Applications and Applied Mathematics: An International Journal (AAM)

Volume 15 | Issue 2 | Article 6

12-2020

On a New Class of Bivariate Survival Distributions Based on the Model of Dependent Lives and its Generalization

Shirin Shoaee
Shahid Beheshti University

Available at
http://pvamu.edu/aam
Appl. Appl. Math.
ISSN: 1932-9466

Applications and Applied Mathematics:
An International Journal
(AAM)

Vol. 15, Issue 2 (December 2020), pp. 801 – 829

On a New Class of Bivariate Survival Distributions Based on the Model of Dependent Lives and its Generalization

Shirin Shoaee

Department of Actuarial Science
Faculty of Mathematical Sciences
Shahid Beheshti University
Tehran, Iran
Sh_Shoaee@sbu.ac.ir

Received: April 30, 2020; Accepted: June 9, 2020

Abstract

In this paper, a new class of survival distributions based on the model of dependent lives and proportional hazard rate family is introduced. This new family of bivariate survival models contains several bivariate lifetime models and is more flexible. The main purpose of this paper is to generalize this family of bivariate survival distributions of dependent lives so that more flexible models can be achieved. These new families of distributions are called the bivariate proportional hazard rate (BPHR) and the bivariate proportional hazard rate-geometric (BPHRG) families, respectively. It is also observed that, if $\theta = 1$, then the BPHR family is a particular state of the BPHRG family. Several features of these new families of distributions such as the multivariate aging properties, the bivariate hazard gradient, and dependency structures are investigated. We design a flexible computational EM algorithm to calculate the maximum likelihood estimation of parameters. Also, several simulation studies are represented to evaluate the efficiency of the EM algorithm. Finally, we analyze three real datasets and compare the BPHRG models with the BPHR models.

Keywords: Dependent Lives; EM Algorithm; Life Insurance; Longevity; Proportional Hazard Rate Family

MSC 2010 No.: 62H10, 62H12, 62E15

1. Introduction

In recent years, the proportional hazard rate models have been considered by many authors. This class of models was principally designed by Cox (1997) and has been spaciously discussed in the statistical literature. The hazard rate and reliability function of this family of distributions have many different forms. Therefore, by selecting the special formation of the baseline cumulative distribution, this family can be applied in many applications of quality control, survival, insurance, and failure time modeling. Also, the simple partial likelihood function is another advantage of this model. For further details on PHRM, we suggest that readers refer to Marshall and Olkin (2007), Ahmadi et al. (2009b), Ahmadi et al. (2009a), and Asgharzadeh and Valiollahi (2010).

As we know, there are various methods for the modeling of longevity. In many fields of science, including statistics and life insurance, it is assumed that the remainder of a person's or components' lives is independent. But applying this assumption is not always correct. Because sometimes there may be common risk factors for a pair of individuals, and individuals are exposed to the same risk factors. For example, in the case of twins, the same risk factors may be genetic, or in the case of married couples, these factors may come from the environment.

One of the classic models of dependent lives, which is very popular and widely used by many researchers recently, is the common shock model. To structure these models, it is assumed that the lifespan of the two individuals is independent of each other unless a common accident causes their death that the time of the disaster should be used in the modeling of survival data. In this regard, we can mention Marshall and Olkin (1967), Sarhan and Balakrishnan (2007), Al-Khedhairi and El-Gohary (2008), and Kundu and Gupta (2010). This structure has many applications and is widely used, especially in actuarial theory, finance, reliability theory in the competing risk and shock models, as well as medical and demographic studies.

In the field of actuarial theory and life insurance, as we know, pensions guarantee the payment of regular periodic income during the life of the pensioner. Therefore, these dependent models are beneficial for estimating the probabilities of a couple's survival and reviewing and evaluating their annual contracts. The reason for applying the dependent model has been described in Frees et al. (1996). In actuarial calculations to determine the insurance benefits, retirement benefits, and premiums payable in the joint life, people's lives are considered as a status and according to the terms of the insurance contract, the random variables of the time-until failure are defined by the first death or the last survivor. Therefore, common shock models can be easily used in calculations related to annuity and life insurance. In actuarial science, competing risk models actually represent multiple reductions, and a person may die from one of several possible causes (cancer, heart disease, accident, etc.). Similar common shock models are used in finance and reliability theory. A detailed study of this issue and modeling of credit and insurance losses was performed by Giesecke (2003) and Lindskog and McNeil (2003).

Another structure for modeling longevity data was provided by Marshall and Olkin (1997) in the case of univariate distributions. In this structure, new univariate distributions were presented through minimizing and maximizing independent and identically distributed continuous random variables by assuming a geometric distribution is selected for the sample size. So, a parameter was added to the models that made the models more flexible. For further details, researchers can refer to the work of Silva et al. (2013), Ghitany et al. (2005), Ghitany et al. (2007), Pham and Lai (2007), and Barreto-Souza (2012), Kundu and Gupta (2014), and Shoaee and Khorram (2019).

According to this content, we assume that the remainder of the lifetime belongs to the proportional hazard rate family. Thus, a family of bivariate survival distributions is introduced based on the dependent lives models. This new family of bivariate distributions is called the bivariate proportional hazard rate (BPHR) family. For this new family of survival distributions, we examine the various properties which are very important and applicable. Also, our main purpose of this research is to generalize the family of bivariate survival distributions of dependent lives. Thus, a new class of family of dependent life distributions is obtained. This extension is obtained by minimizing independent and identically distributed continuous random variables by assuming a geometric distribution is selected for the sample size.

It is necessary to mention, the use of this structure to generalize bivariate models is not as widespread as the use of this structure in univariate models. Therefore, this generalized class contains the family of bivariate proportional hazard rate distributions (BPHR) and is called the bivariate proportional hazard rate-geometric (BPHRG) class of distributions. Also, the marginal and conditional distributions of this generalized family of distributions are obtained, which can be seen that these distributions belong to the univariate proportional hazard rate-geometric distributions (UPHRG) family. Besides, different properties of the BPHRG models have been investigated.

We obtain the estimation of the BPHRG model parameters using the maximum likelihood method, but we find that there are no explicit expressions for them. Because there are nonlinear equations that need to be solved simultaneously, we can use techniques that have been proposed in the past to solve this problem, including Newton-Raphson and Gauss-Newton methods. However, in these methods, it is essential to select the appropriate initial values for the convergence of these algorithms. In this paper, it is suggested to use the missing value method and an EM algorithm for parameter estimation. In this proposed EM algorithm, only one nonlinear one-dimensional equation is solved at each "E-step". We also find that the performance of the proposed algorithm is very desirable and easy to use. Therefore, it is possible to estimate the parameters of this family of distributions using this algorithm and deduce and analyze this family of models. Therefore, a family of alternative survival models is derived that can be fitted better than existing models.

The present paper is organized as follows: A family of bivariate survival distributions based on the structure of dependent lives that are very applicable is presented in Section 2. In the following section, we describe several features of this new model family. A useful and practical extension of this new family of bivariate survival models is provided in Section 3, and the properties and characteristics of this new family of bivariate models are studied in this section. Inference about the unknown parameters of this family, the estimation method, and the proposed algorithm structure are presented in Section 4. In Section 5, the simulation study to evaluate the efficiency of the designed algorithm and analysis of three real datasets for comparing the new models are presented. Finally, the conclusions and results are described in Section 6.

2. Marshall–Olkin Bivariate Proportional Hazard Rate Models

In this section, we will introduce a family of the bivariate proportional hazard rate distributions. Therefore, at the beginning of this section, we will define the structure of the proportional hazard rate family.

Definition 2.1.

The family of random variables is called the proportional hazard rate (PHR) family, if its the hazard rate function has the form of $\{\lambda h_B(.) : \lambda > 0\}$, where $h_B(.)$ is the hazard rate.

In other words, we can say that, if Z is a member of the proportional hazard rate family, then the cumulative distribution function becomes

$$F_{PHRM}(z;\alpha,\lambda) = 1 - [\bar{F}_B(z,\alpha)]^{\lambda}, \quad -\infty \leq b_1 < z < b_2 \leq \infty, \quad \lambda > 0, \tag{1}$$

where $\bar{F}_B(.) = 1 - F_B(.)$ is the baseline survival function with $F_B(b_1) = 0$ and $F_B(b_2) = 1$. Therefore, by selecting the special formation of the baseline cumulative distribution, this family can be applied in many applications of quality control, survival, insurance, and failure time modeling. Another advantage of the proportional hazard rate model is the simple partial likelihood.

We can obtain the probability density function (PDF) from the model (1) as

$$f_{PHRM}(z;\alpha,\lambda) = \lambda f_B(z,\alpha)[\bar{F}_B(z,\alpha)]^{\lambda-1}, \quad -\infty \leq b_1 < z < b_2 \leq \infty, \quad \lambda > 0, \tag{2}$$

where $f_B(.)$ is the PDF of $F_B(.)$. Table 1 shows some useful quantities for some distributions belonging to the proportional hazard rate family (Weibull, Lomax, Chen, and Gompertz).

Now let's assume that T_i follows ($\sim$) $PHRM(\alpha,\lambda_i)$, for $i = 0,1,2$, and also they are independent. Define $X_i = \min\{T_0, T_i\}$, for $i = 1,2$. Then, the random vector $\mathbf{X} = (X_1, X_2)$ belong to the bivariate proportional hazard rate distributions family. This family of distributions has four parameters and is denoted by the symbol $BPHR(\alpha,\lambda_0,\lambda_1,\lambda_2)$. We discuss some of the important functions and consequences of this family of distributions.

Table 1. Some useful quantities for some distributions belonging to the proportional hazard rate family

Distribution	$\bar{F}_B(x,\alpha)$	$f_B(x,\alpha)$	$\ln \bar{F}_B(x,\alpha)$	$\ln \dfrac{f_B(x,\alpha)}{\bar{F}_B(x,\alpha)}$
Weibull	$e^{-x^{\alpha}}$	$\alpha x^{\alpha-1} e^{-x^{\alpha}}$	$-x^{\alpha}$	$\ln\alpha + (\alpha-1)\ln x$
Lomax	$\dfrac{1}{1+\alpha x}$	$\dfrac{\alpha}{(1+\alpha x)^2}$	$-\ln(1+\alpha x)$	$\ln\alpha - \ln(1+\alpha x)$
Chen	$e^{1-e^{x^{\alpha}}}$	$\alpha x^{\alpha-1} e^{x^{\alpha}} e^{1-e^{x^{\alpha}}}$	$1 - e^{x^{\alpha}}$	$\ln\alpha + (\alpha-1)\ln x + x^{\alpha}$
Gompertz	$e^{-(e^{\alpha x}-1)}$	$\alpha e^{\alpha x} e^{-(e^{\alpha x}-1)}$	$-(e^{\alpha x}-1)$	$\ln\alpha + \alpha x$

Theorem 2.1.

Suppose $\mathbf{X} \sim BPHR(\alpha, \lambda_0, \lambda_1, \lambda_2)$. Then, the joint survival function can be obtained for $z = \max\{x_1, x_2\}$ as follows:

$$\bar{F}_{X_1,X_2}(x_1, x_2) = \begin{cases} \bar{F}_{PHRM}(x_1, \alpha, \lambda_1 + \lambda_0)\bar{F}_{PHRM}(x_2, \alpha, \lambda_2), & \text{if } x_2 < x_1, \\ \\ \bar{F}_{PHRM}(x_1, \alpha, \lambda_1)\bar{F}_{PHRM}(x_2, \alpha, \lambda_2 + \lambda_0), & \text{if } x_1 < x_2, \\ \\ \bar{F}_{PHRM}(x, \alpha, \lambda_0 + \lambda_1 + \lambda_2, \lambda), & \text{if } x_1 = x_2 = x. \end{cases} \tag{3}$$

Proof:

To calculate the joint survival function, we have:

$$\begin{aligned} F_{X_1,X_2}(x_1, x_2) &= P(X_1 > x_1, X_2 > x_2) \\ &= P(\min\{T_0, T_1\} > x_1, \min\{T_0, T_2\} > x_2) \\ &= P(T_1 \geq x_1, T_2 \geq x_2, T_0 \geq z) = P(T_1 > x_1)P(T_2 > x_2)P(T_0 > z) \\ &= \bar{F}_{PHRM}(x_1, \alpha, \lambda_1)\bar{F}_{PHRM}(x_2, \alpha, \lambda_2)\bar{F}_{PHRM}(z, \alpha, \lambda_0), \end{aligned}$$

where $z = \max\{x_1, x_2\}$. Therefore, the desired result is obtained. ■

Theorem 2.2.

