Journal of Sciences, Islamic Republic of Iran 32(3): 245 - 258 (2021) http://jsciences.ut.ac.ir
University of Tehran, ISSN 1016-1104

A New Lifetime Model, Stochastic Orders and Kidney Infection Regression Model

A. Khanjari Idenak, M.R. Zadkarami*, S.M.R. Alavi

Department of Statistics, Faculty of Mathematical Sciences and Computer, Shahid Chamran University of Ahvaz, Ahvaz, Iran

Received: 14 April 2021 / Revised: 12 June 2021 / Accepted: 28 June 2021

Abstract

We introduce a method to generate a new class of lifetime models based on the bounded distributions such that the defined models are exclusively a special case of the new class. A new subfamily, Generalized Alpha Power (GAP) is discussed and some stochastic orders in this subfamily are investigated to identify the proposed method effect. The performance of the maximum likelihood estimators based on the simulation is studied and in the end, the importance and flexibility of the new family for the models are illustrated by a real data set. Our results indicate that using the proposed method substantially improves the fitness of any G-family model and can be extended to any real data set. Finally, the GAPTW regression model is applied to the kidney infection data.

Keywords: Bounded Distribution; G-family Model; Lifetime; Regression; Stochastic Orders.

Introduction

There has been a considerable study to introduce new models that are based on bounded distributions, which can describe real-world phenomena. However, using stochastic orders to compare the theoretical effect of generating new models with the existing models is not considering by the authors [1, 2]. In the last decades, stochastic orders have been studied by researchers in many areas of probability, statistics, and other applied sciences include reliability theory, operations research, lifetime and survival analysis, and economics sciences [3-7]. Shaked and Shanthikumar [8] studied some of the most applied stochastic orders (i.e. distributions order, hazard order, and star order). They also investigated the relations between stochastic orders in more detail. Belzunce [9] reviewed some relationships between the main stochastic orders, and Belzunce et al. [10] considered some of the stochastic orders in the classical distribution.

Using stochastic orders to compare the elements of the family of distributions are studied by several authors as follows: Giovagnoli and Wynnb [11] investigated stochastic orderings for the family of the discrete random variables and Yu [12] perused stochastic ordering between exponential family and mixture exponential family distributions. Klenke and Mattner [13] inquired stochastic ordering in the family of the discrete distributions and stochastic ordering in the elliptical random vectors was introduced by Pan et al. [14]. Stochastic orders in the extended additive hazards models were considered by Raeisi and Yari [15] and Catana and Raducan [16] studied stochastic orders in the multivariate uniform distributions.

Many new models based on bounded distributions are developed and studied, however, the most popular

* Corresponding author: Tel: +986133331043; Fax: +98 6133331041; Email: zadkarami_m@scu.ac.ir

models are usually based on some limited bounded distributions with zero and one support such as Beta, Kumaraswamy (Jones [17]), and so on. Eugene et al. [18] proposed the Beta-Generated (B-G) family of distributions using $G(x) = \mathrm{B}\big(F(x)\big)$ where $B(t)$ and $F(x)$ denote the cumulative distributions functions (cdf) of the Beta and the random variable X, respectively. Many B-G distributions have been considered by the following authors. Beta-Weibull (Famoye et al. [19]), Beta-Exponential (Nadarajah and Kotz [20]), Beta-Pareto (Akinsete et al. [21]), Beta-Burr XII (Paranaiba et al. [22]) distributions are some typical examples. Recently, a new family of models is deployed by replacing the Beta with the Kumaraswamy distribution. Kumaraswamy-Weibull (Cordeiro et al. [23]), Kumaraswamy-generalized Gamma (Pascoa et al. [24]), and Kumaraswamy-generalized half-normal (Cordeiro et al. [25]) are some studies in this regard. The Alpha power (AP) distribution is another beta-type distribution which was considered by several authors to introduce the AP-G family of distributions, including Alpha power -Weibull (Dey et al. [2]), Alpha power –Inverse Lindley (Dey et al. [26]), and Alpha power -Lindley (Hassan et al. [27]).

In this research, a new flexible method is proposed to generate a lifetime distribution, when the generator distribution is bounded. The advantage of the proposed method is that the new model is based on $G(x) = \mathrm{R}(W_2[F(x)]) - R(W_1[F(x)])$ where $W_1[.]$ and $W_2[.]$ are linear functions. Moreover, $R(t)$ and $F(x)$ denote the cdf of any distribution defined on the bounded interval $[a, b]$ and the random variable X, respectively. It is worth noting that the defined models are based on $G(x) = \mathrm{R}\big(F(x)\big) - R(a)$ and therefore, the defined models are exclusively a special case of the new class. The proposed method allows researchers to define a flexible model that can apply to fit any real lifetime data. Moreover, we introduce a new subfamily named Generalized Alpha Power (GAPTW) and a theoretical comparison of the aforementioned model with the existing ones is done using stochastic orders, likelihood ratio order, star-shaped order, and so on.

The data are the recurrence times between the insertion of a catheter and the next infection of kidney patients who were using a portable dialysis machine (McGilchrist and Aisbett [28]). Risk factors related to the kidney infection of the catheter insertion is one of the considerable subjects for medical researchers (Erbay et al. [29]; Delistefani et al. [30]). The data set includes a sample of 38 patients and for each patient, first and second recurrence times (in days) of infection from the time of insertion of the catheter until it has to be removed owing to infection are recorded. The data also includes risk variables: age, sex, and disease types: Glomerulo Nephtiritis (GN), Acute Nephtiritis (AN), Polycystic Kidney Disease (PKD), and others. Furthermore, the data includes censoring because the catheter may have to be removed for reasons other than kidney infection. Some authors consider the data and applied various types of statistical modeling literature to model the kidney infection data (Hanagal and Dabade, [31]). We consider only the first recurrence time and due to flexibility, the GAPTW censored regression model is applied to the data to identify the risk factor related to the kidney infection data.

This paper is organized as follows: In Section 2, a new method to generate lifetime models based on the bounded distributions is proposed, and a member of the proposed family, namely Generalized Alpha Power (GAP) is presented along with discussing its general properties. In Section 3, Many stochastic orders between AP and GAP distributions are presented and an application is demonstrated to indicate the flexibility of the GAPTW, and the GAPTW regression model is applied for the kidney infection data. Section 4 concludes the paper.

Materials and Methods

Let X be a random variable with pdf and cdf $f(x)$ and $F(x)$, respectively. Moreover, consider a continuous random variable T with bounded pdf $r(t)$ is defined on $[a, b]$, where a and b can be any real values. The cdf of a new family of models is defined as:

$$G(x) = \int_{W_1(F(x))}^{W_2(F(x))} r(t)dt \qquad (1)$$

where functions $W_1(F(x))$ and $W_2(F(x))$ satisfy the following conditions:

1. $W_1(F(x)) \in [a, c]$ and $W_2(F(x)) \in [c, b]$, where $a \leq c \leq b$.
2. $W_1(x)$ is a differentiable and monotonically nonincreasing function of x, while $W_2(F(x))$ is a differentiable and monotonically nondecreasing function of x.
3. $W_1(F(x)) \to a$ as $x \to -\infty$, whilst $W_1(F(x)) \to c\epsilon[a, b]$ as $x \to \infty$. (2)
4. $W_2(F(x)) \to c\epsilon[a, b]$ as $x \to -\infty$ and $W_2(F(x)) \to b$ as $x \to \infty$.

The pdf $r(t)$ in (1) is transformed into a new cdf $G(x)$ through the functions $W_1(F(x))$ and $W_2(F(x))$ and we refer to the distribution $G(x)$ as a transformed from random variable T through the transformer random variable X. The corresponding pdf associated with (1) is

$$g(x) = r(W_2(F(x)))\frac{\partial}{\partial x}[W_2(F(x))] - r(W_1(F(x)))\frac{\partial}{\partial x}[W_1(F(x))]$$

Note that the random variable X can be either discrete or continuous. Also, in the aforementioned case, $G(x)$ denotes the cdf of a discrete or continuous distribution.

Consider (1), it is convenient to assume $W_1(.)$ and $W_2(.)$ as linear functions, because the pdf r(t) is defined on the bounded interval [a,b].

