AUT Journal of Mathematics and Computing

AUT J. Math. Comput., 7(1) (2026) 1-18
https://doi.org/10.22060/AJMC.2024.23495.1261

Original Article

A bivariate α-power transformed family: Theory and application

Vahid Nekoukhou*[a], Hamid Bidram[b], Ashkan Khalifeh[c]

[a]*Department of Statistics, Khansar Campus, University of Isfahan, Iran*
[b]*Department of Statistics, Faculty of Mathematics and Statistics, University of Isfahan, Isfahan, Iran*
[c]*Department of Statistics, Yazd University, Yazd, Iran*

ABSTRACT: In this paper, some general classes of bivariate semi-parametric continuous distributions are introduced. Some important properties of these family of distributions will be illustrated. It is seen that the bivariate distribution corresponds to the known Ali-Mikhail-Haq copula. Hence, some important properties such as the TP_2 property are justified. It will be shown that the marginals are kind of heavy tailed weighted distributions whose hazard rate functions can take variety of shapes. The behavior of the hazard rate function is mathematically illustrated. In addition, the α-power transformed distributions of a second type, which are introduced for the first time here, can be verified as special cases of the marginals. Some members of the new bivariate classes are studied in details. The estimation of the parameters is illustrated by means of an efficient expectation-maximization algorithm, and some real data sets are also analyzed for illustrative purposes. Finally, we conclude the paper.

Review History:

Received:31 August 2024
Revised:29 September 2024
Accepted:17 November 2024
Available Online:01 January 2026

Keywords:

Bivariate distributions
Copula
Expectation-maximization algorithm
Geometric distribution
Hazard rate function
Maximum likelihood estimation

MSC (2020):

62F10; 62E10; 62H10

1. Introduction

Bivariate distributions are mainly used to analyze the marginals and also to model the dependence structure between the two marginals. Extensive works has been done to propose different bivariate continuous distributions and to develop their properties. Similar to the continuous univariate data, the continuous bivariate data are also analyzed quite often in practice mainly due to analytical tractability. See for example the book by Balakrishnan and Lai [4] and the references cited therein for different bivariate probability distributions and for their various properties and applications. Recently, Nekoukhou et al. [18] investigated some extensions of the generalized exponential distributions, both in the univariate and bivariate cases. In addition, Nekoukhou et al. [19] illustrated the bivariate Rayleigh distribution and considered some interesting properties of the model.

*Corresponding author.
E-mail addresses: v.nekoukhou@khc.ui.ac.ir (V. Nekoukhou), h.bidram@sci.ui.ac.ir (H. Bidram), khalifeh68@yahoo.com (A. Khalifeh)

Recently, Mahdavi and Kundu [15] introduced a univariate class of continuous distributions. Suppose $F(.)$ is the cumulative distribution function (CDF) of an absolutely continuous random variable X. The α-power transformation (APT) of $F(.)$, for $x \in \mathbb{R}$, introduced by Mahdavi and Kundu [15] has the following CDF:

$$F_{APT}(x) = \begin{cases} \frac{\alpha^{F(x)}-1}{\alpha-1} & \text{if } \alpha > 0, \alpha \neq 1 \\ F(x) & \text{if } \alpha = 1. \end{cases} \tag{1}$$

The recent authors introduced the CDF (1) directly and investigated its different properties.

In the present study, an APT distribution of a second type, denoted by APT_2 is firstly introduced. Then, we study some generalizations of APT_2 distributions both in the univariate and bivariate cases. More precisely, the geometric maximum of the APT_2 distributions are studied. Therefore, the APT_2 distribution can be verified as a special case. These new generalizations are kind of semi-parametric family of distributions and it will be shown that they can supply desirable applications in practice. We will see that the marginals are kind of heavy tailed weighted distributions with a general structure. Some distributional properties of the marginals are studied in detail. The APT_2 distribution can be used as an alternative to the well-known Weibull, gamma or generalized exponential (GE) distributions. The CDF of the APT_2 distribution can be expressed in analytical form and therefore can also be used quite conveniently for analyzing censored or truncated data. We will see that the hazard rate function of the marginals can take different shapes such as increasing, decreasing, bathtub-shaped and upside-down bathtub. In addition, it is observed that the generation of random samples from the proposed bivariate models is very simple. Hence, simulation experiments can be performed quite conveniently.

The rest of the paper is organized as follows. First, in Section 2, the APT_2 distribution is introduced. A class of bivariate continuous distributions, which is indeed the class of geometric maximum of APT_2 distributions, is investigated in Section 3 and some of its important futures are studied. In this section, the marginals and some of their important relative properties are studied. We will see that the APT_2 distributions can be verified as special cases of the marginals. In addition, some basic properties of the bivariate class of distributions is investigated. Some examples are also given. In Section 4, the statistical inference is done and some real data sets are also analyzed. Finally, we conclude the paper in Section 5.

2. The APT distribution of a second type

The lifetime of a parallel (series) system with N components, in reliability studies, is defined by $X = \max_{1\leq i\leq N} X_i$ ($\min_{1\leq i\leq N} X_i$), in which X_i denotes the life length of the i-th ($i = 1, 2, \ldots, N$) component. In practice, the number of components may be itself a discrete random variable. In recent years, many researchers have studied such models to illustrate the characteristics and properties of the lifetime of series and parallel systems, where the number of components follows a certain probability distribution function. In recent two decades, some known continuous lifetime distributions such as the exponential, GE, gamma and Weibull have been usually compounded with a classic discrete distribution such as the geometric and zero-truncated Poisson, which are some special cases of the power-series family of distributions . For example, Adamidis and Loukas [2] and Kus [10] introduced the exponential-geometric and exponential-Poisson distributions, respectively, with decreasing failure rates. We can also address the Weibull-Poisson distribution proposed by Hemmati et al. [8], the Weibull-geometric distribution of Barreto-Souza et al. [5], and the extended exponential-geometric distribution of Adamidis et al. [1]. GE-power series class of distributions given by Mahmoudi and Jafari [16] and the compound class of extended Weibull power series distributions proposed by Silva et al. [22] are other examples in this connection.

Let $X_1, X_2, \ldots, X_N$ be independent and identically distributed (iid) random variables from an absolutely continuous CDF, say F. In addition, suppose that N is a zero-truncated Poisson random variable, with the following probability mass function (PMF)

$$P(N = n) = p_n = \frac{e^{-\lambda}\lambda^n}{n!(1-e^{-\lambda})}, \quad n \in \mathbb{N} = \{1, 2, \ldots\}; \lambda > 0. \tag{2}$$

Moreover, suppose that N is independent of X_i's. If $X = \max\{X_1, X_2, \ldots, X_N\}$, then we have the following result.

Theorem 2.1. *Under the above conditions, the marginal CDF of X, is given by*

$$G(x; \alpha, \overline{F}) = \frac{\alpha - \alpha^{\overline{F}(x)}}{\alpha - 1}, \quad x \in \mathbb{R}, \tag{3}$$

where $0 < \alpha = e^{-\lambda} < 1$, and $\overline{F}(x) = P(X \geq x)$ denotes the survival function (SF) of the random variable X.

Proof. It is enough to note that

$$G(x;\alpha,\overline{F}) = P(X \le x) = \sum_{n=1}^{\infty} P(X \le x|N=n)p_n = \sum_{n=1}^{\infty} \frac{F^n(x)e^{-\lambda}\lambda^n}{n!(1-e^{-\lambda})} = \frac{e^{-\lambda}}{1-e^{-\lambda}}\left\{\sum_{n=0}^{\infty}\frac{(\lambda F(x))^n}{n!} - 1\right\}$$
$$= \frac{\alpha - \alpha^{1-F(x)}}{\alpha - 1} = \frac{\alpha - \alpha^{\overline{F}(x)}}{\alpha - 1}. \qquad \square$$

Although, in Theorem 2.1, G has been made based on $0 < \alpha < 1$, one can easily show that it can be a bona fide CDF for all values of $\alpha \in \mathbb{R}^+ - \{1\}$. So, such rigorous constraint on α can be relaxed and G defined for all values of $\alpha \in \mathbb{R}^+ - \{1\}$ in the rest of the paper. In addition, as $\alpha \to 1$, $G(x;\alpha,\overline{F})$ tends to $F(x)$. So, we have the following definition.

Definition 2.2. *A random variable X is said to be APT distributed of a second type (APT_2) if its CDF has the following structure:*

$$G(x;\alpha,\overline{F}) = \begin{cases} \frac{\alpha-\alpha^{\overline{F}(x)}}{\alpha-1} & \text{if } \ \alpha>0,\ \alpha\neq 1 \\ 1-\overline{F}(x) & \text{if } \ \alpha = 1. \end{cases} \tag{4}$$

The APT_2 distribution will be denoted by $APT_2(\alpha,\overline{F})$ in the rest of the paper.

Remark 2.3. *Under the conditions of Theorem 2.1, if $Y = \min\{X_1, X_2, \ldots, X_N\}$, then it is easy to show that the marginal CDF of Y is given by*

$$F_Y(y) = \frac{\alpha^{F(y)}-1}{\alpha-1}, \quad y \in \mathbb{R}, \tag{5}$$

which is known as the APT distribution in the literature, introduced by Mahdavi and Kundu (2017) [15].

The PDF of the random variable $X \sim \mathrm{APT}_2(\alpha,\overline{F})$ becomes

$$g(x;\alpha,\overline{F}) = \frac{d}{dx}G(x;\alpha,\overline{F}) = \frac{f(x)\alpha^{\overline{F}(x)}\log\alpha}{\alpha-1}, \quad x \in \mathbb{R}, \tag{6}$$

where $f(.)$ is the corresponding PDF of $\overline{F}(.)$.

It is interesting to note that PDF (5) is a kind of weighted distribution. In fact,

$$g(x;\alpha,\overline{F}) = \frac{\alpha^{\overline{F}(x)}f(x)}{E(\alpha^{\overline{F}(X)})}, \quad x \in \mathbb{R}, \tag{7}$$

where $E(\alpha^{\overline{F}(X)}) = \frac{\alpha-1}{\log\alpha}$.