Suppose $\mathbf{X} \sim BPHR(\alpha, \lambda_0, \lambda_1, \lambda_2)$. Then, the joint probability density function of $\mathbf{X}$ is expressed as follows:

$$f_{X_1,X_2}(x_1, x_2) = \begin{cases} f_1(x_1, x_2), & \text{if } x_2 < x_1, \\ f_2(x_1, x_2), & \text{if } x_1 < x_2, \\ f_0(x), & \text{if } x_1 = x_2 = x, \end{cases} \tag{4}$$

where

$$\begin{aligned} f_1(x_1, x_2) &= f_{PHRM}(x_1, \alpha, \lambda_1 + \lambda_0)f_{PHRM}(x_2, \alpha, \lambda_2), \\ f_2(x_1, x_2) &= f_{PHRM}(x_1, \alpha, \lambda_1)f_{PHRM}(x_2, \alpha, \lambda_2 + \lambda_0), \\ f_0(x) &= \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2} f_{PHRM}(x, \alpha, \lambda_0 + \lambda_1 + \lambda_2). \end{aligned}$$

Proof:

We obtain the phrases $f_1(x_1, x_2)$ and $f_2(x_1, x_2)$ through $-\frac{\partial^2 \bar{F}_{X_1,X_2}(x_1,x_2)}{\partial x_1 \partial x_2}$ for $x_1 < x_2$ and $x_2 < x_1$ respectively. For the expression $f_0(x)$ we cannot use the same method and to calculate it we need to consider the following relation:

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

We compute,

$$\int_0^\infty \int_{x_2}^\infty f_1(x_1, x_2)dx_1 dx_2 = \lambda_2 \int_0^\infty [\bar{F}_B(x_2, \alpha)]^{\lambda_0+\lambda_1+\lambda_2-1} f_B(x_2, \alpha)dx_2 = \frac{\lambda_2}{\lambda_0 + \lambda_1 + \lambda_2},$$

and similarly,

$$\int_0^\infty \int_{x_1}^\infty f_2(x_1,x_2)dx_2dx_1 = \lambda_1 \int_0^\infty [\bar{F}_B(x_1,\alpha)]^{\lambda_0+\lambda_1+\lambda_2-1} f_B(x_1,\alpha)dx_1 = \frac{\lambda_1}{\lambda_0+\lambda_1+\lambda_2}.$$

Therefore, we should have:

$$\int_0^\infty f_0(x)dx = \lambda_0 \int_0^\infty [\bar{F}_B(x,\alpha)]^{\lambda_0+\lambda_1+\lambda_2-1} f_B(x,\alpha)dx = \frac{\lambda_0}{\lambda_0+\lambda_1+\lambda_2}.$$

So, the desired result is achieved. ■

Regarding Theorem 2.2, we see that this family of models has an absolutely continuous part and a singular part. In other words, Theorem 2.2 expresses the density function of this family. The first two terms represent the density function with respect to the two-dimensional Lebesgue measure. The third expression represents the density function according to the one-dimensional Lebesgue measure, see Bemis et al. (1972). In this family of distributions, the singular part indicates that X_1 and X_2 are continuous random variables, and the probability of being equal is positive, see Marshall and Olkin (1967). Theorem 2.3 provides a better explanation of the absolutely continuous and singular parts in this family of distributions.

Theorem 2.3.

Suposse $\mathbf{X} \sim BPHR(\alpha,\lambda_0,\lambda_1,\lambda_2)$. Then,

$$\bar{F}_{X_1,X_2}(x_1,x_2) = \frac{\lambda_1+\lambda_2}{\lambda_0+\lambda_1+\lambda_2}\bar{F}_a(x_1,x_2) + \frac{\lambda_0}{\lambda_0+\lambda_1+\lambda_2}\bar{F}_s(x_1,x_2), \tag{5}$$

where

$$\bar{F}_s(x_1,x_2) = [\bar{F}_B(z,\alpha)]^{(\lambda_0+\lambda_1+\lambda_2)},$$

and

$$\begin{aligned}\bar{F}_a(x_1,x_2) = {} & \frac{\lambda_0+\lambda_1+\lambda_2}{\lambda_1+\lambda_2}[\bar{F}_B(x_1,\alpha)]^{\lambda_1}[\bar{F}_B(x_2,\alpha)]^{\lambda_2}[\bar{F}_B(z,\alpha)]^{\lambda_0} \\ & - \frac{\lambda_0}{\lambda_0+\lambda_1+\lambda_2}[\bar{F}_B(z,\alpha)]^{(\lambda_0+\lambda_1+\lambda_2)}.\end{aligned}$$

Proof:

We use the same method as provided by Muhammed (2016). Suppose event A is as follows:

$$A = \{T_0 < T_1\} \cap \{T_0 < T_2\}.$$

Therefore, $P(A) = \lambda_0/(\lambda_0+\lambda_1+\lambda_2)$ and $P(A') = 1 - P(A) = (\lambda_1+\lambda_2)/(\lambda_0+\lambda_1+\lambda_2)$. Therefore,

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

In addition, we know that for $z = max\{x_1,x_2\}$

$$\begin{aligned}P(X_1 \geq x_1, X_2 \geq x_2|A) &= [P(A)]^{-1}P(T_1 \geq T_0, T_2 \geq T_0, T_0 \geq z) \\ &= [\bar{F}_B(z,\alpha)]^{(\lambda_0+\lambda_1+\lambda_2)}.\end{aligned}$$

So, the expression $P(X_1 \geq x_1, X_2 \geq x_2|A')$ is calculated through the subtraction. In fact, $[\bar{F}_B(z,\alpha)]^{(\lambda_0+\lambda_1+\lambda_2)}$ represents a singular part because it can be seen that the mixed second partial derivatives are zero when $x_1 \neq x_2$. Therefore, $P(X_1 \geq x_1, X_2 \geq x_2|A')$ represents the absolutely continuous part because its mixed second partial derivatives is a bivariate density function. ∎

Using Theorem 2.3 and 2.2, the joint probability density function for $z = \max\{x_1, x_2\}$ can be written as follows:

$$f_{X_1,X_2}(x_1,x_2) = \frac{\lambda_1+\lambda_2}{\lambda_0+\lambda_1+\lambda_2} f_a(x_1,x_2) + \frac{\lambda_0}{\lambda_0+\lambda_1+\lambda_2} f_s(z), \tag{6}$$

where

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

and

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

where $f_a(x_1,x_2)$ is the absolutely continuous part, and $f_s(z)$ is the singular part, respectively.

From Theorem 2.3 we can conclude that $\bar{F}_{X_1,X_2}(x_1,x_2) \to [\bar{F}_B(x_1,\alpha)]^{\lambda_1}[\bar{F}_B(x_2,\alpha)]^{\lambda_2}$ for fixed λ_1 and λ_2 when $\lambda_0 \to 0$. In fact, we can say that X_1 and X_2 become independent. Also, if we define an event A as before, we have

$$A = (T_0 < T_1) \cap (T_0 < T_2) = \{\min\{T_1,T_2\} > T_0\} = \{X_1 = X_2\}.$$

Therefore, $P(A) = \frac{\lambda_0}{\lambda_0+\lambda_1+\lambda_2} = P(X_1 = X_2) \to 1$ when $\lambda_0 \to \infty$. So, it can be concluded that X_1 and X_2 are asymptotically almost surely equal. The final result is that $corr(X_1, X_2)$ changes from zero to one for fixed λ_1 and λ_2, and when λ_0 changes from zero to infinity.

We can obtain the absolutely continuous bivariate proportional hazard rate (ACBPHR) family of models by removing the singular part and keeping the absolutely continuous part.

$$f_{ACBPHR}(x_1,x_2) = \begin{cases} c f_{PHRM}(x_1,\alpha,\lambda_0+\lambda_1) f_{PHRM}(x_2,\alpha,\lambda_2), & \text{if } x_2 < x_1, \\ \\ c f_{PHRM}(x_2,\alpha,\lambda_1) f_{PHRM}(x_2,\alpha,\lambda_0+\lambda_2), & \text{if } x_1 < x_2, \end{cases} \tag{7}$$

where $c = \frac{\lambda_0+\lambda_1+\lambda_2}{\lambda_1+\lambda_2}$ is the normalizing constant.

2.1. Various Features

In this subsection, we describe the useful features of this family of models. We first obtain the marginal and conditional functions of this family. In the following section, we describe the concepts of positive and negative dependence, see for example, Balakrishnan and Lai (2009). We also explain the properties of aging and the bivariate hazard gradient for this family, see Johnson and Kotz (1975).

Proposition 2.1.

Assume that the random vector $\mathbf{X}$ belongs to the bivariate proportional hazard rate family, i.e., $\mathbf{X} \sim BPHR(\alpha, \lambda_0, \lambda_1, \lambda_2)$. Then,

I: $X_i \sim PHRM(\alpha, \lambda_0 + \lambda_i), \quad i = 1, 2.$
II: $P(X_1 < X_2) = \frac{\lambda_1}{\lambda_0+\lambda_1+\lambda_2}$.
III: $\min\{X_1, X_2\} \sim BPHR(\alpha, \sum_{i=0}^{2} \lambda_i)$.

Proof:

I: To prove this part, we use the cumulative distribution function for variable X_1 as follows:

$$\begin{aligned}
F_{X_i}(x) &= P(X_i < x) = 1 - P(\min\{T_0, T_i\} > x) \\
&= 1 - \bar{F}_{PHRM}(x, \alpha, \lambda_0)\bar{F}_{PHRM}(x, \alpha, \lambda_i) \\
&= 1 - [\bar{F}_B(x, \alpha)]^{\lambda_0}[\bar{F}_B(x, \alpha)]^{\lambda_i} \\
&= F_{PHRM}(x, \alpha, \lambda_0 + \lambda_i).
\end{aligned}$$

II: The proof of this part is as follows:

$$\begin{aligned}
P(X_1 < X_2) &= \int_{x_1}^{\infty}\int_{0}^{\infty} f_{PHRM}(x_1, \alpha, \lambda_1) f_{PHRM}(x_2, \alpha, \lambda_0 + \lambda_2) dx_1 dx_2 \\
&= \int_{0}^{\infty} f_{PHRM}(x_1, \alpha, \lambda_1)\bar{F}_{PHRM}(x_1, \alpha, \lambda_0 + \lambda_2) dx_1 \\
&= \int_{0}^{\infty} \lambda_1 [\bar{F}_B(x_1, \alpha)]^{\lambda_1 - 1}[\bar{F}_B(x_1, \alpha)]^{\lambda_0+\lambda_2} f_B(x_1, \alpha) dx_1 \\
&= \frac{\lambda_1}{\lambda_0 + \lambda_1 + \lambda_2}.
\end{aligned}$$

III: The proof of this part is as follows:

$$\begin{aligned}
P(\min\{X_1, X_2\} < x) &= P(X_1 < x, X_2 < x) = 1 - P(T_0 > x)P(T_1 > x)P(T_2 > x) \\
&= 1 - \bar{F}_{PHRM}(x, \alpha, \lambda_0)\bar{F}_{PHRM}(x, \alpha, \lambda_1)\bar{F}_{PHRM}(x, \alpha, \lambda_2) \\
&= 1 - [\bar{F}_B(x_1, \alpha)]^{\lambda_0+\lambda_1+\lambda_2} \\
&= F_{PHRM}(x, \alpha, \lambda_0 + \lambda_1 + \lambda_2).
\end{aligned}$$

■

Proposition 2.2.