Theorem 1: Let X be a random variable with pdf and cdf $f(x)$ and $F(x)$, respectively. Also consider the continuous random variable T with pdf $r(t)$ defined on $[a, b]$. Suppose $W_1(F(x)) = d + eF(x)$ and $W_2(F(x)) = h + kF(x)$ are linear functions of $F(x)$ such that for $c \in [a, b]$, $W_1(F(x))$ and $W_2(F(x))$ are defined in the intervals $[a, c]$ and $[c, b]$, respectively. Then the cdf of the G-family is given by

$$G(x) = \int_{c-(c-a)F(x)}^{c+(b-c)F(x)} r(t)dt = \int_{aF(x)+c(1-F(x))}^{bF(x)+c(1-F(x))} r(t)dt \tag{3}$$

Proof: Since $W_1(F(x))$ is a monotonically non-increasing function of the cdf $F(x)$ in the interval $[a, c]$, the inequality $a \leq W_1(F(x)) = d + eF(x) \leq c$ satisfies. Moreover, $\lim_{x\to-\infty}[d + eF(x)] = d = c$ and $\lim_{x\to\infty}[d + eF(x)] = d + e = a$, therefore, $d = c$, $e = -(c - a)$ and then, $\frac{\partial W_1(F(x))}{\partial F(x)} = -(c - a) < 0$. Furthermore, $W_2(F(x))$ is a monotonically non-decreasing function of the cdf $F(x)$ results in $c \leq W_2(F(x)) = h + kF(x) \leq b$. As a result, $\lim_{x\to-\infty}[h + kF(x)] = h = c$ and $\lim_{x\to\infty}[h + kF(x)] = h + k = b$. Then, we have $h = c$, $k = b - a$ and $\frac{\partial W_2(F(x))}{\partial F(x)} = (b - a) > 0$ and, (3) satisfies.

Apparently, for $W_1(F(x)) = a = c$, equation (3) changes to $G(x) = \int_a^{a+(b-a)F(x)} r(t)dt$. Besides, if $W_2(F(x)) = b = c$ leads to $G(x) = \int_{b-(b-a)F(x)}^{b} r(t)dt$. Figure 1 displays $W_1(F(x))$ and $W_2(F(x))$ for different values of a, b, and c.

Corollary 1. For a special case, let $c = 0$, hence, $a \leq 0 \leq b$. As a consequent, equation (3) changes to

$$G(x) = \int_{aF(x)}^{bF(x)} r(t)dt \tag{4}$$

If $r(t)$ is a symmetric density function about zero, then $a = -b$ and $G(x) = 2\int_0^{bF(x)} r(t)dt$.

Corollary 2. Let the support of T be $[0, 1]$, then $0 \leq c \leq 1$, and the pdf $r(t)$ belongs to the beta-type distributions such as Beta, Kumaraswamy, and Alpha-power distributions. In this case,

$$G(x) = \int_{c-cF(x)}^{c+(1-c)F(x)} r(t)dt = \int_{c(1-F(x))}^{F(x)+c(1-F(x))} r(t)dt. \tag{5}$$

Moreover, for $c = 0$, depends on Beta, Kumaraswamy, or Alpha-power distributions for $r(t)$, the B-G, KW-G, or APT-G distributions are obtained, respectively. Furthermore, we can rewrite (5) as follows:

$$G(x) = \int_{W(F(x))}^{1-\frac{1-c}{c}W(F(x))} r(t)dt \tag{6}$$

where $W(F(x)) = c - cF(x)$. The corresponding pdf associated with (5) is given by

$$g(x) = \frac{c-1}{c} r(W(F(x)))\frac{\partial}{\partial x}[W(F(x))] - r(W(F(x)))\frac{\partial}{\partial x}[W(F(x))].$$

Corollary 3. In the case $b = 0$, then $a \leq c \leq 0$ and $G(x)$ in (3) is

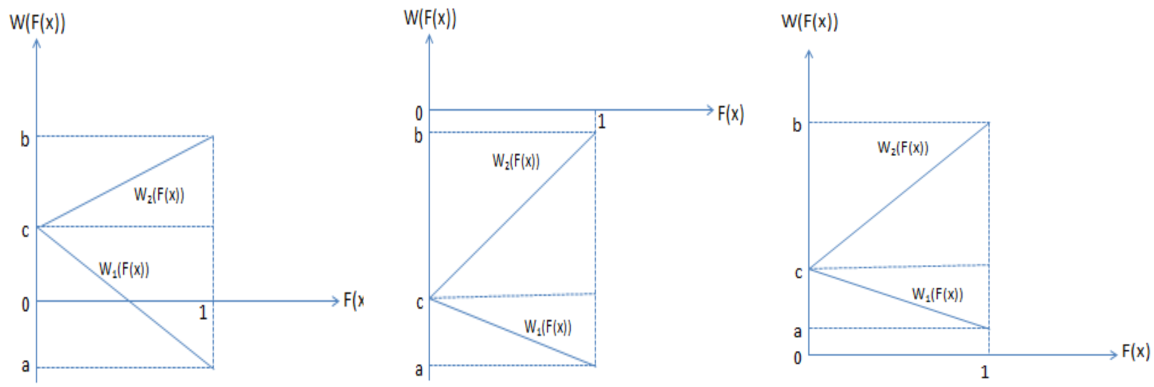

Figure 1. Plots of the functions $W_1(F(x))$ and $W_2(F(x))$ for $a < c < b$.

$$G(x) = \int_{c-(c-a)F(x)}^{c-cF(x)} r(t)dt = \int_{aF(x)+c(1-F(x))}^{c(1-F(x))} r(t)dt$$

Consider $c = b = 0$, thus $G(x) = \int_{aF(x)}^{0} r(t)dt$.

Corollary 4. Consider $a = 0$, then $0 \leq c \leq b$ and the equation (3) changes to

$$G(x) = \int_{c-cF(x)}^{c+(b-c)F(x)} r(t)dt = \int_{c(1-F(x))}^{bF(x)+c(1-F(x))} r(t)dt.$$

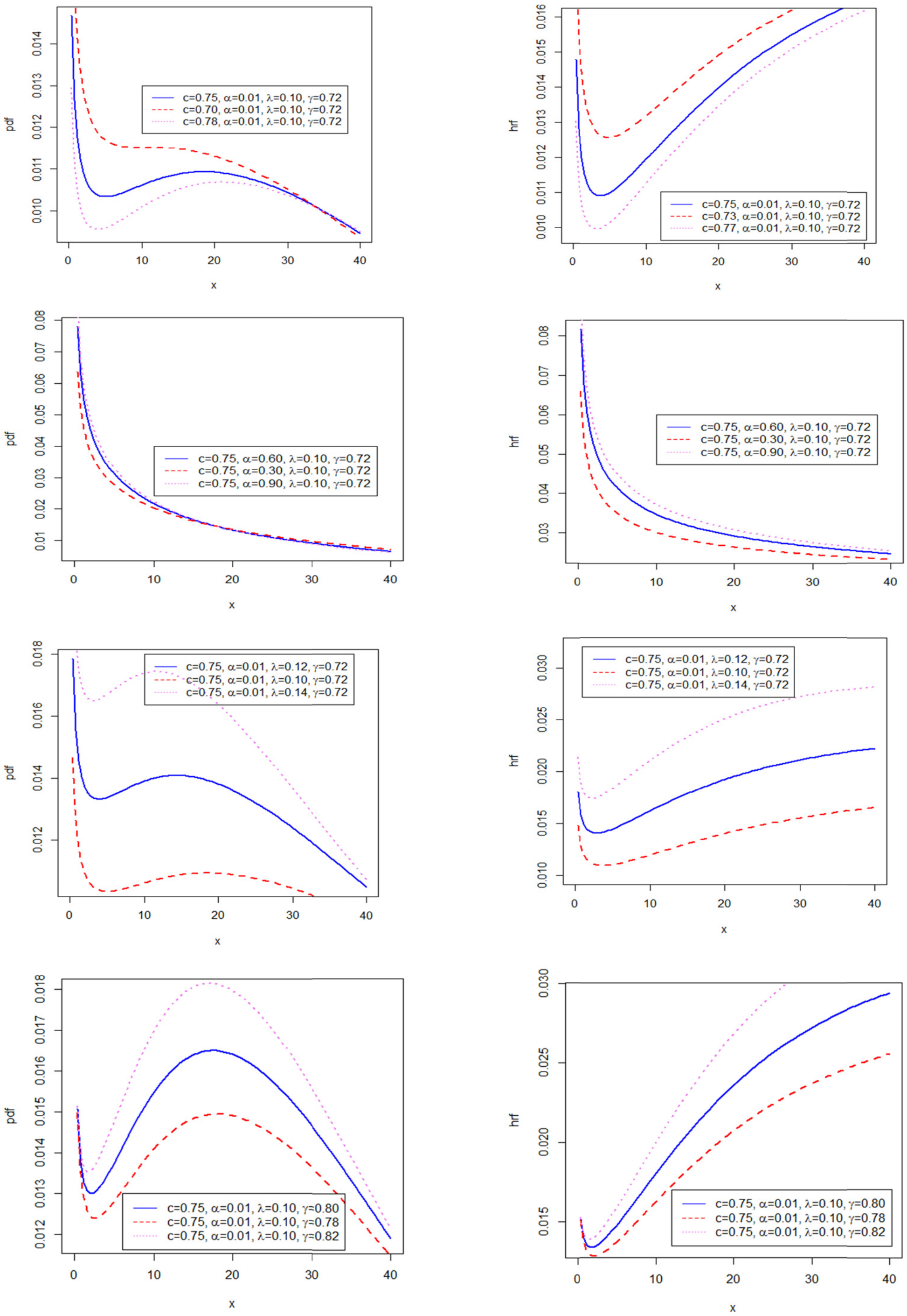

Figure 2. The Graphs of the pdf (left) and HRF (right) of the GAPW for selected values of parameters

