Multivariate Distributions with Proportional Reversed Hazard Marginals

Debasis Kundu[1] & Manuel Franco[2] & Juana-Maria Vivo[3]

Abstract

Several univariate proportional reversed hazard models have been proposed in the literature. Recently, Kundu and Gupta (2010, Sankhya, Ser. B, vol. 72, 236 - 253.) proposed a class of bivariate models with proportional reversed hazard marginals. It is observed that the proposed bivariate proportional reversed hazard models have a singular component. In this paper we introduce the multivariate proportional reversed hazard models along the same manner. Moreover, it is observed that the proposed multivariate proportional reversed hazard model can be obtained from the Marshall-Olkin copula. The multivariate proportional reversed hazard models also have a singular component, and their marginals have proportional reversed hazard distributions. The multivariate ageing and the dependence properties are discussed in details. We further provide some dependence measure specifically for the bivariate case. The maximum likelihood estimators of the unknown parameters cannot be expressed in explicit forms. We propose to use the EM algorithm to compute the maximum likelihood estimators. One trivariate data set has been analyzed for illustrative purposes.

Keywords: Marshall-Olkin copula; Maximum likelihood estimator; failure rate; EM algorithm; Fisher information matrix.

[1] Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India. Corresponding author. e-mail: kundu@iitk.ac.in

[2] Department of Statistics and Operations Research, University of Murcia, 30100 Murcia, Spain. e-mail: mfranco@um.es

[3] Department of Public Finance and Economy, University of Murcia, 30100 Murcia, Spain. e-mail: jmvivomo@um.es

1 Introduction

If X is an absolutely continuous positive random variable with the probability density function (PDF) $g(\cdot)$ and the cumulative distribution function $G(\cdot)$, then the hazard function of X is defined as

$$h(t) = \frac{g(t)}{1 - G(t)}; \quad t \geq 0.$$

The hazard function plays a very important role in the reliability and survival analysis. Extensive work on different aspects of hazard function has been found in the statistical literature, see for example Meeker and Escobar [26].

Recently, proportional reversed hazard model has received considerable attention since it was introduced by Block *et al.* [1]. If X is an absolutely continuous positive random variable as defined above, then the reversed hazard function of X is defined by

$$r(t) = \frac{g(t)}{G(t)}; \quad t \geq 0.$$

Similar to the hazard function, the reversed hazard function also uniquely characterize a distribution function. The reversed hazard function has been used quite extensively in forensic studies and some related areas. Interested readers may look at the original article of Block *et al.* [1] in this respect.

The class of proportional reversed hazard models can be described as follows. If $F_0(\cdot)$ is any distribution function, then define the class of distribution functions $F(\cdot;\alpha)$ for $\alpha > 0$ as

$$F(t;\alpha) = (F_0(t))^{\alpha}.$$

It can be easily seen that $F(\cdot;\alpha)$ is a proper distribution. From the definition of the proportional reversed hazard function, it is immediate that if $F_0(\cdot)$ has a reversed hazard function $r_0(\cdot)$, then $F(\cdot;\alpha)$ has the proportional reversed hazard function $\alpha r_0(\cdot)$. Recently several proportional reversed hazard models have been introduced by several authors, and their properties have been investigated, see for example Crescenzo [7], Gupta and Gupta [12], Gupta and Kundu [13, 14], Sarhan and Kundu [29] and the references cited therein.

Kundu and Gupta [22] recently introduced a bivariate distribution with proportional reversed hazard marginals. It has several interesting properties, and it has been used quite successfully to analyze bivariate lifetime data. The main aim of this paper is to introduce multivariate (p-dimensional) distributions with proportional reversed hazard marginals. It has been done using the same maximization process from $p+1$ independent proportional reversed hazard models. It introduces positive dependence among the variables. The proposed multivariate proportional reversed hazard model can be obtained from the Marshall-Olkin (MO) copula also, using the proportional reversed hazard model as the marginals. Using the copula properties, several dependence measures like Kendall's τ, Spearman's ρ can be computed specifically for the bivariate proportional reversed hazards distribution.

It is observed that for $q < p$ dimensional subset of the p-variate proportional reversed hazards distribution is a q-variate proportional reversed hazards distribution. The cumulative distribution function of the q-variate proportional reversed hazards distribution can be written in a very convenient form. The decomposition of the absolutely continuous part and the singular part is clearly unique. We provide the joint probability density function of the absolute continuous part explicitly. We discuss some distributional, ageing and dependence properties for the proposed p-variate distribution.

It may be mentioned that the importance of the ageing and dependence notions has been well established in the statistical literature, see for example Lai and Xie [23]. In many reliability and survival analysis applications it has been observed that the components are often positively dependent in some stochastic sense. Hence the derivation of ageing and dependence properties for any multivariate distribution has its own importance. Similarly, the extreme order statistics, the minimum and maximum play a great role in several statistical applications, particularly, where the components have some dependence. For example, the minimum and maximum order statistics play important roles in the competing risks model, and the complementary risks model, respectively. So the distributions of both extreme order statistics for the proposed multivariate distributions and some stochastic ageing results are studied in this paper.

It is observed that the maximum likelihood estimators (MLEs) of the unknown parameters cannot be obtained in explicit form, as expected. Non-linear optimization problem needs to be solved to compute the MLEs. We propose to use the EM algorithm to compute the MLEs, and we provide the implementation details for several multivariate proportional models. Finally, we analyze one simulated data set for illustrative purposes.

Rest of the paper is organized as follows. In Section 2, we briefly discuss about the different dependence concept, some basic copula properties and provide different examples of proportional reversed hazards models which are available in the literature. In Section 3, we introduce the multivariate proportional reversed hazards models. Different dependence and ageing properties are discussed in Section 4. In Section 5, we provide different dependence measures for bivariate proportional reversed hazards models. In Section 6, we apply the EM algorithm. The analysis of a data set has been presented in Section 7, and finally we conclude the paper in Section 8.

2 Preliminaries

2.1 Dependence and Stochastic Order

Several notions of positive or negative dependence for a multivariate distribution, of varying degree of strengths, are available in the literature, see for example Boland *et al.* [3], Colangelo *et al.* [5, 6] and see the references cited therein.

A random vector $\boldsymbol{X}$ is said to be positively lower orthant dependent (PLOD) if the joint cumulative distribution function of $\boldsymbol{X}$ satisfies the following property;

$$F_{\boldsymbol{X}}(x_1, \cdots, x_p) \geq \prod_{i=1}^{p} F_i(x_i), \quad \forall \boldsymbol{x} = (x_1, \cdots, x_p), \tag{1}$$

here F_i's for $i = 1, \cdots, p$, are the marginal distribution functions. Further we will be using the following notation. For $\boldsymbol{x} \in \mathbf{R}^p$, a phrase such as 'non-decreasing in $\boldsymbol{x}$', means non-decreasing in each component x_i, for $i = 1, \cdots, p$. If A is any subset of $\{1, \cdots, p\}$, then $\boldsymbol{X}_A$ denote the vectors, $(X_i | i \in A)$, similarly, $\boldsymbol{x}_A$ is also defined. The following definition are from Brindley and Thompson [4], see also Joe [15].

A p-dimensional random vector $\boldsymbol{X}$ is said to be left tail decreasing (LTD), if

$$P(\boldsymbol{X}_{A_2} \leq \boldsymbol{x}_{A_2} | \boldsymbol{X}_{A_1} \leq \boldsymbol{x}_{A_1}) \tag{2}$$

is a non-increasing in $\boldsymbol{x}_{A_1}$ for all $\boldsymbol{x}_{A_2}$, where the sets A_1 and A_2 are a disjoint partition of $\{1, \cdots, p\}$. In particular, this LTD notion implies the multivariate left tail decreasing property given by Colangelo *et al.* [5].

Another multivariate dependence notion is the multivariate left corner set decreasing property. A random vector $\boldsymbol{X}$ is said to have left corner set decreasing property (LCSD), if

$$P(X_1 \leq x_1, \cdots, X_p \leq x_p | X_1 \leq x'_1, \cdots, X_p \leq x'_p) \tag{3}$$

is non-increasing in $x'_1, \cdots, x'_p$ for every choice of $\boldsymbol{x} = (x_1, \cdots, x_p)$. Equivalently, (3) can be written as

$$\frac{F_{\boldsymbol{X}}(\boldsymbol{x} \wedge \boldsymbol{x}')}{F_{\boldsymbol{X}}(\boldsymbol{x}')} \quad \text{non-decreasing in } \; x'_1, \cdots, x'_p, \tag{4}$$

where $\boldsymbol{x}' = (x'_1, \cdots, x'_p)$, and $\boldsymbol{x} \wedge \boldsymbol{x}' = (\min(x_1, x'_1), \cdots, \min(x_p, x'_p))$.