3. A bivariate class of semi-parametric continuous distributions

3.1. Definition and Interpretations

Suppose that $X_1, X_2, \ldots$ and $Y_1, Y_2, \ldots$ are two sequences of random variables. It is assumed that X_i's and Y_i's are iid $\mathrm{APT}_2(\alpha,\overline{F}_1)$ and $\mathrm{APT}_2(\alpha,\overline{F}_2)$ random variables, respectively. In addition, X_i's and Y_j's are independent. Let N be a geometric random variable with PMF

$$p_n = p(1-p)^{n-1}, \quad n \in \mathbb{N};\ 0 < p < 1. \tag{8}$$

The above geometric distribution will be denoted by $\mathrm{GM}(p)$ in the rest of the paper. Moreover, N is independent of X_i's and Y_j's. Consider the bivariate random variable (X,Y), where

$$X = \max\{X_1, X_2, \ldots, X_N\} \quad \text{and} \quad Y = \max\{Y_1, Y_2, \ldots, Y_N\}.$$

(X,Y) defines the bivariate geometric maximum of APT_2 (BG in short) distributions. The following interpretations can be consider for the BG class of distributions.

Parallel Systems (cf. Kundu [11]): Consider two systems, say 1 and 2, each having N number of independent and identical components attached in parallel. Here N is itself a random variable. If $X_1, X_2, \ldots$ denote the lifetime of

the components of System 1, and in a same manner, $Y_1, Y_2, \ldots$ denote the lifetime of the components of System 2. Then, the lifetime of the two systems becomes (X, Y).
Random Stress Model (cf. Kundu [11]): Suppose a system has two components. Each component is subject to random number of individual independent stresses, say $\{X_1, X_2, \ldots\}$ and $\{Y_1, Y_2, \ldots\}$, respectively. If N is the number of stresses, then the observed stresses at the two components are $X = \max\{X_1, \ldots, X_N\}$ and $Y = \max\{Y_1, \ldots, Y_N\}$, respectively.

Theorem 3.1. *The semi-parametric joint CDF of the BG distributions, for $\alpha \in \mathbb{R}^+ - \{1\}$, is given by*

$$Q_{X,Y}(x, y; \mathbf{\Omega}) = \frac{p(\alpha - \alpha^{\overline{F}_1(x)})(\alpha - \alpha^{\overline{F}_2(y)})}{(\alpha-1)^2 - (1-p)(\alpha - \alpha^{\overline{F}_1(x)})(\alpha - \alpha^{\overline{F}_2(y)})}, \quad x \in \mathbb{R}, y \in \mathbb{R}, \tag{9}$$

where $\mathbf{\Omega}$ consists of α, p and some parameters of the SFs $\overline{F}_1$ and $\overline{F}_2$. For $\alpha = 1$, the CDF of the class of BG distributions becomes

$$Q_{X,Y}(x, y; \mathbf{\Omega}^*) = \frac{p\{1 - \overline{F}_1(x)\}\{1 - \overline{F}_2(y)\}}{1 - (1-p)\{1 - \overline{F}_1(x)\}\{1 - \overline{F}_2(y)\}},$$

where $\mathbf{\Omega}^*$ is the same as that of $\mathbf{\Omega}$ with $\alpha = 1$.

Proof. The CDF of (X, Y), for $\alpha \in \mathbb{R}^+ - \{1\}$, is given by

$$\begin{aligned} Q_{X,Y}(x, y; \mathbf{\Omega}) &= \sum_{n=1}^{\infty} P(X \le x, Y \le y | N = n) P(N = n) \\ &= \frac{p}{1-p} \sum_{n=1}^{\infty} \left\{ \frac{(1-p)(\alpha - \alpha^{\overline{F}_1(x)})(\alpha - \alpha^{\overline{F}_2(y)})}{(\alpha-1)^2} \right\}^n. \end{aligned}$$

After some calculations the result is immediately obtained. A similar argument can be performed for case of $\alpha = 1$. $\square$

It is interesting to note that if $p = 1$, then we have

$$Q_{X,Y}(x, y; \mathbf{\Omega}) = G(x; \alpha, \overline{F}_1) G(y; \alpha, \overline{F}_2),$$

i.e., X and Y become independent. Therefore, the parameter p plays the role of the correlation parameter.
The joint PDF of (X, Y) can be found as $q(x, y; \mathbf{\Omega}) = \frac{\partial^2}{\partial x \partial y} Q_{X,Y}(x, y; \mathbf{\Omega})$, and it is

$$q(x, y; \mathbf{\Omega}) = p \frac{(\alpha-1)^2 f_1(x) f_2(y) \alpha^{\overline{F}_1(x) + \overline{F}_2(y)} (\log \alpha)^2 \left\{ (\alpha-1)^2 + (1-p)(\alpha - \alpha^{\overline{F}_1(x)})(\alpha - \alpha^{\overline{F}_2(y)}) \right\}}{\left\{ (\alpha-1)^2 - (1-p)(\alpha - \alpha^{\overline{F}_1(x)})(\alpha - \alpha^{\overline{F}_2(y)}) \right\}^3}. \tag{10}$$

The BG distribution with PDF (10) will be denoted by $\mathrm{BG}(\alpha, p, \overline{F}_1, \overline{F}_2)$ in the sequel.

3.2. Marginal distributions

Here, the marginal distributions of X and Y are discussed. We only supply the results for the random variable X. The distribution of Y, can be achieved in a same manner. First, note that the joint CDF of (X, N) is given by

$$\begin{aligned} Q_{X,N}(x, n) &= \sum_{j=1}^{n} P(X \le x | N = j) P(N = j) \\ &= \frac{p}{1-p} \sum_{j=1}^{n} \left\{ \frac{1-p}{\alpha-1} (\alpha - \alpha^{\overline{F}_1(x)}) \right\}^j = \frac{p(\alpha - \alpha^{\overline{F}_1(x)}) \left\{ 1 - \left[\frac{1-p}{\alpha-1} (\alpha - \alpha^{\overline{F}_1(x)}) \right]^n \right\}}{\alpha - 1 - (1-p)(\alpha - \alpha^{\overline{F}_1(x)})}. \end{aligned}$$

Hence, the marginal CDF of X, for $0 < p < 1$ and $\alpha \in \mathbb{R}^+ - \{1\}$ can be obtained as follows:

$$Q_X(x; \mathbf{\Omega}_1) = \lim_{n \to \infty} Q_{X,N}(x, n) = \frac{p(\alpha - \alpha^{\overline{F}_1(x)})}{\alpha - 1 - (1-p)(\alpha - \alpha^{\overline{F}_1(x)})}, \quad x \in \mathbb{R}, \tag{11}$$

where $\boldsymbol{\Omega}_1$ consists of α, p and the parameters of the SF $\overline{F}_1$.

The class of univariate geometric maximum of APT_2 (UG in short) distributions, with CDF (11), will be denoted by $\mathrm{UG}(\alpha, p, \overline{F}_1)$.

It must be mentioned that the marginal CDF of X can also be achieved directly from (9), as $y \to \infty$. In addition, for a fixed value of x, as $p \to 0$, we conclude that $P(X > x) \to 1$. Therefore, it can be considered as a heavy-tailed distribution.

If $p = 1$, the distribution of X coincides with an $\mathrm{APT}_2(\alpha, \overline{F}_1)$ distribution, whose CDF is given by (3). This means that for values of p close to 1, the behaviour of the PDF of the random variable X is quite similar to that of $\mathrm{APT}_2(\alpha, \overline{F}_1)$ distribution. Therefore, some basic distributional properties of the $\mathrm{APT}_2(\alpha, \overline{F}_1)$ distribution can be studied with the choice of $p = 1$ in a $\mathrm{UG}(\alpha, p, \overline{F}_1)$ distribution.

The following result indicates that the class of UG distributions is closed under geometric maximum.

Theorem 3.2. *If $\{U_i : i \geq 1\}$ is a sequence of iid $UG(\alpha, p_1, \overline{F}_1)$ random variables, $M \sim GM(p_2)$, and moreover U_i's and M are independent, then $U = \max\{U_1, \ldots, U_M\}$ has a $UG(\alpha, p_1 p_2, \overline{F}_1)$ distribution.*

Proof. It is seen that

$$P(U \leq u) = \sum_{m=1}^{\infty} P(U_1 \leq u, \ldots, U_M \leq u | M = m) P(M = m)$$
$$= \frac{p_1 p_2 (\alpha - \alpha^{\overline{F}_1(x)})}{\alpha - 1 - (1 - p_1 p_2)(\alpha - \alpha^{\overline{F}_1(x)})},$$

which adapts the CDF of the $\mathrm{UG}(\alpha, p_1 p_2, \overline{F}_1)$ distribution. □

Theorem 3.3. *If ξ_γ is the γ-th quantile of a UG family of distributions, then it satisfies*

$$\overline{F}_1(\xi_\gamma) = \log\left(\frac{p\alpha(1-\gamma) + \gamma}{\gamma(1-p) + p}\right) / \log \alpha.$$

Proof. The result is straightforward and hence the details are avoided. □

The PDF of the random variable $X \sim \mathrm{UG}(\alpha, p, \overline{F}_1)$, say $q(x; \boldsymbol{\Omega}_1)$, can be obtained as

$$q(x; \boldsymbol{\Omega}_1) = \frac{d}{dx} Q(x; \boldsymbol{\Omega}_1),$$

where $Q(x; \boldsymbol{\Omega}_1)$ is given by (11). More precisely, we see that

$$q(x; \boldsymbol{\Omega}_1) = w(x; \boldsymbol{\Omega}_1) g(x; \alpha, \overline{F}_1), \tag{12}$$

where $g(x; \alpha, \overline{F}_1)$ is given by (6), and

$$w(x; \boldsymbol{\Omega}_1) = \frac{p(\alpha - 1)^2}{\left\{\alpha - 1 - (1-p)(\alpha - \alpha^{\overline{F}_1(x)})\right\}^2}.$$

The weight function $w(x; \boldsymbol{\Omega}_1)$ increases from $(\alpha - 1)p$ to $\frac{1}{p}$, as x varies from $-\infty$ to ∞.