The Class of $GAP(\alpha, c)$ Models

Consider the random variable T follows the α-power distribution with parameters α, then $r(t) = \frac{\ln(\alpha)}{\alpha-1}\alpha^t$, $0 \le t \le 1$ and $r(t) = 1$, $0 \le t \le 1$ for $\alpha = 1$. According to theorem 1, we have $a = 0$ and $b = 1$. With regard to from (3), The generalized alpha-power (GAP) cdf and pdf are respectively given by

$$G(x) = \int_{c-cF(x)}^{c+(1-c)F(x)} \frac{\ln(\alpha)}{\alpha-1}\alpha^t dt = \alpha^{c(1-F(x))}\frac{\alpha^{F(x)}-1}{\alpha-1} \quad (7)$$

and

$$g(x) = \frac{\ln(\alpha)}{\alpha-1} f(x)\alpha^{c(1-F(x))}[(1-c)\alpha^{F(x)} + c] \quad (8)$$

where $0 \le c \le 1$, $\alpha > 0, \alpha \ne 1$ are both the shape parameters and the new model is called the $GAP(\alpha, c)$ family of the distributions. The hazard function of the model is given by

$$h(x) = ln(\alpha)f(x)\alpha^{c(1-F(x))}\frac{[(1-c)\alpha^{F(x)}+c]}{\alpha^{c(1-F(x))}[1-\alpha^{F(x)}]+\alpha-1}. \quad (9)$$

For $c = 0$ in equation (8) the GAP distribution reduces to GPT, proposed by Mahdavi and Kundu [32] with cdf

$$G(x) = \begin{cases} \frac{\alpha^{F(x)}-1}{\alpha-1} & \alpha > 0, \alpha \ne 1 \\ F(x) & \alpha = 1 \end{cases}$$

For $c = \frac{1}{2}$ and $c = 1$, the GAP distribution reduces to $G(x) = \alpha^{\frac{1-F(x)}{2}}\frac{\alpha^{F(x)}-1}{\alpha-1}$ and $G(x) = \alpha^{1-F(x)}\frac{\alpha^{F(x)}-1}{\alpha-1}$, respectively. One of the popular distributions to use for modeling lifetime data is the Weibull (W) distribution, which is applied by several researchers in the last few years. Such that various types of the Weibull extensions are studied. In the GAP, consider the Weibull distribution as the baseline distribution with cdf $F(x) = 1 - exp(-\lambda x^\gamma)$, where λ and γ denote the scale and shape parameters, respectively. The cdf and pdf of the GAPTW distribution are respectively given by

$$G(x) = \alpha^{c.e^{-\lambda x^\gamma}}\frac{\alpha^{1-e^{-\lambda x^\gamma}}-1}{\alpha-1} \quad (10)$$

and

$$g(x) = \frac{\log(\alpha)\lambda\gamma x^{\gamma-1}e^{-\lambda x^\gamma}}{\alpha-1}\alpha^{c.e^{-\lambda x^\gamma}}[(1-c)\alpha^{1-e^{-\lambda x^\gamma}} + c \quad (11)$$

where $0 \le c \le 1$, $\alpha > 0, \alpha \ne 1$ are both shape parameters. The hazard rate function (hrt) is

$$h(x) = \log(\alpha)\lambda\gamma x^{\gamma-1}e^{-\lambda x^\gamma}\alpha^{c.e^{-\lambda x^\gamma}}\frac{[(1-c)\alpha^{1-e^{-\lambda x^\gamma}}+c]}{\alpha^{c.e^{-\lambda x^\gamma}}[1-\alpha^{1-e^{-\lambda x^\gamma}}]+\alpha-1}.$$

For $\alpha = 1$ from (3), the cdf of the GAPTW is:

$$G(x) = \int_{c-cF(x)}^{c+(1-c)F(x)} 1dt = [c + (1-c)F(x)] - [c - cF(x)] = F(x)$$

Which is independent of c and $G(x) = F(x)$ for all values $c \in [0,1]$. Figure 2 indicates that the shape of the GAPTW density function would be bimodal and decreasing. Moreover, the hazard rate shape can be decreasing and decreasing-increasing. For $\gamma = 1$, GAPTW distribution reduces to the GAPE distribution.

Some Statistical Properties of the GAP

In this section, we discuss some statistical properties of the GAPTW family, e.g. quantile, moment generating function, and moments. Consider the power series

$$\alpha^{-z} = \sum_{n=0}^{\infty}\frac{(-\ln(\alpha))^n}{n!}z^n . \quad (12)$$

The GAP density is expanded using the power series (12) such that for $\alpha > 0$, $\alpha \ne 1$, the density of the GAP is given by

$$g(x) = \frac{\ln(\alpha)\alpha^c}{\alpha-1}f(x)\left[(1-c)\alpha^{(1-c)F(x)} + c\alpha^{-cF(x)}\right] = \frac{\alpha^c\ln(\alpha)}{\alpha-1}f(x)\sum_{i=0}^{\infty}\frac{(\ln(\alpha)F(x))^i}{i!}[(c-1)^{i+1} - (-c)^{i+1}] = \sum_{i=0}^{\infty} b_i h_{i+1}(x) \quad (13)$$

where, $b_i = \frac{\alpha^c}{\alpha-1}\frac{(\ln(\alpha))^{i+1}}{(i+1)!}[(c-1)^{i+1} - (-c)^{i+1}]$ and $h_{i+1}(x) = (i+1)f(x)(F(x))^i$.

Consider the Exp-G random variable Y_{i+1} with density function $h_{i+1}(x)$. The cdf of the GAP family can be determined by $G(x) = \sum_{i=0}^{\infty} b_i H_{i+1}(x)$ where $H_{i+1}(x)$ denotes the cdf of the Exp-G random variable Y_{i+1}. Therefore, the explanation of the GAPTW density function is given by

$$g(x)=\frac{\alpha\gamma}{\alpha-1}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{[-(1-c)\ln(\alpha)]^{i+1}}{(i+1)!}\frac{[-(i+1)\lambda]^{j+1}}{j!}x^{\gamma(j+1)-1}+\frac{\alpha\gamma}{\alpha-1}\sum_{k=0}^{\infty}\sum_{l=0}^{\infty}\frac{\left(c.\ln(\alpha)\right)^{k+1}}{(k+1)!}\frac{[-(k+1)\lambda]^{l+1}}{l!}x^{\gamma(l+1)-1}$$
$$=\frac{\alpha\gamma}{\alpha-1}\sum_{i=0}^{\infty}\sum_{j=0}^{\infty}\frac{[\ln(\alpha)]^{i+1}}{(i+1)!}\frac{[-(i+1)\lambda]^{j+1}}{j!}[(c-1)^{i+1}+c^{i+1}]$$

Moments and Moments Generating Function: Using equation (13), the rth moment of X, say $\mu_r'=E(X^r)$, is obtained by

$$\mu_r'=\sum_{i=0}^{\infty}b_iE(Y_{i+1}^r)=\sum_{i=0}^{\infty}(i+1)b_i\delta_{r,i}$$

where $\delta_{r,i}=\int_0^1(Q_F(u))^r u^i du$ and $Q_F(u)=F^{-1}(u)$ denotes the baseline quantile function (qf). The moment generating function (mgf) $M_X(t)$ of X can be expressed from (12) as

$$M_X(t)=E(e^{tX})=\sum_{i=0}^{\infty}b_iM_{i+1}(x)=\sum_{i=0}^{\infty}(i+1)b_i\psi_{t,i}$$

where $M_{i+1}(x)$ denotes the mgf of Y_{i+1} for $(i\ge 0)$ and $\psi_{t,i}=\int_0^1 exp(tQ_F(u))\,u^i du$ can be evaluated numerically using the baseline qf.

Quantile: We can compute the quantile by equation $u=\alpha^{c(1-F(Q(u)))}\frac{\alpha^{F(Q(u))}-1}{\alpha-1}$. Therefore, the quantiles are given by $Q(u)=F^{-1}(\frac{\ln((\alpha-1)u+1)}{\ln(\alpha)})$ and $Q(u)=F^{-1}(1-\frac{\ln(\alpha-(\alpha-1)u+1)}{\ln(\alpha)})$ for $c=0$ and $c=1$, respectively.