Further, a function $h(\boldsymbol{x})$ defined on $\mathbf{R}^p$ is called multivariate totally positive of order two, (MTP_2) if it satisfies

$$h(\boldsymbol{x} \vee \boldsymbol{y}) h(\boldsymbol{x} \wedge \boldsymbol{y}) \geq h(\boldsymbol{x}) h(\boldsymbol{y}), \qquad \text{for all } \boldsymbol{x}, \boldsymbol{y} \in \mathbf{R}^p.$$

Here for $\boldsymbol{x} = (x_1, \cdots, x_p)$ and $\boldsymbol{y} = (y_1, \cdots, y_p)$, $\boldsymbol{x} \vee \boldsymbol{y} = (\max(x_1, y_1), \cdots, \max(x_p, y_p))$.

Analogously to the hazard gradient by Johnson and Kotz [16], the reversed hazard gradient of a p-variate random vector $\boldsymbol{X} = (X_1, \cdots, X_p)$ is defined as extension of the univariate case, see e.g. Roy [28] and Domma [10] for $p = 2$. If $X_1, \cdots, X_p$ are p absolutely continuous random variables, then the reversed hazard gradient of $\boldsymbol{X}$ for $\boldsymbol{x} = (x_1, \cdots, x_p)$ is

$$r_{\boldsymbol{X}}(\boldsymbol{x}) = \left(\frac{\partial}{\partial x_1}, \cdots, \frac{\partial}{\partial x_p}\right) \ln F_{\boldsymbol{X}}(x_1, \cdots, x_p). \tag{5}$$

The i-th component of (5) is $r_{\boldsymbol{X},i}(\boldsymbol{x}) = \frac{\partial}{\partial x_i} \ln F_{\boldsymbol{X}}(x_1, \cdots, x_p)$ and represents the reversed hazard function of $(X_i \mid X_j < x_j, j \neq i)$, $i = 1, \cdots, p$.

If for all values of $\boldsymbol{x}$, all components of $r_{\boldsymbol{X}}(\boldsymbol{x})$ are decreasing (increasing) functions of the corresponding variables, then the distribution is called multivariate decreasing (increasing) reversed hazard gradient, MDRHG (MIRHG).

Now we will define the following stochastic orderings between two p-dimensional random vectors $\boldsymbol{X}$ and $\boldsymbol{Y}$.

It is said that $\boldsymbol{X}$ is smaller than $\boldsymbol{Y}$ in the stochastic order, $\boldsymbol{X} \leq_{st} \boldsymbol{Y}$, if

$$P(\boldsymbol{X} \in U) \leq P(\boldsymbol{Y} \in U) \quad \text{for all upper sets} \;\; U \subset \mathbf{R}^p.$$

It is well known that $\boldsymbol{X} \leq_{st} \boldsymbol{Y}$ if and only if $E(\phi(\boldsymbol{X})) \leq E(\phi(\boldsymbol{Y}))$, for all non-increasing function ϕ for which the expectations exists.

A random vector $\boldsymbol{X}$ is said to be smaller than $\boldsymbol{Y}$ in the lower orthant order, $\boldsymbol{X} \leq_{lo} \boldsymbol{Y}$, if $F_{\boldsymbol{X}}(\boldsymbol{x}) \geq F_{\boldsymbol{Y}}(\boldsymbol{x})$ for all $\boldsymbol{x}$.

2.2 Copula

It is well known that the dependence among the random variables $X_1, \cdots, X_p$ is completely described by the joint distribution function $F_{\boldsymbol{X}}(x_1, \cdots, x_p)$. The idea of separating

$F_{\boldsymbol{X}}(x_1, \cdots, x_p)$ in two parts, the one which describes the dependence structure, and the other one which describes only the marginal behavior, leads to the concept of copula. A p-variate copula, defined on $[0,1]^p$, is a multivariate distribution with univariate uniform marginals on $[0,1]$. Let $X_1, \cdots, X_p$ be random variables with distribution functions $F_1(\cdot), \cdots, F_p(\cdot)$ respectively, then Sklar's theorem, see for example Nelsen [27], proved that $F_{\boldsymbol{X}}(x_1, \cdots, x_p)$ has a unique copula representation,

$$F_{\boldsymbol{X}}(x_1, \cdots, x_p) = C(F_1(x_1), \cdots, F_p(x_p)),$$

if $F_1(\cdot), \cdots, F_p(\cdot)$ are absolutely continuous. Moreover, from Sklar's theorem, it also follows that if $F(x_1, \cdots, x_p)$ is a joint distribution function with continuous marginals $F_1(\cdot), \cdots, F_p(\cdot)$, and if $F_1^{-1}(\cdot), \cdots, F_p^{-1}(\cdot)$ are the inverse functions of $F_1(\cdot), \cdots, F_p(\cdot)$ respectively, then there exists a unique copula C in $[0,1]^p$, such that

$$C(u_1, \cdots, u_p) = F_{\boldsymbol{X}}(F_1^{-1}(u_1), \cdots, F_p^{-1}(u_p)). \tag{6}$$

It is well known that many dependence properties of a multivariate distribution are obtained from copula properties, and therefore many dependence properties of a multivariate distribution can be obtained by studying the corresponding copula.

3 Multivariate Proportional Reversed Hazard Models

Now we are in a position to define the multivariate proportional reversed hazards (MPRH) model along the same line as the bivariate generalized exponential distribution or bivariate proportional reversed hazard models as defined by Kundu and Gupta [21, 22]. From now on unless otherwise mentioned it is assumed that $\alpha_1 > 0, \cdots, \alpha_{p+1} > 0$, $\theta > 0$. Suppose $U_1 \sim \text{PRH}(\alpha_1, \theta), \cdots, U_{p+1} \sim \text{PRH}(\alpha_{p+1}, \theta)$, and they are independently distributed. Now we define $X_1 = \max\{U_1, U_{p+1}\}, \cdots, X_p = \max\{U_p, U_{p+1}\}$. Then we say that the random vector $\boldsymbol{X} = (X_1, \cdots, X_p)$ has a multivariate proportional reversed hazards distribution with shape parameters $\alpha_1, \cdots, \alpha_{p+1}$ and scale parameter θ. It will be denoted from now on as MPRH$(\alpha_1, \cdots, \alpha_{p+1}, \theta)$.

Now we provide the joint CDF and the joint PDF of the MPRH model.

THEOREM 3.1: If $\boldsymbol{X} = (X_1, \cdots, X_p) \sim$ MPRH$(\alpha_1, \cdots, \alpha_{p+1}, \theta)$ with base distribution $F_B(\cdot; \theta)$, then the joint CDF of $\boldsymbol{X}$ for $x_1 > 0, \cdots, x_p > 0$ is

$$F_{\boldsymbol{X}}(\boldsymbol{x}) = (F_B(x_1; \theta))^{\alpha_1} \cdots (F_B(x_p; \theta))^{\alpha_p} (F_B(z; \theta))^{\alpha_{p+1}}, \tag{7}$$

where $\boldsymbol{x} = (x_1, \cdots, x_p)$, $z = \min\{x_1, \cdots, x_p\}$ and 0, otherwise.

PROOF: It simply follows from the definition of MPRH model as defined above. The details are omitted. ∎

For brevity we assume the parameter $\theta = 1$, unless otherwise mentioned, and it will be denoted by MPRH$(\alpha_1, \cdots, \alpha_{p+1})$. It is clear that the MPRH distribution is not absolutely continuous distribution, except when $p = 1$. For $p > 1$, it has an absolutely continuous part and a singular part. The MPRH distribution function can be written as follows;

$$F_{\boldsymbol{X}}(\boldsymbol{x}) = \alpha F_a(\boldsymbol{x}) + (1 - \alpha) F_s(\boldsymbol{x}).$$

Here $0 < \alpha < 1$, $F_a(\boldsymbol{x})$ and $F_s(\boldsymbol{x})$ denote the absolutely continuous part and the singular part of $F_{\boldsymbol{X}}(\boldsymbol{x})$, respectively. The corresponding probability density function of $\boldsymbol{X}$ also can be written as

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \alpha f_a(\boldsymbol{x}) + (1 - \alpha) f_s(\boldsymbol{x}). \tag{8}$$

In writing (8) it needs to be understood that the $f_a(\boldsymbol{x})$ is a density function with respect to p-dimensional Lebesgue measure, and $f_s(\boldsymbol{x})$ also can be further decomposed and they are density functions with respect to $1, 2, \cdots, (p-1)$ dimensional Lebesgue measure. It is difficult to obtain the explicit expression of $f_s(\boldsymbol{x})$, and it is not pursued here. For $p = 2$, the result is available in Kundu and Gupta [21]. The explicit expressions of $f_a(\boldsymbol{x})$ and α for general p are provided in Appendix A.

We now provide the distribution functions of the marginal, conditional, and the extreme order statistics of the MPRH distribution.

THEOREM 3.2: If $(X_1, \cdots, X_p) \sim \text{MPRH}(\alpha_1, \cdots, \alpha_{p+1})$ with the base distribution $F_B(\cdot)$, then

(a) $X_1 \sim \text{PRH}(\alpha_1 + \alpha_{p+1})$, $\cdots$, $X_p \sim \text{PRH}(\alpha_p + \alpha_{p+1})$.