If the random vector $\mathbf{X}$ belongs to the bivariate proportional hazard rate family. Then,

I: The conditional survival function of X_1 given $X_2 \geq x_2$ can be computed as follows:

$$P(X_1 \geq x_1 | X_2 \geq x_2) = \bar{F}_{X_1|X_2 \geq x_2}(x_1) = \begin{cases} [\bar{F}_B(x_1, \alpha)]^{\lambda_0+\lambda_1}[\bar{F}_B(x_2, \alpha)]^{-\lambda_0}, & \text{if } x_2 < x_1 \\ [\bar{F}_B(x_1, \alpha)]^{\lambda_1}, & \text{if } x_1 < x_2. \end{cases} \tag{8}$$

It is clearly seen that this function is absolutely continuous.

II: The conditional survival function was presented in relation (8) can be rewritten as follows:

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

where,

$$G(x_1) = \frac{1}{p}\begin{cases} [\bar{F}_B(x_1,\alpha)]^{\lambda_0+\lambda_1}[\bar{F}_B(x_2,\alpha)]^{-\lambda_0}, & \text{if } x_2 < x_1, \\ [\bar{F}_B(x_1,\alpha)]^{\lambda_1} - \frac{\lambda_0}{\lambda_0+\lambda_2}[\bar{F}_B(x_2,\alpha)]^{\lambda_1}, & \text{if } x_1 < x_2, \end{cases} \qquad H(x) = \begin{cases} 1, & \text{if } x < x_2, \\ 0, & \text{if } x > x_2, \end{cases}$$

and

$$p = 1 - \frac{\lambda_0}{\lambda_0+\lambda_2}[\bar{F}_B(x_2,\alpha)]^{\lambda_1}.$$

Proof:

I: For $x_2 < x_1$, we express the calculations, and for $x_1 < x_2$ it can be calculated similarly.

$$\begin{aligned} P(X_1 \geq x_1 | X_2 \geq x_2) &= \frac{P(X_1 \geq x_1, X_2 \geq x_2)}{P(X_2 \geq x_2)} = \frac{\bar{F}_{X_1,X_2}(x_1,x_2)}{\bar{F}_{X_2}(x_2)} \\ &= \frac{\bar{F}_{PHRM}(x_1,\alpha,\lambda_0+\lambda_1)\bar{F}_{PHRM}(x_2,\alpha,\lambda_2)}{\bar{F}_{PHRM}(x_2,\alpha,\lambda_0+\lambda_2)} \\ &= \bar{F}_{PHRM}(x_1,\alpha,\lambda_0+\lambda_1)\frac{[\bar{F}_B(x_2,\alpha)]^{\lambda_2}}{[\bar{F}_B(x_2,\alpha)]^{\lambda_0+\lambda_2}} \\ &= \bar{F}_{PHRM}(x_1,\alpha,\lambda_0+\lambda_1)[\bar{F}_B(x_2,\alpha)]^{-\lambda_0}. \end{aligned}$$

II: For $x_2 < x_1$, we see that $H(x_1) = 0$ and the desired result is obtained. Also for the case $x_2 > x_1$ because $H(x_1) = 1$ and $1 - p = \frac{\lambda_0}{\lambda_0+\lambda_2}[\bar{F}_B(x_2,\alpha)]^{\lambda_1}$, the result can be achieved. ∎

Now we will provide the dependency properties between the two variables. Lehmann (1966) explained two random variables X_1 and X_2 to be positive quadrant dependent (PQD), if for all x_1 and x_2,

$$P(X_1 \leq x_1, X_2 \leq x_2) \geq P(X_1 \leq x_1)P(X_2 \leq x_2).$$

Proposition 2.3.

Assume that the random vector $\mathbf{X}$ belongs to the bivariate proportional hazard rate family. Then, the random vector $\mathbf{X}$ is

I: a positive upper orthant dependent.
II: a positive quadrant dependent.

Proof:

I: Since the random vector $\mathbf{X}$ belongs to the bivariate proportional hazard rate family, so $P(X_1 \geq x_1, X_2 \geq x_2) \geq P(X_1 > x_1)P(X_2 > x_2)$ for all $x_1 > 0$ and $x_2 > 0$. Therefore, the random vector $\mathbf{X}$ is positive upper orthant dependent.

II: Since the random vector $\mathbf{X}$ belongs to the bivariate proportional hazard rate family, so $\bar{F}_{\mathbf{X}}(x_1, x_2) \geq \bar{F}_{X_1}(x_1)\bar{F}_{X_2}(x_2)$ for all x_1, x_2. Now, we can conclude that $F_{\mathbf{X}}(x_1, x_2) \geq F_{X_1}(x_1)F_{X_2}(x_2)$ for all x_1, x_2. So, the random vector $\mathbf{X}$ is a positive quadrant dependent. It should be noted that the last inequality is verified by the relation between the joint distribution function and the joint survival in the bivariate distributions as follows:

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

An essential and practical consequence of Part II is that for each pair of increasing functions $h_1(.)$ and $h_2(.)$ we conclude that $Cov(h_1(X_1), h_2(X_2)) > 0$ (see Barlow and Proschan (1981)). ■

Proposition 2.4.

Assume that the random vector $\mathbf{X}$ belongs to the bivariate proportional hazard rate family. Then, the random vector $\mathbf{X}$ has

I: The right tail increasing (RTI) property.

II: The right corner set increasing (RCSI) property.

Proof:

I: Since the random vector $\mathbf{X}$ belongs to the bivariate proportional hazard rate family, then, $P(X_i > x_i | X_j > x_j)$ is a non-decreasing in x_j for all $x_i > 0$ and $i \neq j$. Therefore, the desired result is obtained.

II: Since the random vector $\mathbf{X}$ belongs to the bivariate proportional hazard rate family, then, $P(X_1 > x_1, X_2 > x_2 | X_1 \geq \tilde{x}_1, X_2 \geq \tilde{x}_2)$ increases in $\tilde{x}_1, \tilde{x}_2$ for any value (x_1, x_2). Therefore, the desired result is obtained. ■

Proposition 2.5.

Assume that the random vector $\mathbf{X}$ belongs to the bivariate proportional hazard rate family. Then,

I: The random vector $\mathbf{X}$ has the multivariate increasing failure rate (MIFR) property.

II: The components of the bivariate hazard gradient are increasing functions of x_1 and x_2.

Proof:

I: To prove this property, we can see 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 $t > 0$ and $f_B(x_i, \alpha)/f_B(x_i + t, \alpha) < \bar{F}_B(x_i, \alpha)/\bar{F}_B(x_i + t, \alpha)$, $i = 1, 2$. So, this family of distributions has a MIFR property.

II: Using the structure provided for bivariate gradient by Johnson and Kotz (1975) and the fact that the random vector $\mathbf{X}$ belongs to the bivariate proportional hazard rate family, both the components of $h_{\mathbf{X}}(x_1, x_2) = (-\frac{\partial}{\partial x_1}, -\frac{\partial}{\partial x_2}) \ln P(X_1 > x_1, X_2 > x_2)$ are increasing functions of x_1 and x_2, for $f'_B(x_i, \alpha) > 0, i = 1, 2$ and all value of $x_1, x_2 > 0$. ∎

3. Generalization: Bivariate Proportional Hazard Rate–Geometric Models

In this section, the class of BPHR models is generalized, and the bivariate proportional hazard rate-geometric distributions are presented. For this extension, let $\{(X_{11}, X_{21}), \ldots, (X_{1n}, X_{2n})\}$ be independent and identically distributed (i.i.d) random variables with the joint distribution function $F_{\mathbf{X}}(.,.)$. Also, suppose N is a random variable independent of $\{(X_{11}, X_{21}), \ldots, (X_{1n}, X_{2n})\}$ that represents the number of failures, and it has a geometric distribution. Which is indicated by $N \sim Ge(\theta)$. In the following, we define the random variable $\mathbf{Y} = (Y_1, Y_2)$ as follows:

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

For given $N = n$, the joint survival function of $\mathbf{Y}$ can be computed as follows:

$$\bar{F}_{Y_1,Y_2|N}(y_1, y_2|n) = (\bar{F}_{\mathbf{X}}(y_1, y_2))^n. \tag{9}$$

Using Equation (9), the joint survival function of $\mathbf{Y}$ can be computed as follows:

$$\begin{aligned}\bar{G}_{\mathbf{Y}}(y_1, y_2) &= \sum_{n=1}^{\infty} P(Y_1 > y_1, Y_2 > y_2|N = n)P(N = n) \\ &= \sum_{n=1}^{\infty} [\bar{F}_{\mathbf{X}}(y_1, y_2)]^n \theta(1-\theta)^{n-1} = \frac{\theta \bar{F}_{\mathbf{X}}(y_1, y_2)}{1 - (1-\theta)\bar{F}_{\mathbf{X}}(y_1, y_2)}.\end{aligned} \tag{10}$$

Therefore, the joint survival function of $\mathbf{Y}$ can be written as follows:

$$\bar{G}_{\mathbf{Y}}(y_1, y_2) = \sum_{n=1}^{\infty} p_n \bar{F}_{BPHR}(y_1, y_2; \alpha, n\lambda_0, n\lambda_1, n\lambda_2), \tag{11}$$

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

So, the random vector $\mathbf{Y}$ has a bivariate F-geometric (BFG) distribution. Therefore, the marginal survival function Y_i is as follows:

$$\bar{F}_{Y_i}(y_i) = \frac{\theta \bar{F}_{X_i}(y_i)}{1 - (1-\theta)\bar{F}_{X_i}(y_i)}, \qquad i = 1, 2, \tag{12}$$

where $\bar{F}_{X_i}$ is the marginal survival functions of $\bar{F}_{\mathbf{X}}$. This method was presented in the univariate distributions by Marshall and Olkin (1997). In the proposed method, a parameter was added to the model, which resulted in greater flexibility. Many authors have used this method in a univariate case. But this generalization has not been examined for bivariate distributions.

Now suppose that the distribution F in Equation (10) belongs to the bivariate proportional hazard rate family then, the joint survival function of $\mathbf{Y}$ becomes:

$$\bar{G}_{\mathbf{Y}}(y_1,y_2)=\begin{cases}\dfrac{\theta\bar{F}_{PHRM}(y_1,\alpha,\lambda_0+\lambda_1)\bar{F}_{PHRM}(y_2,\alpha,\lambda_2)}{1-(1-\theta)\bar{F}_{PHRM}(y_1,\alpha,\lambda_0+\lambda_1)\bar{F}_{PHRM}(y_2,\alpha,\lambda_2)}, & \text{if } y_2\le y_1,\\[2ex] \dfrac{\theta\bar{F}_{PHRM}(y_1,\alpha,\lambda_1)\bar{F}_{PHRM}(y_2,\alpha,\lambda_0+\lambda_2)}{1-(1-\theta)\bar{F}_{PHRM}(y_1,\alpha,\lambda_1)\bar{F}_{PHRM}(y_2,\alpha,\lambda_0+\lambda_2)}, & \text{if } y_1< y_2.\end{cases} \tag{13}$$

Therefore, the random variable $\mathbf{Y}$ has a bivariate proportional hazard rate-geometric distribution. This family of new distributions has five parameters, and we represent it with the symbol $BPHRG(\theta,\alpha,\lambda_0,\lambda_1,\lambda_2)$.

Theorem 3.1.