Incomplete Moments: The rth incomplete moment of X, say $\varphi_r(t)$, can be written from (13) as

$$\varphi_r(t)=\int_{-\infty}^{t}x^rg(x)dx=\sum_{i=0}^{\infty}b_i\int_{-\infty}^{t}x^r\,h_{i+1}(x)\,dx \qquad (14)$$

where the last integral in (14) denotes the rth incomplete moment of the Exp-G random variable Y_{i+1}.

Order Statistics. Consider random sample $X_1,\dots,X_n$ from the GAP distribution and let the order statistics denote by $X_{(1)},\dots,X_{(n)}$. Using the generalized binomial expansion and the power series (12), the pdf of the rth order statistic $X_{r:n}$, say $g_{r:n}(x)$, is given by

$$g_{r:n}(x)=\frac{g_{GAP}(x)}{B(r,n-r+1)}\sum_{s=0}^{n-r}(-1)^s\binom{n-r}{s}G_{GAP}(x)^{s+r-1}, \qquad (15)$$

where $B(.,.)$ denotes the beta function. Using (7) and (8), we have

$$g_{GAP}(x)G_{GAP}(x)^{s+r-1}=\frac{\alpha^c\ln(\alpha)}{(1-\alpha)^{s+r}}f(x)\alpha^{c(1-F(x))}\left[(1-c)\alpha^{(1-c)F(x)}+c\right]\left[1-\alpha^{F(x)}\right]^{s+r-1}$$

Applying the generalized binomial expansion and the power series (12), we obtain

$$g_{GAP}(x)G_{GAP}(x)^{s+r-1}=\sum_{i,j=0}^{\infty}\frac{\alpha^c(-1)^j(\ln(\alpha))^{i+1}}{(1-\alpha)^{s+r}i!}\binom{s+r-1}{j}[(1-c)(s+r-c-j)^i+c(s+r-c-j-1)^i]f(x)F(x)^i \qquad (16)$$

Combining (15) and (16), the pdf of the rth order statistic $X_{r:n}$ is $g_{r:n}(x)=\sum_{i=0}^{\infty}k_ih_{i+1}(x)$, where $h_{i+1}(x)$ denotes the density of the Exp-G random variable Y_{i+1} and

$$k_i=\frac{\alpha^c(\ln(\alpha))^{i+1}}{(i+1)!}\sum_{j=0}^{\infty}\sum_{s=0}^{n-r}\frac{(-1)^{s+j}}{B(j,n-j+1)(1-\alpha)^{s+r}}[(1-c)(s+r-c-j)^i+c(s+r-c-j-1)^i]\binom{n-r}{s}\binom{s+r-1}{j}$$

Maximum Likelihood Estimation

We determine the maximum likelihood estimate of the parameters of the GAP distribution. Let

$x_1, x_2, \dots, x_n$ be the observed values from the GAP family distribution with parameters c and α. The log-likelihood function corresponding to (9) is given by

$$\ln(L(x; c, \alpha)) = n\ln\left(\frac{\ln(\alpha)}{\alpha-1}\right) + c\ln(\alpha)\left[n - \sum_{i=1}^{n} F(x_i)\right] + \sum_{i=1}^{n} \left[\ln((1-c)\alpha^{F(x_i)} + c) + \ln f(x_i)\right] \quad (17)$$

The partial derivatives of (17) are:

$$\frac{\partial}{\partial c}\ln(L(x; c, \alpha)) = \ln(\alpha)\left[n - \sum_{i=1}^{n} F(x_i)\right] + \sum_{i=1}^{n} \frac{1-\alpha^{F(x_i)}}{(1-c)\alpha^{F(x_i)}+c}$$

$$\frac{\partial}{\partial \alpha}\ln\big(L(x; c, \alpha)\big) = n\left[\frac{1}{\alpha ln(\alpha)} - \frac{1}{\alpha-1}\right] + \frac{c\left[n-\sum_{i=1}^{n} F(x_i)\right]}{\alpha} + \sum_{i=1}^{n} \frac{[(1-c)F(x_i)\alpha^{F(x_i)-1}]}{(1-c)\alpha^{F(x_i)}+c}$$

Setting $\frac{\partial}{\partial c}\ln(L(x; c, \alpha))$ and $\frac{\partial}{\partial \alpha}\ln(L(x; c, \alpha))$ equal to zero, and solving numerically these expressions simultaneously yields the maximum likelihood estimators (MLEs) of (c, α).

Theorems 2 and 3 indicate that the above equations have unique solutions $(\hat{c}, \hat{\alpha})$ which are the MLE of (c, α), Lemann and Casella [33].

Theorem 2: Equation $h_1(\alpha, c) = \frac{\partial}{\partial c}\ln\big(L(x; c, \alpha)\big) = 0$, as a function of c has only one root for $\alpha > 0, \alpha \neq 1$.

Proof: $h_1(\alpha, c) = \frac{\partial}{\partial c}\ln(L(x; c, \alpha)) = \ln(\alpha)\left[n - \sum_{i=1}^{n} F(x_i)\right] + \sum_{i=1}^{n} \frac{1-\alpha^{F(x_i)}}{(1-c)\alpha^{F(x_i)}+c}$. Consider $0 < \alpha < 1$, then,

$$\ln(\alpha)\left[n - \sum_{i=1}^{n} F(x_i)\right] < \left[n - \sum_{i=1}^{n} \alpha^{-F(x_i)}\right] \text{ and } \ln(\alpha)\left[n - \sum_{i=1}^{n} F(x_i)\right] > \left[n - \sum_{i=1}^{n} \alpha^{F(x_i)}\right]. \quad (18)$$

To show that $h_1(\alpha, c) = 0$ has only one solution for $0 < c < 1$, it is sufficed to show that the function $h_1(\alpha, c)$ is strictly descending because $\frac{\partial}{\partial c} h_1(\alpha, c) = -\sum_{i=1}^{n} \left[\frac{1-\alpha^{F(x_i)}}{(1-c)\alpha^{F(x_i)}+c}\right]^2 < 0$. Therefore, $h_1(\alpha, c)$ is strictly descending relative to

$$M = \lim_{c\to 0} h_1(\alpha, c) = \lim_{c\to 0}\left[\ln(\alpha)\left[n - \sum_{i=1}^{n} F(x_i)\right] + \sum_{i=1}^{n} \frac{1-\alpha^{F(x_i)}}{(1-c)\alpha^{F(x_i)}+c}\right]$$
$$= \ln(\alpha)\left[n - \sum_{i=1}^{n} F(x_i)\right] - \left[n - \sum_{i=1}^{n} \alpha^{-F(x_i)}\right]$$

and

$$m = \lim_{c\to 1} h_1(\alpha, c) = \lim_{c\to 1}\left[\ln(\alpha)\left[n - \sum_{i=1}^{n} F(x_i)\right] + \sum_{i=1}^{n} \frac{1-\alpha^{F(x_i)}}{(1-c)\alpha^{F(x_i)}+c}\right]$$
$$= \ln(\alpha)\left[n - \sum_{i=1}^{n} F(x_i)\right] + \left[n - \sum_{i=1}^{n} \alpha^{F(x_i)}\right]$$

respectively. From (18), we have $\ln(\alpha)\left[n - \sum_{i=1}^{n} F(x_i)\right] < \left[n - \sum_{i=1}^{n} \alpha^{-F(x_i)}\right]$, and $\ln(\alpha)\left[n - \sum_{i=1}^{n} F(x_i)\right] > \left[n - \sum_{i=1}^{n} \alpha^{F(x_i)}\right]$ when $c \to 0^+$ becomes positive. Hence, it refers to given values of $m < 0 < M$ and using the intermediate value of the theorem. The equation $h_1(\alpha, c) = 0$ has only one root. For $\alpha > 1$, the proof is similar.

Theorem 3: Equation $h_2(\alpha, c) = \frac{\partial}{\partial \alpha}\ln\big(L(x; c, \alpha)\big) = 0$, as a function of α has only one root.