From now on, without loss of generality, we assume that $\overline{F}_1$ is the SF of a lifetime distribution. So, the hazard rate function of a $\mathrm{UG}(\alpha, p, \overline{F}_1)$ distribution, denoted by $h_{UG}(.)$, is

$$h_{UG}(x) = w^*(x) h_{APT_2}(x), \quad x \geq 0, \tag{13}$$

where

$$w^*(x) = \frac{p(\alpha - 1)}{\alpha - 1 - (1-p)(\alpha - \alpha^{\overline{F}_1(x)})} \tag{14}$$

and $h_{APT_2}(x)$ is the hazard rate function of the random variable $X \sim \mathrm{APT}_2(\alpha, \overline{F}_1)$ and is given by

$$h_{APT_2}(x) = \frac{f_1(x)\alpha^{\overline{F}_1(x)} \log \alpha}{\alpha^{\overline{F}_1(x)} - 1}, \quad x \geq 0. \tag{15}$$

In the next theorem, it will be shown that for $0 < \alpha < 1$, a $\mathrm{UG}(\alpha, p, \overline{F}_1)$ distribution has an increasing hazard rate (IHR) function, provided that F_1 is IHR.

Theorem 3.4. *For $0 < \alpha < 1$, a $UG(\alpha, p, \overline{F}_1)$ distribution is IHR if F_1 is IHR.*

Proof. For $0 < \alpha < 1$, one can easily show that $w^*(x)$, given by (14), is a positive and increasing function of x. So, $\frac{d}{dx}w^*(x) > 0$. In addition, $h_{APT_2}(x)$ can also be written as

$$h_{APT_2}(x) = h(x)\frac{\overline{F}_1(x)\alpha^{\overline{F}_1(x)}}{\alpha^{\overline{F}_1(x)} - 1}(\log\alpha),$$

where $h(x) = \frac{f_1(x)}{\overline{F}_1(x)}$ is the hazard rate function of f_1. Hence, we see that

$$\frac{d}{dx}h_{APT_2}(x) = (\frac{d}{dx}h(x))\frac{\overline{F}_1(x)\alpha^{\overline{F}_1(x)}}{\alpha^{\overline{F}_1(x)} - 1}(\log\alpha) + h(x)(\log\alpha)\frac{d}{dx}\frac{\overline{F}_1(x)\alpha^{\overline{F}_1(x)}}{\alpha^{\overline{F}_1(x)} - 1}.$$

F_1 is IHR, means that $\frac{d}{dx}h(x) > 0$. Moreover, for $0 < \alpha < 1$, $\alpha^{\overline{F}_1(x)} - 1 < 0$ and $\log\alpha < 0$. In addition, in view of the fact that for any $a > 0$, $a - 1 > \log a$ (logarithm is to the base e), one can conclude that $1 - \alpha^{\overline{F}_1(x)} + \log\alpha^{\overline{F}_1(x)} < 0$, which yields that $\frac{d}{dx}\frac{\overline{F}_1(x)\alpha^{\overline{F}_1(x)}}{\alpha^{\overline{F}_1(x)}-1} < 0$. Therefore, we see that $\frac{d}{dx}h_{APT_2}(x) > 0$, and

$$\frac{d}{dx}h_{UG}(x) = (\frac{d}{dx}w^*(x))h_{APT_2}(x) + w^*(x)(\frac{d}{dx}h_{APT_2}(x)) > 0.$$

That is the UG distribution is IHR. □

Although the above theorem yields that, under some conditions, the UG distribution is IHR, it can also have decreasing, bathtub-shaped and upside-down bathtub hazard rates, which will be illustrated by means of some examples.

Theorem 3.5. *A $UG(\alpha, 1, \overline{F}_1)$ distribution is log-convex, if $\alpha > 1$ and $\overline{F}_1$ has a log-convex PDF.*

Proof. If f is log-convex, then its corresponding SF, $\overline{F}$, is also log-convex (see, [17, pp. 101]). On the other hand, $\overline{F}$ is convex (see, [17, pp. 690]) and hence, for $\alpha > 1$,

$$\frac{d^2}{dx^2}\log g(x;\alpha,\overline{F}) = \frac{d^2}{dx^2}\log f_1(x) + (\log\alpha)\frac{d^2}{dx^2}\overline{F}_1(x) > 0,$$

which completes the proof. □

The joint density function of (X, N), say $q(x, n)$, is given by

$$q(x,n) = np(\alpha-1)^{-n}(\log\alpha)(1-p)^{n-1}f_1(x)\alpha^{\overline{F}_1(x)}\left\{\alpha - \alpha^{\overline{F}_1(x)}\right\}^{n-1}.$$

Therefore, after some calculations, it follows that

$$P(N = n|X = x) = \frac{np(1-p)^{n-1}(\alpha - \alpha^{\overline{F}_1(x)})^{n-1}}{w(x;\mathbf{\Omega}_1)(\alpha-1)^{n-1}} = np^{*^2}(1-p^*)^{n-1},$$

where $p^* = 1 - \frac{(1-p)(\alpha-\alpha^{\overline{F}_1(x)})}{\alpha-1}$ and $n \in \mathbb{N}$. It is clear that $\{N = n|X = x\} \overset{d}{=} N^* + 1$, where N^* is a negative binomial random variable whose PMF is

$$P(N^* = n) = (n+1)p^{*^2}(1-p^*)^n, \quad n \in \mathbb{N}_0 = \{0, 1, 2, \ldots\}.$$

Therefore, one can immediately conclude that

$$E(N = n|X = x) = E(N^* + 1) = \frac{2(1-p^*)}{p^*} + 1 = \frac{\alpha - 1 + (1-p)(\alpha - \alpha^{\overline{F}_1(x)})}{\alpha - 1 - (1-p)(\alpha - \alpha^{\overline{F}_1(x)})}. \quad (16)$$

The above conditional expectation will be useful in developing the expectation-maximization (EM) algorithm, in order to obtain the estimates of the UG parameters.

Here, we make the following note in regards to the famous structural stress-strength parameter. The stress-strength parameter $R = P(X > Y)$ and its estimation is an attractive component reliability measure, which is widely used in the statistical literature, reliability context, medical, economic, and other related fields. Suppose that the random variable X is the strength of a component which is subjected to a random stress Y. The component fails whenever $X < Y$ and there is no failure when $X > Y$. In the next theorem, we obtain this quantity for a UG$(\alpha, 1, \overline{F})$ distribution.

Theorem 3.6. *Let $X \sim UG(\alpha_1, 1, \overline{F})$ and $Y \sim UG(\alpha_2, 1, \overline{F})$ be two independent lifetime random variables. The stress-strength parameter $R = P(X > Y)$ is given by*

$$R = \begin{cases} \dfrac{\alpha_2(\alpha_1 - 1)\log\alpha_2 - (\alpha_2 - 1)\log\alpha_1}{(\alpha_1 - 1)(\alpha_2 - 1)(\log\alpha_1 + \log\alpha_2)}, & \alpha_1, \alpha_2 \in \mathbb{R}^+ - \{1\} \\ \int_0^\infty f_X(x)F_Y(x)dx, & \alpha_1 = \alpha_2 = 1. \end{cases}$$

Proof. It is sufficient to calculate

$$R = P(X > Y) = \int_0^\infty g(x; \alpha_1, \overline{F})G(x; \alpha_2, \overline{F})dx, \tag{17}$$

where $g(x; \alpha_1, \overline{F})$ and $G(x; \alpha_2, \overline{F})$ are the probability density and cumulative distribution functions of the $\mathrm{UG}(\alpha_1, 1, \overline{F})$ and $\mathrm{UG}(\alpha_2, 1, \overline{F})$ distributions, respectively. Using $\overline{F}(x) = t$, the last integral is simplified and the result immediately obtained. □

Remark 3.7. *In the above theorem, if $\alpha_1 = \alpha_2$, then $R = \frac{1}{2}$.*

3.3. BG distributions

Here, some properties of the BG class of distributions are investigated. First of all, note that if $(X, Y) \sim \mathrm{BG}(\alpha, p, \overline{F}_1, \overline{F}_2)$, then $X \sim \mathrm{UG}(\alpha, p, \overline{F}_1)$ and $Y \sim \mathrm{UG}(\alpha, p, \overline{F}_2)$.

In the following, we want to illustrate that the class of BG distributions, similar to the marginals, is closed under geometric maximum. More precisely, we have the following result whose proof is avoided.

Theorem 3.8. *Let $\{(U_i, V_i) : i \geq 1)\}$ be a sequence of iid $BG(\alpha, p_1, \overline{F}_1, \overline{F}_2)$ random variables, and $M \sim GM(p_2)$, $0 < p_2 < 1$. In addition, M is independent of (U_i, V_i)'s. If we consider the random variables*

$$U = \max\{U_1, U_2, \ldots, U_M\} \quad \text{and} \quad \mathrm{V} = \max\{\mathrm{V}_1, \mathrm{V}_2, \ldots, \mathrm{V}_\mathrm{M}\},$$

then $(U, V) \sim BG(\alpha, p_1p_2, \overline{F}_1, \overline{F}_2)$.

