Statistical Analysis of Bivariate Failure Time Data Based on Bathtub-Shaped Failure Rate Model

Shirin Shoaee

Department of Statistics, Faculty of Mathematical Sciences, Shahid Beheshti University

Abstract. Many distributions have been presented with bathtub-shaped failure rates for real-life data. A two-parameter distribution was defined by Chen (2000). This distribution can have a bathtub-shaped or increasing failure rate function. In this paper, we consider two bivariate models based on the proposed distribution by Chen and use the proposed methods of Marshall and Olkin (1967) in the bivariate case and Marshall and Olkin (1997) in the univariate case. In the second case, their method is generalized to the bivariate case and a new bivariate distribution is introduced. These new bivariate distributions have natural interpretations, and they can be applied in fatal shock models or in competing risks models. We call these new distributions as the bivariate Chen (BCH) distribution and bivariate Chen-geometric (BCHG) distribution, respectively. Moreover, the BCH can be obtained as a special case of the BCHG model. Then, the various properties of the new distributions are investigated. The BCHG distribution has five parameters and the maximum likelihood estimators cannot be obtained in a closed form. We suggest using an EM algorithm that is very easy to implement. Also, Monte Carlo simulations are performed to investigate the effectiveness of the proposed algorithm. Finally, we analyze two real data sets for illustrative purposes.

Keywords. Bathtub-Shaped failure rate, Bivariate model, EM algorithm, Monte Carlo simulation, Pseudo likelihood function.

MSC: 62H10; 62H12; 62E15.

Shirin Shoaee (Sh_Shoaee@sbu.ac.ir).

1   Introduction

Many parametric distributions have been presented to analyze real data with bathtub-shaped failure rates. The bathtub-shaped hazard function introduces a suitable model for some electronic, mechanical products and also, the lifetime of humans. The reason for using this model is that many data sets available in reliability theory, clinical trials and biostatistics surveys specify the bathtub-shaped failure rate.

In the area of lifetime analysis, there has been a lot of research on distributions that have bathtub-shaped failure rate function. The exponential power distribution is proposed by Smith and Bain (1975) and studied by Leemis (1986). A four parameter distribution was introduced by Gaver and Acar (1979). Also, a distribution with increasing, decreasing or bathtub-shaped failure rate has been considered by Hjorth (1980). An exponentiated-Weibull distribution with three parameters is investigated by Mudholkar and Srivastava (1993). Xie *et al.* (2002) proposed a modified Weibull extension with three parameters and Bebbington *et al.* (2007) proposed a new two-parameter aging distribution, which is a generalization of the Weibull, and studied its properties. Zhang and Xie (2007) addressed the extended Weibull distribution in Marshall and Olkin (1997) by introducing an additional parameter to an existing model and investigated some properties of this distribution. Complex systems usually have a bathtub-shaped failure rate over the life cycle of the product. Though some distributions are mentioned to have a bathtub-shaped failure rate, they may fail to give a good bathtub-shape of the failure rate. Zhang *et al.* (2005) discussed the parametric analysis of some models which exhibit a good bathtub-shape.

In this paper, we use a two-parameter lifetime distribution with either bathtub-shaped or increasing failure rate investigated by Chen (2000). The cumulative distribution function (CDF) of this distribution is

$$F(x, \alpha, \lambda) = 1 - e^{\lambda(1 - e^{x^\alpha})}, \qquad x > 0, \tag{1.1}$$

where $\lambda > 0$ and $\alpha > 0$ are the scale and shape parameters, respectively. The probability density function (PDF) is

$$f(x, \alpha, \lambda) = \lambda \alpha x^{\alpha-1} e^{x^\alpha} e^{\lambda(1 - e^{x^\alpha})}, \qquad x > 0. \tag{1.2}$$

The corresponding failure rate function of this distribution is

$$h(x, \alpha, \lambda) = \lambda \alpha x^{\alpha-1} e^{x^\alpha}, \qquad x > 0. \tag{1.3}$$

Using the first derivative of $h(x)$, one can easily verify that $h(x)$: $h(x)$ has a bathtub-shape when $\alpha < 1$ and $h(x)$ is an increasing function when $\alpha \geq 1$.

The maximum likelihood and Bayes estimates of parameters are discussed by Sarhan *et al.* (2012). The statistical inference about the shape parameter with type-II right censored data is investigated by Wu *et al.* (2004) and the estimation problem of progressive type-II censored data using the maximum likelihood procedure is studied by Wu (2008). Also, based on progressively Type-II censored samples, the Bayesian and classical inference of $R = P(Y < X)$, are investigated by Shoaee and Khorram (2015). Nadaraja and Kotz (2007) explained that this distribution is a particular case of the distribution considered by Gurvich *et al.* (1997) and offered by Haynatzki *et al.* (2000).

The modeling of a lifetime is an important problem in many technical fields. Several methods for multivariate survival data have been proposed, for more details, see Kotz *et al.* (2000).

One classical model of dependent lives was proposed by Marshall and Olkin (1967). This model assumes that the lifetimes of two persons, say T_1 and T_2, are independent unless a common shock causes the death of both. Thus, T_0 denotes the time until the common disaster. If X and Y are defined as:

$$X = \min(T_1, T_0), \qquad and \qquad Y = \min(T_2, T_0),$$

then, the bivariate distribution of the random vector (X, Y) is defined for $x, y > 0$ by the following joint survival function

$$\bar{F}_{X,Y}(x, y) = P(X > x, Y > y) = P(T_1 > x, T_2 > y, T_0 > \max(x, y))$$
$$= \bar{F}_{T_1}(x)\bar{F}_{T_2}(y)\bar{F}_{T_0}(\max(x, y)).$$

In view of the mutual independence of T_0, T_1 and T_2, this model is also called the bivariate survival model of Marshall-Olkin type.

Also, for modeling survival data, Marshall and Olkin (1997) presented a method to add an extra parameter to a family of univariate distributions. In this method, they introduced a class of univariate distributions which can be obtained by minimum and maximum of independent and identically distributed continuous random variables, where the sample size follows the geometric distribution. Therefore, due to the presence of an extra parameter, the proposed class of univariate distributions is more flexible than the exponential or Weibull class, respectively. Since then extensive work has been done and many researchers have investigated the same approach for different distributions, see for example Ghitany *et al.* (2005), Ghitany *et al.* (2007), Pham and Lai (2007), Chahkandi and Ganjali (2009), Barreto-Souza (2012).

Interestingly, not much work has been done on the bivariate distributions mainly because of its analytical intractability. In this regard, Kundu and Gupta (2014) produced the bivariate Weibull-geometric distribution. They investigated some of its properties and obtained the estimates of parameters. Also, Kundu (2015) introduced a bivariate-geometric (maximum) generalized exponential distribution.

Following these works, we introduce two new bivariate distributions by methods of Marshall and Olkin. In the first case, a new bivariate distribution using the method of Marshall and Olkin (1967) is obtained. This new bivariate distribution is called the bivariate Chen (BCH) distribution and the marginal distributions have CH distributions. Also, we discuss various properties of the new distribution.

In the second case, we generalize the BCH distribution. So, a new bivariate distribution by the method of Marshall and Olkin (1997) is introduced. We call this new bivariate distribution the bivariate Chen-geometric (BCHG) distribution. In fact, this method produces a new bivariate distribution which is analytically quite tractable. Also, the marginals and conditionals are univariate Chen-geometric distributions (UCHG). Different properties of this new distribution have been investigated too. Due to the presence of five parameters, it is a very flexible model and the joint PDF can take different shapes.

The maximum likelihood estimators (MLEs) of the unknown parameters of BCHG distribution cannot be obtained in closed forms and it is a main difficultly. It involves solving five nonlinear equations simultaneously. Finding initial values and the convergence of the algorithm are important issues. To avoid this problem, we investigate it as a missing value problem and propose to use the expectation maximization (EM) algorithm to compute the MLEs. In this case, at each E-step we need to solve one-dimensional nonlinear optimization problems. Also, the corresponding M-step involves maximization of one nonlinear function. Therefore, the performance of the proposed EM algorithm is very simple. For illustrative purposes, we analyze two bivariate data sets using this model, and the performance is quite satisfactory. The remaining part of the paper is arranged as follows: In Section 2, a new bivariate survival model of the Marshall-Olkin type which can be used in shock or competing risks models are obtained. Also, various properties of the new bivariate distribution are discussed in this section. A new bivariate distribution by the proposed method of Marshall and Olkin (1997) are introduced and also, different properties of this model are investigated in Section 3. The maximum likelihood estimators are discussed in Section 4. The results of the simulation experiments and the analysis of two real data sets have been presented in Section 5 and finally, we conclude the paper in Section 6.

2 Marshall-Olkin Type Distribution in Competing Risks and Shock Models

2.1 Shock Model and Competing Risks Model

The bivariate distribution is introduced in this paper has popular descriptions and it can be used in shock models or in competing risks models.

In the context of the competing risks and shock models, suppose that two components labeled 1 and 2 in a system are subject to three types of events called shocks in such a way that if the first (second) type of shock occurs, component 1 (2) fails, whereas at the occurrence of the third type of shock, both 1 and 2 fail. Assume that the occurrences of these shocks are governed by three independent renewal processes with the corresponding inter-arrival times denoted by U_1, U_2 and U_0. The lifetime of the first component is the random variable $X_1 = \min(U_1, U_0)$ and that of the second component is $X_2 = \min(U_2, U_0)$, then, (X_1, X_2) denotes the survival time of the components.

These models are very papular in the reliability literature where the failure of different kinds of system components is constructed as being contingent on independent shocks that may affect one or more components. If X_1 and X_2 are life lengths of components subject to shocks, then (X_1, X_2) denotes the life length of a two-component system.

In the context of life insurance, this model can be used for modeling the dependency between two lives in one life insurance contract. In this case, the lifetimes of two persons, say U_1 and U_2, are independent unless a common shock causes the death of both. For example, a contagious deadly disease, a natural catastrophe or a car accident can affect the lives of the two spouses. So, U_0 denotes the time until the common disaster, the actual ages-at-death are obtained by $X_1 = \min(U_1, U_0)$ and $X_2 = \min(U_2, U_0)$.

The common shock model can also be used in finance, where the shocks such as local or global recessions affecting one or more credit-risky assets at a time may cause joint defaults and in this case, the random variables $U_i, i = 0, 1, 2$ are interpreted as arrival times of the shocks while X_1 and X_2 define the default times.

The competing risks model arises in actuarial science as well as survival analysis, system reliability and medical studies. The subject of competing risks is called the multiple decrement theory in actuarial science. For example, a person may die because of one of several possible causes: cancer, heart disease, accident and so on.

2.2 Marshall-Olkin Bivariate Chen Distribution

In this subsection, we will introduce a new bivariate distribution. Suppose U_i follow ($\sim$) $CH(\alpha, \lambda_i)$, for $i = 0, 1, 2$ and are independent. Also, $CH(\alpha, \lambda)$ denotes the bathtub-shaped failure rate (Chen) distribution with the shape parameter $\alpha > 0$ and scale parameter $\lambda > 0$.

Define $X_1 = \min\{U_0, U_1\}$ and $X_2 = \min\{U_0, U_2\}$. Then, the bivariate vector (X_1, X_2) is a bivariate Chen distribution with the parameters $\alpha, \lambda_0, \lambda_1$ and λ_2. We call this new distribution, a bivariate Chen distribution and it will be denoted from now on as $BCH(\alpha, \lambda_0, \lambda_1, \lambda_2)$. The following result is a consequence of the definition above.

