulae derived throughout the paper can be easily handled in most symbolic computation platforms such as **Maple**, **Mathematica** and **Matlab**. These platforms have currently the ability to deal with analytic expressions of formidable size and complexity. Established explicit expressions to obtain some statistical measures can be more efficient than computing them directly by numerical integration.

5.1. Asymptotics

Let $a = \inf\{x | G(x) > 0\}$, the asymptotics of $F(x)$, $f(x)$ and $\tau(x)$ as $x \to a$ are given by

$$F(x) \sim \frac{(1+\lambda)\,\theta G(x)}{e^{\theta}-1} \qquad \text{as} \qquad x \to a,$$

$$f(x) \sim \frac{(1+\lambda)\,\theta g(x)}{e^{\theta}-1} \qquad \text{as} \qquad x \to a,$$

$$\tau(x) \sim \frac{(1+\lambda)\,\theta g(x)}{e^{\theta}-1} \qquad \text{as} \qquad x \to a.$$

The asymptotics of $F(x)$, $f(x)$ and $\tau(x)$ as $x \to \infty$ are given by

$$1 - F(x) \sim \frac{\theta \overline{G}(x)}{e^{\theta}-1} \qquad \text{as} \qquad x \to \infty,$$

$$f(x) \sim \frac{\theta g(x)}{\mathrm{e}^{\theta}-1} \qquad \text{as} \qquad x \to \infty,$$

$$\tau(x) \sim \frac{g(x)}{\overline{G}(x)} \qquad \text{as} \qquad x \to \infty.$$

These equations show the effect of parameters on tails of distribution.

5.2. Moments

The nth ordinary moment of X, say μ'_n, can be determined from (9) as

$$\mu'_n = E(X^n) = \sum_{k=0}^{\infty} b_k \, E(Y_{k+1}^n). \tag{10}$$

For $\delta > 0$, $E(Y_\delta^n) = \delta \int_{-\infty}^{\infty} x^n \, g(x;\xi) \, G(x;\xi)^{\delta-1} \, dx$, which can be obtained numerically from the baseline quantile function (qf) $Q_G(u;\xi) = G^{-1}(x;\xi)$ as

$$E(Y_\delta^n) = \delta \int_0^1 Q_G(u;\xi)^n \, u^{\delta-1} du.$$

Setting $n = 1$ in (10), we have the mean of X. The central moments (μ_s) and cumulants (κ_s) of X follow from (10) as $\mu_s = \sum_{k=0}^{s} (-1)^k \binom{s}{k} \mu_1'^k \, \mu'_{s-k}$ and $\kappa_s = \mu'_s - \sum_{k=1}^{s-1} \binom{s-1}{k-1} \kappa_k \, \mu'_{s-k}$, respectively, where $\kappa_1 = \mu'_1$. The skewness and kurtosis of X are the third and fourth standardized cumulants given by $\gamma_1 = \kappa_3/\kappa_2^{3/2}$ and $\gamma_2 = \kappa_4/\kappa_2^2$, respectively.

5.3. Incomplete moments

Here, we determine the nth incomplete moment of X defined by $m_n(y) = \int_{-\infty}^{y} x^n \, f(x) dx$. We have

$$m_n(y) = \sum_{k=0}^{\infty} b_k \, m_{n,k+1}(y), \tag{11}$$

where

$$m_{n,\delta}(y) = \int_0^{G(y;\xi)} Q_G(u;\xi)^n \, u^{\delta-1} \, du.$$

The integral $m_{n,\delta}(y)$ can be obtained analytically for special models with closed-form expressions for $Q_G(u;\xi)$ or evaluated at least numerically for most baseline distributions.

An important application of the first incomplete moment of X in (11), say $m_1(y)$, refers to the Bonferroni and Lorenz curves. These curves are very useful in economics, reliability, demography, insurance and medicine.

For a given probability π, the Bonferroni and Lorenz curves are given by $B(\pi) = m_1(p)/(p\mu'_1)$ and $L(p) = m_1(p)/\mu'_1$, where $p = Q(\pi) = F^{-1}(\pi)$ can be determined numerically by inverting $F(x) = 1 - R(x)$ from equation (8).

Another application is related to the mean deviations about the mean ($\delta_1 = E(|X - \mu'_1|)$) and about the median ($\delta_2 = E(|X - M|)$) of X given by

$$\delta_1 = 2\mu'_1 \, F(\mu'_1) - 2m_1(\mu'_1) \qquad \text{and} \qquad \delta_2 = \mu'_1 - 2m_1(M),$$

respectively, where $M = Q(0.5)$ is the median of X, $\mu'_1 = E(X)$ comes from equation (10), $F(\mu'_1)$ is easily evaluated from (8) and $m_1(z)$ is obtained from (11) with $n = 1$.

5.4. Residual and reversed residual life functions

For $n = 1, 2, \ldots$, the nth moment of the residual life of X, say $v_n(t) = E[(X-t)^n \mid X > t]$, uniquely determines $F(x)$ and it is given by

$$v_n(t) = \frac{1}{R(t)} \int_t^{\infty} (x-t)^n \, dF(x).$$

Then,

$$v_n(t) = \frac{1}{R(t)} \sum_{r=0}^{n} (-t)^{n-r} \binom{n}{r} \sum_{k=0}^{\infty} b_k \int_t^{\infty} x^r \, h_{k+1}(x) \, .$$

Another interesting function is the mean residual life (MRL) function or the life expectation at age t just given by $v_1(t) = E\left[(X-t) \mid X > t\right]$, which represents the expected additional life length for a unit which is alive at age t. The MRL of X can be obtained by setting $n = 1$ in the last equation.

The nth moment of the reversed residual life $M_n(t) = E\left[(t-X)^n \mid X \le t\right]$, for $t > 0$ and $n = 1, 2, \ldots$, uniquely determines $F(x)$ and follows from $v_n(t)$. The mean inactivity time (MIT) of X given by $M_1(t)$ represents the waiting time elapsed since the failure of an item on condition that this failure had occurred in $(0, t)$.

5.5. Quantile and generating functions

The qf of X is obtained by inverting $F(x) = 1 - R(x)$ in (6), say $Q(u) = F^{-1}(u)$, for $0 < u < 1$. If U is a uniform variate on the unit interval $(0, 1)$, then the random variable $X = Q(U)$ has density (7).