Now, suppose thet $(X, Y) \sim \mathrm{BG}(\alpha, p, \overline{F}_1, \overline{F}_2)$. The joint PDF of (X, Y, N), say $q(x, y, n)$, is given by

$$q(x, y, n) = \frac{n^2 p\left\{(1-p)(\alpha - \alpha^{\overline{F}_1(x)})(\alpha - \alpha^{\overline{F}_2(y)})\right\}^{n-1} f_1(x)f_2(y)\alpha^{\overline{F}_1(x)+\overline{F}_2(y)}(\log\alpha)^2}{(\alpha-1)^{2n}}. \tag{18}$$

Therefore, we have

$$P(N = n|X = x, Y = y) = \frac{q(x, y, n)}{q(x, y)} = \frac{n^2 a^{n-1}(x, y)\left\{(\alpha-1)^2 - a(x, y)\right\}^3}{(\alpha-1)^{2(n+1)}\left\{(\alpha-1)^2 + a(x, y)\right\}},$$

where

$$a(x, y) = (1-p)(\alpha - \alpha^{\overline{F}_1(x)})(\alpha - \alpha^{\overline{F}_2(y)}). \tag{19}$$

Using the fact that if $M \sim \mathrm{GM}(1-p)$, then $E(M^3) = (p^2 + 4p + 1)/(1-p)^3$, we obtain

$$E(N = n|X = x, Y = y) = \frac{b^2(x, y) + 4b(x, y) + 1}{1 - b^2(x, y)}, \tag{20}$$

where $b(x, y) = a(x, y)/(\alpha - 1)^2$.

Nelsen (2006) [20] indicated that every bivariate CDF, say $Q_{X,Y}$, with continuous marginals, say Q_X and Q_Y, corresponds a unique copula function $C : [0, 1]^2 \to [0, 1]$ such that, for $(x, y) \in \mathbb{R}^2$,

$$Q_{X,Y}(x, y) = C(Q_X(x), Q_Y(y)).$$

Equivalently, the copula $C(u, v)$ can be obtained from the joint CDF, as

$$C(u, v) = Q_{X,Y}(Q_X^{-1}(u), Q_Y^{-1}(v)).$$

Theorem 3.9. *If $(X,Y) \sim BG(\alpha, p, \overline{F}_1, \overline{F}_2)$, then its corresponding copula is*

$$C(u,v) = \frac{uv}{1-(1-p)(1-u)(1-v)}. \tag{21}$$

Proof. Let us consider $Q_X(x) = u$ and $Q_Y(y) = v$. Then, we see that $\alpha - \alpha^{\overline{F}_1(x)} = \frac{u(\alpha-1)}{p+(1-p)u}$ and $\alpha - \alpha^{\overline{F}_2(y)} = \frac{v(\alpha-1)}{p+(1-p)v}$. Hence, after some calculations the result is immediately obtained, and the details are avoided. □

The copula of the class of BG distributions, given by (21), is known as Ali-Mikhail-Haq (AMH) copula, with parameter $(1-p)$, see Ali et al. (1978). The recent authors investigated some interpretations of the above copula in terms of bivariate odds ratio.

Suppose the random variables X and Y have an absolutely continuous joint CDF. Let us recall the following definitions from Nelsen (2006).

Definition 3.10. *If for all x, $P(X > x|Y = y)$ is a non-decreasing function of y, then X is stochastically increasing in Y.*

Definition 3.11. *If for all $(x, y) \in \mathbb{R}^2$, $P(X \le x, Y \le y) \ge P(X \le x)P(Y \le y)$, then X and Y are positively quadrant dependent (PQD).*

Definition 3.12. *If for all x, $P(X \le x|Y \le y)$ is a non-increasing function of y, then X is left tail decreasing (LTD) in Y.*

Definition 3.13. *A function $f : \mathbb{R}^2 \to \mathbb{R}$ is said to be total positivity of order two (TP_2), if for all $(x, y) \in \mathbb{R}^2$, $f(x, y) \ge 0$ and whenever $x_1 \le x_2$, and $y_1 \le y_2$,*

$$f(x_1, y_1)f(x_2, y_2) \ge f(x_1, y_2)f(x_2, y_1).$$

In addition, (X, Y) is said to be TP_2, if its joint CDF is TP_2.

Therefore, the following properties can be established for a BG class of distributions.

Theorem 3.14. *If $(X,Y) \sim BG(\alpha, p, \overline{F}_1, \overline{F}_2)$, then X is stochastically increasing in Y and vice versa.*

Proof. Y is stochastically increasing in X, iff for any $v \in [0, 1]$, $C(u, v)$ is a concave function of u; see Nelsen (2006). In case of Ali-Mikhail-Haq copula, in view of the fact that $\frac{\partial^2 C(u,v)}{\partial u^2} \le 0$, the result is obtained. □

Theorem 3.15. *If $(X,Y) \sim BG(\alpha, p, \overline{F}_1, \overline{F}_2)$, then X and Y are PQD.*

Proof. First note that PQD property is a copula property. In addition, X and Y are PQD iff their copula satisfies $C(u, v) \ge uv$, for all values of $(u, v) \in [0, 1] \times [0, 1]$, which one can easily show that the corresponding copula of the BG distribution justifies. Hence, the result is immediately obtained. □

Theorem 3.16. *If $(X,Y) \sim BG(\alpha, p, \overline{F}_1, \overline{F}_2)$, then X is LTD in Y.*

Proof. First note that LTD property is a copula property. In addition, X is LTD in Y iff for any $u \in [0, 1]$, $C(u, v)/v$, is non-increasing with respect to v, which one can easily show that the corresponding copula of the BG distribution justifies. Hence, the result is immediately obtained. □

Theorem 3.17. *If $(X,Y) \sim BG(\alpha, p, \overline{F}_1, \overline{F}_2)$, then (X, Y) has TP_2 property.*

Proof. First note that TP_2 property is a copula property. In addition, the joint CDF of (X, Y) is TP_2 iff Ali-Mikhail-Haq copula is TP_2. Now, for $u_1 < u_2$ and $v_1 < v_2$, it is seen that

$$(1-p)(u_2 - u_1)(v_2 - v_1) \ge 0 \quad \text{iff} \quad C(u_1, v_1)C(u_2, v_2) \ge C(u_1, v_2)C(u_2, v_1).$$

Hence, the result follows. □

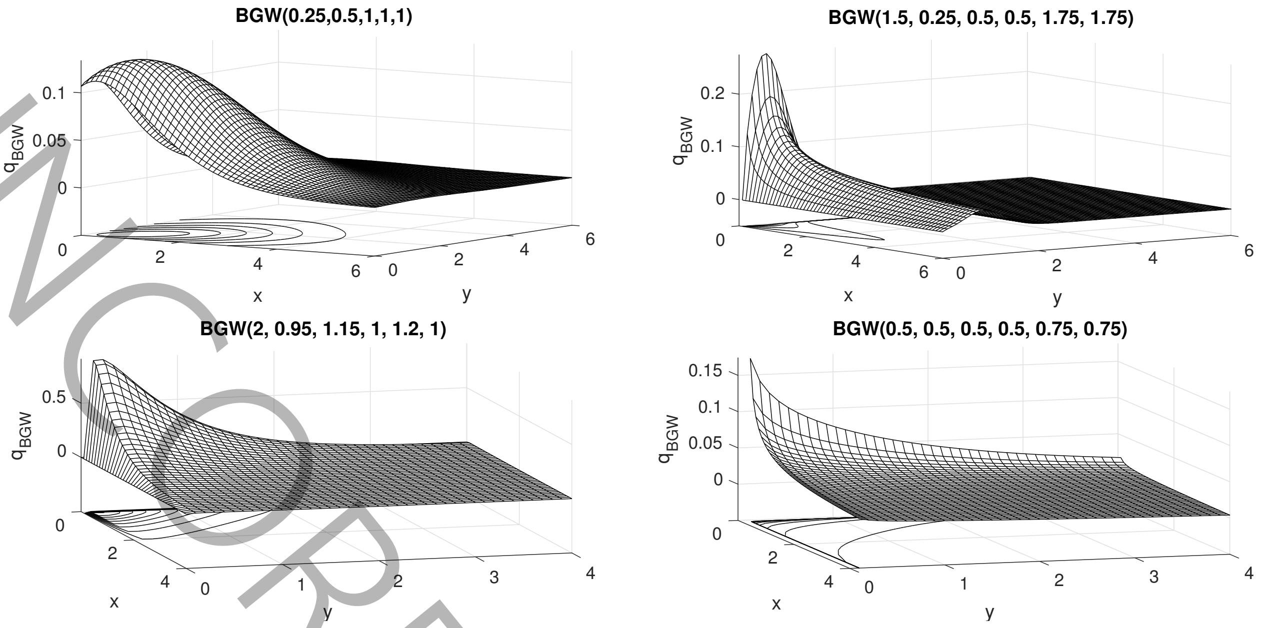

Figure 1: PDF plots of BGW$(\alpha, p, \alpha_1, \beta_1, \alpha_2, \beta_2)$ distributions.

3.4. Exapmle: Bivariate G-Weibull distribution

Let $\overline{F}_1(x) = e^{-\beta_1 x^{\alpha_1}}$ and $\overline{F}_2(y) = e^{-\beta_2 y^{\alpha_2}}$ correspond to the Weibull SFs for $x \geq 0$ and $y \geq 0$, respectively. By inserting these SFs into Eq. (9), a new bivariate G-Weibull distribution, which will be denoted by BGW$(\alpha, p, \alpha_1, \beta_1, \alpha_2, \beta_2)$, is obtained. The BGW$(\alpha, p, \alpha_1, \beta_1, \alpha_2, \beta_2)$ distribution can be reduced to bivariate-G exponential and bivariate-G Rayleigh distributions, when $\alpha_1 = \alpha_2 = 1$ and $\alpha_1 = \alpha_2 = 2$, respectively.