(b) For any non-empty subset $I_q = (i_1, \cdots, i_q) \subset (1, \cdots, p)$, the q-dimensional marginal $\boldsymbol{X}_{I_q} = (X_{i_1}, \cdots, X_{i_q}) \sim \text{MPRH}(\alpha_{i_1}, \cdots, \alpha_{i_q}, \alpha_{p+1})$.

(c) The conditional distribution of $\boldsymbol{X}_{A_2}$ given $\{\boldsymbol{X}_{A_1} \leq \boldsymbol{x}_{A_1}\}$, where the non-empty subsets A_1 and A_2 are a disjoint partition of $\{1, \cdots, p\}$, is an absolutely continuous distribution function as follows:

$$P(\boldsymbol{X}_{A_2} \leq \boldsymbol{x}_{A_2} \mid \boldsymbol{X}_{A_1} \leq \boldsymbol{x}_{A_1}) = \begin{cases} \prod_{i \in A_2} (F_B(x_i))^{\alpha_i} & \text{if} \quad z = v \\ \left(\frac{F_B(z)}{F_B(v)}\right)^{\alpha_{p+1}} \prod_{i \in A_2} (F_B(x_i))^{\alpha_i} & \text{if} \quad z < v, \end{cases}$$

where $z = \min\{x_i : i \in A \cup B\}$ and $v = \min\{x_i : i \in A\}$.

(d) If $T_1 = \min\{X_1, \cdots, X_p\}$, then

$$F_{T_1}(t) = P(T_1 \leq t) = (F_B(t))^{\alpha_{p+1}} \times \left(1 - \prod_{i=1}^{p} (1 - (F_B(t))^{\alpha_i})\right). \tag{9}$$

(e) If $T_n = \max\{X_1, \cdots, X_p\}$, then

$$F_{T_n}(t) = P(T_n \leq t) = (F_B(t))^{\alpha_1 + \cdots + \alpha_{p+1}}. \tag{10}$$

PROOF: The proofs of (a), (b), (c) and (e) are straight forward and they are avoided.

(d) Note that

$$F_{T_1}(t) = \sum_{k=1}^{p}(-1)^{k-1} \sum_{I_k \in S_k} F_{I_k}(t, \cdots, t),$$

where $I_k = (i_1, \cdots, i_k)$, $1 \leq i_1 \neq \cdots \neq i_k \leq n$, is a k-dimensional subset and S_k is the set of all ordered k-dimensional subsets of $\{1, \cdots, n\}$. Further

$$F_{I_k}(t, \cdots, t) = P(X_{i_1} \leq t, \cdots, X_{i_k} \leq t).$$

Therefore, using part (b),

$$F_{T_1}(t) = (F_B(t))^{\alpha_{p+1}} \times \sum_{k=1}^{p}(-1)^{k-1} \sum_{I_k \in S_k} (F_B(t))^{\alpha_{i_1}+\cdots+\alpha_{i_k}}.$$

Now the result follows using the next equality

$$\sum_{k=1}^{p}(-1)^{k-1} \sum_{I_k \in S_k} (F_B(t))^{\alpha_{i_1}+\cdots+\alpha_{i_k}} = 1 - \prod_{i=1}^{p} \left(1 - (F_B(t))^{\alpha_i}\right).$$

∎

Note that, from Theorem 3.2, the maximum order statistic of $\boldsymbol{X} \sim \text{MPRH}(\alpha_1, \cdots, \alpha_{p+1})$ has PRH model, $T_n \sim \text{PRH}\left(\sum_{i=1}^{p+1} \alpha_i\right)$. Consequently, the monotonicity of the reversed hazard function of the base distribution $F_B(\cdot)$ is preserved by the maximum statistic of a MPRH model.

Remark: Note that the *IRH* distributions have upper bounded support (see e.g. Block *et al.* [1]). Thus, if $F_B(\cdot)$ is not upper bounded, then F_B only can be either *DRH* or with peaks and valleys being decreasing at the end, and so T_n does too. Therefore, the *IRH* class is only preserved to T_n of the MPRH model when the base distribution is upper bounded.

Theorem 3.3: If $\boldsymbol{X} = (X_1, \cdots, X_p) \sim \text{MPRH}(\alpha_1, \cdots, \alpha_{p+1})$, with base distribution function $F_B(\cdot)$, then the minimum of order statistics is smaller than the maximum of order statistics in the reversed hazard order, $T_1 \leq_{rh} T_n$.

Proof: From Theorem 3.2, the reversed hazard function of T_1 can be written as

$$r_{T_1}(t) = \alpha_{p+1} r_B(t) + \sum_{i=1}^{p} \alpha_i (F_B(t))^{\alpha_i - 1} f_B(t) \frac{\prod_{j \neq i} \left(1 - (F_B(t))^{\alpha_j}\right)}{1 - \prod_{j=1}^{p} \left(1 - (F_B(t))^{\alpha_j}\right)}$$

where $r_B(t)$ denotes the reversed hazard function of the base distribution model. Now, taking into account that

$$\frac{1-\prod_{j=1}^{p}(1-(F_B(t))^{\alpha_j})}{\prod_{j\neq i}(1-(F_B(t))^{\alpha_j})} = \frac{1-\prod_{j\neq i}(1-(F_B(t))^{\alpha_j})}{\prod_{j\neq i}(1-(F_B(t))^{\alpha_j})} + (F_B(t))^{\alpha_i} \geq (F_B(t))^{\alpha_i}$$

for each $i \in \{1, \cdots, p\}$, the following inequality holds

$$r_{T_1}(t) \leq \sum_{i=1}^{p+1} \alpha_i r_B(t) = r_{T_n}(t),$$

which completes the proof of the theorem. ∎

4 MPRH: Dependence and Ordering Properties

In this section, first we will show that the MPRH distribution as defined in the previous section, is nothing but PRH marginal distributions coupled by an Marshall-Olkin (MO) copula. Hence the MPRH distribution is a generalization of the multivariate Marshall-Olkin exponential distribution (see Marshall and Olkin [25]), because more general marginals are used. Let us consider the following copula for $\theta = (\theta_1, \cdots, \theta_p)$

$$C_\theta(u_1, \cdots, u_p) = u_1^{1-\theta_1} \cdots u_p^{1-\theta_p} \min\{u_1^{\theta_1}, \cdots, u_p^{\theta_p}\}. \tag{11}$$

Now for

$$\theta_1 = \frac{\alpha_{p+1}}{\alpha_1 + \alpha_{p+1}}, \cdots, \theta_p = \frac{\alpha_{p+1}}{\alpha_p + \alpha_{p+1}}$$

it is immediate that the MPRH distribution can be obtained using the Marshall-Olkin copula (11) and when the marginals are PRH distributions with parameters $\alpha_1 + \alpha_{p+1}, \cdots, \alpha_p + \alpha_{p+1}$, respectively. Several dependence measures mainly for the BPRH distribution, can be obtained directly using the above copula structure and it will be discussed later. It may be mentioned that Lin and Li [24] have also introduced a multivariate generalized Marshall-Olkin distribution based on Marshall-Olkin copula, but the two models are quite different.

None of them can be obtained from the other. Now first we discuss some of the dependence properties of MPRH distribution.

THEOREM 4.1: If $\boldsymbol{X} = (X_1, \cdots, X_p) \sim \text{MPRH}(\alpha_1, \cdots, \alpha_{p+1})$, then $\boldsymbol{X}$ is

(a) PLOD, positively lower orthant dependent.

(b) LTD, left tail decreasing.

(c) LCSD, left corner set decreasing.

PROOF: (a) The random vector $\boldsymbol{X}$ is positively lower orthant dependent, if and only if the CDF of $\boldsymbol{X}$ satisfies (1). Since for $\boldsymbol{x} = (x_1, \cdots, x_p)$, $x_i > 0$, $i = 1, \cdots, p$,

$$F_{\boldsymbol{X}}(\boldsymbol{x}) = (F_B(x_1))^{\alpha_1} \cdots (F_B(x_p))^{\alpha_p} (F_B(z))^{\alpha_{p+1}}$$

and due to Theorem 3.2 with the base distribution function $F_B(\cdot)$,

$$F_i(x) = (F_B(x))^{\alpha_i + \alpha_{p+1}}, \quad \text{for} \;\; i = 1, \cdots, p.$$

The result immediately follows as $\prod_{i=1}^{p} a_i \leq \min\{a_1, \cdots, a_p\}$ for $0 \leq a_i \leq 1$.