We can also simulate the CGTP-G distribution as follows: if $u \sim U(0, 1)$, the solution of the nonlinear equation, for $\lambda \neq 0$, is given by

$$G(x_u) = 1 - \frac{1}{\theta} \log \left[\mathrm{e}^{\theta} - \frac{(\mathrm{e}^{\theta} - 1)\left[1 + \lambda - \sqrt{(1+\lambda)^2 - 4\lambda\, u}\right]}{2\,\lambda} \right].$$

For $\lambda = 0$, we obtain

$$G(x_u) = 1 - \frac{1}{\theta} \log \left[u + \mathrm{e}^{\theta}\,(1-u)\right].$$

The moment generating function (mgf) of X, say $M(t) = E(\mathrm{e}^{t\,X})$, can be obtained from (9) as

$$M(t) = \sum_{k=0}^{\infty} b_k \, M_{k+1}(t; \xi),$$

where $M_\delta(t;\xi)$ is the mgf of Y_δ given by

$$M_\delta(t;\xi) = \delta \int_{-\infty}^{\infty} \exp\left(t\,x\right)\, G(x;\xi)^{\delta-1}\, g(x;\ \xi) dx = \delta \int_0^1 \exp[t\, Q_G(u; \delta\,)]\, u^{\delta-1} du.$$

The last two integrals can be computed numerically for most parent distributions.

5.6. Entropies

The Rényi entropy of a random variable X represents a measure of variation of the uncertainty. It is defined by

$$I_\delta\,(X) = \frac{1}{1-\delta} \log \left[\int_{-\infty}^{\infty} f\,(x)^{\delta}\, dx\right], \ \delta > 0 \text{ and } \delta \neq 1.$$

Using the pdf in (7), we can write

$$f\,(x)^{\delta} = \frac{\theta^{\delta} g(x)^{\delta} \mathrm{e}^{\delta\theta\overline{\mathrm{G}}(\mathrm{x})}}{(\mathrm{e}^{\theta}-1)^{\delta}} \left\{ 1 - \lambda \frac{1 - \mathrm{e}^{\theta\overline{\mathrm{G}}(\mathrm{x})}}{\mathrm{e}^{\theta} - 1} - \lambda \frac{\mathrm{e}^{\theta}\left[1 - \mathrm{e}^{-\theta\overline{\mathrm{G}}(\mathrm{x})}\right]}{\mathrm{e}^{\theta} - 1} \right\}^{\delta}.$$

Let $(\delta)_n = \Gamma(\delta)/\Gamma(\delta - n)$ be the falling factorial. Then,

$$f\,(x)^{\delta} = \sum_{k=0}^{\infty} u_k \; g(x)^{\delta} G(x)^k,$$

where

$$u_k = \sum_{i,j,w,m,h=0}^{\infty} \frac{(-1)^{i+j+w+m+h} \; (\delta+1)_i \; (m+1)_i \; (h+1)_j}{i! \, j! \, w! \, m! \, h! \, k!}$$
$$\times \frac{(w+1)_k \; (\delta-m+1)_h \; \lambda^{h+m}\theta^{\delta+w} \; (\delta+j+m)^{w+k} \; (\delta+j-i)^k}{(e^{\theta}-1)^{\delta+m+h}}.$$

Then, the Rényi entropy of the CGTP-G family is given by

$$I_\delta(X) = \frac{1}{1-\delta} \log\left\{ \sum_{k=0}^{\infty} u_k \int_{-\infty}^{\infty} g(x)^\delta G(x)^k dx \right\},$$

where the integral can be determined numerically for any parent distribution.
The δ-entropy, say $H_\delta(X)$, for $\delta > 0$ and $\delta \neq 1$, is defined by

$$H_\delta(X) = \frac{1}{\delta-1} \log\left\{ 1 - \int_{-\infty}^{\infty} f(x)^\delta \, dx \right\},$$

and then

$$H_\delta(X) = \frac{1}{\delta-1} \log\left\{ 1 - \left[\sum_{k=0}^{\infty} u_k \int_{-\infty}^{\infty} g(x)^\delta G(x)^k dx \right] \right\}.$$

The Shannon entropy, say SI, of a random variable X is given by

$$SI = E\left\{ -\left[\log f(X)\right] \right\}.$$

It is the special case of the Rényi entropy when $\delta \uparrow 1$.

6. Order statistics

Order statistics make their appearance in many areas of statistical theory and practice. Let $X_1, \ldots, X_n$ be a random sample from the CGTP-G family. The pdf of the ith order statistic, say $X_{i:n}$, is given by

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

where $K = n!/[(i-1)!(n-i)!]$.
After some algebra, we can write

$$f_{i:n}(x) = K \sum_{j=0}^{n-i} (-1)^j \binom{n-i}{j} \left[\sum_{r=0}^{\infty} (1+r) b_r G(x)^r g(x) \right] \left[\sum_{k=0}^{\infty} b_k G(x)^{k+1} \right]^{j+i-1}.$$

Further, we have

$$\left[\sum_{k=0}^{\infty} b_k \, G(x)^{k+1} \right]^{j+i-1} = \sum_{k=0}^{\infty} \varphi_{j+i-1,k} \; G(x)^{j+i+k-1},$$

where $\varphi_{j+i-1,0} = b_0^{j+i-1}$ and (for $k \geq 1$)

$$\varphi_{j+i-1,k} = (k \, b_0)^{-1} \sum_{m=1}^{k} \left[m(j+i) - k \right] b_m \varphi_{j+i-1,k-m}.$$

Hence,

$$f_{i:n}(x) = \sum_{j=0}^{n-i} \sum_{k,r=0}^{\infty} d_{j,k,r} \quad h_{i+j+k+r}(x), \tag{12}$$

where

$$d_{j,k,r} = \frac{(-1)^j \, n!}{(i-1)!(n-i)!} \binom{n-i}{j} \frac{(1+r) b_r \, \varphi_{j+i-1,k}}{j+i+k+r}.$$

Equation (12) is the main result of this section. Thus, the density function of the CGTP-G order statistics is a triple linear combination of Exp-G distributions. Based on equation (12), we can obtain some structural properties of $X_{i:n}$ from those of the Exp-G model.