Theorem 2.1. *If $(X_1, X_2) \sim BCH(\alpha, \lambda_0, \lambda_1, \lambda_2)$ then the joint survival function takes the following form for $z = \max\{x_1, x_2\}$,*

$$\bar{F}_{X_1,X_2}(x_1, x_2) = P(X_1 \geq x_1, X_2 \geq x_2) = P(U_1 \geq x_1, U_2 \geq x_2, U_0 \geq z)$$

$$= e^{\lambda_1(1-e^{x_1^{\alpha}})} e^{\lambda_2(1-e^{x_2^{\alpha}})} e^{\lambda_0(1-e^{z^{\alpha}})}. \tag{2.1}$$

Corollary 2.1. The joint survival function of the $BCH(\alpha, \lambda_0, \lambda_1, \lambda_2)$ can also be written as:

$$\bar{F}_{X_1,X_2}(x_1, x_2) = \bar{F}_{CH}(x_1, \alpha, \lambda_1)\bar{F}_{CH}(x_2, \alpha, \lambda_2)\bar{F}_{CH}(z, \alpha, \lambda_0)$$

$$= \begin{cases} \bar{F}_{CH}(x_1, \alpha, \lambda_1 + \lambda_0)\bar{F}_{CH}(x_2, \alpha, \lambda_2) & \text{if} \quad x_2 < x_1 \\ \bar{F}_{CH}(x_1, \alpha, \lambda_1)\bar{F}_{CH}(x_2, \alpha, \lambda_2 + \lambda_0) & \text{if} \quad x_1 < x_2 \\ \bar{F}_{CH}(x, \alpha, \lambda_0 + \lambda_1 + \lambda_2, \lambda) & \text{if} \quad x_1 = x_2 = x. \end{cases} \tag{2.2}$$

Theorem 2.2. *If $(X_1, X_2) \sim BCH(\alpha, \lambda_0, \lambda_1, \lambda_2)$, then the joint PDF of (X_1, X_2) is:*

$$f_{X_1,X_2}(x_1, x_2) = \begin{cases} f_1(x_1, x_2) & \text{if} \quad x_2 < x_1 \\ f_2(x_1, x_2) & \text{if} \quad x_1 < x_2 \\ f_0(x) & \text{if} \quad 0 < x_1 = x_2 = x < \infty, \end{cases} \tag{2.3}$$

where

$$f_1(x_1, x_2) = \alpha^2 \lambda_2(\lambda_0 + \lambda_1)(x_1 x_2)^{\alpha-1} e^{x_1^{\alpha}+x_2^{\alpha}} e^{(\lambda_0+\lambda_1)(1-e^{x_1^{\alpha}})} e^{\lambda_2(1-e^{x_2^{\alpha}})},$$

$$f_2(x_1, x_2) = \alpha^2 \lambda_1(\lambda_0 + \lambda_2)(x_1 x_2)^{\alpha-1} e^{x_1^{\alpha}+x_2^{\alpha}} e^{\lambda_2(1-e^{x_1^{\alpha}})} e^{(\lambda_0+\lambda_1)(1-e^{x_2^{\alpha}})},$$

$$f_0(x) = \alpha \lambda_0 x^{\alpha-1} e^{x^{\alpha}} e^{(\lambda_0+\lambda_1+\lambda_2)(1-e^{x^{\alpha}})}.$$

Proof. The expressions of $f_1(.,.)$ and $f_2(.,.)$ can be obtained by taking $-\frac{\partial^2 \bar{F}_{X_1,X_2}(x_1,x_2)}{\partial x_1 \partial x_2}$ for $x_1 < x_2$ and $x_2 < x_1$ respectively. But $f_0(.)$ can not be obtained in the same way. First observe that $f_0(.)$, $f_1(.,.)$ and $f_2(.,.)$ must satisfy

$$\int_0^\infty \int_{x_2}^\infty f_1(x_1,x_2)dx_1 dx_2 + \int_0^\infty \int_{x_1}^\infty f_2(x_1,x_2)dx_2 dx_1 + \int_0^\infty f_0(x)dx = 1.$$

After some lengthy calculations, it can be shown that

$$\int_0^\infty \int_{x_2}^\infty f_1(x_1,x_2)dx_1 dx_2 = \lambda_2 \int_0^\infty \alpha x_2^{\alpha-1} e^{x_2^\alpha} e^{(\lambda_0+\lambda_1+\lambda_2)(1-e^{x_2^\alpha})}dx_2 = \frac{\lambda_2}{\lambda_0 + \lambda_1 + \lambda_2}.$$

and similarly,

$$\int_0^\infty \int_{x_1}^\infty f_2(x_1,x_2)dx_2 dx_1 = \lambda_1 \int_0^\infty \alpha x_1^{\alpha-1} e^{x_1^\alpha} e^{(\lambda_0+\lambda_1+\lambda_2)(1-e^{x_1^\alpha})}dx_1 = \frac{\lambda_1}{\lambda_0 + \lambda_1 + \lambda_2}.$$

Hence, $f_0(x)$ must satisfy

$$\int_0^\infty f_0(x)dx = \lambda_0 \int_0^\infty \alpha x^{\alpha-1} e^{x^\alpha} e^{(\lambda_0+\lambda_1+\lambda_2)(1-e^{x^\alpha})}dx = \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2}.$$

Therefore, the result follows. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

It should be mentioned that the BCH distribution has both an absolute continuous part and a singular part, similar to Marshall-Olkin bivariate exponential or bivariate Weibull model. In Theorem 2.2 the function $f_{X_1,X_2}(.,.)$ is considered to be the joint PDF of the BCH distribution if it is understood that the first two terms are densities with respect to two-dimensional Lebesgue measure and the third term is a density function with respect to one dimensional Lebesgue measure, see Bemis *et al.* (1972). It is well known that although in one dimension the practical use of a distribution with this property is usually pathological, but they do arise quite naturally in higher dimension. In case of BCH distribution, the presence of a singular part means that if X_1 and X_2 follow BCH distribution, then $X_1 = X_2$ has a positive probability. In many practical situations it may happen that X_1 and X_2 both are continuous random variables, but $X_1 = X_2$ has a positive probability, see Marshall and Olkin (1967). The following result will provide explicitly the absolute continuous part and the singular part of the BCH distribution function.

Theorem 2.3. *If* $(X_1, X_2) \sim BCH(\alpha, \lambda_0, \lambda_1, \lambda_2)$. *Then,*

$$\bar{F}_{X_1,X_2}(x_1, x_2) = \frac{\lambda_1 + \lambda_2}{\lambda_0 + \lambda_1 + \lambda_2} \bar{F}_a(x_1, x_2) + \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2} \bar{F}_s(x_1, x_2), \qquad (2.4)$$

where for $z = \max\{x_1, x_2\}$,

$$\bar{F}_s(x_1, x_2) = \bar{F}_{CH}(x_1, x_2) = e^{(\lambda_0+\lambda_1+\lambda_2)(1-e^{z^\alpha})},$$

and

$$\bar{F}_a(x_1, x_2) = \frac{\lambda_0 + \lambda_1 + \lambda_2}{\lambda_1 + \lambda_2} e^{\lambda_1(1-e^{x_1^\alpha})} e^{\lambda_2(1-e^{x_2^\alpha})} e^{\lambda_0(1-e^{z^\alpha})} - \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2} e^{(\lambda_0+\lambda_1+\lambda_2)(1-e^{z^\alpha})}.$$

Proof. Suppose A is the following event

$$A = \{U_0 < U_1\} \cap \{U_0 < U_2\}.$$

Then $P(A) = \lambda_0/(\lambda_0 + \lambda_1 + \lambda_2)$ and $P(A') = (\lambda_1 + \lambda_2)/(\lambda_0 + \lambda_1 + \lambda_2)$. Therefore,

$$\bar{F}_{X_1,X_2}(x_1, x_2) = P(X_1 \geq x_1, X_2 \geq x_2|A)P(A) + P(X_1 \geq x_1, X_2 \geq x_2|A')P(A').$$

Moreover, for z as defined before,

$$P(X_1 \geq x_1, X_2 \geq x_2|A) = [P(A)]^{-1} P(U_1 \geq U_0, U_2 \geq U_0, U_0 \geq z)$$
$$= e^{(\lambda_0+\lambda_1+\lambda_2)(1-e^{z^\alpha})},$$

and $P(X_1 \geq x_1, X_2 \geq x_2|A')$ can be obtained by subtraction. It is immediate that $e^{(\lambda_0+\lambda_1+\lambda_2)(1-e^{z^\alpha})}$ is the singular part as its mixed second partial derivatives is zero when $x_1 \neq x_2$, and $P(X_1 \geq x_1, X_2 \geq x_2|A')$ is the absolute continuous part as its mixed second partial derivatives is a bivariate density function. □

Corollary 2.2. The joint PDF of (X_1, X_2) can be written as follows for $z = \max\{x_1, x_2\}$:

$$f_{X_1,X_2}(x_1, x_2) = \frac{\lambda_1 + \lambda_2}{\lambda_0 + \lambda_1 + \lambda_2} f_a(x_1, x_2) + \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2} f_s(z), \qquad (2.5)$$

where

$$f_a(x_1, x_2) = \frac{\lambda_1 + \lambda_2}{\lambda_0 + \lambda_1 + \lambda_2} \begin{cases} f_{CH}(x_1, \alpha, \lambda_0 + \lambda_1) f_{CH}(x_2, \alpha, \lambda_2) & \text{if} \quad x_2 < x_1 \\ f_{CH}(x_2, \alpha, \lambda_1) f_{CH}(x_2, \alpha, \lambda_0 + \lambda_2) & \text{if} \quad x_1 < x_2, \end{cases}$$

and

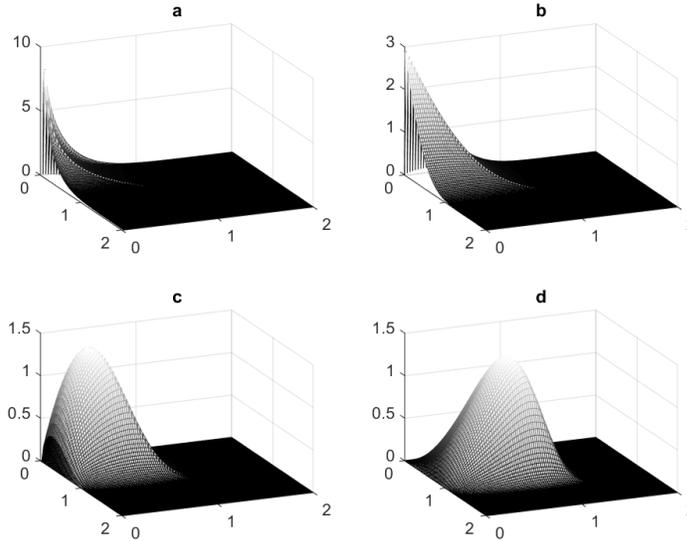$$f_s(z) = f_{CH}(z, \alpha, \lambda_0 + \lambda_1 + \lambda_2).$$

Figure 1: The shape of the joint probability density function of BCH for different values of parameters $\Theta = (\alpha, \lambda_0, \lambda_1, \lambda_2)$: (a) $\Theta = (0.75, 2, 1, 1)$ (b) $\Theta = (1, 2, 1, 1)$ (c) $\Theta = (1.5, 2, 1, 1)$ (d) $\Theta = (2, 0.7, 0.7, 0.7)$.

Clearly, here $f_a(x_1, x_2)$ and $f_s(z)$ are the absolutely continuous and singular part, respectively. Also, the surface plot of the joint probability density function of the BCH is drawn in Figure 1.

It is clear from Theorem 2.3 that for fixed λ_1 and λ_2, as $\lambda_0 \to 0$, $\bar{F}_{X_1, X_2}(x_1, x_2) \to e^{\lambda_1(1 - e^{x_1^\alpha})} e^{\lambda_2(1 - e^{x_2^\alpha})}$, i.e., X_1 and X_2 become independent. Moreover, since

$$A = (U_1 < U_0) \cap (U_2 < U_0) = \{\max\{U_1, U_2\} > U_0\} = \{X_1 = X_2\},$$

and $P(A) = \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2}$, therefore, as $\lambda_0 \to \infty$, $P(A) = P(X_1 = X_2) \to 1$, i.e. X_1 and X_2 are asymptotically almost surely equal. This implies that for fixed λ_1 and λ_2, as λ_0 varies from 0 to ∞, the correlation between X_1 and X_2 varies from 0 to 1.

Note that the absolutely continuous bivariate Chen (ACBCH) distribution can be obtained from the BCH distribution by removing the singular part and keeping only the continuous part. The joint PDF of ACBCH can be written as:

$$f_{ACBCH}(x_1, x_2) = \begin{cases} c f_{CH}(x_1, \alpha, \lambda_0 + \lambda_1) f_{CH}(x_2, \alpha, \lambda_2) & \text{if} \quad x_2 < x_1 \\ c f_{CH}(x_2, \alpha, \lambda_1) f_{CH}(x_2, \alpha, \lambda_0 + \lambda_2) & \text{if} \quad x_1 < x_2, \end{cases} \tag{2.6}$$

here c is the normalizing constantand and $c = \frac{\lambda_0 + \lambda_1 + \lambda_2}{\lambda_1 + \lambda_2}$.

2.3 Different Properties

In this subsection, we provide different basic properties of the BCH model. First, we provide the marginal and conditional distributions of the BCH model. An algorithm is presented to generate a random sample of BCH distribution. Also, the aging properties and the bivariate hazard gradient are discussed.

Proposition 2.1. *If* $(X_1, X_2) \sim BCH(\alpha, \lambda_0, \lambda_1, \lambda_2)$. *Then,*

1) $X_1 \sim CH(\alpha, \lambda_0 + \lambda_1)$ *and* $X_2 \sim CH(\alpha, \lambda_0 + \lambda_2)$.

2) $P(X_1 < X_2) = \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2}$.

3) $\min\{X_1, X_2\} \sim BCH(\alpha, \lambda_0 + \lambda_1 + \lambda_2)$.

Proof. They can be obtained by routine calculation. □

An algorithm to generate from BCH:

Step 1: Generate v_0, v_1 and v_2 from $U(0, 1)$,

Step 2: Compute

$$U_0 = [\ln(1 - \frac{\ln(1 - v_0)}{\lambda_0})]^{\frac{1}{\alpha}}, \qquad U_1 = [\ln(1 - \frac{\ln(1 - v_1)}{\lambda_1})]^{\frac{1}{\alpha}}, \qquad U_2 = [\ln(1 - \frac{\ln(1 - v_2)}{\lambda_2})]^{\frac{1}{\alpha}},$$

Step 3: Obtain
$$X_1 = \min\{U_0, U_1\}, \qquad \text{and} \qquad X_2 = \min\{U_0, U_2\}.$$

Since the marginal distributions of the bivariate vector (X_1, X_2) are the CH distributions, therefore, they can have a bathtub-shaped or increasing failure rate function.