In order to prove (b), without loss of generality, let us take $A_1 = \{1, \cdots, q\}$ and $A_2 = \{q+1, \cdots, p\}$. If $\boldsymbol{x} = (x_1, \cdots, x_p)$, $x_i > 0$, $i = 1, \cdots, p$, then

$$P(\boldsymbol{X}_{A_2} \leq \boldsymbol{x}_{A_2} | \boldsymbol{X}_{A_1} \leq \boldsymbol{x}_{A_1}) = \frac{(F_B(x_{q+1}))^{\alpha_{q+1}} \cdots (F_B(x_p))^{\alpha_p} (F_B(z))^{\alpha_{p+1}}}{(F_B(w))^{\alpha_{p+1}}}, \tag{12}$$

here $z = \min\{x_1, \cdots, x_p\}$ and $w = \min\{x_1, \cdots, x_q\}$. Now the proof will be complete if we can show that (12) is a non-increasing function of $x_1, \cdots, x_q$, for fixed $x_{q+1}, \cdots, x_p$. Observe that $z \leq w$, $\forall \, x_1, \cdots, x_p$. First we will show that (12) is a non-increasing function of x_1, when $x_2, \cdots, x_q$ are kept fixed. Without loss of generality let us assume $x_2 = \min\{x_2, \cdots, x_q\}$, and suppose $v = \min\{x_{q+1}, \cdots, x_p\}$. Thus, in the case $v < x_2$, we consider $g_1(x_1) =$

$P\left(\boldsymbol{X}_{A_2} \le \boldsymbol{x}_{A_2} | \boldsymbol{X}_{A_1} \le \boldsymbol{x}_{A_1}\right)$ as a function of x_1 only, we have

$$g_1(x_1) = \begin{cases} c & \text{if} \quad 0 < x_1 < v \\ \\ c \times \left(\frac{F_B(v)}{F_B(x_1)}\right)^{\alpha_{p+1}} & \text{if} \quad v \le x_1 < x_2 \\ \\ c \times \left(\frac{F_B(v)}{F_B(x_2)}\right)^{\alpha_{p+1}} & \text{if} \quad x_2 \le x_1. \end{cases}$$

Here c is a constant with respect to x_1. Clearly, $g_1(\cdot)$ is a non-increasing function, and the result follows. For other cases the results can be obtained along the same line.

Let us see now the proof (c). The random vector $\boldsymbol{X}$ is said to have left corner set decreasing property, LCSD, if it satisfies (4). First we will show that (3) is a non-increasing function of x'_1, when rest of the variables are kept fixed. Now consider for $\boldsymbol{x} = (x_1, \cdots, x_p)$, and $\boldsymbol{x}' = (x'_1, \cdots, x'_p)$. Without loss of generality, we assume that $x_2 < x'_2, \cdots, x_q < x'_q$ and $x_{q+1} > x'_{q+1}, \cdots, x_p > x'_p$. We treat $\dfrac{F_{\boldsymbol{X}}(\boldsymbol{x} \wedge \boldsymbol{x}')}{F_{\boldsymbol{X}}(\boldsymbol{x}')}$ as a function of x'_1 only. Let us use the following notations:

$$\begin{aligned} v &= \min\{x_2, \cdots, x_q\}, & v' &= \min\{x'_2, \cdots, x'_q\}, \\ w &= \min\{x_{q+1}, \cdots, x_p\}, & w' &= \min\{x'_{q+1}, \cdots, x'_p\}. \end{aligned}$$

Clearly, $v \le v'$ and $w' \le w$. Consider different cases separately.

Case 1: $x_1 < v < w' \le v'$. In this case $g_2(x'_1)$ becomes

$$g_2(x'_1) = \begin{cases} c & \text{if} \quad 0 < x'_1 < x_1 \\ \\ c \times \left(\frac{F_B(x_1)}{F_B(x'_1)}\right)^{\alpha_1+\alpha_{p+1}} & \text{if} \quad x_1 \le x'_1 < w' \\ \\ c \times \left(\frac{F_B(x_1)}{F_B(w')}\right)^{\alpha_1+\alpha_{p+1}} & \text{if} \quad w' \le x'_1. \end{cases}$$

Here c is a constant with respect to x'_1. It is immediate that $g_2(\cdot)$ is a non-decreasing function as $F_{PRH}(x; \alpha_1 + \alpha_{p+1})$ is continuous and non-decreasing in $x_1 < x < w'$.

Case 2: $v < x_1 < w' \le v'$. In this case $g_2(x_1')$ can be written as

$$g_2(x_1') = \begin{cases} c & \text{if } \ 0 < x_1' < v \\ c \times \left(\frac{F_B(v)}{F_B(x_1')}\right)^{\alpha_{p+1}} & \text{if } \ v \le x_1' < x_1 \\ c \times \left(\frac{F_B(x_1)}{F_B(x_1')}\right)^{\alpha_1} \left(\frac{F_B(v)}{F_B(x_1')}\right)^{\alpha_{p+1}} & \text{if } \ x_1 \le x_1' < w' \\ c \times \left(\frac{F_B(x_1)}{F_B(x_1')}\right)^{\alpha_1} \left(\frac{F_B(v)}{F_B(w')}\right)^{\alpha_{p+1}} & \text{if } \ w' \le x_1' \end{cases}$$

where c is a constant with respect to x_1'. It is immediate that $g_2(\cdot)$ is a non-increasing function, since it is continuous and piecewise non-increasing function.

For other cases the results can be obtained exactly along the same line. ∎

THEOREM 4.2: Let $\boldsymbol{X} = (X_1, \cdots, X_p) \sim \text{MPRH}(\alpha_1, \cdots, \alpha_{p+1})$, then $F_{\boldsymbol{X}}(\boldsymbol{x})$ has MTP_2 property.

PROOF: Recall that $F_{\boldsymbol{X}}(\boldsymbol{x})$ has MTP_2 property, if and only if

$$\frac{F_{\boldsymbol{X}}(\boldsymbol{x})F_{\boldsymbol{X}}(\boldsymbol{y})}{F_{\boldsymbol{X}}(\boldsymbol{x} \vee \boldsymbol{y})F_{\boldsymbol{X}}(\boldsymbol{x} \wedge \boldsymbol{y})} \le 1. \tag{13}$$

Here $\boldsymbol{x} = (x_1, \cdots, x_p)$, $\boldsymbol{y} = (y_1, \cdots, y_p)$, $\boldsymbol{x} \vee \boldsymbol{y} = \{x_1 \vee y_1, \cdots, x_p \vee y_p\}$ and $\boldsymbol{x} \wedge \boldsymbol{y} = \{x_1 \wedge y_1, \cdots, x_p \wedge y_p\}$, where $c \vee d = \max\{c, d\}$, $c \wedge d = \min\{c, d\}$.

We will use the following notations:

$$\begin{array}{ll} u = \min\{x_1, \cdots, x_p\}, & v = \min\{y_1, \cdots, y_p\}, \\ a = \min\{x_1 \vee y_1, \cdots, x_p \vee y_p\}, & b = \min\{x_1 \wedge y_1, \cdots, x_p \wedge y_p\}. \end{array}$$

Therefore, observe that

$$b = \min\{u, v\} \le \max\{u, v\} \le a.$$

First consider the case when $u \le v$, therefore,

$$b = u \le v \le a.$$

Now the left hand side of (13) can be written as

$$\frac{F_{\boldsymbol{X}}(\boldsymbol{x})F_{\boldsymbol{X}}(\boldsymbol{y})}{F_{\boldsymbol{X}}(\boldsymbol{x} \vee \boldsymbol{y})F_{\boldsymbol{X}}(\boldsymbol{x} \wedge \boldsymbol{y})} = \left(\frac{F_B(v)}{F_B(a)}\right)^{\alpha_{p+1}}, \tag{14}$$

where $F_B(\cdot)$ is the base distribution function. Since $v \leq a$, the right hand side of (14) is less than or equal to 1. For the other case $u > v$, it can be shown along the same line. ∎

THEOREM 4.3: Let $\boldsymbol{X} = (X_1, \cdots, X_p) \sim$ MPRH$(\alpha_1, \cdots, \alpha_{p+1})$ with the base distribution $F_B(\cdot)$. If $F_B \in DRH(IRH)$, then $\boldsymbol{X}$ has MDRHG (MIRHG).

PROOF: From (5) and (7), the i-th component of the reversed hazard gradient of the random vector $\boldsymbol{X}$ can be written as

$$r_{\boldsymbol{X},i}(\boldsymbol{x}) = \frac{\partial}{\partial x_i} \ln F_{\boldsymbol{X}}(\boldsymbol{x}) = \begin{cases} r(x_i; \alpha_i) & \text{if} \quad x_i = \min\{x_1, \cdots, x_p\} \\ \\ r(x_i; \alpha_i + \alpha_{p+1}) & \text{if} \quad x_i > \min\{x_1, \cdots, x_p\} \end{cases}$$

Here $r(\cdot; \alpha)$ denotes the reversed hazard function of PRH(α). Therefore, the monotonicity of the reversed hazard function of $F_B(\cdot)$ is preserved by each component of the multivariate reversed hazard gradient, and then the result immediately follows. ∎

Now we have the following stochastic ordering results.