Suppose $\mathbf{Y}\sim BPHRG(\theta,\alpha,\lambda_0,\lambda_1,\lambda_2)$. Then, the joint probability density function of $\mathbf{Y}$ is

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

where

$$\begin{aligned} g_1(y_1,y_2)=&\frac{\theta f_{PHRM}(y_1,\alpha,\lambda_0+\lambda_1)f_{PHRM}(y_2,\alpha,\lambda_2)}{[1-(1-\theta)\bar{F}_{PHRM}(y_1,\alpha,\lambda_0+\lambda_1)\bar{F}_{PHRM}(y_2,\alpha,\lambda_2)]^3}\\ &\times\{1+(1-\theta)\bar{F}_{PHRM}(y_1,\alpha,\lambda_0+\lambda_1)\bar{F}_{PHRM}(y_2,\alpha,\lambda_2)\},\end{aligned}$$

$$\begin{aligned} g_2(y_1,y_2)=&\frac{\theta f_{PHRM}(y_1,\alpha,\lambda_1)f_{PHRM}(y_2,\alpha,\lambda_0+\lambda_2)}{[1-(1-\theta)\bar{F}_{PHRM}(y_1,\alpha,\lambda_1)\bar{F}_{PHRM}(y_2,\alpha,\lambda_0+\lambda_2)]^3}\\ &\times\{1+(1-\theta)\bar{F}_{PHRM}(y_1,\alpha,\lambda_1)\bar{F}_{PHRM}(y_2,\alpha,\lambda_0+\lambda_2)\},\end{aligned}$$

$$g_0(y)=\frac{\lambda_0}{\lambda_0+\lambda_1+\lambda_2}\times\frac{\theta f_{PHRM}(y,\alpha,\lambda_0+\lambda_1+\lambda_2)}{[1-(1-\theta)\bar{F}_{PHRM}(y,\alpha,\lambda_0+\lambda_1+\lambda_2)]^2}.$$

Proof:

We obtain the expressions $g_1(y_1,y_2)$ and $g_2(y_1,y_2)$ through $-\frac{\partial^2\bar{G}_{\mathbf{Y}}(y_1,y_2)}{\partial y_1\partial y_2}$ for $y_2\neq y_1$ and the expression $g_0(x)$ we cannot use the same method and to calculate it we need to consider the following relation:

$$\int_0^\infty g_0(y)dy+\int_0^\infty\int_{y_1}^\infty g_2(y_1,y_2)dy_2dy_1+\int_0^\infty\int_{y_2}^\infty g_1(y_1,y_2)dy_1dy_2=1.$$

We compute,

$$\int_0^\infty\int_{x_2}^\infty g_1(y_1,y_2)dy_1dy_2=\lambda_2\int_0^\infty\frac{\theta[\bar{F}_B(y_2,\alpha)]^{\lambda_0+\lambda_1+\lambda_2-1}f_B(y_2,\alpha)}{[1-(1-\theta)[\bar{F}_B(y_2,\alpha)]^{\lambda_0+\lambda_1+\lambda_2]^2}}dy_2=\frac{\lambda_2}{\lambda_0+\lambda_1+\lambda_2},$$

and similarly,

$$\int_0^\infty \int_{x_2}^\infty g_1(y_1,y_2)dy_2dy_1 = \lambda_1 \int_0^\infty \frac{\theta[\bar{F}_B(y_1,\alpha)]^{\lambda_0+\lambda_1+\lambda_2-1} f_B(y_1,\alpha)}{[1-(1-\theta)[\bar{F}_B(y_1,\alpha)]^{\lambda_0+\lambda_1+\lambda_2}]^2} dy_1 = \frac{\lambda_1}{\lambda_0+\lambda_1+\lambda_2}.$$

Therefore, we should have:

$$\int_0^\infty g_0(y)dy = \lambda_0 \int_0^\infty \frac{\theta f_{PHRM}(y,\alpha,\lambda_0+\lambda_1+\lambda_2)}{(\lambda_0+\lambda_1+\lambda_2)[1-(1-\theta)\bar{F}_{PHRM}(y,\alpha,\lambda_0+\lambda_1+\lambda_2)]^2} dy = \frac{\lambda_0}{\lambda_0+\lambda_1+\lambda_2}.$$

So, the desired result is achieved. ■

Using Equation (11) and Theorem 3.1, the following relation can be immediately obtained for the joint probability density function of $\mathbf{Y}$.

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

where

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

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

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

and $p_n = P(N=n) = \theta(1-\theta)^{n-1}$ and $f_{PHRM}(.,n\lambda)$ is the probability density function of PHR models. It should be noted, if the independent random variables U_i's belong to the proportional hazard rate family with parameters α and λ, then the random variable $\min(U_1,\ldots,U_n) \sim f_{PHRM}(.,n\lambda)$.

The joint probability density function of the BPHRG distribution provided in Theorem 3.1 can be rewritten as

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

here

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

and

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

where, $g_a(y_1,y_2)$ represents the absolutely continuous part, and $g_s(y)$ denotes the singular part. Now, if $\lambda_0 = 0$, then the distribution function of this family will be absolutely continuous. Also, if

$\theta = 1$, the bivariate proportional hazard rate distributions family (BPHR) will be a particular state of the bivariate proportional hazard rate-geometric distribution family (BPHRG).

Theorem 3.2.

Suppose $\mathbf{Y} \sim BPHRG(\theta, \alpha, \lambda_0, \lambda_1, \lambda_2)$ and $N \sim Ge(\theta)$. Then, the joint probability density function of $\mathbf{Y}$ and N can be obtained as follows:

$$f_{Y_1,Y_2,N}(y_1, y_2, n) = \begin{cases} \theta(1-\theta)^{n-1} f_{1n}(y_1, y_2), & \text{if } y_2 < y_1, \\ \theta(1-\theta)^{n-1} f_{2n}(y_1, y_2), & \text{if } y_1 < y_2, \\ \theta(1-\theta)^{n-1} f_{0n}(y), & \text{if } y_1 = y_2 = y, \end{cases} \tag{14}$$

where,

$$\begin{aligned} f_{1n}(y_1, y_2) &= n^2 f_{PHRM}(y_1, \alpha, \lambda_0 + \lambda_1) f_{PHRM}(y_2, \alpha, \lambda_2) \\ &\quad \times [\bar{F}_{PHRM}(y_1, \alpha, \lambda_0 + \lambda_1) \bar{F}_{PHRM}(y_2, \alpha, \lambda_2)]^{n-1}, \end{aligned}$$

$$\begin{aligned} f_{2n}(y_1, y_2) &= n^2 f_{PHRM}(y_1, \alpha, \lambda_1) f_{PHRM}(y_2, \alpha, \lambda_0 + \lambda_2) \\ &\quad \times [\bar{F}_{PHRM}(y_1, \alpha, \lambda_1) \bar{F}_{PHRM}(y_2, \alpha, \lambda_0 + \lambda_2)]^{n-1}, \end{aligned}$$

$$\begin{aligned} f_{0n}(y) &= \frac{\lambda_0}{\lambda_0 + \lambda_1 + \lambda_2} n f_{PHRM}(y, \alpha, \lambda_0 + \lambda_1 + \lambda_2) \\ &\quad \times [\bar{F}_{PHRM}(y, \alpha, \lambda_0 + \lambda_1 + \lambda_2)]^{n-1}. \end{aligned} \tag{15}$$

Proof:

Note that for the random vector $\mathbf{Y}$, we have

$$\begin{aligned} &P(Y_1 > y_1, Y_2 > y_2, N = n) = P(Y_1 > y_1, Y_2 > y_2 | N = n) P(N = n) \\ &= \{\bar{F}(y_1, y_2)\}^n \theta(1-\theta)^{n-1} \\ &= \begin{cases} \theta(1-\theta)^{n-1} \{\bar{F}_{PHRM}(y_1, \alpha, \lambda_0 + \lambda_1) \bar{F}_{PHRM}(y_2, \alpha, \lambda_2)\}^n, & \text{if } y_2 \le y_1 \\ \theta(1-\theta)^{n-1} \{\bar{F}_{PHRM}(y_1, \alpha, \lambda_1) \bar{F}_{PHRM}(y_2, \alpha, \lambda_0 + \lambda_2)\}^n, & \text{if } y_1 < y_2. \end{cases} \end{aligned}$$

Therefore, the desired joint function is obtained by simple calculations. ■

Theorem 3.3.

Suppose $\mathbf{Y} \sim BPHRG(\theta, \alpha, \lambda_0, \lambda_1, \lambda_2)$ and $N \sim Ge(\theta)$. Then, the conditional probability mass function of N given $Y_1 = y_1$ and $Y_2 = y_2$ can be obtained as follows:

$$f_N(n|y_1,y_2) = \begin{cases} c_1(y_1,y_2)n^2(1-\theta)^{n-1}[\bar{F}_{PHRM}(y_1,\alpha,\lambda_0+\lambda_1)\bar{F}_{PHRM}(y_2,\alpha,\lambda_2)]^{n-1}, & \text{if } y_2 < y_1, \\ c_2(y_1,y_2)n^2(1-\theta)^{n-1}[\bar{F}_{PHRM}(y_1,\alpha,\lambda_1)\bar{F}_{PHRM}(y_2,\alpha,\lambda_0+\lambda_2)]^{n-1}, & \text{if } y_1 < y_2, \\ c_0(y)n(1-\theta)^{n-1}[\bar{F}_{PHRM}(y,\alpha,\lambda_0+\lambda_1+\lambda_2)]^{n-1}, & \text{if } y_1 = y_2 = y, \end{cases} \tag{16}$$

where,

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

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

$$c_0(y) = [1-(1-\theta)\bar{F}_{PHRM}(y,\alpha,\lambda_0+\lambda_1+\lambda_2)]^2.$$

Proof:

To prove this case, the structure of the conditional probability mass function must be used. The conditional probability mass function of N given $Y_1 = y_1$ and $Y_2 = y_2$ is $f_N(n|y_1,y_2) = \frac{f_{Y_1,Y_2,N}(y_1,y_2,n)}{f_{Y_1,Y_2}(y_1,y_2)}$. Now the numerator and the denominator are replaced by the results presented in Theorems 3.2 and 3.1, respectively. Finally, using simple calculations and simplifications, the result is obtained. ∎

So, the conditional probability mass function expressed in Equation (16) can be rewritten as follows:

$$f_N(n|y_1,y_2) = \begin{cases} \dfrac{[1-\xi_1(y_1,y_2,\theta,\gamma)]^3}{[1+\xi_1(y_1,y_2,\theta,\gamma)]} n^2\xi_1^{n-1}(y_1,y_2,\theta,\gamma), & \text{if } y_2 < y_1, \\ \dfrac{[1-\xi_1(y_1,y_2,\theta,\gamma)]^3}{[1+\xi_2(y_1,y_2,\theta,\gamma)]} n^2\xi_2^{n-1}(y_1,y_2,\theta,\gamma), & \text{if } y_1 < y_2, \\ [1-\xi_0(y_1,y_2,\theta,\gamma)]^2 n\xi_0^{n-1}(y_1,y_2,\theta,\gamma), & \text{if } 0 < y_1 = y_2 = y, \end{cases}$$

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

$$\xi_1(y_1,y_2,\theta,\gamma) = (1-\theta)\bar{F}_{PHRM}(y_1,\alpha,\lambda_0+\lambda_1)\bar{F}_{PHRM}(y_2,\alpha,\lambda_2),$$

$$\xi_2(y_1,y_2,\theta,\gamma) = (1-\theta)\bar{F}_{PHRM}(y_1,\alpha,\lambda_1)\bar{F}_{PHRM}(y_2,\alpha,\lambda_0+\lambda_2),$$

$$\xi_0(y_1,y_2,\theta,\gamma) = (1-\theta)\bar{F}_{PHRM}(y,\alpha,\lambda_0+\lambda_1+\lambda_2).$$

Proposition 3.1.