Proposition 2.2. *If* $(X_1, X_2) \sim BCH(\alpha, \lambda_0, \lambda_1, \lambda_2)$. *Then,*

1) *The conditional survival function of* X_1 *given* $X_2 \geq x_2$, *say* $\bar{F}_{X_1|X_2 \geq x_2}(x_1)$ *is an absolute continuous survival function as follows:*

$$P(X_1 \geq x_1 | X_2 \geq x_2) = \bar{F}_{X_1|X_2 \geq x_2}(x_1) = \begin{cases} e^{(\lambda_0 + \lambda_1)(1 - e^{x_1^\alpha})} e^{-\lambda_0(1 - e^{x_2^\alpha})} & \text{if} \quad x_2 < x_1 \\ e^{\lambda_1(1 - e^{x_1^\alpha})} & \text{if} \quad x_1 < x_2. \end{cases}$$
(2.7)

2) *The conditional survival function in (2.7) has a representation*

$$\bar{F}_{X_1|X_2 \geq x_2}(x_1) = pG(x_1) + (1-p)H(x_1),$$

where,

$$G(x_1) = \frac{1}{p} \begin{cases} e^{(\lambda_0 + \lambda_1)(1 - e^{x_1^\alpha})} e^{-\lambda_0(1 - e^{x_2^\alpha})} & if \quad x_2 < x_1 \\ e^{\lambda_1(1 - e^{x_1^\alpha})} - \frac{\lambda_0}{\lambda_0 + \lambda_2} e^{\lambda_1(1 - e^{x_2^\alpha})} & if \quad x_1 < x_2, \end{cases}$$

$$H(x) = \begin{cases} 1 & if \quad x < x_2 \\ 0 & if \quad x > x_2, \end{cases}$$

and

$$p = 1 - \frac{\lambda_0}{\lambda_0 + \lambda_2} e^{\lambda_1(1 - e^{x_2^\alpha})}.$$

Proof. The proofs can be obtained in a routine manner, hence they are avoided. □

Moreover, it is interesting to observe the following properties of (X_1, X_2), if $(X_1, X_2) \sim BCH(\alpha, \lambda_0, \lambda_1, \lambda_2)$,

(a) If $(X_1, X_2) \sim BCH(\alpha, \lambda_0, \lambda_1, \lambda_2)$, then for $\alpha > 1$, it can be shown that

$$\frac{P(X_1 > x_1 + t, X_2 > x_2 + t)}{P(X_1 > x_1, X_2 > x_2)},$$

decreases in x_1 and x_2 for $t > 0$. Therefore, in this case (X_1, X_2) has the multivariate increasing failure rate (MIFR) property.

(b) Johnson and Kotz (1975) defined the bivariate hazard gradient as follows:

$$h_{X_1, X_2}(x_1, x_2) = (-\frac{\partial}{\partial x_1}, -\frac{\partial}{\partial x_2}) \ln P(X_1 > x_1, X_2 > x_2).$$

If $(X_1, X_2) \sim BCH(\alpha, \lambda_0, \lambda_1, \lambda_2)$, then for $\alpha \geq 1$ and all value of $x_1 > 0, x_2 > 0$, both the components of $h_{X_1, X_2}(x_1, x_2)$ are increasing functions of x_1 and x_2.

(c) A random vector (X_1, X_2) is said to be positive upper orthant dependent if for all $x_1 > 0$ and $x_2 > 0$,

$$P(X_1 \geq x_1, X_2 \geq x_2) \geq P(X_1 > x_1)P(X_2 > x_2), \tag{2.8}$$

(see Lehmann (1966)). If $(X_1, X_2) \sim BCH(\alpha, \lambda_0, \lambda_1, \lambda_2)$, then (X_1, X_2) satisfies (2.8). Therefore, (X_1, X_2) is positive upper orthant dependent.

(d) Since the joint survival function and the joint CDF have the following relation

$$\bar{F}_{X_1,X_2}(x_1, x_2) = 1 - F_{X_1}(x_1) - F_{X_2}(x_2) + F_{X_1,X_2}(x_1, x_2).$$

therefore, the joint survival function of X_1 and X_2 also can be expressed in a compact form. If $(X_1, X_2) \sim BCH(\alpha, \lambda_0, \lambda_1, \lambda_2)$, since $F_{X_1,X_2}(x_1, x_2) \geq F_{X_1}(x_1)F_{X_2}(x_2)$ for all x_1, x_2, therefore, they will be positive quadrant dependent, i.e., for every pair of increasing functions $h_1(.)$ and $h_2(.)$ (Barlow and Proschan , 1981) the following relation is satisfied:

$$Cov(h_1(X_1), h_2(X_2)) > 0.$$

(e) A random vector (X_1, X_2) is said to have the right tail increasing (RTI) property if for $i \neq j$,

$$P(X_i > x_i | X_j > x_j), \tag{2.9}$$

is a non-decreasing in x_j for all $x_i > 0$. If $(X_1, X_2) \sim BCH(\alpha, \lambda_0, \lambda_1, \lambda_2)$, then (X_1, X_2) satisfies 2.9. Therefore, (X_1, X_2) has the right tail increasing property.
Another bivariate dependence notion is the bivariate right corner set increasing (RCSI).

(f) A random vector (X_1, X_2) is said to have the right corner set increasing (RCSI) property if

$$P(X_1 > x_1, X_2 > x_2 | X_1 \geq \tilde{x}_1, X_2 \geq \tilde{x}_2), \tag{2.10}$$

increases in $\tilde{x}_1, \tilde{x}_2$ for every choice of (x_1, x_2). If $(X_1, X_2) \sim BCH(\alpha, \lambda_0, \lambda_1, \lambda_2)$, then, (X_1, X_2) satisfies (2.10). Therefore, (X_1, X_2) has the RCSI property.

3 Bivariate Chen-Geometric Distribution

In this section, we discuss the method of Marshall and Olkin (1997) in the bivariate case. Suppose $\{(X_{1n}, X_{2n}); n = 1, 2, \ldots\}$ is a sequence of i.i.d non-negative bivariate random variables with joint distribution function $F_\mathbf{X}(.,.)$, where $\mathbf{X} = (X_1, X_2)$ and N is a geometric random variable independent of $\{(X_{1i}, X_{2i})\}$. Also, a geometric random variable N with probability mass function

$$P(N = n) = \theta(1 - \theta)^{n-1}, \qquad n = 1, 2, 3, \ldots, \quad 0 < \theta < 1, \tag{3.1}$$

will be denoted by $G(\theta)$.

Consider the following bivariate random variable $\mathbf{Y} = (Y_1, Y_2)$

$$Y_i = \min\{X_{i1}, \ldots, X_{iN}\}, \qquad i = 1, 2. \tag{3.2}$$

For given $N = n$, the joint survival function of $\mathbf{Y} = (Y_1, Y_2)$ is:

$$\bar{F}_{Y_1, Y_2|N}(y_1, y_2|n) = (\bar{F}_{\mathbf{X}}(y_1, y_2))^n. \tag{3.3}$$

Therefore, the joint survival function of $\mathbf{Y} = (Y_1, Y_2)$ becomes:

$$P(Y_1 > y_1, Y_2 > y_2) = \sum_{n=1}^{\infty} P(Y_1 > y_1, Y_2 > y_2|N = n)P(N = n)$$

$$= \sum_{n=1}^{\infty} [\bar{F}_{\mathbf{X}}(y_1, y_2)]^n \theta(1-\theta)^{n-1} = \frac{\theta \bar{F}_{\mathbf{X}}(y_1, y_2)}{1 - (1-\theta)\bar{F}_{\mathbf{X}}(y_1, y_2)}. \tag{3.4}$$

In this case, we say $\mathbf{Y}$ has a bivariate F-geometric (BFG) distribution. Then, the survival function of Y_i is:

$$\bar{F}_{Y_i}(y_i) = \frac{\theta \bar{F}_{\mathbf{X}_i}(y_i)}{1 - (1-\theta)\bar{F}_{\mathbf{X}_i}(y_i)}, \qquad i = 1, 2, \tag{3.5}$$

where, $\bar{F}_{\mathbf{X}_i}, i = 1, 2$ are the marginal survival functions of $\bar{F}_{\mathbf{X}}$, i.e. $\bar{F}_{\mathbf{X}_1}(x) = \bar{F}_{\mathbf{X}}(x, 0)$, $\bar{F}_{\mathbf{X}_2}(x) = \bar{F}_{\mathbf{X}}(0, x)$. As mentioned, the structure is proposed by Marshall and Olkin (1997) in the univariate case. After that, their structure was used to add an extra parameter to model and build more flexible models by other authors. In this regard, Adamidis and Loukas (1998) introduced the two-parameter exponential-geometric distribution with decreasing failure rate. The Weibull-geometric distribution is obtained by Barreto-Souza *et al.* (2011). Louzada *et al.* (2012) proposed the exponentiated exponential-geometric. Also, the exponentiated Lomax-geometric distribution is discussed by Hassan and Abd-Allah (2017).

A random variable $\mathbf{Y} = (Y_1, Y_2)$ is said to have the bivariate Chen-geometric distribution with parameters $\theta, \alpha, \lambda_0, \lambda_1$ and λ_2, if the distribution F in (3.4) is $BCH(\alpha, \lambda_0, \lambda_1, \lambda_2)$. Therefore, the joint survival function of (Y_1, Y_2) becomes:

$$\bar{G}_{\mathbf{Y}}(y_1, y_2) = P(Y_1 > y_1, Y_2 > y_2)$$

$$= \begin{cases} \dfrac{\theta e^{(\lambda_0 + \lambda_1)(1 - e^{y_1^{\alpha}})} e^{\lambda_2(1 - e^{y_2^{\alpha}})}}{1 - (1-\theta)e^{(\lambda_0 + \lambda_1)(1 - e^{y_1^{\alpha}})} e^{\lambda_2(1 - e^{y_2^{\alpha}})}} & \text{if} \quad y_2 \leq y_1 \\[4ex] \dfrac{\theta e^{(\lambda_0 + \lambda_2)(1 - e^{y_2^{\alpha}})} e^{\lambda_1(1 - e^{y_1^{\alpha}})}}{1 - (1-\theta)e^{(\lambda_0 + \lambda_2)(1 - e^{y_2^{\alpha}})} e^{\lambda_1(1 - e^{y_1^{\alpha}})}} & \text{if} \quad y_1 < y_2. \end{cases} \tag{3.6}$$

It will be denoted by $(Y_1, Y_2) \sim BCHG(\theta, \alpha, \lambda_0, \lambda_1, \lambda_2)$.

Proposition 3.1. *If* $(Y_1, Y_2) \sim BCHG(\theta, \alpha, \lambda_0, \lambda_1, \lambda_2)$, *then,*

$$\bar{G}_{\mathbf{Y}}(y_1, y_2) = \sum_{n=1}^{\infty} p_n \bar{F}_{BCH}(y_1, y_2; \alpha, n\lambda_0, n\lambda_1, n\lambda_2), \tag{3.7}$$

where $p_n = P(N = n) = \theta(1 - \theta)^{n-1}$.

Theorem 3.1. *Let* $(Y_1, Y_2) \sim BCHG(\theta, \alpha, \lambda_0, \lambda_1, \lambda_2)$, *then the joint PDF of* $\mathbf{Y} = (Y_1, Y_1)$ *is:*

$$g_{\mathbf{Y}}(y_1, y_2) = \begin{cases} g_1(y_1, y_2) & if \quad y_2 < y_1 \\ g_2(y_1, y_2) & if \quad y_1 < y_2 \\ g_0(y) & if \quad 0 < y_1 = y_2 = y, \end{cases} \tag{3.8}$$

where

$$g_1(y_1, y_2) = \frac{\theta\alpha^2\lambda_2(\lambda_0 + \lambda_1)(y_1 y_2)^{\alpha-1} e^{(y_1^\alpha + y_2^\alpha)} e^{(\lambda_0+\lambda_1)(1-e^{y_1^\alpha})} e^{\lambda_2(1-e^{y_2^\alpha})}}{[1 - (1-\theta)e^{(\lambda_0+\lambda_1)(1-e^{y_1^\alpha})} e^{\lambda_2(1-e^{y_2^\alpha})}]^3}$$
$$\times [1 + (1-\theta)e^{(\lambda_0+\lambda_1)(1-e^{y_1^\alpha})} e^{\lambda_2(1-e^{y_2^\alpha})}].$$

$$g_2(y_1, y_2) = \frac{\theta\alpha^2\lambda_1(\lambda_0 + \lambda_2)(y_1 y_2)^{\alpha-1} e^{(y_1^\alpha + y_2^\alpha)} e^{\lambda_1(1-e^{y_1^\alpha})} e^{(\lambda_0+\lambda_2)(1-e^{y_2^\alpha})}}{[1 - (1-\theta)e^{\lambda_1(1-e^{y_1^\alpha})} e^{(\lambda_0+\lambda_2)(1-e^{y_2^\alpha})}]^3}$$
$$\times [1 + (1-\theta)e^{\lambda_1(1-e^{y_1^\alpha})} e^{(\lambda_0+\lambda_2)(1-e^{y_2^\alpha})}].$$

$$g_0(y) = \frac{\theta\alpha\lambda_0 y^\alpha e^{y^\alpha} e^{(\lambda_0+\lambda_1+\lambda_2)(1-e^{y^\alpha})}}{[1 - (1-\theta)e^{(\lambda_0+\lambda_1+\lambda_2)(1-e^{y^\alpha})}]^2}.$$