Figure 1 illustrates the PDF plots of BGW$(\alpha, p, \alpha_1, \beta_1, \alpha_2, \beta_2)$ distributions for some possible values of their parameters. Many researchers such as Basu [6], Puri and Rubin [21] and Kundu and Gupta [13] considered the bivariate hazard rate function as a scalar quantity. The scalar hazard rate function, for a bivariate random variable (X, Y) with the joint PDF $f(x, y)$ and the joint CDF $\overline{F}(x, y)$, is defined as

$$h(x,y) = \frac{f(x,y)}{\overline{F}(x,y)} = \frac{f(x,y)}{1 - F(x) - F(y) + F(x,y)}, \tag{22}$$

where $F(x)$ and $F(y)$ are the marginal CDFs. Since the above definition of bivariate hazard rate function does not uniquely determine the joint PDF, the joint bivariate hazard rate function, defined by Johnson and Kotz [9], is usually considered as

$$((h_1(x,y), h_2(x,y)) = \left(-\frac{\partial}{\partial x}\log \overline{F}(x,y), -\frac{\partial}{\partial y}\log \overline{F}(x,y)\right). \tag{23}$$

The scalar hazard rate function of the BGW$(\alpha, p, \alpha_1, \beta_1, \alpha_2, \beta_2)$ distribution, using (22), is

$$h_{BGW}(x,y) = \frac{q_{BGW}(x,y)}{\overline{Q}_{BGW}(x,y)} = \frac{q_{BGW}(x,y)}{1 - Q_X(x) - Q_Y(y) + Q_{BGW}(x,y)},$$

where, $Q_{BGW}(x, y)$ and $q_{BGW}(x, y)$ are, respectively, the CDF and PDF of a BGW distribution. In addition, $Q_X(x)$ and $Q_Y(y)$ are the corresponding marginals which can be obtained by (11). The scalar hazard rate functions of BGW distributions have been illustrated by Figure 2.

Based on

$$((h_1(x,y), h_2(x,y)) = \left(-\frac{\partial}{\partial x}\log \overline{Q}_{BWG}(x,y), -\frac{\partial}{\partial y}\log \overline{Q}_{BWG}(x,y)\right),$$

the plots of the joint hazard rate functions of the BGW distributions have been illustrated by Figure 3.

The marginal PDF of a BGW distribution, i.e., the univariate-G Weibull distribution with parameters α, p, α_1 and β_1, is denoted by UGW$(\alpha, p, \alpha_1, \beta_1)$ and is given by

$$q_{UGW}(x;\alpha,p,\alpha_1,\beta_1) = \frac{p(\alpha-1)\alpha_1\beta_1 x^{\alpha_1-1}\exp(-\beta_1 x^{\alpha_1})\alpha^{\exp(-\beta_1 x^{\alpha_1})}\log\alpha}{\left\{\alpha - 1 - (1-p)(\alpha - \alpha^{\exp(-\beta_1 x^{\alpha_1})})\right\}^2}, \quad x > 0. \tag{24}$$

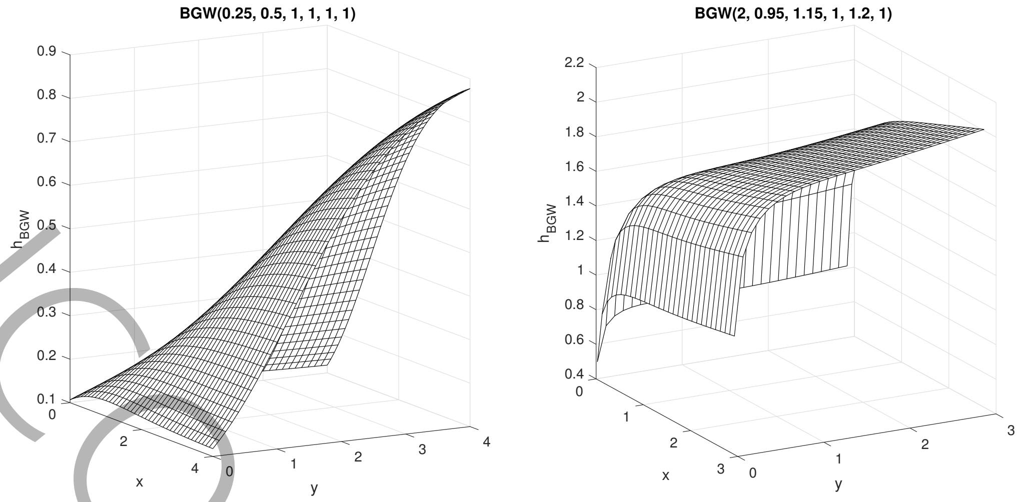

Figure 2: The scalar hazard rate function plots of BGW$(\alpha, p, \alpha_1, \beta_1, \alpha_2, \beta_2)$ distributions.

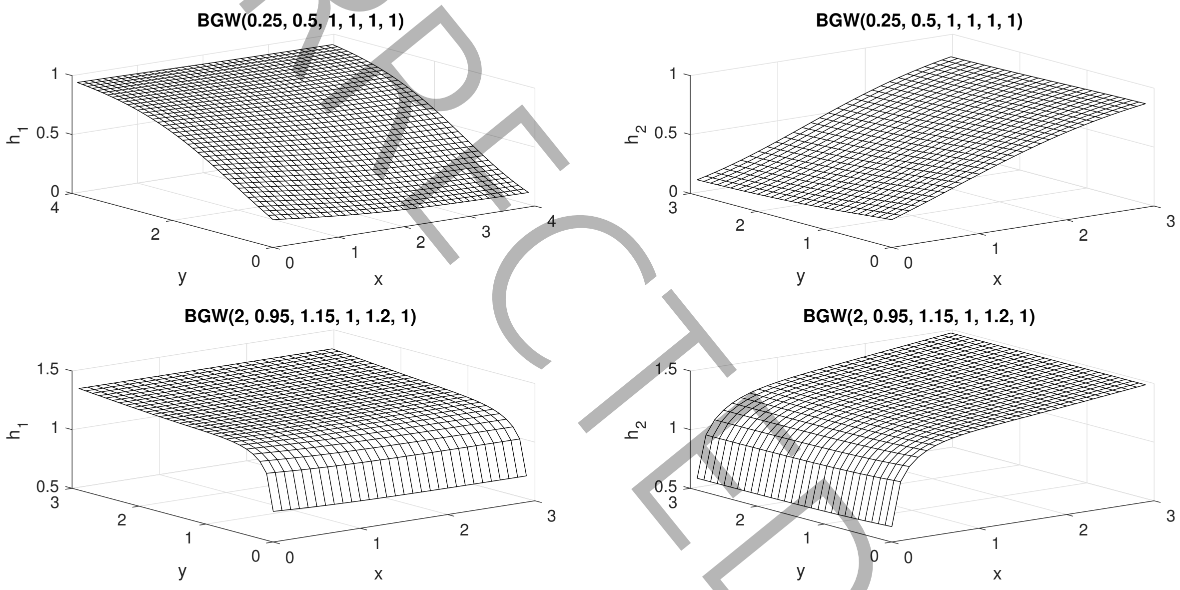

Figure 3: The joint hazard rate function plots of BGW$(\alpha, p, \alpha_1, \beta_1, \alpha_2, \beta_2)$ distributions.

The hazard rate function of a UGW$(\alpha, p, \alpha_1, \beta_1)$ distribution is also given by

$$h_{UGW}(x; \alpha, p, \alpha_1, \beta_1) = \frac{p(\alpha-1)\alpha_1\beta_1 x^{\alpha_1 - 1}\exp(-\beta x^{\alpha_1})\alpha^{\exp(-\beta_1 x_1^{\alpha})}\log\alpha}{\alpha - 1 - (1-p)(\alpha - \alpha^{\exp(-\beta_1 x^{\alpha_1})})(\alpha^{\exp(-\beta_1 x^{\alpha_1})} - 1)}, \quad x > 0.$$

The plots of the density and hazard rate functions of a UGW$(\alpha, p, \alpha_1, 1)$ distribution have been illustrated by Figure 4.

Since, in a UGW model, the hazard rate function can be decreasing, increasing, bathtub-shaped and upside-down bathtub, it is quite appropriate for analyzing different lifetime data. In Figure 4, e.g., in the case of $\alpha = 0.5$, $p = 0.5$ and $\alpha_1 = 1.5$, it is seen that the hazard rate function of the UGW distribution is increasing and these graphical results confirm the proof of Theorem 3.4 and get our stamp of approvals. This is so because when $\alpha_1 > 1$, the Weibull distribution is IHR.

Remark 3.18. *The UGW$(\alpha, p, \alpha_1, \beta_1)$ distribution, similar to the bivariate case, can be reduced to the univariate-G exponential and univariate-G Rayleigh distributions, when $\alpha_1 = 1$ and $\alpha_1 = 2$, respectively.*

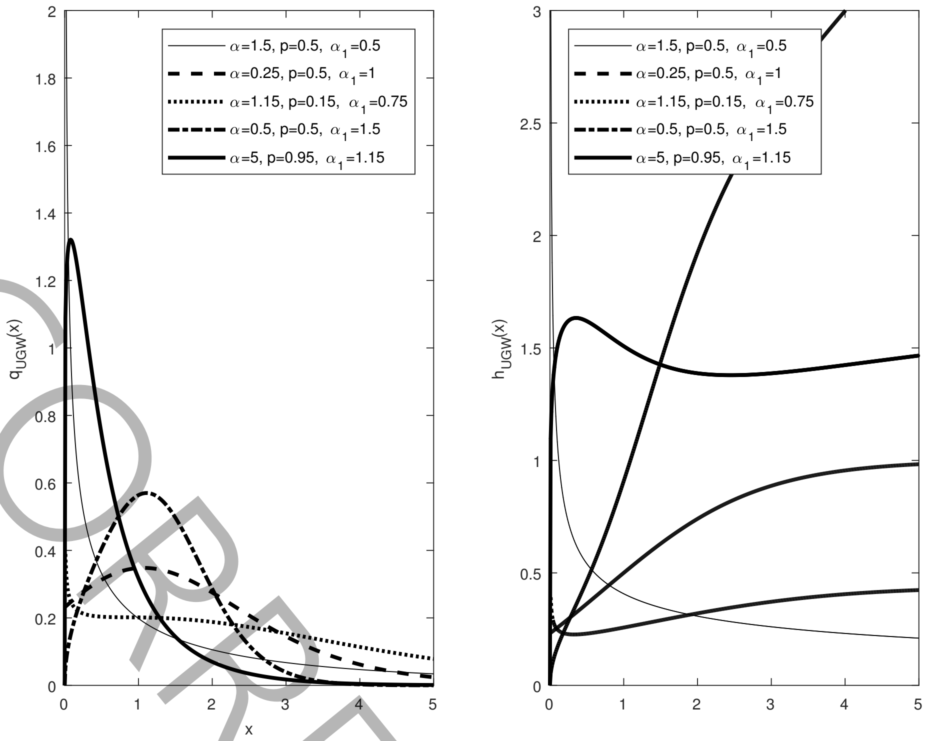

Figure 4: Density (left) and hazard rate function (right) of UGW$(\alpha, p, \alpha_1, \beta_1)$ distributions for $\beta_1 = 1$ and some other parameters values.