The qth moment of $X_{i:n}$ is given by

$$E(X_{i:n}^q) = \sum_{j=0}^{n-i} \sum_{k,r=0}^{\infty} d_{j,k,r} \;\; E(Y_{i+j+k+r}^q). \tag{13}$$

Based upon the moments in equation (13), we can derive explicit expressions for the L-moments of X as infinite weighted linear combinations of the means of suitable Exp-G densities. These moments are analogous to the ordinary moments but can be estimated by linear combinations of order statistics. They are given as linear functions of expected order statistics, namely

$$\lambda_r = \frac{1}{r} \sum_{d=0}^{r-1} (-1)^d \binom{r-1}{d} E(X_{r-d:r}), \;\; r \geq 1.$$

7. Maximum likelihood estimation

Several approaches for parameter estimation were proposed in the literature but the maximum likelihood method is the most commonly employed. We consider the estimation of the unknown parameters of the proposed family from complete samples only by maximum likelihood, although we provide a summary of the least squares method. Let $x_1, \ldots, x_n$ be a random sample from the CGTP-G family with parameters λ, θ and ξ. Let $\zeta = (\lambda, \theta, \xi^{intercal})^{intercal}$ be the $p \times 1$ parameter vector. Then, the log-likelihood function for ζ is given by

$$\ell = \ell(\zeta) = n \log \theta - n \log \left(\mathrm{e}^{\theta} - 1 \right) + \sum_{i=0}^{n} \log g(x_i; \xi) + \theta \sum_{i=0}^{n} \overline{G}(x_i; \xi) + \sum_{i=0}^{n} \log q_i,$$

where $q_i = \left(1 - \frac{\lambda p_i}{\mathrm{e}^{\theta} - 1}\right) - \frac{\lambda \mathrm{e}^{\theta} s_i}{\mathrm{e}^{\theta} - 1}$, $p_i = 1 - \mathrm{e}^{\theta \overline{\mathrm{G}}(\mathrm{x_i};\xi)}$ and $s_i = 1 - \mathrm{e}^{-\theta \mathrm{G}(\mathrm{x_i};\xi)}$.

The equation for $\ell(\zeta)$ can be maximized either directly by using the `MATH-CAD` program, `SAS` (`PROC NLMIXED`), `R` (`optim` function) and `Ox` program (sub-routine `MaxBFGS`), or by solving the nonlinear likelihood equations obtained by differentiating this equation.

The components of the score vector

$$\mathbf{U}(\zeta) = \left(\frac{\partial \ell}{\partial \lambda}, \frac{\partial \ell}{\partial \theta}, \frac{\partial \ell}{\partial \xi_k} \right)^{intercal} = (U_\lambda, U_\theta, U_{\xi_k})^{intercal}$$

are given by

$$U_\lambda = \sum_{i=1}^{n} \frac{-1}{q_i} \left(\frac{p_i + s_i \mathrm{e}^{\theta}}{\mathrm{e}^{\theta} - 1} \right), \quad U_\theta = \frac{n}{\theta} - \frac{n}{\mathrm{e}^{\theta} - 1} + \sum_{i=0}^{n} \overline{G}(x_i; \xi) + \sum_{i=0}^{n} \frac{\lambda \mathrm{e}^{\theta} \, (p_i + s_i)}{q_i \, (\mathrm{e}^{\theta} - 1)^2}$$

and

$$U_{\xi_k} = \sum_{i=0}^{n} \frac{g_k'(x_i; \xi)}{g(x_i; \xi)} - \theta \sum_{i=0}^{n} G_k'(x_i; \xi) - 2\lambda\theta \sum_{i=0}^{n} \frac{G_k'(x_i; \xi)}{q_i \, (\mathrm{e}^{\theta} - 1)} \mathrm{e}^{\theta \overline{\mathrm{G}}(\mathrm{x_i};\xi)},$$

where

$$g'_k(x_i;\xi) = \partial g(x_i;\xi)/\partial\xi_k \, \text{and} \, G'_k(x_i;\xi) = \partial G(x_i;\xi)/\partial\xi_k.$$

Setting the nonlinear system of equations $U_\lambda = U_\theta = 0$ and $U_{\boldsymbol{\xi}_\mathbf{k}} = 0$ (for the components of ξ) and solving them simultaneously yields the maximum likelihood estimate (MLE) $\widehat{\zeta} = (\widehat{\lambda}, \widehat{\theta}, \widehat{\xi}^{intercal})^{intercal}$. It is usually more convenient to adopt nonlinear optimization methods such as the quasi-Newton algorithm to maximize ℓ numerically. For interval estimation of the parameters, we obtain the $p \times p$ observed information matrix $J(\zeta) = \{-\frac{\partial^2 \ell}{\partial r\,\partial s}\}$ (for $r, s = \lambda, \theta$ and varying on the components of ξ), whose elements can be evaluated numerically.

Under standard regularity conditions when $n \to \infty$, the distribution of $\widehat{\zeta}$ can be approximated by a multivariate normal $N_p(0, J(\widehat{\zeta})^{-1})$ distribution to construct confidence intervals for the parameters. Here, $J(\widehat{\zeta})$ is the total observed information matrix evaluated at $\widehat{\zeta}$, whose elements are given in Appendix A.

An alternative approach to the maximum likelihood method is the least square estimation. For the CGTP-G model, the least square estimates (LSEs) $\widetilde{\lambda}, \widetilde{\theta}$ and $\widetilde{\xi}$ of λ, θ and ξ are defined as those arguments that minimize the objective function:

$$Q(\lambda,\theta,\xi) = \sum_{i=1}^{n}\left[F(x_{i:n}) \; - \; \frac{i}{n+1}\right]^2,$$

where $x_{i:n}$ is a possible outcome of the ith order statistic based on a n-points random sample and $F(x_{i:n}) = 1 - R(x_{i:n})$ is obtained from (8).

The minimum point $\widetilde{\lambda}, \widetilde{\theta}$ and $\widetilde{\xi}$ can also be given as a solution in the following system of non-linear equations:

$$\frac{\partial Q(\lambda,\theta,\xi)}{\partial\lambda} = \frac{\partial Q(\lambda,\theta,\xi)}{\partial\theta} = \frac{\partial Q(\lambda,\theta,\xi)}{\partial\xi_k} = 0,$$

where k varies over the components of ξ.

8. Simulation study

We evaluate the performance of the maximum likelihood method to estimate the parameters using Monte Carlo simulations. We choose the CGTPW model, a total of twelve parameter combinations and sample sizes (SSs) n=50, 100 and 300. The process is repeated 1,000 times and the biases (estimate - actual) and mean square errors (MSEs) of the estimates are reported in Table 1. We perform the simulations using the `R` software. The small values of the biases and MSEs indicate that the maximum likelihood method performs quite well to estimate the model parameters.