THEOREM 4.4: Suppose $\boldsymbol{X}$ and $\boldsymbol{Y}$ are p-variate random vectors, and $\boldsymbol{X} \sim$ MPRH$(\alpha_1, \cdots, \alpha_{p+1})$, $\boldsymbol{Y} \sim$ MPRH$(\beta_1, \cdots, \beta_{p+1})$ with the same base distribution $F_B(\cdot)$. If $\alpha_i \leq \beta_i$ for $i = 1, \cdots, p+1$, then $\boldsymbol{X} \leq_{lo} \boldsymbol{Y}$.

PROOF: The results immediately follows from Theorem 3.1, since $(F_B(x))^{\alpha_i} \geq (F_B(x))^{\beta_i}$ when $\alpha_i \leq \beta_i$. ∎

THEOREM 4.5: Suppose $\boldsymbol{X}$ and $\boldsymbol{Y}$ are p-variate random vectors, and $\boldsymbol{X} \sim$ MPRH$(\alpha_1, \cdots, \alpha_{p+1})$, $\boldsymbol{Y} \sim$ MPRH$(\beta_1, \cdots, \beta_{p+1})$ with the same base distribution $F_B(\cdot)$. If $\alpha_i \leq \beta_i$ for $i = 1, \cdots, p+1$, then $\boldsymbol{X} \leq_{st} \boldsymbol{Y}$.

PROOF: Since $\alpha_i \leq \beta_i$, $X_i \leq_{st} Y_i$, for $i = 1, \cdots, p$. Now the result follows using Theorem

4.2, and observing the fact $(X_i|X_{i+1} = x_{i+1}, \cdots, X_p = x_p) \leq_{st} (Y_i|Y_{i+1} = x_{i+1}, \cdots, Y_p = x_p)$ (part (c), Theorem 3.2). ∎

5 BPRH Models: Dependence Measures

In this section, we provide some dependence measure and stochastic ordering results specifically for the BPRH distribution. It is well known that the copula provides a natural way to measure the dependence between two random variables. We explore the copula property of the BPRH distribution to compute some measures of dependence namely the Kendall's tau, Spearman's rho and the medial correlation coefficients. We further study the dependence of extreme events.

The Kendall's τ is defined as the probability of concordance minus the probability of discordance between two pairs of random vectors (X_1, X_2) and (Y_1, Y_2), where (X_1, X_2) and (Y_1, Y_2) are independent and identically distributed random vectors. It is defined as

$$\tau = P[(X_1 - Y_1)(X_2 - Y_2) > 0] - P[(X_1 - Y_1)(X_2 - Y_2) < 0]. \tag{15}$$

It has been shown in Nelsen [27] that Kendall's *tau* index is also a copula property. Moreover, MO copula (11) has the Kendall's tau index as $\frac{\theta_1\theta_2}{\theta_1 - \theta_1\theta_2 + \theta_2}$. Now we have the following result. If $(X_1, X_2) \sim \text{BPRH}(\alpha_1, \alpha_2, \alpha_3)$, then the Kendall's τ index between X_1 and X_2 is

$$\tau_{X_1,X_2} = \frac{\theta_1\theta_2}{\theta_1 - \theta_1\theta_2 + \theta_2} = \frac{\alpha_3}{\alpha_1 + \alpha_2 + \alpha_3}.$$

Moreover, it has been shown in Nelsen [27] that Spearman's ρ index is also a copula property. Let (X_1, X_2), (Y_1, Y_2) and (Z_1, Z_2) be three independent pairs of random variables with a common distribution function, it is defined as

$$\rho = 3\left(P((X_1 - Y_1)(X_2 - Z_2) > 0) - P((X_1 - Y_1)(X_2 - Z_2) < 0)\right).$$

In this case, Spearman's ρ index between X_1 and X_2 having $(X_1, X_2) \sim \text{BPRH}(\alpha_1, \alpha_2, \alpha_3)$ is

$$\rho_{X_1,X_2} = \frac{3\theta_1\theta_2}{2\theta_1 - \theta_1\theta_2 + 2\theta_2} = \frac{3\alpha_3}{2\alpha_1 + 2\alpha_2 + 3\alpha_3}.$$

Therefore, for fixed α_1 and α_2, as α_3 varies from 0 to ∞, both τ_{X_1,X_2} and ρ_{X_1,X_2} vary between 0 and 1.

Blomqvist [2] defined another measure of dependence known as medial correlation coefficient. The medial correlation coefficient, say M_{X_1,X_2}, between two continuous random variables X_1 and X_2 are defined as follows. If M_{X_1} and M_{X_2} denote the median of X_1 and X_2 respectively, then

$$M_{X_1,X_2} = P[(X_1 - M_{X_1})(X_2 - M_{X_2}) > 0] - P[(X_1 - M_{X_1})(X_2 - M_{X_2}) < 0].$$

Again it has been observed that, see Domma [9], that Blomqvist's medial correlation coefficient is a copula property and it can be easily verified that

$$M_{X_1,X_2} = 4F_{\boldsymbol{X}}(M_{X_1}, M_{X_2}) - 1 = 4C_{\theta_1,\theta_2}\left(\frac{1}{2}, \frac{1}{2}\right) - 1.$$

Therefore, if $(X_1, X_2) \sim \text{BPRH}(\alpha_1, \alpha_2, \alpha_3)$, then M_{X_1,X_2} is

$$M_{X_1,X_2} = \begin{cases} 2^{\theta_2} - 1 = 2^{\frac{\alpha_3}{\alpha_2+\alpha_3}} - 1 & \text{if} \quad \alpha_1 < \alpha_2 \\ 2^{\theta_1} - 1 = 2^{\frac{\alpha_3}{\alpha_1+\alpha_3}} - 1 & \text{if} \quad \alpha_1 > \alpha_2 \end{cases} \tag{16}$$

It is clear from (16) that, for fixed α_1 and α_2, as α_3 varies from 0 to ∞, both θ_1 and θ_2 vary between 0 and 1, and hence M_{X_1,X_2} varies between 0 and 1.

The bivariate tail dependence provides the amount of dependence in the upper quadrant (or lower quadrant) tail of a bivariate distribution, see Joe [15] in this respect. For bivariate random vectors (X_1, X_2), the upper tail dependence is defined as follows (if it exists)

$$\lambda_U = \lim_{z \to 1^-} P\left(X_2 > F_2^{-1}(z) |\ X_1 > F_1^{-1}(z)\right).$$

Intuitively, the upper tail dependence exists when there is a positive probability that some positive outliers may occur jointly. If $\lambda_U \in (0,1]$, then X_1 and X_2 are said to be asymptotically dependent, if $\lambda_U = 0$, then they are asymptotically independent. Similarly, the lower tail dependence λ_L is defined as follows (if it exists)

$$\lambda_L = \lim_{z \to 0^+} P(X_2 \leq F_2^{-1}(z) |\, X_1 \leq F_1^{-1}(z)).$$

It is also well known that these indexes are non-parametric and they both depend only on the copula C of X_1 and X_2 as follows:

$$\lambda_U = 2 - \lim_{t \to 1^-} \frac{1 - C(t,t)}{1-t} \qquad \text{and} \qquad \lambda_L = \lim_{t \to 0^+} \frac{C(t,t)}{t}. \tag{17}$$

It can be easily observed from (17) that if $(X_1, X_2) \sim \text{BPRH}(\alpha_1, \alpha_2, \alpha_3)$, then

$$\lambda_U = \begin{cases} \theta_1 & \text{if} \quad \theta_1 < \theta_2 \\ \theta_2 & \text{if} \quad \theta_2 < \theta_1 \end{cases} \quad \Leftrightarrow \quad \lambda_U = \begin{cases} \frac{\alpha_3}{\alpha_1+\alpha_3} & \text{if} \quad \alpha_1 > \alpha_2 \\ \frac{\alpha_3}{\alpha_2+\alpha_3} & \text{if} \quad \alpha_1 < \alpha_2, \end{cases}$$

and $\lambda_L = 0$.

6 Maximum Likelihood Estimators

In this section we discuss about the estimation of the unknown parameters of the multivariate proportional reversed hazards distribution based on a random sample of size n from $\text{MPRH}(\alpha_1, \cdots, \alpha_{p+1}, \theta)$. Note that in presence of the base distribution parameter θ, this model has total $p + 1 + m$ parameters, where m denotes the number of elements in the parameter vector θ. Before discussing the maximum likelihood estimators, first let us see the possible available data. In general, for all $\boldsymbol{x} = (x_1, \cdots, x_p) \in \mathbf{R}^p$, there exists a permutation $\mathcal{P}_k = (i_1, \cdots, i_p)$ of $I = (1, \cdots, p)$, such that exactly $p - k$ components are distinct, i.e. $x_{i_{k+1}} < \cdots < x_{i_p}$ and k components are equal to the minimum, $x_{i_1} = \cdots = x_{i_{k+1}}$, for $0 \leq k \leq p-1$, since $(X_i = X_j > X_k)$ has null probability for $i \neq j \neq k$, i.e. there are no possible ties $x_i = x_j > x_k$ for $i \neq j \neq k$. If $k = 0$ all the components are distinct, and if

$k=p-1$, all the components are equal. Therefore, for $0 \le k \le p-1$, the possible outcomes will be of the form

$$\{x_{i_1} = \cdots = x_{i_{k+1}} = x^* < x_{i_{k+2}} < \cdots < x_{i_p}\}. \tag{18}$$

It can be easily seen that based on the observations (18) the MLEs of the unknown parameters can be obtained by solving a $(p+1)$ optimization problem, and which can be computationally quite challenging if p is large. To avoid that we propose to use the expectation maximization (EM) algorithm to compute the MLEs of the unknown parameters as in Karlis [17], Kundu and Dey [19] or Franco *et al.* [11].