4. Statistical inference

In this section, we demonstrate how to employ the EM type algorithms in order to estimate the unknown parameters of the bivariate G-Weibull distributions.

Let $\{(x_1, y_1), \ldots, (x_m, y_m)\}$ be a bivariate sample of size m from the BGW distribution with parameters vector $\mathbf{\Omega} = (\alpha, p, \alpha_1, \alpha_2, \beta_1, \beta_2)^T$. It must be mentioned that generating random samples from a BGW distribution is quite simple. More precisely, the following algorithm can help us for that purpose:

1. Generate n from the GM(p) distribution.
2. Generate $x_1, \ldots, x_n$ from the UGW$(\alpha, 1, \alpha_1, \beta_1)$ distribution, and $y_1, \ldots, y_n$ from the UGW$(\alpha, 1, \alpha_2, \beta_2)$ distribution.
3. Compute the desired (x, y) as $x = \max\{x_1, \ldots, x_n\}$ and $y = \max\{y_1, \ldots, y_n\}$.

It is obvious that computing the MLEs, using the direct likelihood function maximization, involves a six-dimensional optimization process. So, to avoid that we propose to treat this model as an incomplete data problem and use the EM algorithm. Along with (X, Y), suppose we also observe the associated value of N. Therefore, $\{(x_1, y_1, n_1), \ldots, (x_m, y_m, n_m)\}$ performs the complete observations.

The complete log-likelihood function of $\mathbf{\Omega} = (\alpha, p, \alpha_1, \alpha_2, \beta_1, \beta_2)$, using (18), is given by

$$\begin{aligned}
\ell(\mathbf{\Omega}) = m\{\log p + \log\alpha_1 + \log\beta_1 + \log\alpha_2 + \log\beta_2 + 2\log\log\alpha\} + \sum_{i=1}^{m}(n_i - 1)\log(1-p) \\
+ \sum_{i=1}^{m}(n_i - 1)\log(\alpha - \alpha^{\exp(-\beta_1 x_i{}^{\alpha_1})}) + \sum_{i=1}^{m}(n_i - 1)\log(\alpha - \alpha^{\exp(-\beta_2 y_i{}^{\alpha_2})}) \\
+ (\alpha_1 - 1)\sum_{i=1}^{m}\log x_i - \beta_1\sum_{i=1}^{m} x_i{}^{\alpha_1} + (\alpha_2 - 1)\sum_{i=1}^{m}\log y_i - \beta_2\sum_{i=1}^{m} y_i{}^{\alpha_2} \\
+ \sum_{i=1}^{m}(\exp(-\beta_1 x_i{}^{\alpha_1}) + \exp(-\beta_2 y_i{}^{\alpha_2})\log\alpha - \sum_{i=1}^{m} 2n_i\log(\alpha - 1),
\end{aligned} \tag{25}$$

which is free of ineffective constants.

Suppose $\hat{\mathbf{\Omega}}^{(k)} = (\hat{\alpha}^{(k)}, \hat{p}^{(k)}, \hat{\alpha_1}^{(k)}, \hat{\alpha_2}^{(k)}, \hat{\beta_1}^{(k)}, \hat{\beta_2}^{(k)})$ is the current estimate of $\mathbf{\Omega}$, in the k^{th} iteration. Using (20), and considering $\hat{\eta_i}^{(k)} = E(N_i|X_i = x_i, Y_i = y_i; \hat{\mathbf{\Omega}}^{(k)})$, it follows that the conditional expectation of the log-likelihood function, $Q(\mathbf{\Omega}|\mathbf{\Omega}^{(k)})$, based on complete data, has the following structure:

$$
\begin{aligned}
Q(\mathbf{\Omega}|\mathbf{\Omega}^{(k)}) = m\left\{\log p + \log\alpha_1 + \log\beta_1 + \log\alpha_2 + \log\beta_2 + 2\log\log\alpha\right\} + \sum_{i=1}^{m}(\hat{\eta_i}^{(k)} - 1)\log(1-p) \\
+ \sum_{i=1}^{m}(\hat{\eta_i}^{(k)} - 1)\log(\alpha - \alpha^{\exp(-\beta_1 x_i{}^{\alpha_1})}) + \sum_{i=1}^{m}(\eta_i - 1)\log(\alpha - \alpha^{\exp(-\beta_2 y_i{}^{\alpha_2})}) \\
+ (\alpha_1 - 1)\sum_{i=1}^{m}\log x_i - \beta_1\sum_{i=1}^{m} x_i{}^{\alpha_1} + (\alpha_2 - 1)\sum_{i=1}^{m}\log y_i - \beta_2\sum_{i=1}^{m} y_i{}^{\alpha_2} \\
+ \sum_{i=1}^{m}\exp(-\beta_1 x_i{}^{\alpha_1}) + \exp(-\beta_2 y_i{}^{\alpha_2})\log\alpha - \sum_{i=1}^{m} 2\hat{\eta_i}^{(k)}\log(\alpha - 1).
\end{aligned}
\tag{26}
$$

Now, we propose the following EM-type algorithm:

E-step: Given $\mathbf{\Omega} = \hat{\mathbf{\Omega}}^{(k)}$, compute $\hat{\eta_i}^{(k)}$ for $i = 1\dots m$, using (20).
CM-step 1: Update $\hat{p}^{(k)}$ by maximizing (26) w.r.t. p, which yields that

$$
\hat{p}^{(k+1)} = \frac{m}{\sum\limits_{i=1}^{m}\hat{\eta}_i^{(k)}}. \tag{27}
$$

CM-step 2: Obtain $\alpha_1{}^{(k+1)}$ as the solution of the following equation (w.r.t. α_1):

$$
\begin{aligned}
\sum_{i=1}^{m}(\hat{\eta_i}^{(k)} - 1)\frac{\beta_1\alpha^{\exp(-\beta_1 x_i{}^{\alpha_1})} x_i{}^{\alpha_1}\exp(-\beta_1 x_i{}^{\alpha_1})\log(\alpha)\log(x_i)}{\alpha - \alpha^{\exp(-\beta_1 x_i{}^{\alpha_1})}} + \sum_{i=1}^{m}\frac{\alpha_1\log(x_i) - \alpha_1\beta_1 x_i{}^{\alpha_1}\log(x_i) + 1}{\alpha_1} \\
- \sum_{i=1}^{m}\beta x_i{}^{\alpha_1}\exp(-\beta_1 x_i{}^{\alpha_1})\log(x_i)\log\alpha = 0.
\end{aligned}
\tag{28}
$$

CM-step 3: Fix $\hat{\alpha_1}^{(k+1)}$ and obtain $\hat{\beta_1}^{(k+1)}$ as the solution of the following equation (w.r.t. β_1):

$$
\begin{aligned}
\sum_{i=1}^{m}(\eta_i{}^{(k)} - 1)\frac{\alpha^{\exp(-\beta_1 x_i{}^{\alpha_1{}^{(k+1)}})} x_i{}^{\alpha_1{}^{(k+1)}}\exp(-\beta_1 x_i{}^{\alpha_1{}^{(k+1)}})\log(\alpha)}{\alpha - \alpha^{\exp(-\beta_1 x_i{}^{\alpha_1{}^{(k+1)}})}} - \sum_{i=1}^{m}\frac{\beta_1 x_i{}^{\alpha_1{}^{(k+1)}} - 1}{\beta_1} \\
- \sum_{i=1}^{m}\log(\alpha) x_i{}^{\alpha^{(k+1)}}\exp(-\beta_1 x_i{}^{\alpha_1{}^{(k+1)}}) = 0.
\end{aligned}
$$

CM-step 4: Obtain $\alpha_2{}^{(k+1)}$ as the solution of the following equation (w.r.t. α_2):

$$
\begin{aligned}
\sum_{i=1}^{m}(\hat{\eta_i}^{(k)} - 1)\frac{\beta_1\alpha^{\exp(-\beta_1 x_i{}^{\alpha_1})} x_i{}^{\alpha_1}\exp(-\beta_1 x_i{}^{\alpha_1})\log(\alpha)\log(x_i)}{\alpha - \alpha^{\exp(-\beta_1 x_i{}^{\alpha_1})}} + \sum_{i=1}^{m}\frac{\alpha_1\log(x_i) - \alpha_1\beta_1 x_i{}^{\alpha_1}\log(x_i) + 1}{\alpha_1} \\
- \sum_{i=1}^{m}\beta x_i{}^{\alpha_1}\exp(-\beta_1 x_i{}^{\alpha_1})\log(x_i)\log\alpha = 0.
\end{aligned}
\tag{29}
$$

CM-step 5: Fix $\hat{\alpha_2}^{(k+1)}$ and obtain $\hat{\beta_2}^{(k+1)}$ as the solution of the following equation (w.r.t. β_2):

$$
\begin{aligned}
\sum_{i=1}^{m}(\eta_i{}^{(k)} - 1)\frac{\alpha^{\exp(-\beta_1 x_i{}^{\alpha_1{}^{(k+1)}})} x_i{}^{\alpha_1{}^{(k+1)}}\exp(-\beta_1 x_i{}^{\alpha_1{}^{(k+1)}})\log(\alpha)}{\alpha - \alpha^{\exp(-\beta_1 x_i{}^{\alpha_1{}^{(k+1)}})}} - \sum_{i=1}^{m}\frac{\beta_1 x_i{}^{\alpha_1{}^{(k+1)}} - 1}{\beta_1} \\
- \sum_{i=1}^{m}\log(\alpha) x_i{}^{\alpha^{(k+1)}}\exp(-\beta_1 x_i{}^{\alpha_1{}^{(k+1)}}) = 0.
\end{aligned}
$$

CM-step 6: Update $\hat{\alpha}^{(k)}$ by optimizing the following constrained actual log-likelihood function

$$
\hat{\alpha}^{(k+1)} = \arg\max_{\alpha}\sum_{i=1}^{m}\log q_{BGW}\left(x_i, y_i; \hat{p}^{(k+1)}, \alpha, \hat{\alpha_1}^{(k+1)}, \hat{\alpha_2}^{(k+1)}, \hat{\beta_1}^{(k+1)}, \hat{\beta_2}^{(k+1)}\right). \tag{30}
$$

When the absolute difference between the two consecutive log-likelihood values is less than 10^{-3}, the algorithm is stopped.