Proof: Consider the equation $h_2(\alpha, c) = 0$, we have

$$\frac{\partial}{\partial \alpha} h_2(c, \alpha) = -n\left[\frac{(\alpha-1)^2[ln(\alpha)+1] - [\alpha\, ln(\alpha)]^2}{[\alpha(\alpha-1)\, ln(\alpha)]^2}\right] - \frac{c\left[n - \sum_{i=1}^{n} F(x_i)\right]}{\alpha^2} + \left[\sum_{i=1}^{n}\left[\frac{-c(1-c)[1-F(x_i)]\alpha^{F(x_i)-2} - (1-c)^2\alpha^{2F(x_i)-2}}{[(1-c)\alpha^{F(x_i)}+c]^2}\right] F(x_i)\right] < 0$$

Therefore, $h_2(c, \alpha)$ is a strictly decreasing function of α and, the absolute maximum and minimum values of $h_2(\alpha, c)$ are equal to

$$M = \lim_{\alpha\to 0^+} h_2(c, \alpha) = \lim_{\alpha\to 0^+}\left[n\left[\frac{1}{\alpha ln(\alpha)} - \frac{1}{\alpha-1}\right] + \frac{c\left[n - \sum_{i=1}^{n} F(x_i)\right]}{\alpha} + \sum_{i=1}^{n} \frac{[(1-c)F(x_i)\alpha^{F(x_i)-1}]}{(1-c)\alpha^{F(x_i)}+c}\right] = +\infty$$

and, $m = \lim_{\alpha\to\infty} h_2(c, \alpha) = \lim_{\alpha\to\infty}\left[n\left[\frac{1}{\alpha ln(\alpha)} - \frac{1}{\alpha-1}\right] + \frac{c\left[n-\sum_{i=1}^{n} F(x_i)\right]}{\alpha} + \sum_{i=1}^{n} \frac{[(1-c)F(x_i)\alpha^{F(x_i)-1}]}{(1-c)\alpha^{F(x_i)}+c}\right] = 0^-$

respectively, since for each value of α and $i = 1,2,3,\dots,n$ we have $n - \sum_{i=1}^{n} \alpha^{-F(x_i)} < 0$ and $n - \sum_{i=1}^{n} \alpha^{F(x_i)} > 0$.

Results and Discussion

Stochastic order relationships between $G_{AP}(x)$ and $G_{GAP}(x)$

Stochastic orders are utilized to compare distribution functions because they include various forms of possible knowledge about distribution functions. There are many concepts of the stochastic ordering between distributions which are defined in Probability and Statistics (for more detail see Shaked and Shanthikumar [8]). In this section, some stochastic orders between $G_{AP}(x)$ and $G_{GAP}(x)$ distributions are studied to identify the effect of the proposed method on the new distributions ordering when using the α-power distribution as a baseline distribution.

Let $F(x)$ and $G(x)$ denote the cdfs of two nonnegative random variables X and Y with absolutely continuous distribution functions respectively. Consider the survival function $\bar{F}(x) = 1 - F(x)$, the hazard rate $h_F(x) = \frac{f(x)}{\bar{F}(x)}$ and the mean residual life $m_X(t) = E[X - t|X > 0]$ of the random variable X where $t < t^*$ and $t^* = Sup\{t: \bar{F}(x) > 0\}$. Over the union of the supports of X and Y, the random variable X is said to be

i) smaller than Y in the *(usual) stochastic order* ($X \le_{st} Y$), if $F(x) \ge G(x), \forall x \in (-\infty, \infty)$.
ii) smaller than Y in the *star-shaped order* ($X \le_{*} Y$), if $\frac{G^{-1}(x)}{F^{-1}(x)}$ is an increasing function of $x \in (0,1)$.
iii) smaller than Y in the *hazard rate order* ($X \le_{hr} Y$), if $h_F(x) \ge h_G(x)$.
iv) smaller than Y in the *mean residual life order* ($X \le_{mrl} Y$) if $m_X(t) \le m_Y(t)$ for all t.
v) smaller than Y in the *likelihood ratio order* ($X \le_{lr} Y$) if $\frac{g(x)}{f(x)}$ increases in x where $f(x)$ and $g(x)$ denote the pdfs of the random variables X and Y, respectively.
vi) smaller than Y in the *convex transform order* ($X \le_{c} Y$) if $G^{-1}F(x)$ is convex in x on the support of F.

Some authors studied the relationships between different stochastic orders. For example, Muller and Stoyan [4] shown that

$$X \le_{lr} Y \Rightarrow X \le_{hr} Y \Rightarrow X \le_{st} Y$$

and Shaked and Shanthikumar [8] proven that

$$X \le_{hr} Y \Rightarrow X \le_{mrl} Y \text{ and } X <_{c} Y \Rightarrow X <_{*} Y.$$

Therefore, the likelihood ratio and convex transform orders are considered in this paper.

Theorem 4. Let Y and Z denote the random variables from $AP(\alpha)$ and $GAP(c,\alpha)$ respectively, then

i) $Y \le_{lr} Z$ for $\alpha \in (0,1)$
ii) $Z \le_{lr} Y$ for $\alpha \in (1,\infty)$

Proof. We have

$$w = \frac{g_{GAP}(x)}{g_{AP}(x)} = \frac{\frac{f(x)\ln(\alpha)\alpha^{c(1-F(x))}[(1-c)\alpha^{F(x)}+c]}{\alpha-1}}{\frac{f(x)\ln(\alpha)\alpha^{F(x)}}{\alpha-1}} = \alpha^{c-(1+c)F(x)}[(1-c)\alpha^{F(x)}+c]$$

It can see that

$$\frac{\partial w}{\partial x} = -cf(x)\ln(\alpha)\alpha^{c-(1+c)F(x)}\big[(1-c)\alpha^{F(x)} + (1+c)\big]$$

For $0 < \alpha < 1$, $ln(\alpha)$ is negative and because $c \in (0,1)$, then $\frac{\partial w}{\partial x}$ is positive and $\frac{g_{GAP}(x)}{g_{AP}(x)}$ is an increasing function of x. Therefore, $Y \le_{lr} Z$ for $\alpha \in (0,1)$.

For $1 < \alpha < \infty$, $ln(\alpha)$ is a positive value and thus, $\frac{\partial w}{\partial x}$ is negative. Then, $\frac{g_{GAP}(x)}{g_{AP}(x)}$ decreases in x and, therefore $\frac{g_{AP}(x)}{g_{GAP}(x)}$ is an increasing function of x. So, $Z \le_{lr} Y$ for $\alpha \in (1,\infty)$.

Theorem 5. Suppose random variables X and Z are from distributions $F(x)$ and $GAP(c,\alpha)$ based on $F(x)$ respectively, then

i) $Z \le_{lr} X$ if $\alpha \in (0,1)$ and $\frac{c^2}{(1-c)^2} < \alpha$.
ii) $X \le_{lr} Z$ if $\alpha \in (1,\infty)$ and $c < \frac{1}{2}$.

Proof. We have

$$w_1 = \frac{g_{GAP}(x)}{f(x)} = \frac{ln(\alpha)\alpha^{c(1-F(x))}[(1-c)\alpha^{F(x)}+c]}{\alpha-1}$$

Thus

$$\frac{\partial w_1}{\partial x} = \frac{f(x)(ln(\alpha))^2\alpha^{c(1-F(x))}}{\alpha-1}\big[(1-c)^2\alpha^{F(x)} - c^2\big].$$

The value $\alpha - 1$ is negative for $\alpha \in (0,1)$ and $\alpha < \alpha^{F(x)} < 1$, then $(1-c)^2\alpha^{F(x)} - c^2$ is positive if $\frac{c^2}{(1-c)^2} < \alpha$. Thus $\frac{\partial w_1}{\partial x}$ is negative and $\frac{g_{GAP}(x)}{f(x)}$ is a decreasing function of x. Therefore, $\frac{f(x)}{g_{GAP}(x)}$ is an increasing function of x and thus, $Z \le_{lr} X$.

Besides, for $1 < \alpha < \infty$, we have $1 < \alpha^{F(x)} < \alpha$ and therefore, the value of $(1-c)^2\alpha^{F(x)} - c^2$ is positive if $c < \frac{1}{2}$. In this case, $\frac{\partial w_1}{\partial x}$ is positive and $\frac{g_{GAP}(x)}{f(x)}$ increases for $x \in (0,1)$ and then, $X \le_{lr} Z$.

Theorem 6. Assume X and Y are two random variables from distributions $F(x)$ and $AP(\alpha)$, based on $F(x)$, respectively, then

i) $Y \le_{lr} X$ for $\alpha \in (0,1)$
ii) $X \le_{lr} Y$ for $\alpha \in (1,\infty)$

Proof. In this case, $w_2 = \frac{g_{AP}(x)}{f(x)} = \frac{ln(\alpha)\alpha^{F(x)}}{\alpha-1}$ and

$\frac{\partial w_2}{\partial x} = \frac{f(x)(ln(\alpha))^2 \alpha^{F(x)}}{\alpha-1}$. The value of $\frac{\partial w_2}{\partial x}$ is negative for $\alpha \in (0,1)$, because $\alpha - 1$ is negative. Therefore, $\frac{g_{AP}(x)}{f(x)}$ is a decreasing function of x and $\frac{f(x)}{g_{AP}(x)}$ is an increases function of x and thus, $Y \leq_{lr} X$.