Suppose $\mathbf{Y} \sim BPHRG(\theta, \alpha, \lambda_0, \lambda_1, \lambda_2)$ and $N \sim Ge(\theta)$. Then, the conditional expectation of N given $Y_1 = y_1$ and $Y_2 = y_2$ is

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

Proof:

This result can be obtained using the conditional function presented in Theorem 3.3 and the structure of conditional expectation. For $y_2 < y_1$ and $y_1 = y_2 = y$, we express the calculations, and for $y_1 < y_2$ it can be calculated in a similar way. For $y_2 < y_1$, we have:

$$\begin{aligned} E(N|y_1, y_2) &= \sum_{n=1}^{\infty} n f_N(n|y_1, y_2) = \sum_{n=1}^{\infty} n^3 \frac{(1-\xi_1(y_1, y_2, \theta, \gamma))^3}{1+\xi_1(y_1, y_2, \theta, \gamma)} (\xi_1(y_1, y_2, \theta, \gamma))^{n-1} \\ &= \frac{(1-\xi_1(y_1, y_2, \theta, \gamma))^3}{1+\xi_1(y_1, y_2, \theta, \gamma)} \sum_{n=1}^{\infty} n^3 (\xi_1(y_1, y_2, \theta, \gamma))^{n-1} \\ &= \frac{(1-\xi_1(y_1, y_2, \theta, \gamma))^3}{1+\xi_1(y_1, y_2, \theta, \gamma)} \times \frac{(1-\xi_1(y_1, y_2, \theta, \gamma))^2 - 6(1-\xi_1(y_1, y_2, \theta, \gamma) + 6}{(1-\xi_1(y_1, y_2, \theta, \gamma))^4} \\ &= \frac{(1-\xi_1(y_1, y_2, \theta, \gamma))^2 - 6(1-\xi_1(y_1, y_2, \theta, \gamma) + 6}{1-(\xi_1(y_1, y_2, \theta, \gamma))^2}. \end{aligned}$$

Also, we can show that for the case of $y_2 = y_1 = y$:

$$\begin{aligned} E(N|y_1, y_2) &= \sum_{n=1}^{\infty} n f_N(n|y_1, y_2) = \sum_{n=1}^{\infty} n^2 (1-\xi_0(y_1, y_2, \theta, \gamma))^2 (\xi_0(y_1, y_2, \theta, \gamma))^{n-1} \\ &= (1-\xi_0(y_1, y_2, \theta, \gamma))^2 \sum_{n=1}^{\infty} n^2 (1-\xi_0(y_1, y_2, \theta, \gamma))^{n-1} \\ &= (1-\xi_0(y_1, y_2, \theta, \gamma))^2 \frac{1+\xi_0(y_1, y_2, \theta, \gamma)}{(1-\xi_0(y_1, y_2, \theta, \gamma))^3} = \frac{1+\xi_0(y_1, y_2, \theta, \gamma)}{1-\xi_0(y_1, y_2, \theta, \gamma)}. \end{aligned}$$ ■

Proposition 3.2.

Let $\mathbf{Y} \sim BPHRG(\theta, \alpha, \lambda_0, \lambda_1, \lambda_2)$. Then,

I: Each Y_i belongs to the univariate proportional hazard rate-geometric family (UPHRG) with parameters $\alpha, \lambda_0 + \lambda_i$ and θ.

II: The random variable $Y = \min(Y_1, Y_2)$ belongs to the UPHRG family with parameters $\sum_{i=0}^{2} \lambda_i$, α, and θ.
III: $P(Y_1 < Y_2) = \frac{\lambda_1}{\lambda_0+\lambda_1+\lambda_2}$.

Proof:

To illustrate these parts, we can use the relationships described in the article and the basic expressions.

I: To prove this part, we use Equation (12). The function $\bar{F}_{X_i}$ in the Equation (12) must be calculated as follows:

$$\begin{aligned}\bar{F}_{\mathbf{X}_1}(y_1) &= P(\min\{T_0, T_1\} > y_1) = P(T_0 > y_1, T_1 > y_1)\\ &= \bar{F}_{PHRM}(y_1, \alpha, \lambda_0)\bar{F}_{PHRM}(y_1, \alpha, \lambda_1) = \bar{F}_{PHRM}(y_1, \alpha, \lambda_0 + \lambda_1).\end{aligned}$$

II: To prove this part, we use Equation (10). The function $\bar{F}_{\mathbf{X}}$ in the Equation (10) has the univariate proportional hazard rate models with parameters α and $\lambda_0 + \lambda_1 + \lambda_2$.
III: The proof of this part is as follows:

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

4. Estimation of Parameters

In this section, we estimate the parameters of the bivariate proportional hazard rate-geometric distributions family by the maximum likelihood method. But it can be seen that this method cannot obtain explicit expressions for parameter estimations. Therefore, we recommend using an EM algorithm to estimate our parameters. Now, we describe this algorithm in detail.

To estimate the parameters of this family, suppose that $\{(y_{11}, y_{21}), \ldots, (y_{1m}, y_{2m})\}$ is a random sample from the bivariate proportional hazard rate-geometric distributions family. We also know that this family has five parameters as $\Theta = (\theta, \alpha, \lambda_0, \lambda_1, \lambda_2)$. Next, consider 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 $m = m_0 + m_1 + m_2$.

Given the notation above and the functions $g_0(y)$, $g_1(y_1, y_2)$ and $g_2(y_1, y_2)$ which were presented in Theorem 3.1, we can obtain the maximum likelihood function as follows:

$$\ell(\Theta) = \sum_{i \in I_0} \ln g_0(y_i) + \sum_{i \in I_1} \ln g_1(y_{1i}, y_{2i}) + \sum_{i \in I_2} \ln g_2(y_{1i}, y_{2i}), \tag{17}$$

To estimate the parameters in this method, we need to maximize the likelihood function $\ell(\Theta)$ in (17) in terms of parameters. Given the likelihood function, it can be seen that explicit expressions for parameter estimation cannot be obtained. Therefore, to obtain parameter estimates, five non-linear equations must be solved simultaneously. This is a complicated procedure, and also it is very difficult to select the initial values for each parameter to converge algorithms such as Newton-Rephson or Gauss-Newton. Therefore, in this case, the use of the EM algorithm for parameter estimation is suggested, and it can be seen that the proposed algorithm performs well.

We use the idea of the missing data problem for this algorithm. For this purpose, using the joint probability density function of (Y_1, Y_2, N) which is presented in Equation (14), we can conclude that

$$(Y_1, Y_2|N) \sim BPHR(\alpha, n\lambda_0, n\lambda_1, n\lambda_2).$$

So, let's take a random sample of complete data as $\{(y_{1i}, y_{2i}, n_i); i = 1, ..., m\}$. In the following, the conditional likelihood function is obtained using the functions f_{0n_i}, f_{1n_i}, and f_{2n_i} which were introduced in Equation (15) as follows:

$$\ell_1(\alpha, \lambda_0, \lambda_1, \lambda_2) = \sum_{i \in I_0} \ln f_{0n_i}(y_i) + \sum_{i \in I_1} \ln f_{1n_i}(y_{1i}, y_{2i}) + \sum_{i \in I_2} \ln f_{2n_i}(y_{1i}, y_{2i}).$$

Then, the parameter estimation is computed by maximizing the conditional log-likelihood function. To investigate the missing data by the EM algorithm, we need to consider the new random variable. For given $N = n$, the independent random variables $\{U_i|N = n\}$ for $i = 0, 1, 2$ belong to the proportional hazard rate family with parameters α and $n\lambda_i$. Therefore,

$$\{U_i|N = n\} \sim PHRM(\alpha, n\lambda_i), \qquad i = 0, 1, 2. \tag{18}$$

Hence,

$$\{Y_i|N = n\} = \min\{T_0, T_i\}|N = n, \qquad i = 1, 2.$$

Now, a corresponding random vector for $\mathbf{Y} = (Y_1, Y_2)$ can be defined as follows:

$$(\Delta_1, \Delta_2) = \begin{cases} (0,0), & \text{if } Y_1 = T_0, Y_2 = T_0\ , \\ (0,1), & \text{if } Y_1 = T_0, Y_2 = T_2\ , \\ (1,0), & \text{if } Y_1 = T_1, Y_2 = T_0\ , \\ (1,1), & \text{if } Y_1 = T_1, Y_2 = T_2\ . \end{cases} \tag{19}$$

Now to continue the computation, we must consider a sample size m of the complete observations from $(Y_1, Y_2, \Delta_1, \Delta_2, N)$. Then, the parameter estimation in this method is obtained by one-dimensional optimization.

As mentioned, a vector (Δ_1, Δ_2) is related to every (Y_1, Y_2). But it may not always be known. Note that, if $Y_1 = Y_2$, then $\Delta_1 = \Delta_2 = 0$, is known. If $Y_1 > Y_2$ or $Y_2 > Y_1$, then (Δ_1, Δ_2) is not known and is missing. In other words, 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 values of (Δ_1, Δ_2) are (1, 0) or (1, 1), with positive probabilities, see for example Kundu and Dey (2009). Using this information we can describe how the EM algorithm is implemented. In this algorithm, the conditional 'pseudo' log-likelihood function must first be computed. We compute this function by conditioning on N, and the next step, N will be replaced through $E(N|Y_1, Y_2)$. This step of the algorithm is called "E-step". In "E-step," we should consider the following points:

I: If all observations belong to I_0 because the associated (Δ_1, Δ_2) with them is entirely known, then the log-likelihood contribution is complete. Then, the log-likelihood contribution of an observation $y \in I_0$ can be obtained as follows:

$$\ln n + \ln \lambda_0 + [n(\lambda_0 + \lambda_1 + \lambda_2) - 1] \ln \bar{F}_B(y, \alpha) + \ln f_B(y, \alpha).$$

As mentioned, n and $\ln n$ are missing and at this point need to be replaced by $E(N|y_1, y_2)$ and $E(\ln N|y_1, y_2)$, respectively.

II: If observations belong to I_1 because the associated (Δ_1, Δ_2) with them is unknown, therefore, all observations are missing. Then, "pseudo-observations" are created. In this case, (y_1, y_2) should be divided to two partially complete "pseudo-observations" of the form $(y_1, y_2, u_i(\Theta))$ for $i = 1, 2$, where $u_i(\Theta)$ are the conditional probabilities that (Δ_1, Δ_2) takes the values of (0,1) and (1,1), respectively. Therefore,

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

So, the "pseudo" log-likelihood contribution of an observation $(y_1, y_2) \in I_1$ becomes

$$\begin{aligned} &2 \ln n + \ln \lambda_2 + u_1 \ln \lambda_0 + u_2 \ln \lambda_1 + (n\lambda_2 - 1) \ln \bar{F}_B(y_2, \alpha) + \ln f_B(y_2, \alpha) \\ &\quad + [n(\lambda_0 + \lambda_1) - 1] \ln \bar{F}_B(y_1, \alpha) + \ln f_B(y_1, \alpha). \end{aligned}$$

As before, the missing values of n and $\ln n$ are replaced.

III: If observations belong to I_2 because the associated (Δ_1, Δ_2) with them is unknown, therefore, all observations are missing. Then, "pseudo-observations" are created. In this case, (y_1, y_2) should be divided to two partially complete "pseudo-observations" of the form $(y_1, y_2, v_i(\Theta))$ for $i = 1, 2$, where $v_i(\Theta)$ are the conditional probabilities that (Δ_1, Δ_2) takes the values of (1,0) and (1,1), respectively. Then,

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

As before, the "pseudo" log-likelihood contribution of an observation $(y_1, y_2) \in I_2$ can be computed

$$\begin{aligned} &2 \ln n + \ln \lambda_1 + v_1 \ln \lambda_0 + v_2 \ln \lambda_2 + (n\lambda_1 - 1) \ln \bar{F}_B(y_1, \alpha) + \ln f_B(y_1, \alpha) \\ &\quad + [n(\lambda_0 + \lambda_2) - 1] \ln \bar{F}_B(y_2, \alpha) + \ln f_B(y_2, \alpha). \end{aligned}$$

Similar to the previous entries, the missing values of n and $\ln n$ should be replaced.