9. Applications

In this section, we provide two applications to real data to prove empirically the flexibility of the CGTPLi model introduced in Section 3.2. The goodness-of-fit statistics for this model are compared with other competitive models and the MLEs of the model parameters are determined numerically. For the two real data sets, we compare the fits of the CGTPLi distribution with the Kumaraswamy Lindley (KwLi) (Cakmakyapan and Kadilar, 2014), beta Lindley (BLi) (Merovci and Sharma, 2014), McDonald modified Weibull (McMW) (Merovci and Elbatal, 2013), Kumaraswamy-transmuted exponentiated modified Weibull (KwTEMW) (Al-Babtain et al., 2015), transmuted modified Weibull (TMW) (Khan and King, 2013) and power Lindley (PoLi) (Ghitany et al., 2013) models with corresponding densities (for $x > 0$):

Table 1: Biases and MSEs for simulation data

SS	Actual values			Bias			MSE		
n	β	λ	θ	$\widehat{\beta}$	$\widehat{\lambda}$	$\widehat{\theta}$	$\widehat{\beta}$	$\widehat{\lambda}$	$\widehat{\theta}$
50	0.5	0.5	1.0	0.992	0.433	1.876	0.997	0.217	1.326
	1.5	0.7	1.0	0.362	−0.787	0.274	0.304	1.217	0.515
	0.5	−0.5	1.0	1.719	1.478	1.121	1.986	1.195	2.631
	0.5	−1.0	1.0	0.331	1.353	1.0974	0.124	1.279	1.112
100	0.5	0.5	1.0	0.972	0.423	1.764	0.953	0.216	0.455
	1.5	0.7	1.0	0.188	−0.569	0.278	0.117	0.921	0.422
	0.5	−0.5	1.0	0.689	0.494	0.502	2.869	0.235	0.642
	0.5	−1.0	1.0	0.316	0.329	0.107	0.107	0.206	0.179
300	0.5	0.5	1.0	0.099	0.041	0.028	0.003	0.156	0.017
	1.5	0.7	1.0	0.032	−0.036	0.089	0.014	0.029	0.014
	0.5	−0.5	1.0	0.674	0.500	0.053	0.007	0.025	0.096
	0.5	−1.0	1.0	0.309	0.058	0.064	0.098	0.051	0.003

- The KwLi density is given by

$$f(x) = \frac{ab\alpha^2(1+x)}{(1+\alpha)}\mathrm{e}^{-\alpha \mathrm{x}}\left(1-\frac{1+\alpha+\alpha x}{1+\alpha}\mathrm{e}^{-\alpha \mathrm{x}}\right)^{a-1}$$
$$\times\left[1-\left(1-\frac{1+\alpha+\alpha x}{1+\alpha}\mathrm{e}^{-\alpha \mathrm{x}}\right)^{a}\right]^{b-1}.$$

- The BLi density is given by

$$f(x) = \frac{\alpha^2(1+x)}{B(a,b)(1+\alpha)}\mathrm{e}^{-\alpha x}\left(\frac{1+\alpha+\alpha x}{1+\alpha}\mathrm{e}^{-\alpha x}\right)^{b-1}.$$

- The McMW density is given by

$$f(x) = \frac{c\alpha+c\gamma\beta x^{\beta-1}}{B(a/c,b)}\mathrm{e}^{-\alpha \mathrm{x}-\gamma \mathrm{x}^{\beta}}\left(1-\mathrm{e}^{-\alpha \mathrm{x}-\gamma \mathrm{x}^{\beta}}\right)^{ac-1}$$
$$\times\left[1-\left(1-\mathrm{e}^{-\alpha \mathrm{x}-\gamma \mathrm{x}^{\beta}}\right)^{c}\right]^{b-1}.$$

- The KwTEMW density is given by

$$f(x) = \left(\alpha+\gamma\beta x^{\beta-1}\right)\mathrm{e}^{-\alpha \mathrm{x}-\gamma \mathrm{x}^{\beta}}\left[1+\lambda-2\lambda\left(1-\mathrm{e}^{-\alpha \mathrm{x}-\gamma \mathrm{x}^{\beta}}\right)^{\delta}\right]$$
$$\times ab\delta\left(1-\mathrm{e}^{-\alpha \mathrm{x}-\gamma \mathrm{x}^{\beta}}\right)^{a\delta-1}\left[1+\lambda-\lambda\left(1-\mathrm{e}^{-\alpha \mathrm{x}-\gamma \mathrm{x}^{\beta}}\right)^{\delta}\right]^{a-1}$$
$$\times\left\{1-\left(1-\mathrm{e}^{-\alpha \mathrm{x}-\gamma \mathrm{x}^{\beta}}\right)^{a\delta}\left[1+\lambda-\lambda\left(1-\mathrm{e}^{-\alpha \mathrm{x}-\gamma \mathrm{x}^{\beta}}\right)^{\delta}\right]^{a}\right\}^{b-1}.$$

- The TMW density is given by

$$f(x) = \left(\alpha+\gamma\beta x^{\beta-1}\right)\mathrm{e}^{-\alpha \mathrm{x}-\gamma \mathrm{x}^{\beta}}\left(1-\lambda+2\lambda\mathrm{e}^{-\alpha \mathrm{x}-\gamma \mathrm{x}^{\beta}}\right).$$

- The PoLi density is given by

$$f(x) = \frac{\beta\alpha^2}{1+\alpha}\left(1+x^{\beta}\right)x^{\beta-1}\mathrm{e}^{-\alpha \mathrm{x}^{\beta}}.$$

The parameters of the above densities are all positive real numbers except for the KwTEMW and TMW distributions for which $|\lambda| \leq 1$.

In order to compare the fitted models, we consider some goodness-of-fit statistics, namely the Akaike information criterion (AIC), consistent Akaike information criterion ($CAIC$), Hannan-Quinn information criterion ($HQIC$), Bayesian information criterion (BIC) and $-2\widehat{\ell}$, where $\widehat{\ell}$ is the maximized log-likelihood. Moreover, we use the Anderson-Darling (A^*) and the Cramr-von Mises (W^*) statistics in order to compare the fits of the two new models with