Proof. The expressions for $g_1(.,.)$ and $g_2(.,.)$ can be obtained simply by taking $-\frac{\partial^2 \bar{G}(y_1, y_2)}{\partial y_1 \partial y_2}$ for $y_2 < y_1$ and $y_1 < y_2$ respectively. But $g_0(.)$ cannot be obtained in the same way. Using the facts that

$$\int_0^\infty \int_{y_2}^\infty g_1(y_1, y_2) dy_1 dy_2 + \int_0^\infty \int_{y_1}^\infty g_2(y_1, y_2) dy_2 dy_1 + \int_0^\infty g_0(y) dy = 1.$$

So,

$$\int_0^\infty \int_{x_2}^\infty g_1(y_1, y_2) dy_1 dy_2 = \lambda_2 \int_0^\infty \frac{\alpha \theta e^{(\lambda_0+\lambda_1+\lambda_2)(1-e^{y_2^\alpha})} y_2^{\alpha-1} e^{y_2^\alpha}}{[1-(1-\theta)e^{(\lambda_0+\lambda_1+\lambda_2)(1-e^{y_2^\alpha})}]^2} dy_2 = \frac{\lambda_2}{\lambda_0+\lambda_1+\lambda_2}.$$

And similarly,

$$\int_0^\infty \int_{x_1}^\infty g_2(y_1, y_2) dy_2 dy_1 = \lambda_1 \int_0^\infty \frac{\alpha \theta e^{(\lambda_0+\lambda_1+\lambda_2)(1-e^{y_1^\alpha})} y_1^{\alpha-1} e^{y_1^\alpha}}{[1-(1-\theta)e^{(\lambda_0+\lambda_1+\lambda_2)(1-e^{y_1^\alpha})}]^2} dy_1 = \frac{\lambda_1}{\lambda_0+\lambda_1+\lambda_2}.$$

Note that

$$\int_0^\infty g_0(y) dy = \lambda_0 \int_0^\infty \frac{\theta \alpha y^\alpha e^{y^\alpha} e^{(\lambda_0+\lambda_1+\lambda_2)(1-e^{y^\alpha})}}{[1-(1-\theta)e^{(\lambda_0+\lambda_1+\lambda_2)(1-e^{y^\alpha})}]^2} dy = \frac{\lambda_0}{\lambda_0+\lambda_1+\lambda_2}.$$

Therefore, the result follows. □

Proposition 3.2. *If $(Y_1, Y_2) \sim BCHG(\theta, \alpha, \lambda_0, \lambda_1, \lambda_2)$, then,*

$$g_1(y_1, y_2) = \sum_{n=1}^\infty p_n f_{CH}(y_1, \alpha, n(\lambda_0+\lambda_1)) f_{CH}(y_2, \alpha, n\lambda_2),$$

$$g_2(y_1, y_2) = \sum_{n=1}^\infty p_n f_{CH}(y_1, \alpha, n\lambda_1) f_{CH}(y_2, \alpha, n(\lambda_0+\lambda_2)),$$

$$g_0(y) = \frac{\lambda_0}{\lambda_0+\lambda_1+\lambda_2} \sum_{n=1}^\infty p_n f_{CH}(y, \alpha, n(\lambda_0+\lambda_1+\lambda_2)).$$

where $p_n = P(N = n) = \theta(1-\theta)^{n-1}$ and $f_{CH}(., n\lambda)$ is the PDF of CH distribution with parameters α and $n\lambda$. Note that $f_{CH}(., n\lambda)$ is the PDF of the random variable $\min(U_1, \ldots, U_n)$ where U_i's are independent random variables from a CH distribution with parameters α and λ.

Corollary 3.1. The joint PDF of the BCHG distribution provided in Theorem 3.1 can be written as:

$$g(y_1, y_2) = \frac{\lambda_1+\lambda_2}{\lambda_0+\lambda_1+\lambda_2} g_a(y_1, y_2) + \frac{\lambda_0}{\lambda_0+\lambda_1+\lambda_2} g_s(y),$$

where

$$g_a(y_1, y_2) = \frac{\lambda_0+\lambda_1+\lambda_2}{\lambda_1+\lambda_2} \times \begin{cases} g_1(y_1, y_2) & \text{if} \quad y_2 < y_1 \\ g_2(y_1, y_2) & \text{if} \quad y_1 < y_2, \end{cases}$$

and

$$g_s(y) = \frac{\theta \alpha \lambda_0 y^\alpha e^{y^\alpha} e^{(\lambda_0+\lambda_1+\lambda_2)(1-e^{y^\alpha})}}{[1-(1-\theta)e^{(\lambda_0+\lambda_1+\lambda_2)(1-e^{y^\alpha})}]^2}, \qquad \text{if} \quad y_1 = y_2 = y,$$

and 0 otherwise. Clearly, $g_a(.,.)$ is the absolute continuous part and $g_s(.)$ is the singular part. If $\lambda_0 = 0$, it does not have any singular part, and it becomes an absolute continuous distribution. For $\theta = 1$, it is immediate that BCH can be obtained as a special case of BCHG.

The joint PDF of the BCHG distribution can take different shapes. We have provided the contour plots of the PDF of BCHG distribution in Figure 1 for different parameter values.
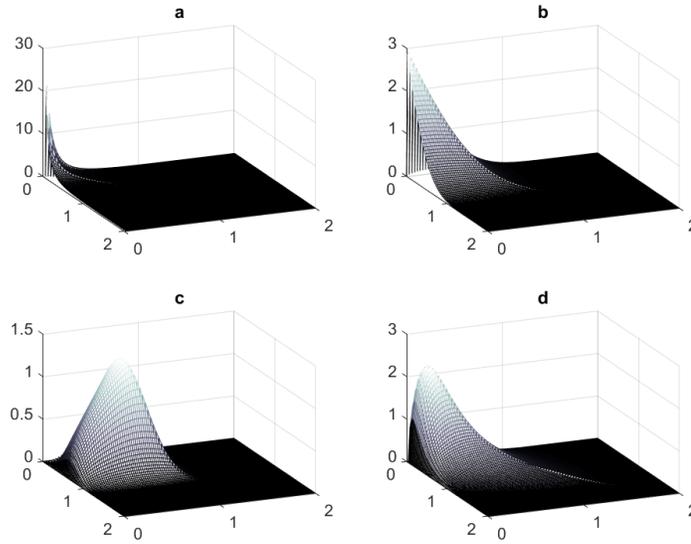
Figure 2: The shape of the joint probability density function of BCHG for different values of parameters $\Theta = (\theta, \alpha, \lambda_0, \lambda_1, \lambda_2)$: (a) $\Theta = (0.5, 0.75, 2, 1, 1)$ (b) $\Theta = (1, 1, 2, 1, 1)$ (c) $\Theta = (2, 1.5, 2, 1, 1)$ (d) $\Theta = (0.2, 1.5, 0.4, 0.4, 0.4)$.

Now, the joint PDF of (Y_1, Y_2, N) is obtained. Based on equation (3.3), we know that:

$$P(Y_1 > y_1, Y_2 > y_2, N = n) = P(Y_1 > y_1, Y_2 > y_2 | N = n)P(N = n)$$

$$= \begin{cases} \theta(1-\theta)^{n-1} e^{-n(\lambda_0+\lambda_1)(1-e^{y_1^\alpha})} e^{-n\lambda_2(1-e^{y_2^\alpha})} & \text{if} \quad y_2 \leq y_1 \\ \theta(1-\theta)^{n-1} e^{-n\lambda_1(1-e^{y_1^\alpha})} e^{-n(\lambda_0+\lambda_2)(1-e^{y_2^\alpha})} & \text{if} \quad y_1 < y_2. \end{cases}$$

So, the joint the joint PDF of Y_1, Y_2 and N is

$$f_{Y_1,Y_2,N}(y_1, y_2, n) = \begin{cases} \theta(1-\theta)^{n-1} f_{1n}(y_1, y_2) & \text{if} \quad y_2 < y_1 \\ \theta(1-\theta)^{n-1} f_{2n}(y_1, y_2) & \text{if} \quad y_1 < y_2 \\ \theta(1-\theta)^{n-1} f_{0n}(y) & \text{if} \quad y_1 = y_2 = y, \end{cases} \tag{3.9}$$

where,

$$f_{1n}(y_1, y_2) = n^2 \alpha^2 \lambda_2 (\lambda_0 + \lambda_1)(y_1 y_2)^{\alpha-1} e^{y_1^\alpha + y_2^\alpha} e^{n(\lambda_0 + \lambda_1)(1 - e^{y_1^\alpha})} e^{n\lambda_2(1 - e^{y_2^\alpha})}.$$

$$f_{2n}(y_1, y_2) = n^2 \alpha^2 \lambda_1 (\lambda_0 + \lambda_2)(y_1 y_2)^{\alpha-1} e^{y_1^\alpha + y_2^\alpha} e^{n(\lambda_0 + \lambda_2)(1 - e^{y_2^\alpha})} e^{n\lambda_1(1 - e^{y_1^\alpha})}.$$

$$f_{0n}(y) = n\alpha\lambda_0 y^{\alpha-1} e^{y^\alpha} e^{n(\lambda_0 + \lambda_1 + \lambda_2)(1 - e^{y^\alpha})}. \tag{3.10}$$

Also, the conditional probability mass function of N given $Y_1 = y_1$ and $Y_2 = y_2$ can be calculated as follows:

$$f_N(n|y_1, y_2) = \begin{cases} c_1(y_1, y_2) n^2 (1-\theta)^{n-1} e^{(n-1)(\lambda_0 + \lambda_1)(1 - e^{y_1^\alpha})} e^{(n-1)\lambda_2(1 - e^{y_2^\alpha})} & \text{if} \quad y_2 < y_1 \\ \\ c_2(y_1, y_2) n^2 (1-\theta)^{n-1} e^{(n-1)\lambda_1(1 - e^{y_1^\alpha})} e^{(n-1)(\lambda_0 + \lambda_2)(1 - e^{y_2^\alpha})} & \text{if} \quad y_1 < y_2 \\ \\ c_0(y) n (1-\theta)^{n-1} e^{(n-1)(\lambda_0 + \lambda_1 + \lambda_2)(1 - e^{y^\alpha})} & \text{if} \quad y_1 = y_2 = y, \end{cases}$$
$$\tag{3.11}$$

where,

$$c_1(y_1, y_2) = \frac{[1 - (1-\theta)e^{(\lambda_0 + \lambda_1)(1 - e^{y_1^\alpha})} e^{\lambda_2(1 - e^{y_2^\alpha})}]^3}{[1 + (1-\theta)e^{(\lambda_0 + \lambda_1)(1 - e^{y_1^\alpha})} e^{\lambda_2(1 - e^{y_2^\alpha})}]},$$

$$c_2(y_1, y_2) = \frac{[1 - (1-\theta)e^{(\lambda_0 + \lambda_2)(1 - e^{y_2^\alpha})} e^{\lambda_1(1 - e^{y_1^\alpha})}]^3}{[1 + (1-\theta)e^{(\lambda_0 + \lambda_2)(1 - e^{y_2^\alpha})} e^{\lambda_1(1 - e^{y_1^\alpha})}]},$$

$$c_0(y) = [1 - (1-\theta)e^{(\lambda_0 + \lambda_1 + \lambda_2)(1 - e^{y^\alpha})}]^2.$$

We will be using the following equations for (3.11)

$$
f_N(n|y_1, y_2) = \begin{cases} \frac{[1-\xi_1(y_1,y_2,\theta,\gamma)]^3}{[1+\xi_1(y_1,y_2,\theta,\gamma)]} n^2 \xi_1^{n-1}(y_1, y_2, \theta, \gamma) & \text{if} \quad y_2 < y_1 \\[2ex] \frac{[1-\xi_1(y_1,y_2,\theta,\gamma)]^3}{[1+\xi_2(y_1,y_2,\theta,\gamma)]} n^2 \xi_2^{n-1}(y_1, y_2, \theta, \gamma) & \text{if} \quad y_1 < y_2 \\[2ex] [1 - \xi_0(y_1, y_2, \theta, \gamma)]^2 n \xi_0^{n-1}(y_1, y_2, \theta, \gamma) & \text{if} \quad y_1 = y_2 = y, \end{cases}
$$

where $\gamma = (\alpha, \lambda_0, \lambda_1, \lambda_2)$ and

$$
\xi_1(y_1, y_2, \theta, \gamma) = (1 - \theta)e^{(\lambda_0+\lambda_1)(1-e^{y_1^\alpha})}e^{\lambda_2(1-e^{y_2^\alpha})},
$$

$$
\xi_2(y_1, y_2, \theta, \gamma) = (1 - \theta)e^{\lambda_1(1-e^{y_1^\alpha})}e^{(\lambda_0+\lambda_2)(1-e^{y_2^\alpha})},
$$

$$
\xi_0(y_1, y_2, \theta, \gamma) = (1 - \theta)e^{(\lambda_0+\lambda_1+\lambda_2)(1-e^{y_2^\alpha})}.
$$