Consider $\alpha \in (1, \infty)$, then $\frac{g_{AP}(x)}{f(x)}$ is an increasing function of x and so, $X \leq_{lr} Y$.

Theorem 7. Let Y and Z denote two random variables from $AP(\alpha)$ and $GAP(c, \alpha)$ distributions respectively, then

i) $Y \leq_c Z$ if $\frac{[1-2c]\ln(\alpha)}{(1-c)\alpha+c} < H'(x) < ln(\alpha)$
ii) $Z \leq_c Y$ if $ln(\alpha) < H'(x) < ln(\alpha)[(1-c)^2\alpha - c^2]$

where $H(x) = \frac{1}{f(x)}$ and $H'(x)$ denotes the derivative of $H(x)$.

Proof. It is sufficient to show that $G_{GAP}^{-1} G_{AP}(x)$ is a convex function and it should indicate that $G_{AP}(x)$ is convex and $G_{GAP}^{-1}(x)$ is a non-decreasing and convex function (Mrsevi [34]; Boyd and Vandenberghe [35]). The second derivative of $G_{AP}(x)$ is given by

$$G''_{GAP}(x) = \frac{\partial^2 G_{GAP}}{\partial x^2} = \frac{ln(\alpha)}{\alpha - 1} \alpha^{F(x)} [f'(x) + f(x)^2 ln(\alpha)]$$

$f'(x)$ denotes the derivative of $f(x)$ and $G''_{GAP}(x)$ is positive if $f'(x) + f(x)^2 ln(\alpha) > 0$ or $-\frac{f'(x)}{f(x)^2} < ln(\alpha)$, then $H'(x) < ln(\alpha)$. (19)

For convexity of $F^{-1}(x)$, it is sufficient to show that $F(x)$ is concave, because $F(x)$ and $F^{-1}(x)$ are increasing functions (Boyd and Vandenberghe [35]). Therefore, it is should be shown the second derivative of $G_{GAP}(x)$ is negative. But,

$$G''_{GAP}(x) = \frac{ln(\alpha)}{\alpha - 1} \alpha^{c(1-F(x))} [f'(x)[(1-c)\alpha^{F(x)} + c] + f(x)^2 ln(\alpha)[(1-c)^2 \alpha^{F(x)} - c^2]]$$

Therefore, $G_{GAP}''(x)$ is negative if

$$f'(x)[(1-c)\alpha^{F(x)} + c] + f(x)^2\, ln(\alpha)\, [(1-c)^2 \alpha^{F(x)} - c^2] < 0$$

or

$$f'(x)[(1-c)\alpha^{F(x)} + c] < -f(x)^2\, ln(\alpha)\, [(1-c)^2 \alpha^{F(x)} - c^2]$$

Therefore, $-\frac{f'(x)}{f(x)^2} > ln(\alpha) \frac{[(1-c)^2\alpha^{F(x)} - c^2]}{[(1-c)\alpha^{F(x)} + c]}$. But, $ln(\alpha) \frac{[(1-c)^2\alpha^{F(x)} - c^2]}{[(1-c)\alpha^{F(x)} + c]} > \frac{[1-2c]}{(1-c)\alpha + c} ln(\alpha)$. Then

$\frac{[1-2c]}{(1-c)\alpha+c} ln(\alpha) < H'(x)$
(20)

Thus, $G_{AP}(x)$ is a concave function. Because $G_{GAP}^{-1}(x)$ is convex, then by (19) and (20) when

$$\frac{[1-2c]ln(\alpha)}{(1-c)\alpha + c} < H'(x) < ln(\alpha)$$

$G_{GAP}^{-1} G_{AP}(x)$ is a convex function and then, $Y \leq_c Z$.

The proof of (ii) is similar.

Data Analysis

In this section, the performance of GAPTW distribution is considered by applying real data sets. The model parameters are estimated using the MLE. The well-known goodness-of-fit criteria Akaike information criterion (AIC) and Kolmogorov-Smirnov (K-S) are also used to compare the models. In general, small values of AIC statistics and a large p-value of K-S indicate a good model fit for the data. For the sake of visual comparison, the plots of the pdfs and cdfs of the fitted distributions are presented in Figure 3. The regarded computations were carried out in the R and MATHLAB software. The data set illustrates the strength of the Alumina (Al2O3) material, which can be found in Nadarajah and Kotz [36]. We applied to sub-models of the proposed G-family based on the bounded distribution: GB(EW), KUW, and APTW distributions then with three other well-known competing distributions. The Weibull is the baseline distribution of all six G-family models. Therefore, we can identify the effect of the proposed method on the popular three G-families. The distribution functions of the competing models are:

•Beta-Weibull or Exponentiated Weibull (EW), Mudholkar et al. [37]

$$G(x) = (1 - e^{-\lambda x^{\gamma}})^{\alpha}, \quad x, \alpha, \lambda, \gamma > 0$$

•Kumaraswamy Weibull (Ku-W), Cordeiro et al. [23]

$$G(x) = 1 - (1 - (1 - e^{-\lambda x^{\gamma}})^{\alpha})^{\beta}, \quad x, \alpha, \beta, \lambda, \gamma > 0$$

•Alpha power transformed Weibull (APTW), Dey et

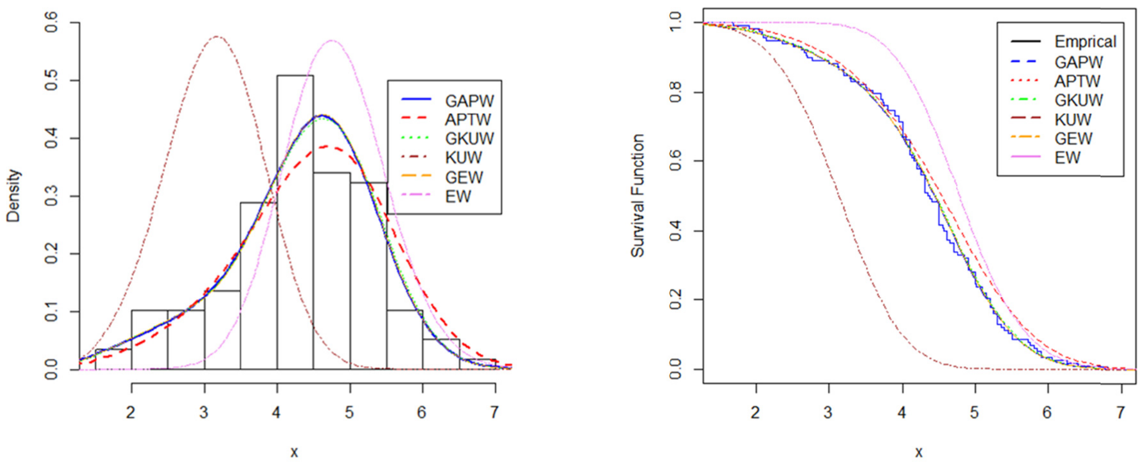

Figure 3. Histogram (left) and empirical and fitted survival function (right) to the Al2O3 data set.

al. [2]

$$G(x) = \frac{\alpha^{\left(1-e^{-\lambda x^{\gamma}}\right)}-1}{\alpha-1}, \quad x, \alpha, \lambda, \gamma > 0, \alpha \neq 1\cdot$$

Figure 3 depicts the histogram and the fitted GAPTW distribution density with five other competing distributions for the data set. We can see that the GAPTW distribution provides a perfect fit for the data set. Table 1 indicates the MLEs of the parameters, AIC, and K–S test statistics for all fitted models. With respect to Table 1, it is observed that using the proposed method improves the model fitting substantially for either of the following G-families: B-G, KU-G, or APT-G. Moreover, the values of AIC and K–S statistics are implied on the GAPTW distribution provides the perfect fit among all fitted distributions.

The GAPTW Regression Model

Let response variable x be distributed as GAPTW distribution with the hazard function $h(x)$ which is defined in equation (9), then

$$h(x) = ln(\alpha)\frac{f(x)}{1-F(x)}\alpha^{c(1-F(x))}\frac{(1-F(x))[(1-c)\alpha^{F(x)}+c]}{\alpha^{c(1-F(x))}[1-\alpha^{F(x)}]+\alpha-1} = \lambda(x)h_0(x)$$

where $\lambda(x)$ denotes the hazard function of the baseline density function $f(x)$ and

$$h_0(x) = ln(\alpha)\,\alpha^{c(1-F(x))}\frac{(1-F(x))[(1-c)\alpha^{F(x)}+c]}{\alpha^{c(1-F(x))}[1-\alpha^{F(x)}]+\alpha-1} \geq 0\ .$$

Because, for $0 < \alpha < 1$, $ln(\alpha) < 0$ and $1-\alpha^{F(x)} < 1-\alpha$, and then $\alpha^{c(1-F(x))}\left[1-\alpha^{F(x)}\right]+\alpha-1 < 0$ however, for $1 < \alpha < \infty$, $ln(\alpha) > 0$ and $1-\alpha^{F(x)} > 1-\alpha$ thus $\alpha^{c(1-F(x))}\left[1-\alpha^{F(x)}\right]+\alpha-1 > 0$ and therefore, $h_0(x) > 0$ for $0 < \alpha < \infty$.