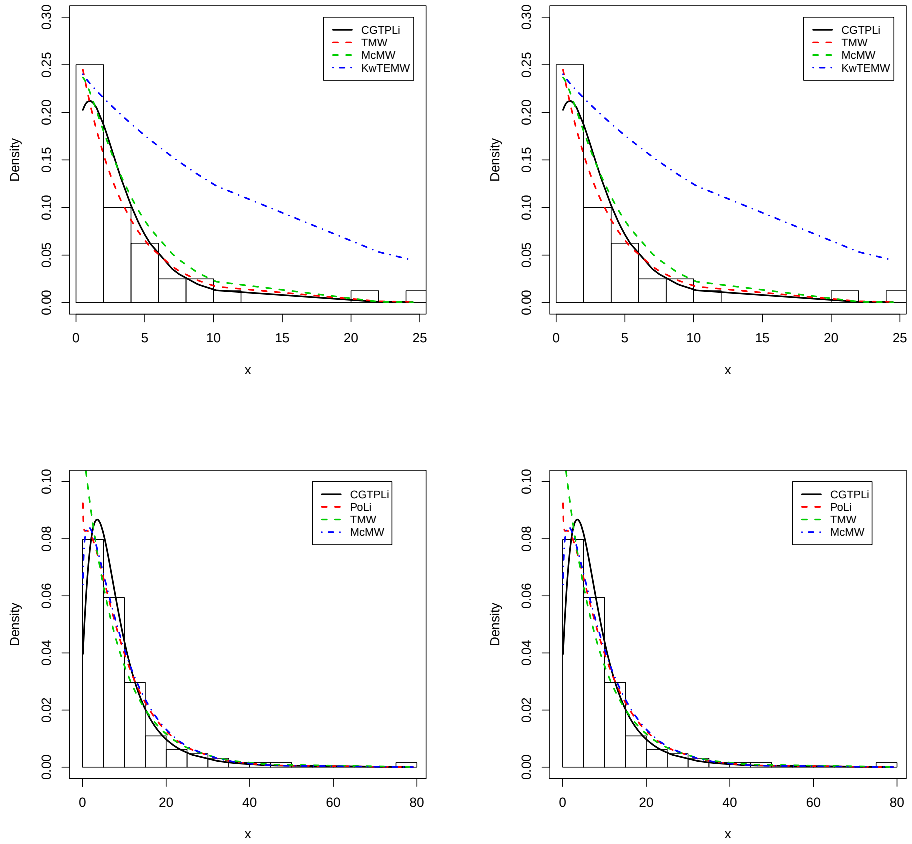

Figure 3: The estimated CGTPLi pdf and other estimated pdfs for active repair times (left panel) and cancer data (right panel).

other nested and non-nested models. The statistics are widely used to determine how closely a specific cdf fits the empirical distribution of the data set. The smaller these statistics are, the better the fit is. Upper tail percentiles of the asymptotic distributions of these goodness-of-fit statistics were tabulated by Nichols and Padgett (2006).

These following numerical results are obtained using the `MATH-CAD` program. Furthermore, the `R` codes to compute the cdf, pdf and maximum likelihood estimates for the CGTPLi distribution are given in Appendix B.

9.1. Active repair times data

The first data set represents the active repair times (h) (Jorgensen, 1982) for an airborne communication transceiver. These data consist of 40 observations (see Appendix C).

Tables 2 and 3 provide the MLEs and their standard errors (in parentheses) of the model parameters for some distributions and the goodness-of-fit statistics for the current data, respectively. The plots of the fitted CGTPLi pdf and other fitted pdfs defined before, for the two data sets, are displayed in Figure 3.

Table 2: MLEs and their standard errors (in parentheses) for active repair times

Model	Estimates		
CGTPLi	$\widehat{\alpha}= 0.2194$	$\widehat{\theta}= 2.8054$	$\widehat{\lambda}= 0.4213$
	(0.065)	(1.616)	(0.413)
TMW	$\widehat{\alpha}= 0.1705$	$\widehat{\beta}= 1.225$	$\widehat{\gamma}= 3.5176 \cdot 10^{-5}$
	(0.047)	(0.484)	$(3.824 \cdot 10^{-3})$
	$\widehat{\lambda}= 0.6733$		
	(0.338)		
McMW	$\widehat{\alpha}= 0.2498$	$\widehat{\beta}= 0.6972$	$\widehat{\gamma}= 1.4439 \cdot 10^{-6}$
	(0.087)	(0.266)	$(5.671 \cdot 10^{-4})$
	$\widehat{a}= 1.2604$	$\widehat{b}= 1.0682$	$\widehat{c}= 0.8792$
	(1.039)	(0.401)	(0.63)
KwTEMW	$\widehat{\alpha}= 0.072$	$\widehat{\beta}= 0.837$	$\widehat{\gamma}= 2.908 \cdot 10^{-7}$
	(0.11)	(0.238)	$(2.552 \cdot 10^{-4})$
	$\widehat{\lambda}= -0.381$	$\widehat{a}= 7.44$	$\widehat{\delta}= 0.1$ $\widehat{b}= 3.547$
	(0.682)	(7.334)	(0.091) (5.285)
KwLi	$\widehat{\alpha}= 0.037$	$\widehat{a}= 0.6046$	$\widehat{b}= 11.7528$
	(0.061)	(0.072)	(21.385)
BLi	$\widehat{\alpha}= 0.0303$	$\widehat{a}= 0.5015$	$\widehat{b}= 30.1304$
	(0.017)	(0.092)	(28.342)
PoLi	$\widehat{\alpha}= 0.5867$	$\widehat{\beta}= 0.7988$	
	(0.099)	(0.082)	

Table 3: Goodness-of-fit statistics for active repair times

Model	$-2\widehat{\ell}$	AIC	$CAIC$	$HQIC$	BIC	W^*	A^*
CGTPLi	188.7	194.7	195.4	196.6	199.8	0.12264	0.86344
TMW	189.2	197.2	198.4	199.7	203.9	0.14761	1.04929
McMW	190.9	202.9	205.5	206.6	213.1	0.14948	1.0737
KwTEMW	191.7	205.7	209.2	209.9	217.5	0.15634	1.12275
PoLi	191.9	195.9	196.2	197.1	199.3	0.16151	1.15216
KwLi	197.2	203.2	203.9	205.1	208.3	0.21112	1.47435
BLi	203.7	209.7	210.3	211.5	214.7	0.28765	1.93883

9.2. Cancer patient data

The second data set refers to the remission times (in months) of a random sample of 128 bladder cancer patients given in Appendix C (Lee and Wang, 2003). These data have been used by Nofal et al. (2016) and Mead and Afify (2016) to fit the generalized transmuted log-logistic and Kumaraswamy exponentiated Burr XII distributions, respectively.

The MLEs and their corresponding standard errors (in parentheses) of the model parameters and the values of $-2\widehat{\ell}$, AIC, $CAIC$, $HQIC$, BIC, W^* and A^* are given in Tables 4 and 5, respectively.