Therefore, we can obtain the conditional expectation of N given $Y_1 = y_1$ and $Y_2 = y_2$ as:

$$
E(N|y_1, y_2) = \begin{cases} \frac{(1-\xi_1(y_1,y_2,\theta,\gamma))^2 - 6(1-\xi_1(y_1,y_2,\theta,\gamma))+6}{(1-\xi_1(y_1,y_2,\theta,\gamma))^2} & \text{if} \quad y_2 < y_1 \\[2ex] \frac{(1-\xi_2(y_1,y_2,\theta,\gamma))^2 - 6(1-\xi_2(y_1,y_2,\theta,\gamma))+6}{(1-\xi_2(y_1,y_2,\theta,\gamma))^2} & \text{if} \quad y_1 < y_2 \\[2ex] \frac{1+\xi_0(y_1,y_2,\theta,\gamma)}{1-\xi_0(y_1,y_2,\theta,\gamma)} & \text{if} \quad y_1 = y_2 = y. \end{cases}
$$

Proposition 3.3. *Let $(y_1, y_2) \sim BCHG(\theta, \alpha, \lambda_0, \lambda_1, \lambda_2)$. Then*

1. *Each Y_i has a univariate Chen-geometric distribution (UCHG) with parameters α, $\lambda_0 + \lambda_i$ and θ.*

2. *The random variable $Y = \min(Y_1, Y_2)$ has an UCHG distribution with parameters $\lambda_0 + \lambda_1 + \lambda_2$, α and θ.*

3. $P(Y_1 < Y_2) = \frac{\lambda_1}{\lambda_0+\lambda_1+\lambda_2}.$

Proof. The proofs of the parts 1 and 2 are similar. The proof of 1 is immediate from (3.5) with $\bar{F}_{X_i}$ having the Chen distribution. Also, we have

$$
\bar{F}_{X_1}(y_1) = P(\min\{U_0, U_1\} > y_1) = P(U_0 > y_1, U_1 > y_1)
$$
$$
= \bar{F}_{CH}(y_1, \alpha, \lambda_0)\bar{F}_{CH}(y_1, \alpha, \lambda_1) = \bar{F}_{CH}(y_1, \alpha, \lambda_0 + \lambda_1).
$$

The proof of 2 is obtained by (3.5). Finally, the proof of 3 can be obtained as follows:

$$
\begin{aligned}
P(Y_1 < Y_2) &= \sum_{n=1}^{\infty} P(Y_1 < Y_2, N = n) \\
&= \sum_{n=1}^{\infty} \int_0^{\infty} \int_{y_1}^{\infty} \theta(1-\theta)^{n-1} f_{2n}(y_1, y_2) dy_2 dy_1 \\
&= \sum_{n=1}^{\infty} \theta(1-\theta)^{n-1} \int_0^{\infty} \int_{y_1}^{\infty} f_{2n}(y_1, y_2) dy_2 dy_1 \\
&= \theta \sum_{n=1}^{\infty} (1-\theta)^{n-1} \frac{\lambda_1}{\lambda_0 + \lambda_1 + \lambda_2} = \frac{\lambda_1}{\lambda_0 + \lambda_1 + \lambda_2}.
\end{aligned}
$$

$\square$

4   Estimation

In this section, we pay attention to the maximum likelihood estimators of the unknown parameters of the BCHG distribution using the EM algorithm.

4.1   EM Algorithm

Suppose $\{(y_{11}, y_{21}), \ldots, (y_{1m}, y_{2m})\}$ is a random sample from BCHG with parameters $\Theta = (\theta, \alpha, \lambda_0, \lambda_1, \lambda_2)$, and our aim is to compute the MLE's of the unknown parameters. Let us define the following notation,
$I_0 = \{i : y_{1i} = y_{2i} = y_i\}$, $I_1 = \{i : y_{1i} > y_{2i}\}$ and $I_2 = \{i : y_{1i} < y_{2i}\}$. Also, $|I_0| = m_0$, $|I_1| = m_1$, $|I_2| = m_2$ and $m = m_0 + m_1 + m_2$. Therefore, the log-likelihood function can be written as:

$$
\ell(\Theta) = \sum_{i \in I_0} \ln g_0(y_i) + \sum_{i \in I_1} \ln g_1(y_{1i}, y_{2i}) + \sum_{i \in I_2} \ln g_2(y_{1i}, y_{2i}), \tag{4.1}
$$

where g_0, g_1 and g_2 are defined in Theorem 3.1. We can obtain the MLEs of the parameters by maximizing $\ell(\Theta)$ in (4.1) with respect to the unknown parameters. This is clearly a five-dimensional problem. They can be obtained only by solving five equations with five unknowns. We propose to use the EM algorithm, which is computationally more tractable. We treat this as a missing value problem.

For given n, consider the independent random variables

$$
\{U_i | N = n\} \sim CH(\alpha, n\lambda_i), \qquad i = 0, 1, 2. \tag{4.2}
$$

Also, It is clear that

$$\{Y_i|N = n\} = \min\{U_0, U_i\}|N = n, \qquad i = 1, 2.$$

Assume that for the bivariate random vector (Y_1, Y_2), there is an associated random vector

$$(\Delta_1, \Delta_2) = \begin{cases} (0,0) & \text{if} \quad Y_1 = U_0, Y_2 = U_0 \\ (0,1) & \text{if} \quad Y_1 = U_0, Y_2 = U_2 \\ (1,0) & \text{if} \quad Y_1 = U_1, Y_2 = U_0 \\ (1,1) & \text{if} \quad Y_1 = U_1, Y_2 = U_2 . \end{cases} \tag{4.3}$$

Here Y_i's are same as defined above. Therefore, a sample is obtained from $(Y_1, Y_2, \Delta_1, \Delta_2 , N)$ which is the complete observation. Note that if we have a random sample of size m from $(Y_1, Y_2, \Delta_1, \Delta_2, N)$, then the MLEs of the unknown parameters can be computed as a one-dimensional optimization problem.

Note that if $Y_1 = Y_2$, then $\Delta_1 = \Delta_2 = 0$, is known. But if $Y_1 < Y_2$ or $Y_1 > Y_2$, then (Δ_1, Δ_2) is missing. If $(Y_1, Y_2) \in I_1$ then the possible values of (Δ_1, Δ_2) are (0, 1) or (1, 1) and if $(Y_1, Y_2) \in I_2$ then the possible values of (Δ_1, Δ_2) are (1, 0) or (1, 1) with non-zero probabilities, see for example Kundu and Dey (2009). We compute the pseudo log-likelihood function for the performance of the EM algorithm. For this purpose, we use the proposed structure of Dinse (1982) or Kundu (2004).

In this method, the conditional pseudo log-likelihood function is formed by conditioning on N and then N is replaced by $E(N|Y_1, Y_2)$.

In the *E*-step, we kept the log-likelihood contribution of all the observations belonging to I_0 intact, as in this case the corresponding (Δ_1, Δ_2) are known completely. Also, the observations are treated as missing observations, if the observations belong to I_1 or I_2.

If $(y_1, y_2) \in I_1$, the pseudo observations is formed by fractioning (y_1, y_2) to two partially complete pseudo observation of the form $(y_1, y_2, u_1(\Theta))$ and $(y_1, y_2, u_2(\Theta))$. The fractional mass $u_1(\Theta)$ and $u_2(\Theta)$ assigned to the pseudo observations are the conditional probabilities that (Δ_1, Δ_2) take values (0, 1) or (1, 1), respectively, given that $(Y_1, Y_2) \in I_1$. Similarly, if $(Y_1, Y_2) \in I_2$, pseudo observations are formed as $(y_1, y_2, v_1(\Theta))$ and $(y_1, y_2, v_2(\Theta))$, where $v_1(\Theta)$ and $v_2(\Theta)$ are the conditional probabilities that (Δ_1, Δ_2) take values (1, 0) and (1, 1) respectively. The observations in Table 1 will be used for constructing *E*-step of the EM algorithm. Therefore,

$$v_1(\Theta) = \frac{\lambda_0}{\lambda_0 + \lambda_2}, \qquad v_2(\Theta) = \frac{\lambda_2}{\lambda_0 + \lambda_2}, \qquad u_1(\Theta) = \frac{\lambda_0}{\lambda_0 + \lambda_1}, \qquad u_2(\Theta) = \frac{\lambda_1}{\lambda_0 + \lambda_1}.$$

Table 1: All possible cases of U_0, U_1, U_2, corresponding probabilities and (Δ_1, Δ_2).

Different Case	Probability	(Δ_1, Δ_2)	$Y_1 \& Y_2$	Set
$U_0 < U_1 < U_2$	$\frac{\lambda_1 \lambda_0}{(\lambda_1+\lambda_2)(\lambda_0+\lambda_1+\lambda_2)}$	$(0, 0)$	$Y_1 = Y_2$	I_0
$U_0 < U_2 < U_1$	$\frac{\lambda_2 \lambda_0}{(\lambda_1+\lambda_2)(\lambda_0+\lambda_1+\lambda_2)}$	$(0, 0)$	$Y_1 = Y_2$	I_0
$U_1 < U_0 < U_2$	$\frac{\lambda_1 \lambda_0}{(\lambda_0+\lambda_2)(\lambda_0+\lambda_1+\lambda_2)}$	$(1, 0)$	$Y_1 < Y_2$	I_2
$U_1 < U_2 < U_0$	$\frac{\lambda_1 \lambda_2}{(\lambda_0+\lambda_2)(\lambda_0+\lambda_1+\lambda_2)}$	$(1, 1)$	$Y_1 < Y_2$	I_2
$U_2 < U_0 < U_1$	$\frac{\lambda_0 \lambda_2}{(\lambda_0+\lambda_1)(\lambda_0+\lambda_1+\lambda_2)}$	$(0, 1)$	$Y_2 < Y_1$	I_1
$U_2 < U_1 < U_0$	$\frac{\lambda_1 \lambda_2}{(\lambda_0+\lambda_1)(\lambda_0+\lambda_1+\lambda_2)}$	$(1, 1)$	$Y_2 < Y_1$	I_1

In the following, we consider the pseudo log-likelihood contribution of an observation $y \in I_0$, with conditioning on $N = n$. Then, the log-likelihood contribution becomes:

$$\ln n + \ln \alpha + \ln \lambda_0 + n(\lambda_0 + \lambda_1 + \lambda_2)(1 - e^{y^\alpha}) + (\alpha - 1) \ln y + y^\alpha.$$

The above equation follows from (4.2), and in this case, $U_0 = y$, $U_1 > y$ and $U_2 > y$.

Also, when $(y_1, y_2) \in I_1$ and $(y_1, y_2) \in I_2$, the pseudo log-likelihood contribution conditioning on $N = n$, becomes:

$$2 \ln n + 2 \ln \alpha + \ln \lambda_2 + u_1 \ln \lambda_0 + u_2 \ln \lambda_1 + (\alpha - 1)[\ln y_1 + \ln y_2]$$
$$+ y_1^\alpha + y_2^\alpha + n\lambda_2(1 - e^{y_2^\alpha}) + n(\lambda_0 + \lambda_1)(1 - e^{y_1^\alpha}),$$

and

$$2 \ln n + 2 \ln \alpha + \ln \lambda_1 + v_1 \ln \lambda_0 + v_2 \ln \lambda_2 + (\alpha - 1)[\ln y_1 + \ln y_2]$$
$$+ y_1^\alpha + y_2^\alpha + n\lambda_1(1 - e^{y_1^\alpha}) + n(\lambda_0 + \lambda_2)(1 - e^{y_2^\alpha}),$$

respectively. We know that $\ln n$ and n are missing. So, we will replace them with $E(\ln N | y_1, y_2)$ and $E(N | y_1, y_2)$, respectively, to compute the pseudo log-likelihood function at the E-step of the EM algorithm. Also, we will use the following notations in the k-th step of the EM algorithm.

- $\Theta^{(k)} = (\alpha^{(k)}, \lambda_0^{(k)}, \lambda_1^{(k)}, \lambda_2^{(k)})$ is defined for the estimates of the parameters in the k-th step.

- $E(N|y_{1i}, y_{2i}, \Theta) = a_i.$

- $E(N|y_{1i}, y_{2i}, \Theta^{(k)}) = a_i^{(k)}.$

- $u_1(\Theta^{(k)}) = u_1^{(k)}, \ \ u_2(\Theta^{(k)}) = u_2^{(k)}, \ \ v_1(\Theta^{(k)}) = u_1^{(k)} \ \ \text{and} \ \ v_2(\Theta^{(k)}) = u_2^{(k)}.$

Now, we are ready to provide the EM algorithm.