Consider the Weibull density function $f(x) = \gamma\lambda x^{\lambda-1}e^{-\gamma x^{\lambda}}$, $x > 0, \lambda, \gamma > 0$ with the hazard function $\gamma\lambda x^{\lambda-1}$. Then, the Weibull distribution belongs to the class of the Cox-proportional hazard distributions when it is assumed that the scale parameter is defined as $\gamma = exp(\gamma_0+\gamma_1 Z_1+\cdots+\gamma_p Z_p)$, where $Z = (Z_1, Z_2, \ldots, Z_p)$ denotes the vector of explanatory variables. Therefore, the hazard of the GAPTW distribution is given by

$$h(x) = h_0(x)\lambda x^{\lambda-1}e^{\gamma_0+\gamma_1 Z_1+\cdots+\gamma_p Z_p}.$$

Then, the hazard function of the GAPTW distribution.is affected by the explanatory variables. Let d_i denote the indicator variable taking value zero if the ith subject (i = 1,...,n) censored and value 1 otherwise. Consider (x_i, cen_i, Z_i), $i = 1,2,\ldots,n$ where x_i follows the GAPTW distribution, cen_i denotes the censore time and $Z_i = (Z_{i1}, \ldots, Z_{ip})$ is the set of the explanatory variables related to the ith individual. Assume x_i and cen_i are conditionally on Z_i are independent and $y_i = \min(x_i, cen_i)$, $i = 1,2,\ldots,n.$ The likelihood function is given by

$$L(\theta) = \prod_{i=1}^{n} g_{GAPTW}^{d_i}(y_i)[1-G_{GAPTW}(y_i)]^{1-d_i}$$

where θ denotes the parameters set and g_{GAPTW} and G_{GAPTW} are the density and cumulative distribution of GAPTW distribution which is given by equations (11) and (10), respectively.

Table 1. The MLEs and the goodness of fit statistics for the Alumina (Al2O3) material data.

Model	Parameters Estimates					Statistics	
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\gamma}$	$\hat{c}$	AIC	$K-S$
$GAPTW$	0.00013	-	0.00379	3.95735	0.31084	342.684	0.0553
$APTW$	4.958	-	0.003	3.931	0	347.21	0.144
GEW	8.05	-	0.0028	4.1118	0.702	342.1986	0.0552
EW	4.867	-	0.007	3.620	0	347.92	0.082
$GKUW$	0.3865	3.044	0.0037	4.444	0.272	344.163	0.0552
KUW	1.633	2.301	0.009	0.657	0	348.10	0.59

Application: Kidney Infection Data

The kidney infection data, which firstly was applied by McGilchrist and Aisbett [28] is related to the recurrence time to infection at the point of insertion of the catheter for 38 kidney patients using portable dialysis equipment. For each patient, first and second recurrence times (in days) of infection from the time of insertion of the catheter until it has to be removed owing to infection are recorded. The data includes censoring because the catheter may have to be removed for reasons other than kidney infection.

We consider only the first recurrence time with risk variables: age, sex (1 = male, 2 = female), censored indicator (1 = infection occurs, 0 = censored) and disease types which are coded as 0 = GN, 1 = AN, 2 = PKD and 3 = others. After the occurrence or censoring of the first infection, sufficient time (10 weeks interval) was allowed for the infection to be cured before the second time the catheter was inserted. Thus, the first and second recurrence times can be considered independent.

Firstly, the models based on Weibull distribution: GAPTW, APTW, GEW, EW, GKuW, and KuW are fitted for the first recurrence time to infection at the point of insertion of the catheter to identify the best distribution in modeling the data. Table 2 reports the result of fitting the comparative models. As can see, applying the new method improves the models fitting effectively for all models: APTW, EW, and KuW, and the best model is produced by the GAPTW.

A censored regression model based on the best distribution (GAPTW) is applied to the data to identify the risk factors that affect the hazard rate of the first recurrence time to infection and the results of the fitting censoring regression. Comparing the goodness statistic AIC for the GAPTW model in Tables 2 and 3 shows that explanatory variables improve the model fitting. Table 3 indicates that the explanatory variable age affects the recurrence time negatively. We also found that a lower infection rate for female patients. Moreover, disease types AN, PKD, and others have less infection rate than that of the GN type of the disease.

Simulation

In this section, we study the performance of the maximum likelihood estimators (MLE) for the GAPTW distributions with different sample sizes n = 50, 100, 200, 500, 1000, and 2000. We simulate 10000 samples for each sample size with two sets of parameters combinations; Data Set (1): $c = 0.2$, $\alpha = 0.2$, $\lambda = 1.3$, $\gamma = 1.3$ and Data Set (2): $c = 0.04$, $\alpha = 0.04$, $\lambda = 2.4$, $\gamma = 3.5$ for GAPTW distribution. The simulation study was conducted with R-software. Furthermore, various

Table 2. The MLEs and the goodness statistic for the kidney infection data.

Model	Parameters Estimates (SD in parenthesis)					Statistics
	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	$\hat{\gamma}$	$\hat{c}$	AIC
$GAPTW$	0.000199	-	0.001539	1.1748	0.06258	308.1
	(0.001045)		(0.001189)	(0.2079)	(0.03926)	
$APTW$	0.5570	-	0.006672	0.9987	0	311.0
	(0.0617)		(0.001075)	(0.1207)		
GEW	0.4923	-	0.000424	1.4271	0.000056	311.5
	(0.1575)		(0.000426)	(0.1669)	(0.000310)	
EW	1.0986	-	0.01446	0.8773	0	313.7
	(0.4284)		(0.01428)	(0.1569)		
$GKuW$	3.2030	0.8541	0.001507	1.0642	0.9987	311.5
	(1.9966)	(0.3486)	(0.001082)	(0.1688)	(0.01081)	
KuW	1.6735	2.2881	0.02569	0.7018	0	313.9
	0.6580)	(1.0273)	(0.02285)	(0.1409)		

Table 3. The MLEs of the Regression model for the kidney infection data.

Variables	Categories	Parameter (SD)
Constant		-8.3395 (0.9130)
Age		-0.00899 (0.007187)
Sex	Male (reference)	-
	Female	-0.5477(0.5134)
Disease type	GN (reference)	-
	AN	-0.6644 (0.5817)
	PKD	-2.0536 (0.7820)
	Others	-0.1796 (0.5539)
α		1E-8 (1E-14)
c		0.03298 (0.01068)
λ		1.8395 (0.1849)
AIC		302.3

Table 4. Bias and MSEs (in prentice) of estimators for the GAPTW distribution.

n	Data Set 1				Data Set 2			
	$\hat{c}$	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\gamma}$	$\hat{c}$	$\hat{\alpha}$	$\hat{\lambda}$	$\hat{\gamma}$
50	-0.0640	-0.1655	-0.0657	0.2343	0.0284	-0.0120	1.0940	0.3139
	(0.0281)	(0.0397)	(0.3718)	(0.1497)	(0.0133)	(0.0096)	(7.2498)	(0.3588)
100	-0.0255	-0.1355	-0.0086	0.1244	0.0276	0.0017	0.5744	0.1683
	(0.0380)	(0.0411)	(0.3028)	(0.0737)	(0.0134)	(0.0130)	(3.3836)	(0.1396)
200	-0.0010	-0.0946	0.0013	0.0659	0.0160	0.0112	0.2705	0.0863
	(0.0498)	(0.0423)	(0.2119)	(0.0384)	(0.0085)	(0.0127)	(1.7136)	(0.0624)
500	0.0017	-0.0219	0.0047	0.0310	0.0018	0.0101	0.0932	0.0362
	(0.0499)	(0.0490)	(0.0951)	(0.0157)	(0.0019)	(0.0063)	(0.6852)	(0.0224)
1000	-0.0188	0.0201	0.0002	0.0198	-0.0012	0.0072	0.0543	0.0173
	(0.0354)	(0.0529)	(0.0409)	(0.0072)	(0.0011)	(0.0026)	(0.3320)	(0.0109)
2000	-0.0392	0.0452	-0.0016	0.0117	-0.0025	0.0052	0.0274	0.0079
	(0.0207)	(0.0497)	(0.0179)	(0.0033)	(0.0006)	(0.0011)	(0.1565)	(0.0057)

criteria such as the bias of the estimators (i.e. $Bias(\theta) = \frac{1}{n}\sum_{i=1}^{n}[\theta_i - \widehat{\theta}]$) and the mean square errors (MSEs) of the estimators of the parameters (i.e. $MSE(\theta) = \frac{1}{n}\sum_{i=1}^{n}[\theta_i - \widehat{\theta}]^2$) are obtained which are reported in Table 4.