Table 4: MLEs and their standard errors (in parentheses) for the cancer patient data

Model	Estimates		
CGTPLi	$\widehat{\alpha}$= 0.0989 (0.022)	$\widehat{\theta}$= 2.7639 (1.204)	$\widehat{\lambda}$= 0.4131 (0.301)
TMW	$\widehat{\alpha}$= 0.0612 (0.01)	$\widehat{\beta}$= 1 $(5.522 \cdot 10^{-4})$	$\widehat{\gamma}$= $7.0091 \cdot 10^{-7}$ $(6.158 \cdot 10^{-5})$
	$\widehat{\lambda}$= 0.8616 (0.175)		
McMW	$\widehat{\alpha}$= 0.0649 (0.029)	$\widehat{\beta}$= 0.6719 (0.217)	$\widehat{\gamma}$= 0.0008 $(4.837 \cdot 10^{-3})$
	$\widehat{a}$= 0.8583 (0.322)	$\widehat{b}$= 1.9183 (0.781)	$\widehat{c}$= 1.3349 (0.322)
KwTEMW	$\widehat{\alpha}$= 0.044 (0.028)	$\widehat{\beta}$= 0.618 (0.146)	$\widehat{\gamma}$= $1.437 \cdot 10^{-7}$ $(3.436 \cdot 10^{-5})$
	$\widehat{\lambda}$= −0.47 (0.267)	$\widehat{a}$= 2.287 (1.627)	$\widehat{\delta}$= 0.37 (0.242) $\widehat{b}$= 3.036 (2.065)
KwLi	$\widehat{\alpha}$= 0.0238 (0.028)	$\widehat{a}$= 0.6201 (0.039)	$\widehat{b}$= 8.5854 (11.134)
BLi	$\widehat{\alpha}$= 0.0232 (0.007305)	$\widehat{a}$= 0.5357 (0.055)	$\widehat{b}$= 14.4662 (7.422)
PoLi	$\widehat{\alpha}$= 0.2944 (0.037)	$\widehat{\beta}$= 0.8301 (0.047)	

Table 5: Goodness-of-fit statistics for the cancer patient data

Model	$-2\widehat{\ell}$	AIC	$CAIC$	$HQIC$	BIC	W^*	A^*
CGTPLi	821.3	827.3	827.5	830.7	835.8	0.06923	0.45551
PoLi	826.7	830.7	830.8	833.0	836.4	0.10248	0.62961
TMW	826.9	834.9	835.3	839.6	846.4	0.06054	0.5588
McMW	827.1	839.1	839.8	846.1	856.3	0.12897	0.77078
KwTEMW	829.5	843.5	844.4	851.6	863.5	0.1502	0.89817
KwLi	834.9	840.9	841.2	844.4	849.5	0.19746	1.19401
BLi	846.9	852.9	853.2	856.4	861.5	0.37053	2.19822

In Tables 4 and 5, we compare the fits of the CGTPLi model with the TMW, McMW, KwTEMW, PoLi, KwLi and BLi models. We note that the CGTPLi model has the lowest values for the $-2\widehat{\ell}$, AIC, $CAIC$, $HQIC$, BIC, W^* and A^* statistics (for the two real data sets) among the fitted models. So, the CGTPLi model could be chosen as the best model. It is quite clear from the values in Tables 3 and 5 that the CGTPLi model provides the best fits to these data sets. So, we prove empirically that this distribution can be a better model than other competitive models. Further, the plots in Figure 3 reveal that the CGTPLi distribution provide the best fits. In fact, it can be considered a very competitive model to other distributions with positive support.

10. Conclusions

The idea of generating new extended models from classic ones has been of great interest among researchers in the past decade. We propose a new *complementary generalized transmuted Poisson-G* (CGTP-G) family of distributions, which extends the transmuted class (Shaw and Buckley, 2007) by adding one extra shape parameter. Many well-known distributions emerge as special cases of the proposed family. We provide some mathematical properties of the new family including explicit expressions for the ordinary and incomplete moments, mean deviations, generating function, Rényi and q-entropies and order statistics. The maximum likelihood estimation of the model parameters is investigated and the observed information matrix is determined. By means of two real data sets, we verify that a special case of the CGTP-G family can provide better fits than other models generated from well-known families.

Appendix A

The elements of the observed matrix $J(\zeta)$ are given below:

$$U_{\lambda\lambda} = \sum_{i=1}^{n} q_i^{-2} \left(\frac{p_i + s_i \mathrm{e}^{\theta}}{\mathrm{e}^{\theta} - 1} \right)^2, \; U_{\lambda\xi_k} = -\sum_{i=1}^{n} \frac{p_i + s_i \mathrm{e}^{\theta}}{q_i^2 \left(\mathrm{e}^{\theta} - 1\right)},$$

$$U_{\lambda\theta} = \sum_{i=1}^{n} \frac{-\mathrm{e}^{\theta} \left(s_i + p_i\right)}{q_i \left(\mathrm{e}^{\theta} - 1\right)^2} - \sum_{i=1}^{n} \frac{p_i + s_i \mathrm{e}^{\theta}}{q_i^2 \left(\mathrm{e}^{\theta} - 1\right)} \left[\frac{\lambda \mathrm{e}^{\theta} \left(p_i + s_i - 2 s_i \mathrm{e}^{\theta}\right)}{\left(\mathrm{e}^{\theta} - 1\right)^2} \right],$$

$$\begin{aligned} U_{\theta\theta} \;=\; & \frac{-n}{\theta^2} + \frac{n \mathrm{e}^{\theta}}{\left(\mathrm{e}^{\theta} - 1\right)^2} - \lambda \sum_{i=0}^{n} \frac{\mathrm{e}^{\theta} \left(\mathrm{e}^{\theta} - 1\right) \left(1 - 3\mathrm{e}^{\theta}\right) p_i}{q_i \left(\mathrm{e}^{\theta} - 1\right)^4} + \lambda \sum_{i=0}^{n} \frac{s_i \mathrm{e}^{\theta}}{q_i \left(\mathrm{e}^{\theta} - 1\right)^2} \\ & + \lambda \sum_{i=0}^{n} \frac{2 s_i \left(\mathrm{e}^{\theta} - 1\right)^2 - s_i \mathrm{e}^{\theta}}{q_i \mathrm{e}^{-2\theta} \left(\mathrm{e}^{\theta} - 1\right)^4} - \sum_{i=0}^{n} q_i^{-2} \left[\frac{\lambda \mathrm{e}^{\theta} \left(p_i + 2 s_i \mathrm{e}^{\theta} - s_i\right)}{\left(\mathrm{e}^{\theta} - 1\right)^2} \right]^2, \end{aligned}$$