E-Step: At the *E*-step of the EM algorithm, the pseudo log-likelihood function without the additive constant can be written as follows:

$$
\begin{aligned}
\ell_{pseudo}(\Theta) &= (m_0 + 2m_1 + 2m_2)\ln \lambda_0 + (m_2 + m_1 u_2^{(k)})\ln \lambda_1 + (m_2 v_2^{(k)} + m_1)\ln \lambda_2 \\
&+ \lambda_0\{\sum_{i\in I_0} a_i^{(k)}(1 - e^{y_i^{\alpha}}) + \sum_{i\in I_2} a_i^{(k)}(1 - e^{y_{2i}^{\alpha}}) + \sum_{i\in I_1} a_i^{(k)}(1 - e^{y_{1i}^{\alpha}})\} \\
&+ \lambda_1\{\sum_{i\in I_0} a_i^{(k)}(1 - e^{y_i^{\alpha}}) + \sum_{i\in I_2} a_i^{(k)}(1 - e^{y_{1i}^{\alpha}}) + \sum_{i\in I_1} a_i^{(k)}(1 - e^{y_{1i}^{\alpha}})\} \\
&+ \lambda_2\{\sum_{i\in I_0} a_i^{(k)}(1 - e^{y_i^{\alpha}}) + \sum_{i\in I_2} a_i^{(k)}(1 - e^{y_{2i}^{\alpha}}) + \sum_{i\in I_1} a_i^{(k)}(1 - e^{y_{2i}^{\alpha}})\} \\
&+ \sum_{i\in I_0} y_i^{\alpha} + \sum_{i\in I_2} y_{1i}^{\alpha} + \sum_{i\in I_2} y_{2i}^{\alpha} + \sum_{i\in I_1} y_{2i}^{\alpha} + \sum_{i\in I_1} y_{1i}^{\alpha} \\
&+ (\alpha - 1)\{\sum_{i\in I_0} \ln y_i + \sum_{i\in I_2} \ln y_{1i} + \sum_{i\in I_2} \ln y_{2i} + \sum_{i\in I_1} \ln y_{2i} + \sum_{i\in I_1} \ln y_{2i}\} \\
&+ (m_0 + 2m_1 + 2m_2)\ln \alpha + m \ln \frac{\theta}{1 - \theta} + \ln(1 - \theta)\sum_{i=1}^{m} a_i^{(k)}. \quad (4.4)
\end{aligned}
$$

M-Step: The *M*-step involves maximizing $\ell_{pseudo}(\Theta)$ with respect to the unknown parameters. Note that for fixed α, the maximization with respect to the unknown parameters can be obtained as follows:

$$
\widehat{\lambda}_0(\alpha) = -\frac{m_0 + m_2 v_1^{(k)} + m_1 u_1^{(k)}}{\sum_{i\in I_0} a_i^{(k)}(1 - e^{y_i^{\alpha}}) + \sum_{i\in I_2} a_i^{(k)}(1 - e^{y_{2i}^{\alpha}}) + \sum_{i\in I_1} a_i^{(k)}(1 - e^{y_{1i}^{\alpha}})}. \quad (4.5)
$$

$$
\widehat{\lambda}_1(\alpha) = -\frac{m_2 + m_1 u_2^{(k)}}{\sum_{i\in I_0} a_i^{(k)}(1 - e^{\alpha y_i^{\alpha}}) + \sum_{i\in I_2} a_i^{(k)}(1 - e^{y_{1i}^{\alpha}}) + \sum_{i\in I_1} a_i^{(k)}(1 - e^{y_{1i}^{\alpha}})}. \quad (4.6)
$$

$$\widehat{\lambda}_2(\alpha) = -\frac{m_1 + m_2 v_2^{(k)}}{\sum_{i\in I_0} a_i^{(k)}(1 - e^{\alpha y_i^\alpha}) + \sum_{i\in I_2} a_i^{(k)}(1 - e^{y_{2i}^\alpha}) + \sum_{i\in I_1} a_i^{(k)}(1 - e^{y_{2i}^\alpha})}. \tag{4.7}$$

and $\hat{\theta}$ is

$$\hat{\theta} = \frac{m}{\sum_{i=1}^m a_i^{(k)}}. \tag{4.8}$$

The pseudo-maximum likelihood estimate of α at each step can be derived by maximizing the pseudo-profile log-likelihood function $\ell_{pseudo}(\theta, \alpha, \widehat{\lambda}_0(\alpha), \widehat{\lambda}_1(\alpha), \widehat{\lambda}_2(\alpha))$ or by solving the following non-linear equation

$$g(\alpha) = \alpha, \tag{4.9}$$

where $g(\alpha) = \frac{m_0 + 2m_1 + 2m_2}{h(\alpha)}$ and

$$
\begin{aligned}
h(\alpha) = {}& \hat{\lambda}_0(\alpha)\{\sum_{i\in I_0} a_i^{(k)} y_i^\alpha \ln y_i e^{y_i^\alpha} + \sum_{i\in I_2} a_i^{(k)} y_{2i}^\alpha \ln y_{2i} e^{y_{2i}^\alpha} + \sum_{i\in I_1} a_i^{(k)} y_{1i}^\alpha \ln y_{1i} e^{y_{1i}^\alpha}\} \\
& + \hat{\lambda}_1(\alpha)\{\sum_{i\in I_0} a_i^{(k)} y_i^\alpha \ln y_i e^{y_i^\alpha} + \sum_{i\in I_2} a_i^{(k)} y_{1i}^\alpha \ln y_{1i} e^{y_{1i}^\alpha} + \sum_{i\in I_1} a_i^{(k)} y_{1i}^\alpha \ln y_{1i} e^{y_{1i}^\alpha}\} \\
& + \hat{\lambda}_2(\alpha)\{\sum_{i\in I_0} a_i^{(k)} y_i^\alpha \ln y_i e^{y_i^\alpha} + \sum_{i\in I_2} a_i^{(k)} y_{2i}^\alpha \ln y_{2i} e^{y_{2i}^\alpha} + \sum_{i\in I_1} a_i^{(k)} y_{2i}^\alpha \ln y_{2i} e^{y_{2i}^\alpha}\} \\
& - \{\sum_{i\in I_0} \ln y_i (1 + y_i^\alpha) + \sum_{i\in I_2} \ln y_{1i}(1 + y_{1i}^\alpha) + \sum_{i\in I_2} \ln y_{2i}(1 + y_{2i}^\alpha) \\
& + \sum_{i\in I_1} \ln y_{2i}(1 + y_{2i}^\alpha) + \sum_{i\in I_1} \ln y_{1i}(1 + y_{1i}^\alpha)\}.
\end{aligned}
$$

Therefore, a simple iterative procedure can be used to find the solution of equation (4.9), such as the Newton - Raphson method or using the method by Kundu and Dey (2009) and Kundu and Gupta (2006). Now, the following steps can be used to compute the MLEs of the parameters via the EM algorithm:

ALGORITHM

- Step 1: Take some initial value of Θ, say $\Theta^{(0)} = (\theta^{(0)}, \alpha^{(0)}, \lambda_0^{(0)}, \lambda_1^{(0)}, \lambda_2^{(0)})$.

- Step 2: Compute $a_i^{(0)} = E(N|y_{1i}, y_{2i}; \Theta^{(0)})$.

- Step 3: Compute u_1, u_2, v_1, and v_2.

- Step 4: Find $\hat{\alpha}$ by solving the equation (4.9), say $\hat{\alpha}^{(1)}$.

- Step 5: Compute $\hat{\lambda}_i^{(1)} = \hat{\lambda}_i(\hat{\alpha}^{(1)})$, $i = 0, 1, 2$ from (4.5)-(4.7).

- Step 6: Find $\hat{\theta}$ from (4.8).

- Step 7: Replace $\Theta^{(0)}$ by $\Theta^{(1)} = (\theta^{(1)}, \alpha^{(1)}, \lambda_0^{(1)}, \lambda_1^{(1)}, \lambda_2^{(1)})$, go back to step 1 and continue the process until convergence takes place. In fact, the iteration stops when the relative difference between two iterates is less than 10^{-6}.

4.2 Confidance Interval

Since the MLEs of the unknown parameters are not obtained in closed forms, then it is not possible to derive the exact distributions of the maximum likelihood estimators. In this section, we derive asymptotic confidence intervals of these parameters.

Large-sample inference for the parameters is based on the MLEs and their estimated standard errors. Since the expected information matrix does not have a closed form, we consider the observed information matrix for the interval estimation. Therefore, we present the observed Fisher Information Matrix obtained from the EM algorithm using the idea of Louis (1982). Using this method, this Matrix can be obtained as

$$I_{obs} = B - SS^T,$$

where the matrix B denotes the Hessian matrix and the vector S denotes the gradient vector of the pseudo-log-likelihood function. We provide the elements of the matrix B and S as follows:

$$
\begin{aligned}
B_{11} = {}& \frac{m_0 + 2m_1 + 2m_2}{\alpha^2} \\
& + \lambda_0 \{ \sum_{I_0} n(\ln y_i)^2 y_i^\alpha e^{y_i^\alpha}(1 + y_i^\alpha) + \sum_{I_2} n(\ln y_{2i})^2 y_{2i}^\alpha e^{y_{2i}^\alpha}(1 + y_{2i}^\alpha) + \sum_{I_1} n(\ln y_{1i})^2 y_{1i}^\alpha e^{y_{1i}^\alpha}(1 + y_{1i}^\alpha) \} \\
& + \lambda_1 \{ \sum_{I_0} n(\ln y_i)^2 y_i^\alpha e^{y_i^\alpha}(1 + y_i^\alpha) + \sum_{I_2} n(\ln y_{1i})^2 y_{1i}^\alpha e^{y_{1i}^\alpha}(1 + y_{1i}^\alpha) + \sum_{I_1} n(\ln y_{1i})^2 y_{1i}^\alpha e^{y_{1i}^\alpha}(1 + y_{1i}^\alpha) \} \\
& + \lambda_2 \{ \sum_{I_0} n(\ln y_i)^2 y_i^\alpha e^{y_i^\alpha}(1 + y_i^\alpha) + \sum_{I_2} n(\ln y_{2i})^2 y_{2i}^\alpha e^{y_{2i}^\alpha}(1 + y_{2i}^\alpha) + \sum_{I_1} n(\ln y_{2i})^2 y_{2i}^\alpha e^{y_{2i}^\alpha}(1 + y_{2i}^\alpha) \}.
\end{aligned}
$$

$$B_{12} = B_{21} = \sum_{I_0} ny_i^\alpha \ln y_i e^{y_i^\alpha} + \sum_{I_2} ny_{2i}^\alpha \ln y_{2i} e^{y_{2i}^\alpha} + \sum_{I_1} ny_{1i}^\alpha \ln y_{1i} e^{y_{1i}^\alpha}.$$

$$B_{13} = B_{31} = \sum_{I_0} ny_i^\alpha \ln y_i e^{y_i^\alpha} + \sum_{I_2} ny_{1i}^\alpha \ln y_{1i} e^{y_{1i}^\alpha} + \sum_{I_1} ny_{1i}^\alpha \ln y_{1i} e^{y_{1i}^\alpha}.$$

$$B_{14} = B_{41} = \sum_{I_0} ny_i^\alpha \ln y_i e^{y_i^\alpha} + \sum_{I_2} ny_{2i}^\alpha \ln y_{2i} e^{y_{2i}^\alpha} + \sum_{I_1} ny_{2i}^\alpha \ln y_{2i} e^{y_{2i}^\alpha}.$$

$$B_{22} = \frac{m_0 + m_2 v_1 + m_1 v_1}{\lambda_0^2}, \qquad B_{33} = \frac{m_2 + m_1 v_2}{\lambda_1^2}, \qquad B_{44} = \frac{m_2 v_2 + m_1}{\lambda_2^2}.$$

$$B_{15} = B_{51} = B_{23} = B_{32} = B_{24} = B_{42} = B_{25} = B_{52} = B_{34} = B_{43} = B_{35} = B_{53} = B_{45} = B_{54} = 0.$$

$$S_1 = \frac{m_0 + 2m_1 + 2m_2}{\alpha} - \lambda_0 \{ \sum_{I_0} ny_i^\alpha \ln y_i e^{y_i^\alpha} + \sum_{I_2} ny_{2i}^\alpha \ln y_{2i} e^{y_{2i}^\alpha} + \sum_{I_1} ny_{1i}^\alpha \ln y_{1i} e^{y_{1i}^\alpha} \}$$

$$- \lambda_1 \{ \sum_{I_0} ny_i^\alpha \ln y_i e^{y_i^\alpha} + \sum_{I_2} ny_{1i}^\alpha \ln y_{1i} e^{y_{1i}^\alpha} + \sum_{I_1} ny_{1i}^\alpha \ln y_{1i} e^{y_{1i}^\alpha} \}$$

$$- \lambda_2 \{ \sum_{I_0} ny_i^\alpha \ln y_i e^{y_i^\alpha} + \sum_{I_2} ny_{2i}^\alpha \ln y_{2i} e^{y_{2i}^\alpha} + \sum_{I_1} ny_{2i}^\alpha \ln y_{2i} e^{y_{2i}^\alpha} \}$$

$$+ \sum_{I_0} \ln y_i (1 + y_i^\alpha) + \sum_{I_2} \ln y_{1i} (1 + y_{1i}^\alpha) + \sum_{I_2} \ln y_{2i} (1 + y_{2i}^\alpha) + \sum_{I_1} \ln y_{2i} (1 + y_{2i}^\alpha) + \sum_{I_2} \ln y_{1i} (1 + y_{1i}^\alpha).$$

$$S_2 = \frac{m_0 + m_2 v_1 + m_1 v_1}{\lambda_0} + \{ \sum_{I_0} n(1 - e^{y_i^\alpha}) + \sum_{I_2} n(1 - e^{y_{2i}^\alpha}) + \sum_{I_1} n(1 - e^{y_{1i}^\alpha}) \}.$$

$$S_3 = \frac{m_2 + m_1 v_2}{\lambda_1} + \{ \sum_{I_0} n(1 - e^{y_i^\alpha}) + \sum_{I_2} n(1 - e^{y_{1i}^\alpha}) + \sum_{I_1} n(1 - e^{y_{1i}^\alpha}) \}.$$

$$S_4 = \frac{m_2 v_2 + m_1}{\lambda_2} + \{ \sum_{I_0} n(1 - e^{y_i^\alpha}) + \sum_{I_2} n(1 - e^{y_{2i}^\alpha}) + \sum_{I_1} n(1 - e^{y_{2i}^\alpha}) \}.$$

$$S_5 = \frac{m}{\hat{\theta}(1 - \hat{\theta})} - \frac{\sum_{i=1}^m a_i}{(1 - \hat{\theta})}.$$

Under conditions that are fulfilled for the parameter Θ in the interior of the parameter space but not on the boundary, the asymptotic distribution of $\sqrt{n}(\hat{\Theta} - \Theta)$ is a multivariate normal $N_5(0, J(\Theta)^{-1})$, where $J(\Theta) = \lim_{n \to \infty} n^{-1} I(\Theta)$ is the unit information matrix.