Table 4 indicates that the MSEs of parameters decrease as the sample size increases. Furthermore, the results in Table 4 show that the parameters estimates are quite stable; More importantly, the values of the estimates are close to the true values for these sample sizes. Thus, the MLE of the model parameters performs well for the GAPTW distribution.

Comparison of the new G-Families of Distributions

The performance of the proposed method to generate more flexible new G-families than existing G-families of distributions is done using a simulation was studied. To identify the flexibility of GAPTW GKuW, GEW compare to APTW, KuW, and EW respectively, we generated six samples with size 20000 from comparative GAPTW, APTW, GKuW, KuW, GEW, and EW distributions. The results of fitting models to different data sets are presented in Table 5. As it can be seen, the GAPTW, GKuW, GEW in comparison to APTW, KuW, and EW respectively, give better results.

Conclusion

In this paper, a method to generate new families of the models was introduced. We show that the novel method generates a large number of new models as well

Table 5. -Log-likelihood of the GAPTW, APTW, GKuW, KuW, GEW, and EW.

		Fitted Distributions					
		GAPTW	APTW	GKuW	KuW	GEW	EW
Generated Distributions	GAPTW	28088.24	28216.04				
	APTW	29492.27	29492.36				
	GKuW			29743.53	30000.46		
	KuW			21008.71	21009.07		
	GEW					28065.83	28409.95
	EW					21230.24	21130.24

as existing models as special cases. Besides, some properties of the new member of the family, the GAP model, were considered and some of the main stochastic orders are applied to compare the GAP and AP models. We found that the likelihood ratio order between GAP and AP models depends on the parameter value of α however, the convex transform order depends on the value of parameters α and c. The properties and the estimation of the parameters of the new family of models were also presented and, the real data set was applied in order to illustrate the flexibility of the new method. The results showed that the new model is a suitable model to fit data sets with various kinds of shapes. The results in Table 1 show that using the proposed method considerably improves the fitness of any G-family model. In fit the compatible models to the kidney infection data, the GAPTW produces the best model, and in fitting the GAPTW regression model, we found that sex and type of disease (PKD) affect significantly the recurrence times of the kidney infection.

Acknowledgements

The authors express their profound gratitude toward the Associate Editor and referees for their insightful comments and suggestions which have greatly improved this research. This research was supported by the Shahid Chamran University of Ahvaz, Ahvaz, Iran (Grant no. SCU.MS98.5).

References

1. Cordeiro GM, De Castro M. A new family of generalized distributions. J. Stati. Comput. Simul.. 2011;81:883-898.
2. Dey S, Sharma, VK, Mesfioui M. A new extension of Weibull distribution with application to lifetime data. Ann. Data. Sci. 2017;4:31–61.
3. Keilson J, Sumita U. Uniform stochastic ordering and related inequalities. Canad. J. Stat. 1982;10:181-I98.
4. Muller A, Stoyan D. Comparison Methods for Stochastic Models and Risks. 2002;Wiley, New York.
5. Li X, Zuo MJ. Preservation of stochastic orders for random minima and maxima, with applications. Naval Res. Log. 2004;51:332-344.
6. Zhao P. Some new results on convolutions of heterogeneous gamma random variables. J. Multivar. Anal. 2011;102:958-976.
7. Mahdy M. Stochastic ordering and reliability analysis of inactivity lifetime with a cold standby. Am J. Math. Manag. Sci. 2018;DOI: 10.1080/01966324.2018.1502702.
8. Shaked M and Shanthikumar JG. Stochastic orders. 2007; Springer, Berlin, Germany.
9. Belzunce F. An introduction to the theory of stochastic orders. Estadistica. 2010;26,4-18.
10. Belzunce F, Martínez-Riquelme C., Mulero J. An Introduction to Stochastic Orders. Elsevier Ltd; 2016.
11. Giovagnoli A, Wynnb H.P. Stochastic orderings for discrete random variables. Stat. Probab. Lett. 2008;78:2827–2835.
12. Yu Y. Stochastic ordering of exponential family distributions and their mixtures. J. Appl. Probab. 2009;46:244-254.
13. Klenke A, Mattner L. Stochastic ordering of classical discrete distributions. Adv. Appl. Probab. 2010;42:392-410.
14. Pan X, Qiu G, Hu T. Stochastic orderings for elliptical random vectors. J. Multivar. Anal. 2016;148:83-88.
15. Raeisi M and Yari G. Some dependencies, stochastic orders and aging properties in an extended additive hazards model. Iran. J. Sci. Technol. Transaction A: Science. 2018;42:745–752.
16. Catana, LI, Raducan A. Stochastic order for a multivariate uniform distributions family. Mathematics. 2020; 8, DOI: 10.3390/math8091410.
17. Jones M. Kumaraswamy's distribution: A beta-type distribution with some tractability advantages. Stat. Methodol. 2009;6:70–81.
18. Eugene N, Lee C, Famoye F. Beta-normal distribution and its application. Commun. Stat. Theory. Methods. 2002;31:497–512.
19. Famoye F, Lee C, Olumolade O. The beta-Weibull distribution. J. Stat. Theory. Appl. 2005;4:121–136.
20. Nadarajah S, Kotz S. The beta-exponential distribution reliability, engineering and system safety. Reliab. Eng. Syst. Saf. 2006;91:689–697.
21. Akinsete A, Famoye F, Lee C. The Beta-Pareto distribution. Statistics. 2008;42:547-563.
22. Paranaiba PF, Ortega EMM, Cordeiro GM, Pescim RR. The beta Burr XII distribution with application to lifetime data. Comput. Stat. Data. Anal. 2011;55:1118-1136.
23. Cordeiro GM, Ortega, EMM, Nadarajah S. The Kumaraswamy Weibull distribution with application to failure data. J. Franklin Institute. 2010;81:1399-1429.
24. Pascoa AR, Ortega EMM, and Cordeiro GM. The Kumaraswamy generalized gamma distribution with application in survival analysis. Stat. Methodol 2011;8:411–433.
25. Cordeiro GM, Pescim RR, Ortega, EMM. The Kumaraswamy generalized half-normal distribution for skewed positive data. J. Data. Sci. 2012;10: 195-224.
26. Dey S, Nassar M, Kumar D. Alpha power transformed inverse Lindley distribution: A distribution with an upside-down bathtub-shaped hazard function. J. Comput. Appl. Math. 2019;348:130-45.
27. Hassan AS, Elgarhy M, Mohamd RE and Alrajhi S. On the Alpha Power transformed power Lindley distribution. J. Probab. Stat. 2019;Article ID 8024769, 13 pages.
28. McGilchrist CA and Aisbett CW. Regression with frailty in survival analysis. Biometrics. 1991;47(2):461-466.
29. Erbay A, Ergonul O, Stoddard JS and Samore MH. Recurrent catheter-related bloodstream infections: risk factors and outcome. Int. J. Infect. Dis. 2006;10:396-400.
30. Delistefani F, Wallbach M, Müller GA, Koziolek MK and Grupp C. Risk factors for catheter-related infections in patients receiving permanent dialysis catheter. BMC. Nephrol. 2019;20:199.

31. Hanagal DD and Dabade AD. Comparison of Shared Frailty Models for Kidney Infection Data under Exponential Power Baseline Distribution. Commun. Stat.Theory. Methods. 2015;44: 5091–5108.
32. Mahdavi A and Kundu D. A new method for generating distributions with an application to exponential distribution. Commun. Stat. Theory. Methods. 2017;46:6543-6557.
33. Lehmann EL and Casella, G. The Theory of Point Estimation. 1998; Springer, USA.
34. Mrsevi M. Convexity of the inverse function. Teach. Math. 2008;11:21-24.
35. Boyd S, Vandenberghe L. Convex Optimization. Cambridge University Press; 2009.
36. Nadarajah S, Kotz S. Strength modeling using Weibull distributions. J. Mech. Sci. Technol. 2008;22:1247-125.
37. Mudholkar G, Srivastava D, Freimer M. The exponentiated Weibull family: A reanalysis of the bus-motor-failure data. Technometrics. 1995;37:436–445.