A similar procedure can be performed to estimate the parameters of the UGW distribution, which in order to sake of brevity the details are avoided.

4.1. Simulated examples and data analysis for BGW distributions

Here, we firstly present some simulation results to indicate how the proposed EM algorithm works. We generate samples of sizes m =50, 100 and 200 from the BGW distributions for an arbitrary sets of parameters such as $\alpha = 0.5$, $p = 0.5$, $\alpha_1 = 0.2$, $\beta_1 = 1$, $\alpha_2 = 0.2$ and $\beta_2 = 1$. The average estimates (AEs) of the parameters, and the mean squared errors (MSEs) of the estimates have been reported in Table 2 based on 500 repetitions.

Table 1: AEs (MSEs) of the MLEs based on 500 repetitions from the BGW distribution.

Parameter	$m = 50$	$m = 100$	$m = 200$
$\alpha = 0.5$	0.4683 (0.0092)	0.4702 (0.0089)	0.4789 (0.0086)
$p = 0.5$	0.5175 (0.0852)	0.5081 (0.0671)	0.5077 (0.0482)
$\alpha_1 = 0.2$	0.3431 (0.0422)	0.3208 (0.0289)	0.3004 (0.0209)
$\beta_1 = 1$	1.3707 (0.8868)	1.2308 (0.5521)	1.1789 (0.4108)
$\alpha_2 = 0.2$	0.3571 (0.0458)	0.3265 (0.0307)	0.2905 (0.0171)
$\beta_2 = 1$	1.3013 (0.8412)	1.2033 (0.5494)	1.2655 (0.4178)

It is seen that the introduced EM algorithm works quite well and it can be used quite effectively for data analysis purposes.

In the following, for comparing purposes, we will consider another simulated data set drawn from the modified Sarha-Balakrishnan singular bivariate (MSBSB) distribution studied by Kundu and Gupta [12]. Indeed, the last authors generated a sample of size 30 from an MSBSB distribution. This data set has been presented in Table 2.

Table 2: Generated data from an MSBSB distribution by Kundu and Gupta (2010).

No.	x	y	No.	x	y
1	0.418	0.424	16	0.885	0.791
2	0.106	0.851	17	0.049	0.200
3	1.147	1.147	18	1.088	1.128
4	0.529	1.795	19	1.453	1.155
5	0.446	0.446	20	0.878	0.878
6	0.326	0.326	21	0.945	0.945
7	0.205	0.205	22	0.850	0.850
8	1.106	1.106	23	0.354	0.354
9	0.435	0.973	24	0.345	2.308
10	0.935	0.850	25	1.198	0.620
11	0.362	0.645	26	0.525	0.504
12	0.257	1.464	27	0.548	0.548
13	1.608	1.608	28	2.837	1.057
14	0.628	0.628	29	0.212	1.697
15	0.351	0.351	30	2.356	1.348

Now, we fit the BGW distribution to this data set. The parameters involved have been estimated by means of the reliable proposed EM algorithm. To deal with these data, which can be considered as a real data set for the BGW distribution in some sense that they have been produced from another bivariate distribution, we started the algorithm with different initial guesses and observed that the EM algorithm converged to the same point. Figure 5 represents the behaviour of the profile log-likelihood function of α in a BGW distribution.

Table 3 contains the MLEs, the associated 95% confidence intervals which have been constructed by means of a bootstrap method, Akaike information criterion (AIC) and estimated log-likelihood value for the fitted BGW and MSBSB distributions. From the results presented in Table 3, it is obvious the BGW distribution gives a better performance w.r.t. the MSBSB distribution. Moreover, we consider the capacity of the marginals and report the corresponding results, in Table 4, based on the Kolmogorov-Smirnov test statistic and its corresponding p-value.

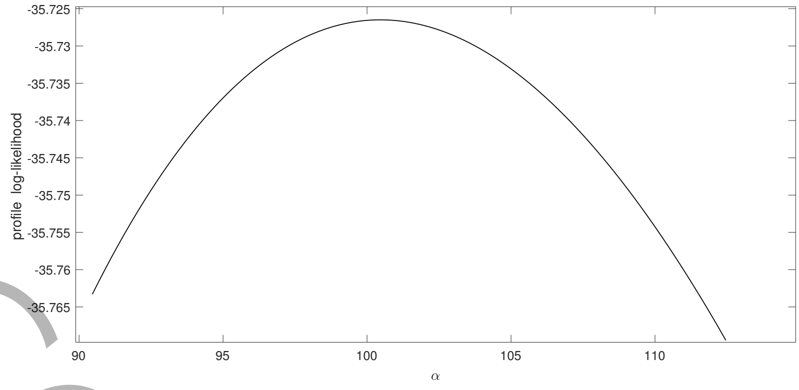

Figure 5: The profile log-likelihood plot of parameter α in a BGW model regarding the data of Table 5.

Although the data have been generated from an MSBSB distribution, based on the p-values reported in Table 4, one cannot reject the statement that states a BGW distribution can adapt with these data.

Table 3: Fitting computations for BGW and MSBSB distributions.

model	MLEs (Confidence Intervals)	$\hat{\ell}$	AIC
BGW	$\hat{\alpha}$ =100.4600 (98.0829, 102.8369)	- 35.7265	83.4529
	$\hat{p}$ = 0.0232 (0.0229, 0.0234)		
	$\hat{\alpha_1}$= 0.8657 (0.8629, 0.8683)		
	$\hat{\beta}_1$= 2.0091 (2.0013, 2.0169)		
	$\hat{\alpha_2}$=1.1851 (1.1813, 1.1889)		
	$\hat{\beta}_2$=1.7987 (1.7925, 1.8048)		
MSBSB	$\hat{\alpha_1}$= 1.5725 (1.0367, 2.8692)	- 64.8979	137.7958
	$\hat{\alpha_2}$= 3.1405 (1.7187, 4.7924)		
	$\hat{\alpha_3}$= 1.9585 (1.3833, 3.8889)		
	$\hat{\lambda}$= 1.0932 (0.7640, 1.4650)		

Table 4: Fitting results of the marginals.

Variable	Model	K-S	p-value
X	UGW(100.4600, 0.0232, 0.8657, 2.0091)	0.1098	0.8244
Y	UGW(100.4600, 0.0232, 1.1851, 1.7987)	0.1177	0.7568

Now, we want to analyze a real data set obtained from the National Basketball Association franchise Boston Celtics during the second half of the 2001 - 2002 session. This data set consists of three star players (Antoine Walker, Paul Pierce, Kenny Anderson) on a basketball team and represents the game times for each player's with second personal foul for a particular game, in which all three players have participated and committed at least two fouls at the end. It is expected that once a player committed at least two fouls, the player might be out of the game for some times and, hence, the foul rate of the other star players might change. Kvam and Pena [14], Deshpande et al. [7] and Asha et al. [3] considered these data. The data set has been presented in Table 5.

Regarding Kvam and Pena [14], once a star player committed two fouls, the foul rate of the other star players changes. So, the last authors assumed that the three star players compose a system, whereas Deshpande et al. [7] assumed that any two star players compose a system. Similar to the last authors, we assume that any two star players compose a system. So, we analyze the data set with the assumption that three different systems, including as System I (Player I and Player II), System II (Player I and Player III) and System III (Player II and Player III), exist.

We fit the BGW distribution to System I, System II and System III. In order to sake of brevity, the results are summarized in Table 6. The confidence intervals, presented in this table, have been obtained by means of a bootstrap method. We also test the marginals for each system. The results of fitting the marginals such as the K-S test statistics and their corresponding p-values have been reported in Table 7. Based on the p-values

Table 5: Time until second foul for the three star players in 28 games of Boston Celtics 2001-2002.

Game	Player I (Pierce)	Player II (Walker)	Player III (Anderson)	Game	Player I (Pierce)	Player II (Walker)	Player III (Anderson)
1	21.02	30.22	43.43	15	42.06	23.21	45.36
2	24.25	45.54	17.19	16	28.51	33.59	16.2
3	6.555	19.47	23.28	17	34.56	32.53	40.44
4	15.35	16.37	25.4	18	40.33	15.35	28.33
5	39.08	30.32	43.53	19	27.56	46.21	28.05
6	16.2	4.16	39.52	20	9.54	36.21	28.12
7	34.59	46.44	16.33	21	27.09	11.11	23.33
8	19.1	38.4	20.17	22	40.36	33.21	17.04
9	28.22	37.43	25.41	23	41.44	36.28	19.13
10	32	45.52	39.11	24	32.23	8.17	41.27
11	11.25	19.09	11.59	25	7.53	37.31	13.43
12	17.39	25.43	22.51	26	28.34	35.58	41.48
13	28.47	31.15	2.41	27	26.32	28.02	29.33
14	23.42	31.28	40.03	28	30.47	40.4	42.13

reported in Table 7, it is obvious that the UGW distributions fit the marginals quit well. So, one cannot reject the statement that states the BGW distribution gives satisfactory fits to these joint data. In addition, based on the log-likelihood values and AICs, presented in Table 6, one can see that System I has recieved the best fit from the BGW distribution.