Note that for $1 \le k \le p-1$, the data will be of the form (18) if U_i's satisfy

$$\max\{U_{i_1}, \cdots, U_{i_{k+1}}\} < U_{p+1} < U_{i_{k+2}} < \cdots < U_{i_p}, \tag{19}$$

and for $k=0$ if U_i's satisfy

$$U_{i_1} < U_{p+1} < U_{i_2} < \cdots < U_{i_p} \quad \text{or} \quad U_{p+1} < U_{i_1} < U_{i_2} < \cdots < U_{i_p}.$$

Note that we do not observe U_i's, we observe only X_i's. Now to compute the MLEs, we treat this problem as a missing value problem, where the complete observations are U_i's which are not observed. First we will show that if U_i's are observed the MLEs of the unknown parameters can be obtained by solving a m-dimensional optimization problem. Note that if the base line distribution is completely known, then the MLEs can be obtained explicitly.

Observe that for $(u_1, \cdots, u_p, u_{p+1})$, the log-likelihood contribution is

$$\sum_{j=1}^{p+1} \ln f_{PRH}(u_j; \alpha_j, \theta) = \sum_{j=1}^{p+1} \ln \alpha_j + \sum_{j=1}^{p+1} (\alpha_j - 1) \ln F_B(u_j; \theta) + \sum_{j=1}^{p+1} \ln f_B(u_j; \theta)$$

If the complete observations (CO) are $\{u_{1i}, \cdots, u_{pi}, u_{(p+1)i}\}$, $i = 1, \cdots, n$, the log-likelihood function is

$$l(\alpha_1, \cdots, \alpha_p, \alpha_{p+1}, \theta \mid CO) = \sum_{i=1}^{n} \sum_{j=1}^{p+1} \ln f_{PRH}(u_{ji}; \alpha_j, \theta).$$

For fixed θ, the MLEs of α_j's can be obtained as

$$\widehat{\alpha}_j(\theta) = -\frac{n}{\sum_{i=1}^{n} \ln(F_B(u_{ji};\theta))}, \tag{20}$$

and the MLE of θ can be obtained by maximizing the profile log-likelihood function of θ, i.e.

$$l(\widehat{\alpha}_1(\theta), \cdots, \widehat{\alpha}_{p+1}(\theta), \theta \mid CO)$$

with respect to θ. It can be obtained by solving a one-dimensional optimization problem.

This is the main motivation behind the EM algorithm. We need the following result to proceed further. If X is a non-negative random variable with the CDF $F_{PRH}(x;\alpha,\theta)$ and the associated PDF $f_{PRH}(x;\alpha,\theta)$, then

$$A(c;\alpha,\theta) = E(X|X \le c) = \frac{1}{F_{PRH}(c;\alpha,\theta)} \int_0^c u f_{PRH}(u;\alpha,\theta)du.$$

Now we will discuss how to construct the 'pseudo' log-likelihood contribution for a given observation of the form (18), for $k = 0, \cdots, (p-1)$. For $k = 1, \cdots, (p-1)$, see (19), the 'pseudo' log-likelihood contribution can be obtained by replacing u_{i_j}, for $j = 1, \cdots, k+1$ with its expectation, i.e. $u^*_{i_j} = A(x^*, \alpha_{i_j}, \theta)$. For $k = 0$, all the entries are distinct, and the original configuration of U_i's given $x_{i_1} < \cdots < x_{i_p}$ is

$$(i)\ \ U_{i_1} < U_{p+1} < U_{i_2} < \cdots < U_{i_p} \quad \text{or} \quad (ii)\ \ U_{p+1} < U_{i_1} < U_{i_2} < \cdots < U_{i_p} \tag{21}$$

with probability $P(U_{i_1} < U_{p+1}) = \dfrac{\alpha_{i_1}}{\alpha_{p+1} + \alpha_{i_1}} = A_{i_1}$ and $P(U_{i_1} > U_{p+1}) = \dfrac{\alpha_{p+1}}{\alpha_{p+1} + \alpha_{i_1}} = B_{i_1}$ respectively. Now using similar idea as of Dinse [8] or Kundu [18], the 'pseudo' log-likelihood contribution of $x_{i_1} < \cdots < x_{i_p}$ is

$$\begin{aligned} & A_{i_1}\left(\sum_{j=2}^{p} \ln f_{PRH}(x_{i_j};\alpha_{i_j},\theta) + \ln f_{PRH}(x_{i_1};\alpha_{p+1},\theta) + \ln f_{PRH}(u^*_{i_1};\alpha_{i_1},\theta)\right) \\ & + B_{i_1}\left(\sum_{j=1}^{p} \ln f_{PRH}(x_{i_j};\alpha_{i_j},\theta) + \ln f_{PRH}(u^*_{p+1};\alpha_{p+1},\theta)\right), \end{aligned}$$

here

$$u_{i_1}^* = A(x_{i_1}, \alpha_{i_1}, \theta) \quad \text{and} \quad u_{p+1}^* = A(x_{i_1}, \alpha_{p+1}, \theta).$$

The 'M'-step involves the maximization of the 'pseudo' log-likelihood function and it can be obtained by first maximizing it with respect to θ, and then obtain the estimates of α_i's from (20) by replacing missing u_{i_j} with $u_{i_j}^*$. The process should be continued until convergence is attained.

7 Illustrative Example

In this section we present one illustrative example when $p = 3$, and the base distribution is an exponential distribution, to show how the proposed EM algorithm can be used in practice. The data represents the total marks out of 100, of 28 students in three courses: Probability-I (X_1), Probability-II (X_2) and Real Analysis (X_3) in M.Sc. (Statistics) first year of one of the premier Institute in India. The data set is presented below.

Table 1: Marks of 28 students in three courses.

S. no.	X_1	X_2	X_3	S. no.	X_1	X_2	X_3	S. no.	X_1	X_2	X_3
1.	91	84	73	11.	76	86	75	21.	70	71	82
2.	78	69	88	12.	74	85	84	22.	78	76	82
3.	72	82	74	13.	79	70	74	23.	81	87	96
4.	76	82	74	14.	83	69	77	24.	86	81	78
5.	79	89	71	15.	80	84	75	25.	75	71	73
6.	72	83	79	16.	76	69	87	26.	91	72	73
7.	80	84	97	17.	77	71	78	27.	77	73	78
8.	66	72	72	18.	82	86	74	28.	71	78	82
9.	79	74	77	19.	88	75	73				
10.	71	73	78	20.	92	69	85				

We have made the transformation $(X - 50)/100$ to each data point mainly for fitting purposes. It is not going to make any difference to the statistical inference. We fit trivariate

PRH model when the base distribution is exponential. Although, for illustrative purposes we have chosen the exponential base distribution, any other distributions also can be easily incorporated along the same line.

To start the EM algorithm, we need some initial guesses of the unknown parameters. We get some idea about the initial guesses by fitting GE distributions to the marginals and also to the maximum, see e.g. Gupta and Kundu [14]. We have taken the initial guesses of λ, α_1, α_2, α_3 and α_4 as 1.80, 58.0, 44.0, 60.0, 2.0, respectively.

We start the EM algorithm with these initial guesses and stop the iteration process when the difference between two consecutive log-likelihood values is less than 10^{-6}. The EM algorithm stops after 10 steps and produces the following estimates: $\widehat{\lambda} = 1.49001$, $\widehat{\alpha}_1 = 41.47528$, $\widehat{\alpha}_2 = 34.79624$, $\widehat{\alpha}_3 = 48.88793$, $\widehat{\alpha}_4 = 7.67161$. It may be mentioned that we have started with some other initial guesses also, in all the cases EM algorithm converges to the same estimates. We may use some other stopping criteria also, but it is not going to make any difference to the final estimates.

In Figure 1 we provide the profile 'pseudo' log-likelihood function of λ at the 1-st iterate. Since it is an unimodal function the maximization with respect to λ can be performed very easily. The unimodality of the profile 'pseudo' log-likelihood function is observed at each iteration step. We obtain 95% confidence intervals of λ, α_1, α_2, α_3 and α_4 based on bootstrapping and they are (0.8351, 2.1121), (35.6568, 46.7786), (29.7865, 38.6462), (40.6548, 55.6452), (4.5673, 9.9875), respectively.