For a better understanding of this step of the algorithm, see Dinse (1982) or Kundu (2004) articles. We can also use the observations in Table 2 for more comfortable and better implementation of

Table 2. All possible states for observations (Δ_1, Δ_2), and associated probabilities

Different Case	probability	(Δ_1, Δ_2)	$Y_1 \& Y_2$	Set
$T_0 < T_1 < T_2$	$\frac{\lambda_1\lambda_0}{(\lambda_1+\lambda_2)(\lambda_0+\lambda_1+\lambda_2)}$	(0, 0)	$Y_1 = Y_2$	I_0
$T_0 < T_2 < T_1$	$\frac{\lambda_2\lambda_0}{(\lambda_1+\lambda_2)(\lambda_0+\lambda_1+\lambda_2)}$	(0, 0)	$Y_1 = Y_2$	I_0
$T_1 < T_0 < T_2$	$\frac{\lambda_1\lambda_0}{(\lambda_0+\lambda_2)(\lambda_0+\lambda_1+\lambda_2)}$	(1, 0)	$Y_1 < Y_2$	I_2
$T_1 < T_2 < T_0$	$\frac{\lambda_1\lambda_2}{(\lambda_0+\lambda_2)(\lambda_0+\lambda_1+\lambda_2)}$	(1, 1)	$Y_1 < Y_2$	I_2
$T_2 < T_0 < T_1$	$\frac{\lambda_0\lambda_2}{(\lambda_0+\lambda_1)(\lambda_0+\lambda_1+\lambda_2)}$	(0, 1)	$Y_2 < Y_1$	I_1
$T_2 < T_1 < T_0$	$\frac{\lambda_1\lambda_2}{(\lambda_0+\lambda_1)(\lambda_0+\lambda_1+\lambda_2)}$	(1, 1)	$Y_2 < Y_1$	I_1

the "E"-step algorithm. Therefore, the steps of the proposed EM algorithm can be presented as follows:

E-Step: The "pseudo" log-likelihood function in the k-th step is as follows:

$$\begin{aligned}
\ell_{pseudo}(\Theta) &= constant + (m_0 + 2m_1 + 2m_2)\ln\lambda_0 + (m_2 + m_1 u_2^{(k)})\ln\lambda_1 + (m_2 v_2^{(k)} + m_1)\ln\lambda_2 \\
&+ \lambda_0\{\sum_{i\in I_0} a_i^{(k)} \ln \bar{F}_B(y_i,\alpha) + \sum_{i\in I_2} a_i^{(k)} \ln \bar{F}_B(y_{2i},\alpha) + \sum_{i\in I_1} a_i^{(k)} \ln \bar{F}_B(y_{1i},\alpha)\} \\
&+ \lambda_1\{\sum_{i\in I_0} a_i^{(k)} \ln \bar{F}_B(y_i,\alpha) + \sum_{i\in I_2} a_i^{(k)} \ln \bar{F}_B(y_{1i},\alpha) + \sum_{i\in I_1} a_i^{(k)} \ln \bar{F}_B(y_{1i},\alpha)\} \\
&+ \lambda_2\{\sum_{i\in I_0} a_i^{(k)} \ln \bar{F}_B(y_i,\alpha) + \sum_{i\in I_2} a_i^{(k)} \ln \bar{F}_B(y_{2i},\alpha) + \sum_{i\in I_1} a_i^{(k)} \ln \bar{F}_B(y_{2i},\alpha)\} \\
&+ \sum_{i\in I_0} \ln\frac{f_B(y_i,\alpha)}{\bar{F}_B(y_i,\alpha)} + \sum_{i\in I_2} \ln\frac{f_B(y_{1i},\alpha)}{\bar{F}_B(y_{1i},\alpha)} + \sum_{i\in I_2} \ln\frac{f_B(y_{2i},\alpha)}{\bar{F}_B(y_{2i},\alpha)} \\
&+ \sum_{i\in I_1} \ln\frac{f_B(y_{2i},\alpha)}{\bar{F}_B(y_{2i},\alpha)} + \sum_{i\in I_1} \ln\frac{f_B(y_{1i},\alpha)}{\bar{F}_B(y_{1i},\alpha)} + m\ln\frac{\theta}{1-\theta} + \ln(1-\theta)\sum_{i=1}^{m} a_i^{(k)}, \qquad (20)
\end{aligned}$$

where, $\Theta^{(k)} = (\alpha^{(k)}, \lambda_0^{(k)}, \lambda_1^{(k)}, \lambda_2^{(k)}), E(N|y_{1i}, y_{2i}, \Theta^{(k)}) = a_i^{(k)}, u_1(\Theta^{(k)}) = u_1^{(k)}, u_2(\Theta^{(k)}) = u_2^{(k)}, v_1(\Theta^{(k)}) = v_1^{(k)}$ and $v_2(\Theta^{(k)}) = v_2^{(k)}$.

M-Step: The "M-step" in the EM algorithm involves maximizing the 'pseudo' log-likelihood function in terms of parameters. Therefore, if we assume that α is fixed, then the estimation of the parameters that maximize Equation (20) is obtained as follows:

$$\widehat{\lambda}_0(\alpha) = -\frac{m_0 + m_2 v_1^{(k)} + m_1 u_1^{(k)}}{\sum_{i\in I_0} a_i^{(k)} \ln \bar{F}_B(y_i,\alpha) + \sum_{i\in I_2} a_i^{(k)} \ln \bar{F}_B(y_{2i},\alpha) + \sum_{i\in I_1} a_i^{(k)} \ln \bar{F}_B(y_{1i},\alpha)}, \qquad (21)$$

$$\widehat{\lambda}_1(\alpha) = -\frac{m_2 + m_1 u_2^{(k)}}{\sum_{i\in I_0} a_i^{(k)} \ln \bar{F}_B(y_i,\alpha) + \sum_{i\in I_2} a_i^{(k)} \ln \bar{F}_B(y_{1i},\alpha) + \sum_{i\in I_1} a_i^{(k)} \ln \bar{F}_B(y_{1i},\alpha)}, \quad (22)$$

$$\widehat{\lambda}_2(\alpha) = -\frac{m_1 + m_2 v_2^{(k)}}{\sum_{i\in I_0} a_i^{(k)} \ln \bar{F}_B(y_i,\alpha) + \sum_{i\in I_2} a_i^{(k)} \ln \bar{F}_B(y_{2i},\alpha) + \sum_{i\in I_1} a_i^{(k)} \ln \bar{F}_B(y_{2i},\alpha)}, \quad (23)$$

and $\hat{\theta}$ is

$$\hat{\theta} = \frac{m}{\sum_{i=1}^{m} a_i^{(k)}}. \quad (24)$$

Finally, to estimate the parameter α, we need to maximize the pseudo-profile log-likelihood function in terms of parameter α. For this purpose, a nonlinear equation with respect to α can be solved. In other words, the $\hat{\alpha}$ can be computed by solving the following Equation,

$$\frac{\partial}{\partial\alpha}\ell_{pseudo}(\alpha, \widehat{\lambda}_0(\alpha), \widehat{\lambda}_1(\alpha), \widehat{\lambda}_2(\alpha)) = 0. \quad (25)$$

It should be noted that, if we obtain the second derivative of the 'pseudo" log-likelihood function with respect to α, it is seen that the "pseudo" log-likelihood function is a unimodal function, so the estimation of the parameter α is obtained uniquely.

Finally, the proposed EM algorithm for estimating the parameters of this family of distributions is as follows:

ALGORITHM

- Step 1: Select the initial values for each parameter, 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: Calculate each of the values of u_i, and v_i for $i = 1, 2$.
- Step 4: Determine the value of $\hat{\alpha}$ by solving Equation (25), and say $\hat{\alpha}^{(1)}$.
- Step 5: Calculate $\hat{\lambda}_i^{(1)} = \hat{\lambda}_i(\hat{\alpha}^{(1)})$, $i = 0, 1, 2$ using Equations (21)-(23).
- Step 6: Determine the value of $\hat{\theta}$ using Equation (24).
- Step 7: Replace $\Theta^{(0)}$ by $\Theta^{(1)} = (\theta^{(1)}, \alpha^{(1)}, \lambda_0^{(1)}, \lambda_1^{(1)}, \lambda_2^{(1)})$ then, go back to step 1 and repeat the process to converge the algorithm.

5. Data Analysis and Comparison Study

In this section, we investigate the family of introduced distributions as well as the efficiency of the proposed EM algorithm for parameter estimations. In this regard, the Monte-Carlo simulation method and three real datasets are used.

5.1. Numerical Experiments

At the beginning of this section, the Monte-Carlo simulation studies are used to evaluate the efficiency of the proposed algorithm. As we know, the proportional hazard rate-geometric distribution family contains different distributions. So for better consideration, we use the different distributions in this family. Also, the following algorithm can be used to simulate this family.

ALGORITHM

- Generate a value of n from $Ge(\theta)$.
- Generate $\{u_{i1}, \ldots, u_{in}\}$ from $PHRM(\alpha, \lambda_i)$, for $i = 0, 1, 2$.
- Obtain $x_{ik} = \min\{u_{0k}, u_{ik}\}$ for $k = 1, \ldots, n$ and $i = 1, 2$.
- Compute the desired (y_1, y_2) as $y_i = \min\{x_{i1}, \ldots, x_{in}\}$ for $i = 1, 2$.

As mentioned, this family comprises a large class of distributors. Different models of this family can be considered to simulate this family of distributions. We examine four sub-models in this simulation, which are: (i) Bivariate Weibull-geometric (BWG), (ii) Bivariate Lomax-geometric (BLG), (iii) Bivariate Chen-geometric (BCHG) and (iv) Bivariate Gompertz-geometric (BGG). Therefore, the estimate of parameters is computed by the presented EM algorithm in Section 4. We use the following parameter values for our simulation studies:

$$\Theta_i = (\theta_i, \alpha, \lambda_0, \lambda_1, \lambda_2) = (\theta_i, 3, 1, 1, 1), \qquad i = 1, 2, 3,$$

where $\theta_i = 0.3, 0.5, 0.7$ for $i = 1, 2, 3$, respectively. To start the implementation of the EM algorithm, consider the initial values as follows:

$$\Theta_i^{(0)} = (\theta_i^{(0)}, \alpha^{(0)}, \lambda_0^{(0)}, \lambda_1^{(0)}, \lambda_2^{(0)}) = (\theta_i^{(0)}, 1, 1, 1, 1), \qquad i = 1, 2, 3,$$

where, $\theta_i^{(0)} = 0.2, 0.4, 0.6$ for $i = 1, 2, 3$, respectively. It should also be noted that when the absolute value of the difference between the sequence of values of two steps is less than 10^{-5}, the repetition of the algorithm stops. This process is repeated 1000 times and the absolute value of the bias estimates and the associated mean squared errors (MSEs) are calculated at each iteration. The results of these simulations are presented in Table 3. The results of this simulation study are useful, and it is observed that as n increases, the biases and the MSEs decrease. This represents that the maximum likelihood estimators are consistent.

5.2. Data Analysis

In the second part of this section, we examine three real datasets. Similar to the previous section, four different sub-models of this family are considered. These four distributions are fitted to these three real datasets. The parameter estimates in these distributions are calculated using the maximum likelihood method and the proposed EM algorithm. For each of these distributions, the parameter estimates, the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), the Kolmogorov-Smirnov (K-S) test and the corresponding p-values are calculated. We also

Table 3. Simulation study results (|Bias| and MSE) for four sub-models of the BPHRG family