$$\begin{aligned} U_{\theta\xi_k} \;=\; & -\sum_{i=0}^{n} G_k'(x_i;\xi) + \sum_{i=0}^{n} \frac{\lambda\theta \mathrm{e}^{\theta} \left(3 - \mathrm{e}^{\theta}\right) G_k'(x_i;\xi) \mathrm{e}^{\theta \overline{\mathrm{G}}(\mathrm{x_i};\xi)}}{q_i \left(\mathrm{e}^{\theta} - 1\right)^2} \\ & + 2\lambda^2 \sum_{i=0}^{n} \frac{\mathrm{e}^{\theta} \left(p_i + s_i\right) \left[\theta G_k'(x_i;\xi) \mathrm{e}^{\theta \overline{\mathrm{G}}(\mathrm{x_i};\xi)}\right]}{q_i \left(\mathrm{e}^{\theta} - 1\right)^3} \end{aligned}$$

and

$$\begin{aligned} U_{\xi_k\xi_r} \;=\; & -\theta \sum_{i=0}^{n} \left[\partial G_k'(x_i;\xi)/\partial\xi_r\right] + \frac{2\lambda\theta}{\left(\mathrm{e}^{\theta} - 1\right)} \sum_{i=0}^{n} \frac{G_k'(x_i;\xi) \mathrm{e}^{\theta \overline{\mathrm{G}}(\mathrm{x_i};\xi)} \left(\partial q_i/\partial\xi_r\right)}{q_i^2} \\ & + \sum_{i=0}^{n} \frac{g(x_i;\xi) \left[\partial g_k'(x_i;\xi)/\partial\xi_r\right]}{\left[g(x_i;\xi)\right]^2} - \sum_{i=0}^{n} \frac{g_k'(x_i;\xi) \left[\partial g(x_i;\xi)/\partial\xi_r\right]}{\left[g(x_i;\xi)\right]^2} \\ & - \frac{2\lambda\theta}{\left(\mathrm{e}^{\theta} - 1\right)} \sum_{i=0}^{n} \frac{\left\{\theta G_k'(x_i;\xi) \left[\partial \overline{G}(x_i;\xi)/\partial\xi_r\right] + \left[\partial G_k'(x_i;\xi)/\partial\xi_r\right]\right\}}{q_i \mathrm{e}^{-\theta \overline{\mathrm{G}}(\mathrm{x_i};\xi)}}. \end{aligned}$$

Appendix B

In this appendix we provide the `R` codes to compute cdf, pdf and maximum likelihood estimates:

```
# define cdf of CGTPLi distribution

cdf_CGTPLi <- function(alpha, theta, lambda,x) {

G = 1-(1+(alpha*x/(1+alpha)))*exp(-alpha*x)

g = alpha\symbol{94}2/(alpha+1)*(1+x)*exp(-alpha*x)

F = ((exp(theta)-exp(theta*(1-G)))/(exp(theta)-1))
```

```
*(1+lambda*((exp(theta*(1-G))-1)/(exp(theta)-1)))

f =theta*g*exp(theta*(1-G))*(1-lambda+2*lambda

*((exp(theta*(1-G))-1)/(exp(theta)-1)))/(exp(theta)-1)

return(F)

}

# define pdf of CGTPLi distribution

pdf_CGTLi <- function(alpha, theta, lambda,x) {

G = 1-(1+(alpha*x/(1+alpha)))*exp(-alpha*x)

g = alpha\symbol{94}2/(alpha+1)*(1+x)*exp(-alpha*x)

F = ((exp(theta)-exp(theta*(1-G)))/(exp(theta)-1))

*(1+lambda*((exp(theta*(1-G))-1)/(exp(theta)-1)))

f =theta*g*exp(theta*(1-G))*(1-lambda+2*lambda

*((exp(theta*(1-G))-1)/(exp(theta)-1)))/(exp(theta)-1)

return(f)

}

# Calculate the maximum likelihood estimators of CGTPLi distribution

library(bbmle)

x <- c(X) # X is the data set

n <- length(x)

z <- (1+alpha+alpha*x)/(1+alpha)

s <- exp(-alpha*x)

p <- theta*z*s

q <- ((exp(theta)+1)-2*exp(p))/(exp(theta)-1)

ll_CGTPLi <- function(alpha, theta, lambda){

n*log(theta)-n*log(exp(theta)-1)+2*n*log(1+alpha)-alpha*sum(x)

+ sum(log(1+x))+sum(log(p))+sum(log(1-lambda*q))

}
```

```
mle.res <- mle2(ll_CGTPLi, start=list(alpha=alpha, theta=theta, lambda=lambda),

 hessian.opt=TRUE)

summary(mle.res)

# variance covariance matrix of CGTPLi distribution

vcov(mle.res)
```

Appendix C

Active repair times data are: 0.50, 0.60, 0.60, 0.70, 0.70, 0.70, 0.80, 0.80, 1.00, 1.00, 1.00, 1.00, 1.10, 1.30, 1.50, 1.50, 1.50, 1.50, 2.00, 2.00, 2.20, 2.50, 2.70, 3.00, 3.00, 3.30, 4.00, 4.00, 4.50, 4.70, 5.00, 5.40, 5.40, 7.00, 7.50, 8.80, 9.00, 10.20, 22.00, 24.50.

Cancer patient data are: 0.08, 2.09, 3.48, 4.87, 6.94, 8.66, 13.11, 23.63, 0.20, 2.23, 3.52, 4.98, 6.97, 9.02, 13.29, 0.40, 2.26, 3.57, 5.06, 7.09, 9.22, 13.80, 25.74, 0.50, 2.46, 3.64, 5.09, 7.26, 9.47, 14.24, 25.82, 0.51, 2.54, 3.70, 5.17, 7.28, 9.74, 14.76, 26.31, 0.81, 2.62, 3.82, 5.32, 7.32, 10.06, 14.77, 32.15, 2.64, 3.88, 5.32, 7.39,10.34, 14.83, 34.26, 0.90, 2.69, 4.18, 5.34, 7.59, 10.66, 15.96, 36.66, 1.05, 2.69, 4.23, 5.41, 7.62, 10.75, 16.62, 43.01, 1.19, 2.75, 4.26, 5.41, 7.63, 17.12, 46.12, 1.26, 2.83, 4.33, 5.49, 7.66, 11.25, 17.14, 79.05, 1.35, 2.87, 5.62, 7.87, 11.64, 17.36, 1.40, 3.02, 4.34, 5.71, 7.93, 11.79, 18.10, 1.46, 4.40, 5.85, 8.26, 11.98, 19.13, 1.76, 3.25, 4.50, 6.25, 8.37, 12.02, 2.02, 3.31, 4.51, 6.54, 8.53, 12.03, 20.28, 2.02, 3.36, 6.76, 12.07, 21.73, 2.07, 3.36, 6.93, 8.65, 12.63, 22.69.