Figure 6 represents the unimodality of the profile-log-likelihood function of α in the fitted BGW distributions to each system.

Table 6: Fitting computations for the BGW distribution for Basketball data.

Model	MLEs (Confidence Intervals)	$\hat{\ell}$	AIC
Player I and Player II (System I)	$\downarrow$ $\hat{\alpha}$ =0.2728 (0.2630, 0.2826) $\hat{p}$ =0.3487 (0.3451, 0.3524) $\hat{\alpha_1}$=1.8585 (1.8568, 1.8603) $\hat{\beta}_1$= 0.0042 (0.0041, 0.0042) $\hat{\alpha_2}$=1.9733 (1.9715, 1.9752) $\hat{\beta}_2$= 0.0021 (0.0021, 0.0021)	-211.322	434.643
Player I and Player III (System II)	$\downarrow$ $\hat{\alpha}$ =0.4627 (0.4476, 0.4778) $\hat{p}$ = 0.3731 (0.3692, 0.3770) $\hat{\alpha_1}$= 2.0034 (2.0016, 2.0053) $\hat{\beta}_1$= 0.0022 (0.0022, 0.0022) $\hat{\alpha_2}$=1.8937 (1.8917, 1.8958) $\hat{\beta}_2$=0.0028 (0.0028, 0.002)	-211.966	435.932
Player II and Player III (System III)	$\downarrow$ $\hat{\alpha}$ =0.1023 (0.0998, 0.1049) $\hat{p}$ =0.9864 (0.9859, 0.9869) $\hat{\alpha_1}$=2.1726 (2.1706, 2.1746) $\hat{\beta}_1$= 0.0008 (0.0008, 0.0008) $\hat{\alpha_2}$=1.8646 (1.8620, 1.8672) $\hat{\beta}_2$= 0.0028 (0.0027, 0.0028)	-216.492	444.9846

Table 7: Marginal computations for the Basketball data.

System	Variable	K-S statistic	p-value
System I	Player I (X)	0.1089	0.8585
	Player II (Y)	0.1319	0.6665
System II	Player I(X)	0.1075	0.8685
	Player III (Y)	0.1768	0.3079
System III	Player II(X)	0.1462	0.5395
	Player III(Y)	0.1835	0.2681

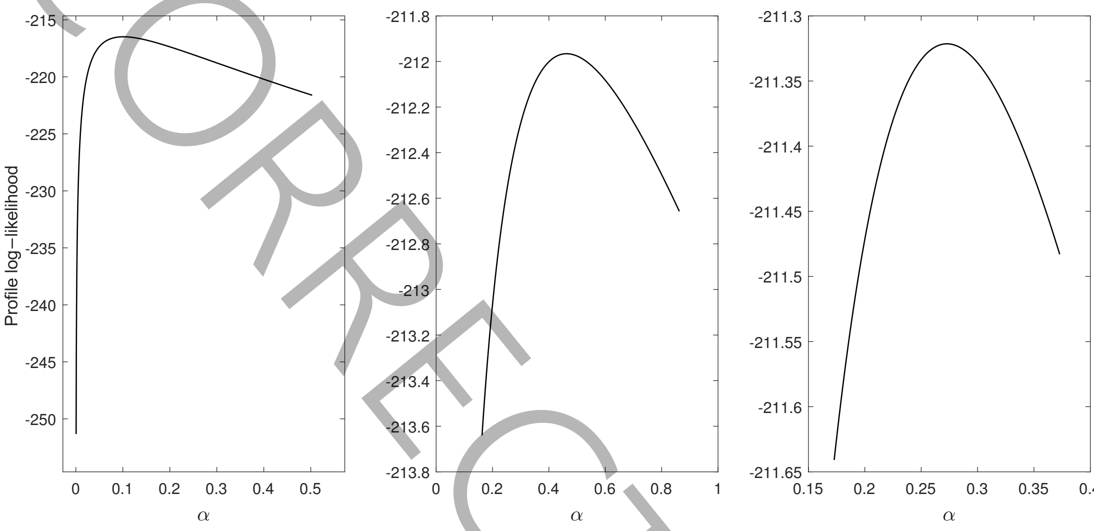

Figure 6: The profile-log-likelihood function of α in fitted BGW distributions to each system.

5. Conclusions

In this paper, some new univariate and bivariate family of semi-parametric distributions were discussed. In the univariate case, the family of UG distributions was introduced and some of its important futures were studied. Specially, the behaviour of the hazard rate function was investigated. In addition, the APT distribution of a second type, was studied as a special member of the UG family of distributions.

In the bivariate case, the BG family of distributions were discussed motivated by the fact that it provides greater flexibility in order to analyze various bivariate data. Different aspects of this family of bivariate distributions, which are important in their own sake, were investigated. We also saw that the introduced EM algorithm had a good performance to estimate the parameters of the distributions of this bivariate family.

Acknowledgments

The authors sincerely thank anonymous referees and also the Editor-in-Chief for their valuable comments that led to improvement of the paper.

Data Availability

The first data set, presented in Table 2, used in this study were obtained from the previously published article: https://doi.org/10.1016/j.jspi.2009.07.026.
The second data set, presented in Table 5, used in this study were obtained from the previously published article: https://doi.org/10.1198/016214504000000863.

References

[1] K. Adamidis, T. Dimitrakopoulou, and S. Loukas, *On an extension of the exponential-geometric distribution*, Statist. Probab. Lett., 73 (2005), pp. 259–269.

[2] K. Adamidis and S. Loukas, *A lifetime distribution with decreasing failure rate*, Statist. Probab. Lett., 39 (1998), pp. 35–42.

[3] G. Asha, J. K. K. M., and D. Kundu, *An extension of the freund's bivariate distribution to model load-sharing systems*, American Journal of Mathematical and Management Sciences, 35 (2016), pp. 207–226.

[4] N. Balakrishnan and C.-D. Lai, *Continuous bivariate distributions*, Springer, Dordrecht, second ed., 2009.

[5] W. Barreto-Souza, A. L. de Morais, and G. M. Cordeiro, *The Weibull-geometric distribution*, J. Stat. Comput. Simul., 81 (2011), pp. 645–657.

[6] A. P. Basu, *Bivariate failure rate*, Journal of the American Statistical Association, 66 (1971), pp. 103–104.

[7] J. V. Deshpande, I. Dewan, and U. V. Naik-Nimbalkar, *Two component load sharing systems with applications to Biology*, tech. rep., Indian Statistical Institute, New Delhi, India, 2007. Available at `https://www.isid.ac.in/~statmath/eprints/2007/isid200706.pdf`.

[8] F. Hemmati, E. Khorram, and S. Rezakhah, *A new three-parameter ageing distribution*, J. Statist. Plann. Inference, 141 (2011), pp. 2266–2275.

[9] N. L. Johnson and S. Kotz, *A vector multivariate hazard rate*, J. Multivariate Anal., 5 (1975), pp. 53–66.

[10] C. Kuş, *A new lifetime distribution*, Comput. Statist. Data Anal., 51 (2007), pp. 4497–4509.

[11] D. Kundu, *Bivariate geometric (maximum) generalized exponential distribution*, Journal of Data Science, 13 (2015), pp. 693–712.

[12] D. Kundu and R. D. Gupta, *Modified Sarhan-Balakrishnan singular bivariate distribution*, J. Statist. Plann. Inference, 140 (2010), pp. 526–538.

[13] ——, *Power-normal distribution*, Statistics, 47 (2013), pp. 110–125.

[14] P. H. Kvam and E. A. Peña, *Estimating load-sharing properties in a dynamic reliability system*, Journal of the American Statistical Association, 100 (2005), pp. 262–272.

[15] A. Mahdavi and D. Kundu, *A new method for generating distributions with an application to exponential distribution*, Comm. Statist. Theory Methods, 46 (2017), pp. 6543–6557.

[16] E. Mahmoudi and A. A. Jafari, *Generalized exponential-power series distributions*, Comput. Statist. Data Anal., 56 (2012), pp. 4047–4066.

[17] A. W. Marshall and I. Olkin, *Life Distributions: Structure of Nonparametric, Semiparametric, and Parametric Families*, Springer New York, NY, 1 ed., 2007.

[18] V. Nekoukhou, A. Khalifeh, and H. Bidram, *Univariate and bivariate extensions of the generalized exponential distributions*, Math. Slovaca, 71 (2021), pp. 1581–1598.

[19] V. Nekoukhou, A. Khalifeh, and E. a. Mahmoudi, *Bivariate rayleigh-geometric distribution*, Journal of Statistical Sciences, 13 (2020).

[20] R. B. Nelsen, *An introduction to copulas*, Springer Series in Statistics, Springer, New York, second ed., 2006.

[21] P. S. Puri and H. Rubin, *On a characterization of the family of distributions with constant multivariate failure rates*, Ann. Probability, 2 (1974), pp. 738–740.

[22] R. B. Silva, M. Bourguignon, C. R. B. Dias, and G. M. Cordeiro, *The compound class of extended Weibull power series distributions*, Comput. Statist. Data Anal., 58 (2013), pp. 352–367.

Please cite this article using:

Vahid Nekoukhou, Hamid Bidram, Ashkan Khalifeh, A bivariate α-power transformed family: Theory and application, AUT J. Math. Comput., 7(1) (2026) 1-18
https://doi.org/10.22060/AJMC.2024.23495.1261