This asymptotic behavior remains valid if $J(\Theta)$ is replaced by the average sample information matrix evaluated at $\hat{\Theta}$, i.e., $n^{-1} I(\hat{\Theta})$. We can use the asymptotic multivariate normal $N_5(0, I(\hat{\Theta})^{-1})$ distribution of $\hat{\Theta}$ to construct approximate confidence regions for some parameters and for the hazard and survival functions.

5   Data Analysis and Comparison Study

In this section, we present some results based on Monte Carlo simulations and two real data sets to see and compare how the proposed model and the EM algorithm work in practice. We use the Matlab software for data analysis.

5.1   Numerical Experiments

In this subsection, we present some simulation results to show how the proposed EM algorithm performs for different sample sizes and various parameters. The following algorithm can be used for this purpose.

ALGORITHM

- Generate n from $GM(\theta)$.

- Generate $\{u_{01}, \ldots, u_{0n}\}$ from $CH(\alpha, \lambda_0)$, similarly, $\{u_{11}, \ldots, u_{1n}\}$ from $CH(\alpha, \lambda_1)$ and $\{u_{21}, \ldots, u_{2n}\}$ from $CH(\alpha, \lambda_2)$.

- Obtain $x_{1k} = \min\{u_{0k}, u_{1k}\}$ and $x_{2k} = \min\{u_{0k}, u_{2k}\}$, for $k = 1, \ldots, n$.

- Compute the desired (y_1, y_2) as $y_1 = \min\{x_{11}, \ldots, x_{1n}\}$ and $y_2 = \min\{x_{21}, \ldots, x_{2n}\}$.

We fit the BCHG model to the simulated data set and compute the MLEs of the unknown parameters, and also, we use the EM algorithm as suggested in the previous section. We used three sets of parameter values $\Theta_i = (\theta_i, \alpha, \lambda_0, \lambda_1, \lambda_2) = (\theta_i, 3, 1, 1, 1)$, where $\theta_i = 0.3, 0.5, 0.7$ for $i = 1, 2, 3$, respectively. In each case, we started the EM algorithm with $\theta_i = 0.2, 0.4, 0.6$ for $i = 1, 2, 3$ and also, $\alpha = \lambda_0 = \lambda_1 = \lambda_2 = 1$. The iteration stops when the absolute value of the difference of the two consecutive iterates for all the five parameters are less than 10^{-5}. We replicate the process 1000 times and report the bias estimates and the associated mean squared errors (MSEs).

Some of the points are quite clear from the simulation results. It is observed in all cases that as the sample size increases, the biases and the mean squared errors decrease. This verifies the consistency properties of the MLEs. The results are presented in Table 2. In Table 3, we present the average confidence lengths and the corresponding coverage percentages. The nominal level for the confidence intervals is 0.95 in each case. From Table 3, it is evident that as the sample size increases, the average confidence lengths decrease and the corresponding coverage percentage increases.

Table 2: The Bias estimates and the associated mean squared errors (MSEs).

Θ	n	α		λ_0		λ_1		λ_2		θ	
		|Bias|	MSE	|Bias|	MSE	|Bias|	MSE	|Bias|	MSE	|Bias|	MSE
Θ_1	20	0.2021	0.3043	0.2245	0.3298	0.2071	0.3268	0.2087	0.3701	0.1030	0.0116
	30	0.1125	0.1798	0.2058	0.1416	0.1987	0.1930	0.2046	0.2012	0.1022	0.0116
	50	0.953	0.1092	0.1755	0.1135	0.1715	0.1379	0.1782	0.1400	0.1020	0.0115
	100	0.0488	0.0459	0.1180	0.1111	0.1102	0.1226	0.1130	0.1240	0.1013	0.0107
	200	0.0309	0.0229	0.1059	0.1071	0.1073	0.1087	0.1000	0.1197	0.0098	0.0095
Θ_2	20	0.2065	0.3140	0.1729	0.1518	0.1231	0.2583	0.1206	0.2708	0.1126	0.0142
	30	0.1703	0.2116	0.1188	0.0969	0.0762	0.1424	0.0707	0.1488	0.1036	0.0139
	50	0.0986	0.1032	0.1070	0.0712	0.0432	0.0862	0.0472	0.0879	0.0971	0.0133
	100	0.0641	0.0523	0.0730	0.0459	0.0330	0.0566	0.0269	0.0594	0.0774	0.0094
	200	0.0375	0.0249	0.0620	0.0389	0.0255	0.0420	0.0183	0.0507	0.0683	0.0082
Θ_3	20	0.2202	0.3707	0.1036	0.1681	0.1022	0.2509	0.1018	0.2379	0.1122	0.0168
	30	0.2040	0.2332	0.0823	0.1018	0.0819	0.1222	0.0758	0.1361	0.1040	0.0155
	50	0.1429	0.1260	0.0719	0.0561	0.0570	0.0735	0.0518	0.0669	0.0966	0.0151
	100	0.1009	0.0569	0.0515	0.0282	0.0500	0.0356	0.0437	0.0384	0.0925	0.0146
	200	0.0777	0.0412	0.0113	0.0200	0.0351	0.0251	0.0362	0.0250	0.0900	0.0145

Table 3: The average confidence lengths and the corresponding coverage percentage.

Θ	n	α		λ_0		λ_1		λ_2		θ	
		Length	CP	Length	CP	Length	CP	Length	CP	Length	CP
Θ_1	20	1.4220	0.8430	1.6502	0.8550	1.8289	0.8170	1.8567	0.8440	0.3429	0.9080
	30	1.2568	0.8880	1.3616	0.8600	1.5302	0.8460	1.5190	0.8490	0.3087	0.9150
	50	1.0920	0.9200	1.1621	0.8920	1.2870	0.8550	1.2926	0.8530	0.2717	0.9150
	100	0.9024	0.9500	0.9219	0.9010	1.0304	0.8930	1.0253	0.8820	0.2276	0.9240
	200	0.7505	0.9650	0.7604	0.9290	0.8392	0.9070	0.8392	0.9060	0.1911	0.9540
Θ_2	20	1.6873	0.8550	1.8248	0.8900	2.0413	0.8830	2.0456	0.8850	0.6192	0.9000
	30	1.5144	0.8990	1.6211	0.9030	1.7965	0.8940	1.8008	0.8910	0.5632	0.9110
	50	1.3070	0.9270	1.3509	0.9290	1.4882	0.9390	1.4818	0.9330	0.4956	0.9400
	100	1.0875	0.9430	1.1127	0.9410	1.2236	0.9500	1.2188	0.9500	0.4158	0.9580
	200	0.9128	0.9680	0.9277	0.9520	1.0240	0.9550	1.0213	0.9600	0.3498	0.9640
Θ_3	20	1.9372	0.8800	2.0428	0.8870	2.2511	0.8810	2.2848	0.8830	0.9613	0.8960
	30	1.7142	0.8950	1.7413	0.8980	1.9302	0.8920	1.9229	0.8980	0.8678	0.9130
	50	1.4854	0.9190	1.4643	0.9150	1.6241	0.9180	1.6167	0.9280	0.7581	0.9480
	100	1.2411	0.9320	1.2203	0.9390	1.3519	0.9390	1.3558	0.9400	0.6337	0.9610
	200	1.0347	0.9490	1.0116	0.9420	1.1176	0.9420	1.1121	0.9510	0.5318	0.9680

5.2 Real Data Set

We consider the BCHG distribution for fitting the two data sets. Using the proposed EM algorithm in the previous section, the MLEs and their corresponding log-likelihood values are calculated. Also, the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) are calculated. We also obtain the Kolmogorov-Smirnov (K-S) distances between the fitted distribution, the empirical distribution function and the corresponding p-values for Y_1, Y_2 and $\min(Y_1, Y_2)$. Finally, we make use of the likelihood ratio test (LRT) and the corresponding p-values to choose a better model.

First data set: The data set has been obtained from Meintanis (2007) and it is presented in Table 4. The data represent the football (soccer) data where at least one goal scored by the home team and at least one goal scored directly from a penalty kick, foul kick or any other direct kick (all of them together will be called as kick goal) by any team has been considered. Here Y_1 represents the time in minutes of the first kick goal scored by any team and Y_2 represents the first goal of any type scored by the home team. We would like to analyze the data using the BCHG model. Before analyzing the data, we have divided all the data points by 100.

Table 4: UEFA Champion's League data.

2005-2006	X_1	X_2	2004-2005	X_1	X_2
Lyon-Real Madrid	26	20	Internazionale-Bremen	34	34
Milan-Fenerbahce	63	18	Real Madrid-Roma	53	39
Chelsea-Anderlecht	19	19	Man. United-Fenerbahce	54	7
Club Brugge-Juventus	66	85	Bayern-Ajax	51	28
Fenerbahce-PSV	40	40	Moscow-PSG	76	64
Internazionale-Rangers	49	49	Barcelona-Shakhtar	64	15
Panathinaikos-Bremen	8	8	Leverkusen-Roma	26	48
Ajax-Arsenal	69	71	Arsenal-Panathinaikos	16	16
Man. United-Benfica	39	39	Dynamo Kyiv-Real Madrid	44	13
Real Madrid-Rosenborg	82	48	Man. United-Sparta	25	14
Villarreal-Benfica	72	72	Bayern-M. TelAviv	55	11
Juventus-Bayern	66	62	Bremen-Internazionale	49	49
Club Brugge-Rapid	25	9	Anderlecht-Valencia	24	24
Olympiacos-Lyon	41	3	Panathinaikos-PSV	44	30
Internazionale-Porto	16	75	Arsenal-Rosenborg	42	3
Schalke-PSV	18	18	Liverpool-Olympiacos	27	47
Barcelona-Bremen	22	14	M. Tel-Aviv-Juventus	28	28
Milan-Schalke	42	42	Bremen-Panathinaikos	2	2
Rapid-Juventus	36	52			

Second data set: This data set contains 32 claims (claims for compensation) from motorcycle accident insurance. This data includes the cost of property damage (Y_1) and medical expenses (Y_2). These data are rounded up to the nearest 100,000 units, and then divided by this number. This data set is reported in Table 5. Before analyzing the data we have divided all the data points by 1000.

Table 5: Number of insurance claims for motorcycle accidents.

S.N.	Y_1	Y_2	S.N.	Y_1	Y_2	S.N.	Y_1	Y_2	S.N.	Y_1	Y_2
1	144	793	9	300	417	17	298	271	25	470	470
2	134	945	10	665	456	18	114	489	26	368	368
3	500	500	11	199	243	19	335	807	27	171	999
4	720	400	12	412	198	20	449	499	28	106	974
5	230	784	13	720	183	21	160	542	29	529	202
6	374	881	14	591	784	22	224	349	30	423	375
7	175	175	15	305	222	23	323	103	31	500	198
8	252	252	16	292	214	24	704	522	32	350	707

In order to get the hazard function of the marginals, we provide the scaled TTT plot as suggested by Aarset (1987), which provides a pattern of the shape of the hazard function of a distribution. It has been shown by Aarset (1987) that the scaled TTT transform is convex (concave) if the hazard rate is decreasing (increasing) and for bathtub (unimodal) hazard rates, the scaled TTT transform is initially convex (concave) and then concave (convex). We plot the scaled TTT transform for two data sets, in Figure 3. From Figure 3, it is quite clear that both Y_1 and Y_2 have an increasing hazard function.
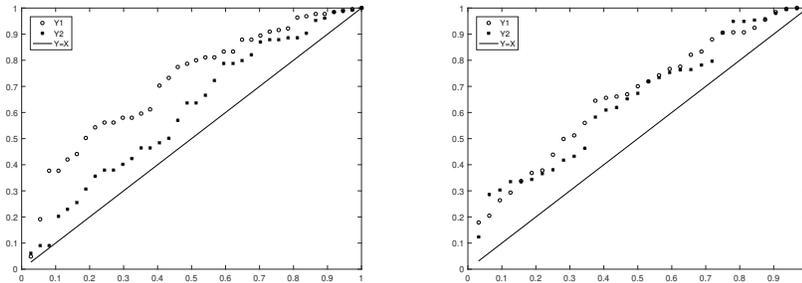

Figure 3: The scaled TTT plots of the marginals for data set 1 (Left), data set 2 (Right).