References

Afify A, Alizadeh M, Yousof H, Aryal G, Ahmad M (2016a). "The Transmuted Geometric-G Family of Distributions: Theory and Applications." *Pakistan Journal of Statistics*, (32), 139–160.

Afify A, Cordeiro G, Yousof H, Alzaatreh A, Nofal ZM (2016b). "The Kumaraswamy Transmuted-G Family of Distributions: Properties and Applications." *Journal of Data Science*, (14), 245–270.

Afify A, Yousof H, Nadarajah S (2016c). "The Beta Transmuted-H Family for Lifetime Data." *Statistics and Its Inference.*

Al-Babtain A, Fattah A, Ahmed A, Merovci F (2015). "The Kumaraswamy Transmuted Exponentiated Modified Weibull Distribution." *Journal of Data Science.*

Alexander C, Cordeiro G, Ortega E, Sarabia J (2012). "Generalized Beta Generated Distributions." *Computational Statistics and Data Analysis*, (56), 1880–1897.

Alizadeh M, Cordeiro G, Nascimento A, Lima M, Ortega E (2016). "Odd-Burr Generalizeed Family of Distributions with Some Applications." *Journal of Statistical Computation and Simulation*, (83), 326–339.

Alizadeh M, Emadi M, Doostparast M, Cordeiro G, Ortega E, Pescim R (2015). "A New Family of Distributions: The Kumaraswamy Odd Log-Logistic, Properties and Applications." *Hacettepa Journal of Mathematics and Statistics*, (83), 326–339.

Alzaatreh A, Lee C, Famoye F (2013). "A New Method for Generating Families of Continuous Distributions." *Metron*, (71), 63–79.

Cakmakyapan S, Kadilar G (2014). "A New Customer Lifetime Duration Distribution: The Kumaraswamy Lindley Distribution." *International Journal of Trade, Economics and Finance*, (5), 441–444.

Cordeiro G, de Castro M (2011). "A New Family of Generalized Distributions." *Journal of Statistical Computation and Simulation*, (81), 883–898.

Cordeiro GM, Alizadeh M, Tahir M, Mansoor M, Bourguignon M, Hamedani G (2015). "The Beta Odd Log-Logistic Generalized Family of Distributions." *Hacettepe Journal of Mathematics and Statistics*.

Eugene N, Lee C, Famoye F (2002). "Beta-Normal Distribution and its Applications." *Communications in Statistics-Theory and Methods*, (31), 497–512.

Flores J, Borges P, Cancho V, Louzada F (2013). "The Complementary Exponential Power Series Distribution." *Brazilian Journal of probability and statistics*, (27), 565–584.

Ghitany M, Al-Mutairi D, Balakrishnan N, Al-Enezi L (2013). "Power Lindley Distribution and Associated Inference." *Computational Statistics and Data Analysis*, (64), 20–33.

Jorgensen B (1982). "Statistical Properties of the Generalized Inverse Gaussian Distribution." *New York: Springer-Verlag*.

K Z, Balakrishnan N (2009). "On Families of Beta and Generalized Gamma Generated Distributions and Associated Inference." *Statistical Methodology*, (6), 344–362.

Khan M, King R (2013). "Transmuted Modified Weibull Distribution: A Generalization of the Mdified Weibull Probability Distribution." *European Journal of Pure and Applied Mathematics*, (6), 66–88.

Lee E, Wang J (2003). *Statistical Methods for Survival Data Analysis*. John Wiley and Sons/New York.

Marshall A, Olkin I (1997). "A New Method for Adding a Parameter to a Family of Distributions with Applications to the Exponential and Weibull Families." *Biometrika*, (84), 641–652.

Mead M, Afify A (2016). "On Five-Parameter Burr XII Distribution: Properties and Applications." *South African Statistical Journal*.

Merovci F, Elbatal I (2013). "The McDonald Modified Weibull Distribution: Properties and Applications." *arXiv preprint arXiv:1309.2961*.

Merovci F, Sharma V (2014). "The Beta Lindley Distribution: Properties and Applications." *Journal of Applied Mathematics*, (51), 1–10.

Nichols M, Padgett W (2006). "A Bootstrap Control Chart for Weibull Percentiles." *Quality and Reliability Engineering International*, (22), 141–151.

Nofal Z, Afify A, Yousof H, Cordeiro G (2016). "The Generalized Transmuted-G Family of Distributions." *Communications in Statistics-Theory and Methods*.

Shaw WT, Buckley IR (2007). "The Alchemy of Probability Distributions: Beyond Gram-Charlier Expansions and a Skew-Kurtotic-Normal Distribution From a Rank Transmutation Map." *arXiv preprint arXiv:0901.0434*.

Yousof H, Afify A, Alizadeh M, Butt N, Hamedani G, Ali M (2015). "The Transmuted Exponentiated Generalized-G Family of Distributions." *Pakistan Journal of Statistics and Operations Research*, (11), 441–464.

Yousof H, Afify A, Hamedani G, Aryal G (2016). "The Burr X Generator of Distributions for Lifetime Data." *Journal of Statistical Theory and Applications*.

Affiliation:

Morad Alizadeh
Department of Statistics, Faculty of Sciences
Persian Gulf University
Bushehr, Iran
E-mail: moradalizadeh78@gmail.com

Haitham M. Yousof
Department of Statistics, Mathematics and Insurance
Benha University
Egypt
E-mail: haitham.yousof@fcom.bu.edu.eg

Ahmed Z. Afify
Department of Statistics, Mathematics and Insurance
Benha University
Egypt
E-mail: AHMED.AFIFY@fcom.bu.edu.eg

Gauss M. Cordeiro
Departamento de Estatística
UniversidadeFederal de Pernambuco
Brazil
E-mail: gauss@de.ufpe.br

M. Mansoor
Department of Statistics
The Islamia University of Bahawalpur
Pakistan E-mail: mansoor.abbasi143@gmail.com

Austrian Journal of Statistics http://www.ajs.or.at/
published by the Austrian Society of Statistics http://www.osg.or.at/
Volume VV *Submitted:* yyyy-mm-dd
MMMMMM YYYY *Accepted:* yyyy-mm-dd