Now the natural question is how good the proposed model fits the data set. It may be mentioned that although several goodness of fit tests are available for an arbitrary univariate distribution function, we do not have a satisfactory goodness of fit test for a general multivariate model. Because of this reason we have tested the marginals only. At least it gives

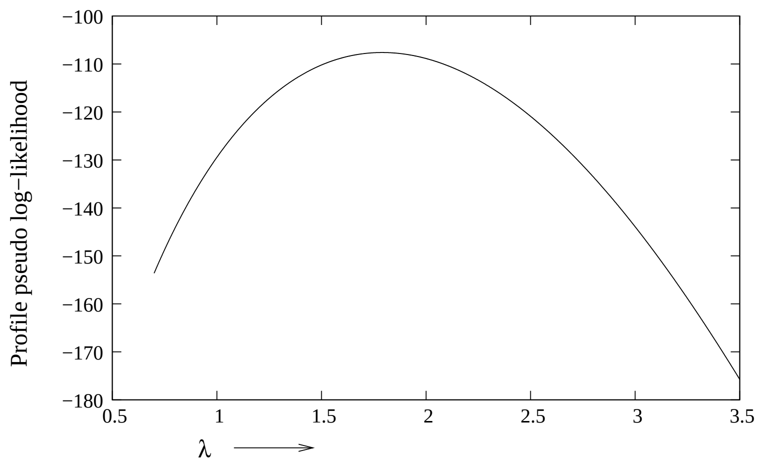

Figure 1: Profile 'pseudo' log-likelihood function at the 1st iterate.

us an indication whether the model can be used or not to analyze that data set. We compute Kolmogorov-Smirnov (KS) distance between the empirical distribution function and the fitted distribution function and the associated p value for all the three marginals. The KS distance and associated p value (reported within brackets) for X_1, X_2 and X_3 are 0.1191 (0.8218), 0.1528 (0.5302) and 0.2059 (0.1856), respectively. Therefore, based on the p values, we can say that the proposed model fits the data reasonably well.

8 Conclusions

In this paper we have developed the multivariate proportional reversed hazards models along the same line as the bivariate proportional reversed hazards model. Several properties of this new multivariate distribution have been established. It has been shown that the proposed distribution can be obtained using Marshall-Olkin copula also. It has helped us to compute several dependence measure specifically for the bivariate proportional reversed hazards model, which has not been established before.

We have further developed the EM algorithm to compute the maximum likelihood estimators of the unknown parameters. The EM algorithm also involves solving a one-dimensional optimization problem, at each 'E'-step. One data analysis has been performed and the performances are quite satisfactory.

Acknowledgements

The authors would like to thank two referees and the associated editor for many constructive suggestions. This work was partially supported by Fundación Seneca of the Regional Government of Murcia (Spain) under Grant 11886/PHCS/09.

Appendix A

In this Appendix A we provide explicitly, the absolutely continuous part of $f_{\boldsymbol{X}}(\cdot)$, namely $f_a(\cdot)$ and α as defined in (8).

The absolutely continuous part $f_a(\boldsymbol{x})$ and α can be obtained from $\dfrac{\partial^p F_{\boldsymbol{X}}(x_1,\cdots,x_p)}{\partial x_1 \cdots \partial x_p}$. It is immediate that $\boldsymbol{x} = (x_1,\cdots,x_p)$ belongs to the set where $F_{\boldsymbol{X}}(\cdot)$ is absolutely continuous if and only if all the x_i's are different. For a given $\boldsymbol{x}$, so that all the x_i's are different, there exists a permutation $\mathcal{P} = (i_1,\cdots,i_p)$, so that $x_{i_1} < x_{i_2} < \cdots < x_{i_p}$. Let us define the following for $x_{i_1} < \cdots < x_{i_p}$

$$f_{\mathcal{P}}(\boldsymbol{x}) = f_{PRH}(x_{i_1};\alpha_{i_1}+\alpha_{p+1}) f_{PRH}(x_{i_2};\alpha_{i_2}) \cdots f_{PRH}(x_{i_p};\alpha_{i_p}). \tag{22}$$

Then from (8) we obtain for $x_{i_1} < \cdots < x_{i_p}$

$$\frac{\partial^p F_{\boldsymbol{X}}(x_1,\cdots,x_p)}{\partial x_1 \cdots \partial x_p} = \alpha f_a(x_1,\cdots,x_p) = f_{\mathcal{P}}(x_1,\cdots,x_p). \tag{23}$$

From (23) we have the following relation;

$$\alpha = \alpha \int_{\mathbf{R}^p} f_a(x_1,\cdots,x_p) dx_1 \cdots dx_p \;=\; \sum_{\mathcal{P}} \int_{x_{i_p}=0}^{\infty} \int_{x_{i_{p-1}}=0}^{x_{i_p}} \cdots \int_{x_{i_1}=0}^{x_{i_2}} f_{\mathcal{P}}(x_1,\cdots,x_p) dx_{i_1} \cdots dx_{i_p}$$

$$= \sum_{\mathcal{P}} J_{\mathcal{P}} \quad \text{(say)}. \tag{24}$$

Since

$$\int_{x_{i_1}=0}^{x_{i_2}} f_{\mathcal{P}}(x_1, ..., x_p) dx_{i_1} = F_{PRH}\left(x_{i_2}; \alpha_{i_1} + \alpha_{p+1}\right) \times \prod_{j=2}^{p} f_{PRH}\left(x_{i_j}; \alpha_{i_j}\right),$$

$$\begin{aligned}
\int_{x_{i_2}=0}^{x_{i_3}} \int_{x_{i_1}=0}^{x_{i_2}} f_{\mathcal{P}}(x_1, ..., x_p) dx_{i_1} dx_{i_2} &= \frac{\alpha_{i_2}}{\alpha_{i_1} + \alpha_{i_2} + \alpha_{p+1}} \\
&\quad \times F_{PRH}\left(x_{i_3}; \alpha_{i_1} + \alpha_{i_2} + \alpha_{p+1}\right) \times \prod_{j=3}^{p} f_{PRH}\left(x_{i_j}; \alpha_{i_j}\right), \\
&\vdots \\
J_{\mathcal{P}} &= \frac{\alpha_{i_2}}{\alpha_{i_1} + \alpha_{i_2} + \alpha_{p+1}} \times ... \times \frac{\alpha_{i_p}}{\alpha_{i_1} + ... + \alpha_{i_p} + \alpha_{p+1}}.
\end{aligned}$$

Therefore, we immediately obtain

$$\alpha = \sum_{\mathcal{P}} \frac{\alpha_{i_2}}{\alpha_{i_1} + \alpha_{i_2} + \alpha_{p+1}} \times \cdots \times \frac{\alpha_{i_p}}{\alpha_{i_1} + \cdots + \alpha_{i_p} + \alpha_{p+1}},$$

and for $x_{i_1} < \cdots < x_{i_p}$

$$f_a(\boldsymbol{x}) = \frac{1}{\alpha} f_{\mathcal{P}}(\boldsymbol{x}),$$

where $f_{\mathcal{P}}$ is same as defined in (22).

Note that when $p = 2$, it matches with the results given in Kundu and Gupta [21].

Now we provide the decomposition of $f_{\boldsymbol{X}}(\boldsymbol{x})$ taking into account that (8) can be written as

$$f_{\boldsymbol{X}}(\boldsymbol{x}) = \alpha f_a(\boldsymbol{x}) + \sum_{k=2}^{p} \sum_{I_k \subset I} \alpha_{I_k} f_{I_k}(\boldsymbol{x}),$$

where $I_k = (i_1, \cdots, i_k) \subset I = (1, \cdots, p)$, such that $i_1 < \cdots < i_k$. Here, it is understood that each $f_{I_k}(\boldsymbol{x})$ is a PDF with respect to $(p - k + 1)$ dimensional Lebesgue measure on the hyperplane $A_{I_k} = \{\boldsymbol{x} \in \mathbf{R}^p : x_{i_1} = \cdots = x_{i_k}\}$. The exact meaning of $f_{\boldsymbol{X}}(\boldsymbol{x})$ is as follows. For any Borel measurable set $B \subset \mathbf{R}^p$

$$P(X \in B) = \alpha \int_B f_a(\boldsymbol{x}) + \sum_{k=2}^{p} \sum_{I_k \subset I} \alpha_{I_k} \int_{B_{I_k}} f_{I_k}(\boldsymbol{x}),$$

where $B_{I_k} = B \cap A_{I_k}$ is the projection of the set B onto the $(p-k+1)$-dimensional hyperplane A_{I_k}. Now we provide the explicit expression of α_{I_k} and $f_{I_k}(\cdot)$.