Before going to analyze the data using BCHG distribution, we fit the Chen distribution to Y_1, Y_2 and min$\{Y_1, Y_2\}$ separately. It will help us to guess the initial values also. The MLEs of parameters, the corresponding Kolmogorov-Smirnov distances between the fitted distribution and the empirical distribution function and the associated p-values are calculated. The results are presented in Table 6. Based on the p-values the Chen distribution cannot be rejected for the marginals and for the minimum also.

Table 6: The MLEs of parameters, the corresponding Kolmogorov-Smirnov (K-S) and the associated p-values for two data sets.

Data Set	Variable	α	λ	K-S	p-value
	Y_1	1.9073	3.7665	0.0937	0.9251
UEFA Champion's League Data	Y_2	1.2579	2.9656	0.1095	0.8122
	min$\{Y_1, Y_2\}$	1.3314	3.8615	0.0936	0.9251
	Y_1	1.6479	4.9098	0.1137	0.7763
Motorcycle Accident Insurance Data	Y_2	1.4448	2.2302	0.1237	0.6838
	min$\{Y_1, Y_2\}$	1.7414	9.5163	0.1422	0.4926

Now, we will fit the BCHG model. Using the proposed EM algorithm, this model is fitted to the bivariate data set. Then, the MLEs and their corresponding log-likelihood values are calculated. For each fitted model, the Akaike Information Criterion (AIC) and the Bayesian information criterion (BIC) are calculated. The results are given in Table 7.

Table 7: The MLEs of parameters, the corresponding log-likelihood, AIC and BIC for two data sets.

Data Set	$\hat{\alpha}$	$\hat{\lambda}_0$	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\theta}$	$log(\ell)$	AIC	BIC
UEFA Champion's League Data	1.6358	1.4440	0.6520	1.6284	0.3712	-18.0147	46.0294	54.0840
Motorcycle Accident Insurance Data	1.6066	0.6200	2.4745	1.2491	0.3463	-12.9549	35.9098	43.2385

We also compute the Kolmogorov-Smirnov (K-S) distances with the corresponding p-values between the fitted distribution and the empirical CDF for three random variables Y_1, Y_2 and min(Y_1, Y_2). The results are given in Table 8. Also, the profile log-likelihood function of θ and α is provided in Figure 4.

Table 8: The Kolmogorov-Smirnov (K-S) distances and the corresponding p-values for three random variables Y_1, Y_2 and $\min(Y_1, Y_2)$ for two data sets.

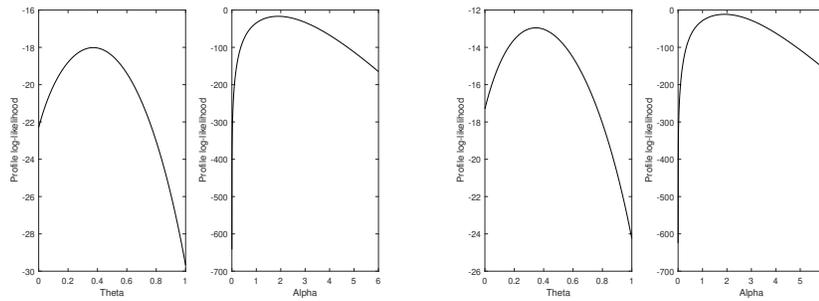| Data Set	Y_1		Y_2		$\min\{Y_1, Y_2\}$	
	K-S	p-value	K-S	p-value	K-S	p-value
UEFA Champion's League Data	0.1850	0.2111	0.1643	0.3353	0.1174	0.7432
Motorcycle Accident Insurance Data	0.1463	0.4767	0.1420	0.5141	0.1028	0.8535

Figure 4: The profile log-likelihood function of θ and α for data set 1 (Left), data set 2 (Right).

Finally, the likelihood ratio test (LRT) and the corresponding p-values are obtained for testing the BCH model against the BCHG model. On the other hand, our purpose is to test the null hypothesis $H_0 : BCH$ against the alternative hypothesis $H_1 : BCHG$. The statistics and the corresponding p-values are given in Table 9. Hence, for any usual significance level, we reject the model in H_0 (BCH) in favor of the alternative model (BCHG).

Table 9: The log-likelihood, AIC, BIC, LRT and the corresponding p-values for different models.

Data Set	Models	Test				
		AIC	BIC	$log\ell$	LRT	p-value
UEFA Champion's League Data	BCHG	46.0294	54.0840	-18.0147		
					16.7856	4.1850×10^{-5}
	BCH	60.8150	67.2587	-26.4075		
Motorcycle Accident Insurance Data	BCHG	35.9098	43.2385	-12.9549		
					17.3398	3.1257×10^{-5}
	BCH	51.2496	57.1125	-21.6248		

6   Conclusions

In this paper, we have proposed two new bivariate distributions using the methods of Marshall and Olkin. In the first structure, a new bivariate model was produced using the method of Marshall and Olkin (1967). In the second case, this model has been generalized. So, a new bivariate distribution by the method of Marshall and Olkin (1997) was introduced. In fact, this bivariate distribution can be obtained by compounding geometric and Chen distributions. We called these distributions as the bivariate Chen (BCH) distribution and bivariate Chen-geometric (BCHG) distribution, respectively. Moreover, the BCH can be obtained as a special case of the BCHG model. Several properties of these distributions have been established. The estimation of unknown parameters by the method of maximum likelihood is obtained. However, it is not directly easy to solve the associated log likelihood equations. Therefore, the corresponding maximum likelihood estimate is numerically computed from the associated nonlinear equation by the EM algorithm. Therefore, we have proposed to use the EM algorithm to obtain the MLEs of the unknown parameters, and it is observed that the proposed EM algorithm works quite well in practice. As shown, the proposed model works quite well for data analysis purposes. Finally, we have compared the BCHG model with BCH model.

Acknowledgements

References

Aarset, M. V. (1987), How to identify a bathtub hazard rate?. *IEEE Transactions on Reliability*, **36**, 106–108.

Adamidis, K. and Loukas, S. (1998), A lifetime distribution with decreasing failure rate. *Statistics & Probability Letters*, **39**, 35–42.

Barlow, R. E. and Proschan, F. (1981), *Statistical Theory of Reliability and Life Testing, Probability Models*. Maryland: Silver Spring.

Barreto-Souza, W., Lemos de Moraisa, A. and Cordeiro, G. M. (2011), The Weibull-geometric distribution. *Journal of Statistical Computation and Simulation*, **81**, 645–657.

Barreto-Souza, W. (2012), Bivariate gamma-geometric law and its induce Levy process. *Journal of Multivariate Analysis*, **109**, 130–145.

Bebbington, M., Lai, C. and Zitikis, R. (2007), A flexible Weibull extension. *Reliability Engineering & System Safety*, **92**, 719–726.

Bemis, B., Bain, L. J. and Higgins, J. J. (1972), Estimation and hypothesis testing for the parameters of a bivariate exponential distribution. *Journal of the American Statistical Association*, **67**, 927–929.

Chahkandi, M. and Ganjali, M. (2009), On some lifetime distributions with decreasing failure rate. *Computational Statistics & Data Analysis*, **53**, 4433–4440.

Chen, Z. (2000), A new two-parameter lifetime distribution with bathtub shape or increasing failure rate function. *Statistics & Probability Letters*, **49**, 155–161.

Dinse, G. (1982), Non-parametric estimation of partially incomplete time and types of failure data. *Biometrics*, **38**, 417–431.

Gaver, D. P. and Acar, M. (1979), Analytical hazard representations for use in reliability, mortality, and simulation studies. *Communications in Statistics*, **8**, 91–111.

Ghitany, M. E., Al-Hussaini, E. K. and Al-Jarallah R. A. (2005), Marshall-Olkin extended Weibull distribution and its application to censored data. *Journal of Applied Statistics*, **32**, 1025–1034.

Ghitany, M. E., Al-Awadhi, M. A. and Alkhalfan, L. A. (2007), Marshall-Olkin extended Lomax distribution and its application to censored data. *Communications in Statistics - Theory and Methods*, **36**, 1855–1866.

Gurvich, M. R., Dibenedetto, A. T. and Rande, S. V. (1997), A new statistical distribution for characterizing the random strength of brittle materials. *Journal of Materials Science*, **32**, 2559–2564.

Hassan, A. S. and Abd-Allah, M. (2017), Exponentiated Lomax Geometric Distribution: Properties and Applications. *Pakistan Journal of Statistics and Operation Research*, **13**, 545–566.

Haynatzki, G. R., Weron, K. and Haynatzka, V. R. (2000), A new statistical model of tumor latency time. *Mathematical and Computer Modelling*, **32**, 251–256.

Hjorth, U. (1980), A reliability distribution with increasing, decreasing, and bathtub-shaped failure rates. *Technometrics*, **22**, 99–107.

Kotz, S., Balakrishnan, N. and Johnson, N. L. (2000), *Continuous multivariate distributions*. New York: John Wiley and Sons.

Johnson, N. L. and Kotz, S. (1975), A vector of multivariate hazard rate. *Journal of Multivariate Analysis*, **5**, 53–66.

Kundu, D. (2004), Parameter estimation for partially complete time and type of failure data. *Biometrical Journal*, **46**, 165–179.

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

Kundu, D. and Dey, A. (2009), Estimating the parameters of the Marshall-Olkin bivariate Weibull distribution by EM algorithm. *Computational Statistics and Data Analysis*, **35**, 956–965.

Kundu, D. and Gupta, R. D. (2006), Estimation of $R = P(Y < X)$ for Weibull distribution. *IEEE Transactions on Reliability*, **55**, 270–280.

Kundu, D. and Gupta, A. K. (2014), On Bivariate Weibull-Geometric Distribution. *Journal of Multivariate Analysis*, **123**, 19–29.

Leemis, L. M. (1986), Relationships among common univariate distributions. *The American Statistician*, **40**, 143–146.

Lehmann, E. L. (1966), Some concepts of dependence. *Annals of Mathematical Statistics*, **37**, 1137–1153.

Louis, T. A. (1982), Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society*, **44**, 226 –233.

Francisco Louzada, F., Marchi, V. A. A. and Roman, M. (2012), The exponentiated exponentialgeometric distribution: a distribution with decreasing, increasing and unimodal failure rate. *Statistics*, **48**, 167–181.

Marshall, A. W. and Olkin, I. (1967), A multivariate exponential distribution. *Journal of the American Statistical Association*, **62**, 30–44.

Marshall, A. W. and Olkin, I. (1997), A new method of adding a parameter to a family of distributions with application to the exponential and Weibull families . *Biometrika*, **84**, 641–652.

Meintanis, S. G. (2007), Test of fit for Marshall-Olkin distributions with applications. *Journal of Statistical Planning and inference*, **137**, 3954–3963.

Mudholkar, G. S. and Srivastava, D. K. (1993), Exponentiated Weibull family for analyzing bathtub failure-rate data. *IEEE Transactions on Reliability*, **42**, 299–302.

Nadaraja, S. and Kotz, S. (2007), The two-parameter bathtub-shaped lifetime distribution. *Quality and Reliability Engineering International*, **23**, 279–280.

Pham, H. and Lai, C. D. (2007), On recent generalizations of the Weibull distribution. *IEEE Transactions on Reliability*, **56**, 454–458.

Sarhan, A., Hamilton, D. C. and Smith, B. (2012), Parameter estimation for a two-parameter bathtub-shaped lifetime distribution. *Applied Mathematical Modelling*, **36**, 5380–5392.

Shoaee, SH. and Khorram, E. (2015), Stress-strength reliability of a two-parameter bathtub-shaped lifetime distribution based on progressively censored samples. *Communications in Statistics–Theory and Methods*, **44**, 5306–5328.

Smith, R. M. and Bain, L. J. (1975), An exponential power life-testing distribution. *Communications in Statistics*, **4**, 469–481.

Wu, S. J. (2008), Estimation of the two-parameter bathtub-shaped lifetime distribution with progressive censoring. *Journal of Applied Statistics*, **35**, 1139–1150.

Wu, J. W., Lu, H. L., Chen, C. H. and Wu, C. H. (2004), Statistical inference about the shape parameter of the new two-parameter bathtub-shaped lifetime distribution. *Quality and Reliability Engineering International*, **20**, 607–616.

Xie, M., Tang, Y. and Goh, T. (2002), A modied Weibull extension with bathtub-shaped failure rate function. *Reliability Engineering & System Safety*, **76**, 279–285.

Zhang, T. and Xie, M. (2007), Failure data analysis with extended Weibull distribution. *Communications in Statistics - Simulation and Computation*, **36**, 579–592.

Zhang, T., Xie, M., Tang, L. C. and Ng, S. H. (2005), Reliability and modeling of systems integrated with firmware and hardware. *International Journal of Reliability, Quality and Safety Engineering*, **12**, 227–239.