Θ	n	α		λ_0		λ_1		λ_2		θ	
		$\|Bias\|$	MSE	$\|Bias\|$	MSE	$\|Bias\|$	MSE	$\|Bias\|$	MSE	$\|Bias\|$	MSE
BWG											
	50	0.0480	0.0385	0.3102	0.1105	0.3014	0.1144	0.3070	0.1182	0.1031	0.0107
Θ_1	100	0.0305	0.0195	0.2143	0.1054	0.2288	0.1112	0.2287	0.1175	0.1006	0.0098
	200	0.0282	0.0130	0.2010	0.1075	0.2008	0.1164	0.2052	0.1195	0.0098	0.0108
	50	0.0454	0.0406	0.1798	0.0487	0.1731	0.0545	0.1801	0.0574	0.1165	0.0139
Θ_2	100	0.0260	0.0198	0.1786	0.0399	0.1641	0.0491	0.1711	0.0514	0.1140	0.0121
	200	0.0260	0.0125	0.1781	0.0372	0.1521	0.0444	0.1539	0.0452	0.1129	0.0120
	50	0.0683	0.0395	0.1031	0.0293	0.1089	0.0354	0.1208	0.0392	0.1164	0.0143
Θ_3	100	0.0652	0.0231	0.1029	0.0215	0.1079	0.0269	0.1162	0.0257	0.1128	0.0135
	200	0.0570	0.0143	0.0096	0.0186	0.0095	0.0218	0.1105	0.0228	0.1096	0.0136
BLG											
	50	0.2062	1.3807	0.1482	0.1734	0.1452	0.1909	0.1516	0.1929	0.1071	0.0115
Θ_1	100	0.1846	0.9930	0.1396	0.1257	0.1374	0.1503	0.1401	0.1461	0.1070	0.0095
	200	0.1543	0.7011	0.1299	0.0094	0.1346	0.1106	0.1355	0.1131	0.1052	0.0095
	50	0.3913	0.8319	0.1078	0.1095	0.0245	0.1186	0.0266	0.1127	0.1232	0.0153
Θ_2	100	0.3275	0.5999	0.0483	0.0546	0.0514	0.0616	0.0547	0.0661	0.1137	0.0154
	200	0.2992	0.4588	0.0394	0.0323	0.0408	0.0415	0.0409	0.0399	0.1037	0.0154
	50	0.4046	0.6781	0.0232	0.0563	0.0190	0.0643	0.0200	0.0653	0.1250	0.0159
Θ_3	100	0.4022	0.4857	0.0090	0.0321	0.0096	0.0423	0.0094	0.0409	0.1245	0.0157
	200	0.4007	0.4143	0.0066	0.0255	0.0059	0.0316	0.0094	0.0306	0.1046	0.0157
BCHG											
	50	0.0489	0.0487	0.3072	0.1112	0.2970	0.1163	0.3027	0.1202	0.1032	0.0107
Θ_1	100	0.0314	0.0244	0.2166	0.1079	0.2298	0.1207	0.2301	0.1203	0.1027	0.0108
	200	0.0307	0.0162	0.2045	0.1101	0.2037	0.1191	0.2081	0.1121	0.1019	0.0108
	50	0.0537	0.0479	0.1710	0.0489	0.1640	0.0553	0.1732	0.0573	0.1169	0.0140
Θ_2	100	0.0361	0.0243	0.1547	0.0394	0.1582	0.0469	0.1528	0.0472	0.1127	0.0140
	200	0.0295	0.0162	0.1502	0.0375	0.1478	0.0442	0.1411	0.0450	0.1081	0.0140
	50	0.0916	0.0520	0.0870	0.0281	0.0872	0.0355	0.0861	0.0375	0.1186	0.0148
Θ_3	100	0.0663	0.0268	0.0831	0.0197	0.0783	0.0246	0.0845	0.0258	0.1091	0.0146
	200	0.0652	0.0202	0.0717	0.0170	0.0700	0.0209	0.0785	0.0210	0.0098	0.0140
BGG											
	50	0.5230	2.1612	0.1979	1.8370	0.2275	1.5689	0.2285	1.5304	0.0736	0.0057
Θ_1	100	0.4391	1.0392	0.1635	0.1999	0.1763	0.1946	0.1756	0.1937	0.0732	0.0055
	200	0.3544	0.7256	0.1508	0.1848	0.1606	0.1783	0.1620	0.1799	0.0732	0.0054
	50	0.6264	1.5728	0.2524	0.2488	0.2738	0.2312	0.2645	0.2373	0.0446	0.0022
Θ_2	100	0.4684	0.7541	0.2420	0.1378	0.2541	0.1339	0.2569	0.1398	0.0443	0.0021
	200	0.4260	0.5540	0.2357	0.1251	0.2476	0.1236	0.2479	0.1226	0.0424	0.0020
	50	0.6306	1.2627	0.2421	0.1766	0.2457	0.1666	0.2414	0.1624	0.2569	0.0299
Θ_3	100	0.4793	0.6624	0.2367	0.1121	0.2436	0.1103	0.2425	0.1090	0.1590	0.0142
	200	0.4204	0.4575	0.2347	0.0932	0.2429	0.0909	0.2406	0.0930	0.0071	0.0099

select the best distribution for these three real datasets from the four distributions that belong to this family. Finally, the existence of an additional parameter in the model is considered. For this purpose, the likelihood ratio (LRT) test and its associated p-value are applied.

First Data Set: These data include the remaining lifetime information of 218 persons from the population of couples in the age range of 30 – 70 years at an insurance company in Tehran. This data set is provided on the *https://www.researchgate.net/publication/341480134_Mortality_Data_Set.*

Second Data Set: This data set contains 50 observations on the burr. In the first component, the hole diameter is 12 mm, and the sheet thickness is 3.15 mm. In the second component, the hole diameter is 9 mm, and the sheet thickness is 2 mm. These two datasets for components are derived from two different machines. This data set was used by Dasgupta (2011), and it is represented in Table 4. Before analyzing this data, all data is multiplied by 10. It should be noted that these changes will not affect our analysis and are for computational reasons only.

Table 4. Burr data set

S.N.	Y_1	Y_2	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	0.04	0.06	11	0.24	0.16	21	0.24	0.12	31	0.24	0.14	41	0.02	0.16
2	0.02	0.12	12	0.04	0.12	22	0.22	0.24	32	0.16	0.06	42	0.18	0.32
3	0.06	0.14	13	0.14	0.24	23	0.12	0.06	33	0.32	0.04	43	0.22	0.18
4	0.12	0.04	14	0.16	0.06	24	0.18	0.02	34	0.18	0.14	44	0.14	0.24
5	0.14	0.14	15	0.08	0.02	25	0.24	0.18	35	0.24	0.22	45	0.06	0.22
6	0.08	0.16	16	0.26	0.18	26	0.32	0.22	36	0.22	0.14	46	0.04	0.04
7	0.22	0.08	17	0.32	0.22	27	0.16	0.14	37	0.16	0.06	47	0.14	0.14
8	0.12	0.26	18	0.28	0.14	28	0.14	0.02	38	0.12	0.04	48	0.26	0.26
9	0.08	0.32	19	0.14	0.22	29	0.08	0.18	39	0.24	0.16	49	0.18	0.18
10	0.26	0.22	20	0.16	0.16	30	0.16	0.22	40	0.06	0.24	50	0.16	0.16

Table 5. Cholesterol levels at 5 and 25 weeks after treatment in 30 patients

S.N.	5-th	25-th	S.N.	5-th	25-th	S.N.	5-th	25-th
1	325	246	2	278	245	3	257	212
4	192	192	5	276	325	6	262	294
7	309	232	8	287	287	9	304	245
10	215	261	11	217	252	12	248	305
13	225	225	14	287	208	15	233	217
16	198	198	17	229	179	18	310	352
19	214	274	20	253	209	21	316	283
22	243	245	23	305	272	24	197	197
25	243	247	26	315	283	27	205	205
28	315	255	29	263	215	30	210	271

Third Data Set: This data set contains cholesterol levels at 5 and 25 weeks after treatment in 30 patients. These data are presented in Table 5. Before analyzing this data, we apply transformation $(X - 150)/100$ to all data. As shown in the secound data description, this transformation will not affect our analysis and are for computational reasons only.

For the purposes of this article, we first draw the proposed TTT plots by Aarset (1987) for each marginal in three datasets. These plots are shown in Figure 1. As can be seen, both diagrams are concave, so it can be concluded that the marginal hazard functions are increasing functions. Another significant result is that the correlation between the marginals is positive. Therefore, the proposed family of bivariate distributions can be used for analyzing these datasets.

The purpose of this section is to examine these four sub-models (BWG, BLG, BCHG, and BGG) for these three datasets. To this purpose, we first fit four models, Weibull, Lomax, Chen and Gompertz to the marginals as well as their minimums for three datasets, separately.

For each data set, the maximum likelihood estimation (MLE), the associated Kolmogorov-Smirnov statistics, and their p-values are obtained. These values are provided in Table 6. Based on this information the Weibull, Chen, and Gompertz distributions cannot be rejected for the marginals and the minimum also. We observe that the Lomax distribution is not suitable for these three datasets.

Now we will fit the BPHRG models. Therefore, three special cases of BPHRG distributions are

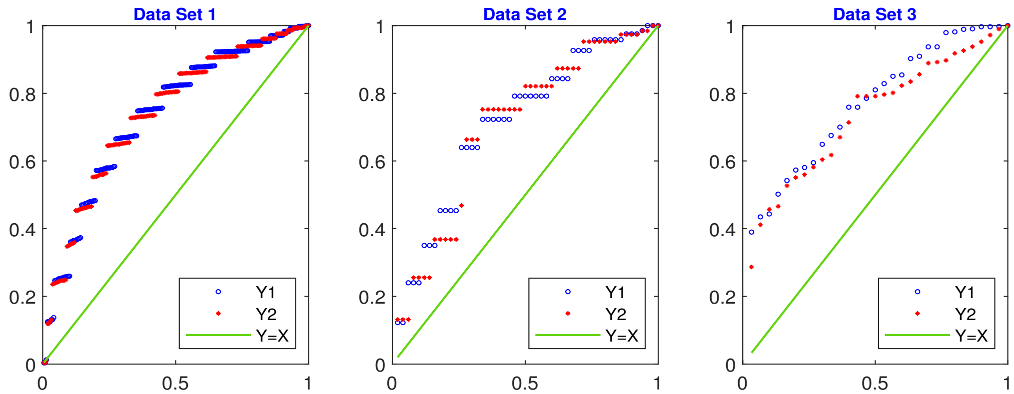

Figure 1. The TTT plots for three datasets

Table 6. The MLEs of parameters, the standard error of estimation in parentheses, the Kolmogorov-Smirnov (K-S), and the corresponding p-values for three datasets

	Model	Variables	α	λ	K-S	P-value
Mortality Data Set	Weibull	Y_1	2.1376 (0.8269)	0.0096 (0.0200)	0.0937	0.1881
		Y_2	2.1768 (0.8030)	0.0079 (0.0141)	0.0823	0.2949
		$\min\{Y_1, Y_2\}$	2.1229 (0.8065)	0.0103 (0.0200)	0.0902	0.2382
	Lomax	Y_1	979.0625 (3.0992)	0.1142 (0.0005)	0.0534	1.9069×10^{-33}
		Y_2	977.5625 (3.0948)	0.1135 (0.0005)	0.5331	6.8134×10^{-35}
		$\min\{Y_1, Y_2\}$	979.0625 (3.0990)	0.1145 (0.0005)	0.5300	2.0984×10^{-32}
	Chen	Y_1	0.6159 (0.0010)	0.0195 (0.0002)	0.0682	0.5520
		Y_2	0.6068 (0.0030)	0.0189 (0.0007)	0.0748	0.4069
		$\min\{Y_1, Y_2\}$	0.6158 (0.0010)	0.0208 (0.0002)	0.0653	0.6277
	Gompertz	Y_1	0.2297 (0.0012)	0.1305 (0.0020)	0.0599	0.7118
		Y_2	0.2141 (0.0011)	0.1388 (0.0021)	0.0680	0.5277
		$\min\{Y_1, Y_2\}$	0.2296 (0.0012)	0.1395 (0.0021)	0.0588	0.7498
Burr Data Set	Weibull	Y_1	2.1195 (0.0346)	0.2754 (0.0100))	0.1994	0.5666
		Y_2	2.0050 (0.0331)	0.3420 (0.0100)	0.1921	0.6556
		$\min\{Y_1, Y_2\}$	1.7812 (0.0294)	0.6283 (0.0153)	0.1571	0.9108
	Lomax	Y_1	503.9063 (6.9132)	0.1529 (0.0307)	0.5068	7.4113×10^{-4}
		Y_2	503.9063 (6.9116)	0.1552 (0.0098)	0.5120	0.0011
		$\min\{Y_1, Y_2\}$	503.9063 (6.7634)	0.1628 (0.0103)	0.5288	0.0021
	Chen	Y_1	1.0174 (0.0109)	0.1599 (0.0051)	0.1801	0.6897
		Y_2	1.0213 (0.0357)	0.1865 (0.0057)	0.1815	0.7210
		$\min\{Y_1, Y_2\}$	1.0619 (0.0140)	0.3144 (0.0084)	0.1441	0.9522
	Gompertz	Y_1	1.0221 (0.0247)	0.1571 (0.0106)	0.1769	0.7102
		Y_2	1.0327 (0.0257)	0.1796 (0.0120)	0.1786	0.7385
		$\min\{Y_1, Y_2\}$	1.0945 (0.0316)	0.2781 (0.0202)	0.1361	0.9703
Cholesterol Data Set	Weibull	Y_1	2.8926 (0.0781)	0.5723 (0.0244)	0.1569	0.4499
		Y_2	2.2805 (0.0565)	0.7355 (0.0282)	0.1238	0.7201
		$\min\{Y_1, Y_2\}$	2.0927 (0.0678)	1.0927 (0.0374)	0.1714	0.3640
	lomax	Y_1	318.7500 (7.2706)	0.1737 (0.0184)	0.5733	7.1559×10^{-9}
		Y_2	329.6250 (7.5178)	0.1754 (0.01861)	0.5514	1.6649×10^{-8}
		$\min\{Y_1, Y_2\}$	321.6250 (7.5147)	0.1796 (0.0190)	0.5599	3.3779×10^{-8}
	Chen	Y_1	1.6586 (0.0374)	0.2906 (0.0141)	0.1584	0.4382
		Y_2	1.1827 (0.0231)	0.4908 (0.0161)	0.1012	0.8991
		$\min\{Y_1, Y_2\}$	1.6038 (0.0360)	0.5723 (0.0200)	0.1544	0.4929
	Gompertz	Y_1	2.3610 (0.0800)	0.0530 (0.0208)	0.1523	0.4877
		Y_2	1.3631 (0.0556)	0.2454 (0.0244)	0.1122	0.8198
		$\min\{Y_1, Y_2\}$	2.2794 (0.0793)	0.1117 (0.0101)	0.1501	0.5287

considered: BWG, BCHG, and BGG. So, these three sub-models are fitted to these three datasets. The MLEs of parameters, the corresponding log-likelihood values, the AIC, and the BIC are calculated. These results are presented in Table 7.

Table 7. The MLEs of parameters, the standard error of estimation in parentheses, the corresponding log-likelihood, AIC, and BIC for three datasets

Data Set	Model	$\hat{\alpha}$	$\hat{\lambda}_0$	$\hat{\lambda}_1$	$\hat{\lambda}_2$	$\hat{\theta}$	$Log(\ell)$	AIC	BIC
Mortality Data Set	BWG	2.2320	0.0042	7.9279×10^{-4}	4.1529×10^{-4}	0.5256	-855.5175	1721	1738
		(0.2740)	(0.0032)	(0.0024)	(0.0015)	(0.2702)			
	BCHG	0.4844	0.0362	0.0065	0.0035	0.7334	-885.7717	1781.5	1798.5
		(0.0469)	(0.0146)	(0.0046)	(0.0032)	(0.0346)			
	BGG	0.2276	0.0709	0.0136	0.0070	0.6018	-841.3176	1692.6	1709.6
		(0.0583)	(0.0031)	(0.0141)	(0.0100)	(0.0223)			
Burr Data Set	BWG	2.1541	0.0512	0.1440	0.1825	0.2668	-137.9229	285.8458	295.4059
		(0.1951)	(0.0360)	(0.0768)	(0.0948)	(0.0911)			
	BCHG	1.0744	0.0259	0.0756	0.0990	0.6158	-136.3488	282.6976	292.2577
		(0.1216)	(0.0173)	(0.0400)	(0.0500)	(0.0806)			
	BGG	1.1516	0.0209	0.0609	0.0796	0.7528	-141.4662	292.9324	302.4925
		(0.1808)	(0.0173)	(0.0387)	(0.0489)	(0.1161)			
Cholesterol Data Set	BWG	2.8984	0.1341	0.2722	0.3066	0.3003	-47.3445	104.6890	111.6950
		(0.3067)	(0.0911)	(0.1539)	(0.1734)	(0.1157)			
	BCHG	1.5812	0.0387	0.0843	0.0887	0.2815	-47.3223	104.6646	111.6706
		(0.1288)	(0.0264)	(0.0458)	(0.0500)	(0.0932)			
	BGG	2.0344	0.0188	0.0405	0.0402	0.6829	-55.9395	121.8790	128.8850
		(0.2088)	(0.0141)	(0.0244)	(0.0264)	(0.1783)			

Table 8. The Kolmogorov-Smirnov (K-S) and the associated p-values for the marginals and the minimum of these three sub-models in three datasets

Data Set	Model	Y_1		Y_2		$\min\{Y_1, Y_2\}$	
		K-S	p-value	K-S	p-value	K-S	p-value
Mortality Data Set	BWG	0.0923	0.2012	0.0920	0.1847	0.1025	0.1293
	BCHG	0.1254	0.0297	0.1252	0.0250	0.1254	0.0336
	BGG	0.0923	0.2011	0.0882	0.2230	0.0865	0.2821
Burr Data Set	BWG	0.2343	0.3676	0.2375	0.3936	0.2829	0.2848
	BCHG	0.1838	0.6660	0.1772	0.7467	0.1489	0.9386
	BGG	0.2085	0.5128	0.2056	0.5728	0.1859	0.7779
Cholesterol Data Set	BWG	0.1627	0.4056	0.1578	0.4223	0.1669	0.3955
	BCHG	0.1239	0.8215	0.1303	0.6610	0.1193	0.7945
	BGG	0.1360	0.6291	0.1730	0.3137	0.1519	0.5134

To evaluate the performance of these three sub-models in three datasets, the Kolmogorov-Smirnov (K-S) and the associated p-values for the marginals and their minimum are obtained, which Proposition 3.2 must be used for this purpose. The results of this study are presented in Table 8.

In the following, the existence and necessity of an additional parameter in the family are examined. So, we test the BPHR models against the BPHRG models. In other words, we test the null hypothesis $H_0 : BPHR$ against the alternative hypothesis $H_1 : BPHRG$. Therefore, three sub-models of the proportional hazard rate distributions family should be considered. These three sub-models are BWE, BCH, and BG. For this purpose, we apply the likelihood ratio (LRT) test. The LRT test statistics and the associated p-values are presented in Table 9. Therefore, based on the results presented in Table 9, it can be concluded that the models presented in the null hypothesis are rejected in favor of the models presented in the alternative hypothesis for each significant level α.

Based on the results presented in Tables 8 and 9, it can be seen that BGG distribution has the best performance for mortality data set, and BCHG distribution is the best option for Burr data set and and cholesterol data set.

Table 9. The information of the likelihood ratio (LRT) test to evaluate the existence of additional parameters in these three sub-models

Data set	Test	Models					
		BWG	BWE	BCHG	BCH	BGG	BG
Mortality Data Set	AIC	1721	1744.6	1781.5	1801.2	1692.6	1724.3
	BIC	1738	1758.1	1798.5	1814.7	1709.6	1737.9
	$\log(\ell)$	-855.5175	-868.3016	-885.7717	-896.5773	-841.3176	-858.1611
	LTR	25.5682		21.6112		33.6870	
	P-value	4.2702×10^{-7}		3.3390×10^{-6}		6.4732×10^{-9}	
Burr Data Set	AIC	285.8458	321.0396	282.6976	313.5012	292.9324	313.5968
	BIC	295.4059	330.5997	292.2577	323.0613	302.4925	323.1569
	$\log(\ell)$	-137.9229	-155.5198	-136.3488	-151.7506	-141.4662	-151.7984
	LRT	35.1938		30.8036		20.6644	
	p-value	2.9847×10^{-9}		2.8551×10^{-8}		5.4724×10^{-6}	
Cholesterol Data Set	AIC	104.6890	127.1864	104.6646	135.5496	121.8790	136.6472
	BIC	111.6950	134.1924	111.6706	142.5556	128.8850	143.6532
	$\log(\ell)$	-47.3445	-58.5932	-47.3223	-62.7748	-55.9395	-63.3236
	LRT	22.4974		30.9050		14.7682	
	p-value	2.1043×10^{-6}		2.7097×10^{-8}		1.2157×10^{-4}	

6. Conclusions

This paper focuses on modeling dependent longevity due to the existence of common risk factors. Accordingly, a new family of distributions was first introduced. Then, assuming that the sample size was geometrically distributed, this family of models was generalized, creating a broader and more flexible class of longevity models for dependent lives. Also, the parameters of this family of distributions are also estimated by the MLE method. However, it is not possible to provide an explicit form for estimating the parameters. Therefore, an EM algorithm was designed and implemented to estimate the parameters. Also, the performance of this proposed algorithm was examined by simulation studies. Also, three data sets were examined and different tests were performed to evaluate and determine the best model.

But the classic method of the maximum likelihood for estimating parameters is not always available. For example, suppose $\{(x_{1i}, x_{2i}); i = 1, 2, ..., n\}$ is a data set and $x_{1i} \leq x_{2i}$ for all $i = 1, ..., n$, then, MLEs do not exist. Another important issue is the convergence of the EM algorithm, which is highly dependent on the initial value selection. Finally, it should be noted that calculating the exact confidence interval for MLEs is not easy. The constructed confidence interval based on the maximum likelihood method is determined using the asymptotic property of MLEs. But in the estimation of parameters by the Bayesian method, it can be seen that the Bayesian method does not require an initial value to estimate the parameters and the convergence of the proposed method is guaranteed by the strong law of large numbers. Also, HPD credible intervals are calculated even for small sample sizes using Bayes estimates. Therefore, it is suggested that in future studies, these models focus on using the Bayesian method to estimate the parameters.

Acknowledgment:

The author is profoundly grateful to the referees and editor for their suggestions, which led to improvement in the quality of this article.

REFERENCES

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

Ahmadi, J., Jafari Jozani, M., Marchand, E. and Parsian, A. (2009a). Bayesian estimation based on k–record data from a general class of distributions under balanced type loss functions, Journal of Statistical Planning and Inference, Vol. 139, pp. 1180–1189.

Ahmadi, J., Jafari Jozani, M., Marchand, E. and Parsian, A. (2009b). Prediction of k–records from a general class of distributions under balanced type loss functions, Metrika, Vol. 70, pp. 19–33.

Al-Khedhairi, A.and El-Gohary, A. (2008). A new class of bivariate gompertz distributions and its mixture, International Journal of Mathematical Analysis, Vol. 2, pp. 235–253.

Asgharzadeh, A. and Valiollahi, R. (2010). Prediction intervals for proportional hazard rate models based on progressively type–II censored samples, Communications of the Korean Statistical Society, Vol. 17, pp. 99–106.

Balakrishnan, N. and Lai, C. D. (2009). *Continuous Bivariate Distributions* (2nd Edition), Springer, New York.

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

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

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, Vol. 67, pp. 927–929.

Cox, D. R. (1997). Regression models and life tables (with discussion), Journal of the Royal Statistical Society. Series B (Methodological), Vol. 34, pp. 187–220.

Dasgupta, R. (2011). On the distribution of Burr with applications, Sankhya B, Vol. 73, pp. 1–19.

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

Frees, E. W., Carriere, J. and Valdez, E. (1996). Annuity valuation with dependent mortality, The Journal of Risk and Insurance, Vol. 63, pp. 229–61.

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, Vol. 32, pp. 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, Vol. 36, pp. 1855–1866.

Giesecke, K. (2003). A simple exponential model for dependent defaults, The Journal of Fixed Income, Vol. 13, pp. 74–83.

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

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

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

Kundu, D. and Gupta, R. D. (2010). Modified Sarhan–Balakrishnan singular bivariate distribution, Journal of Statistical Planning and Inference, Vol. 140, pp. 526–538.

Kundu, D. and Gupta, A. K. (2014). On bivariate Weibull–geometric distribution. Journal of Multivariate Analysis, Vol. 123, pp. 19–29.

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

Lindskog, F. and McNeil, A. J. (2003). Common poisson shock models: applications to insurance and credit risk modelling, ASTIN Bulletin, Vol. 33, pp. 209–38.

Marshall, A. W. and Olkin, I. (1967). A multivariate exponential distribution, Journal of the American Statistical Association, Vol. 62, pp. 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, Vol. 84, pp. 641–652.

Marshall, A. W. and Olkin, I. (2007). *Life Distributions*, New York, Springer.

Muhammed, H. Z. (2016). Bivariate inverse Weibull distribution, Journal of Statistical Computation and Simulation, Vol. 86, pp. 2335–2345.

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

Sarhan, A. M., and Balakrishnan, N. (2007). A new class of bivariate distributions and its mixture, Journal of Multivariate Analysis, Vol. 98, pp. 1508–1527.

Shoaee, S. and Khorram, E. (2019). Survival analysis for a new compounded bivariate failure time distribution in shock and competing risk models via an EM algorithm, Communications in Statistics–Theory and Methods. https://doi.org/10.1080/03610926.2019.1614193

Silva, R., Bourguignon, M., Dias, C. and Cordeiro, G. (2013). The compound class of extended Weibull power series distributions, Computational Statistics and Data Analysis, Vol. 58, pp. 352–367.