Note that if $\boldsymbol{x} \in A_{I_k}$, then $\boldsymbol{x}$ is of the following form:

$$\boldsymbol{x} = (x_1, \cdots, x_{i_1-1}, x^*, x_{i_1+1}, \cdots, x_{i_2-1}, x^*, x_{i_2+1}, \cdots, x_{i_k-1}, x^*, x_{i_k+1}, \cdots, x_p)$$

For a given $\boldsymbol{x} \in \mathbf{R}^p$, we define a function g_{I_k} from the $(p-k+1)$-dimensional hyperplane A_{I_k} to $\mathbf{R}$ as follows,

$$g_{I_k}(\boldsymbol{x}) = f_{PRH}(x^*; \alpha_{p+1}) F_{PRH}\left(x^*; \textstyle\sum_{i \in I_k} \alpha_i\right) \prod_{I-I_k} f_{PRH}(x_i; \alpha_i),$$

if $x_i > x^*$, for $i \in I - I_k$, and zero otherwise, where $\prod_{i \in I-I_k} = 1$, when $k = p$. Now, it can be shown along the same line as before that

$$\begin{aligned} \int_{A_{I_k}} g_{I_k}(\boldsymbol{x}) d\boldsymbol{x} &= \sum_{\mathcal{P}_{I-I_k}} \int_{x_{j_{p-k}}=0}^{\infty} \int_{x_{j_{p-k-1}}=0}^{x_{j_{p-k}}} \cdots \int_{x_{j_1}=0}^{x_{j_2}} \int_{x^*=0}^{x_{x_{j_1}}} g_{I_k}(\boldsymbol{x}) dx^* dx_{j_1} dx_{j_2} \cdots dx_{j_{p-k}} \\ &= \sum_{\mathcal{P}_{I-I_k}} \frac{\alpha_{p+1}}{\sum_{i \in I_k} \alpha_i + \alpha_{p+1}} \times \frac{\alpha_{j_1}}{\sum_{i \in I_k} \alpha_i + \alpha_{p+1} + \alpha_{j_1}} \times \cdots \times \frac{\alpha_{j_{p-k}}}{\sum_{i \in I} \alpha_i + \alpha_{p+1}}, \end{aligned}$$

where $\mathcal{P}_{I-I_k}$ denotes the permutations of $I - I_k$, so that $x_{j_1} < \cdots < x_{j_{p-k}}$. Therefore,

$$\alpha_{I_k} = \sum_{\mathcal{P}_{I-I_k}} \frac{\alpha_{p+1}}{\sum_{i \in I_k} \alpha_i + \alpha_{p+1}} \times \frac{\alpha_{j_1}}{\sum_{i \in I_k} \alpha_i + \alpha_{p+1} + \alpha_{j_1}} \times \cdots \times \frac{\alpha_{j_{p-k}}}{\sum_{i \in I} \alpha_i + \alpha_{p+1}}$$

and

$$f_{I_k}(\boldsymbol{x}) = \frac{1}{\alpha_{I_k}} g_{I_k}(\boldsymbol{x}).$$

Appendix B

In this Appendix B, we present explicitly the EM algorithm when $p = 3$. In this case, we have the following unknown parameters $(\alpha_1, \alpha_2, \alpha_3, \alpha_4, \theta)$. We have the following available data $\{(x_{1i}, x_{2i}, x_{3i}); i = 1, \cdots, n\}$. The following notations; $I_0 = \{i; x_{1i} = x_{2i} = x_{3i} = x_i\}$, $I_{10} = \{i; x_{10i} = x_{2i} = x_{3i} < x_{1i}\}$, $I_{20} = \{i; x_{20i} = x_{3i} = x_{1i} < x_{2i}\}$, $I_{30} = \{i; x_{30i} = x_{1i} =$

$x_{2i} < x_{3i}\}$ and $I_{i_1 i_2 i_3} = \{i; x_{i_1 i} < x_{i_2 i} < x_{i_3 i}\}$ are used. Here $(i_1 i_2 i_3)$ is a permutation of (123). The number of elements in set I_0 will be denoted by n_0, the number of elements in set I_{10} will be denoted by n_{10} and similarly others are also defined.

Note that if $i \in I_0$, then $U_4 = x_i$, and $U_1 < x_i$, $U_2 < x_i$ and $U_3 < x_i$. Similarly, if $i \in I_{10}$, then $U_4 = x_{10i}$, $U_2 < x_{10i}, U_3 < x_{10i}, U_1 > x_{10i}$, and so on. We further denote as $(\alpha_1^{(j)}, \alpha_2^{(j)}, \alpha_3^{(j)}, \alpha_4^{(j)}, \theta^{(j)})$ to the estimates of $\alpha_1, \alpha_2, \alpha_3, \alpha_4$ and θ, respectively, at the j-th step of the EM algorithm. We first present the 'pseudo' log-likelihood function at the j-th stage. The 'pseudo' log-likelihood contribution of the different sets are as follows.

From I_0:

$$
\begin{aligned}
&n_0(\ln\alpha_1 + \ln\alpha_2 + \ln\alpha_3 + \ln\alpha_4) + (\alpha_1 - 1)\sum_{i \in I_0} \ln F_B(x_i(1);\theta) + (\alpha_2 - 1)\sum_{i \in I_0} \ln F_B(x_i(2);\theta) \\
&+(\alpha_3 - 1)\sum_{i \in I_0} \ln F_B(x_i(3);\theta) + (\alpha_4 - 1)\sum_{i \in I_0} \ln F_B(x_i;\theta) + \sum_{i \in I_0} \ln f_B(x_i(1);\theta) + \sum_{i \in I_0} \ln f_B(x_i(2);\theta) \\
&+ \sum_{i \in I_0} \ln f_B(x_i(3);\theta) + \sum_{i \in I_0} \ln f_B(x_i;\theta),
\end{aligned}
$$

here $x_i(1) = A(x_i; \alpha_1^{(j)}, \theta^{(j)})$, $x_i(2) = A(x_i; \alpha_2^{(j)}, \theta^{(j)})$ and $x_i(3) = A(x_i; \alpha_3^{(j)}, \theta^{(j)})$, and they depend on j, but we do not make it explicit for brevity.

From I_{10}:

$$
\begin{aligned}
&n_{10}(\ln\alpha_1 + \ln\alpha_2 + \ln\alpha_3 + \ln\alpha_4) + (\alpha_1 - 1)\sum_{i \in I_{10}} \ln F_B(x_{1i};\theta) + (\alpha_2 - 1)\sum_{i \in I_{10}} \ln F_B(x_{10i}(2);\theta) \\
&+(\alpha_3 - 1)\sum_{i \in I_{10}} \ln F_B(x_{10i}(3);\theta) + (\alpha_4 - 1)\sum_{i \in I_{10i}} \ln F_B(x_{10i};\theta) + \sum_{i \in I_{10i}} \ln f_B(x_{10i};\theta) \\
&+ \sum_{i \in I_{10i}} \ln f_B(x_{10i}(2);\theta) + \sum_{i \in I_{10i}} \ln f_B(x_{10i}(3);\theta) + \sum_{i \in I_{10i}} \ln f_B(x_{10i};\theta),
\end{aligned}
$$

here $x_{10i}(2) = A(x_{10i}; \alpha_2^{(j)}, \theta^{(j)})$ and $x_{10i}(3) = A(x_{10i}; \alpha_3^{(j)}, \theta^{(j)})$. Similarly, the contributions from the sets I_{20} and I_{30} can be obtained. Now we provide the contribution from the set I_{123}.

From I_{123}:

$$
\begin{aligned}
&n_{123}(\ln\alpha_1+\ln\alpha_2+\ln\alpha_3+\ln\alpha_4)+(\alpha_1-1)\left(A_1\sum_{i\in I_{123}}\ln F_B(x_{123i}(1);\theta)+B_1\sum_{i\in I_{123}}\ln F_B(x_{1i};\theta)\right)\\
&+(\alpha_2-1)\sum_{i\in I_{123}}\ln F_B(x_{2i};\theta)+(\alpha_3-1)\sum_{i\in I_{123}}\ln F_B(x_{3i};\theta)\\
&+(\alpha_4-1)\left(A_1\sum_{i\in I_{123}}\ln F_B(x_{1i};\theta)+B_1\sum_{i\in I_{123}}\ln F_B(x_{123i}(4);\theta)\right)\\
&+\sum_{i\in I_{123}}(A_1\ln f_B(x_{123i}(1);\theta)+B_1\ln f_B(x_{1i};\theta))+\sum_{i\in I_{123}}\ln f_B(x_{2i};\theta)+\sum_{i\in I_{123}}\ln f_B(x_{3i};\theta)\\
&+\sum_{i\in I_{123}}(A_1\ln f_B(x_{1i};\theta)+B_1\ln f_B(x_{123i}(4),\theta))
\end{aligned}
$$

here $A_1=\dfrac{\alpha_1}{\alpha_1+\alpha_4}$, $B_1=\dfrac{\alpha_4}{\alpha_1+\alpha_4}$, $x_{123i}(1)=A(x_{1i};\alpha_1^{(j)},\theta^{(j)})$ and $x_{123i}(4)=A(x_{1i};\alpha_4^{(j)},\theta^{(j)})$. Similarly, the contributions from the other $I_{i_1i_2i_3}$ can be obtained.

It is clear that for fixed θ, the 'pseudo' log-likelihood function will be maximized by

$$
\widehat{\alpha}_1^{(j+1)}(\theta)=-\frac{n}{C_1},\quad \widehat{\alpha}_2^{(j+1)}(\theta)=-\frac{n}{C_2},\quad \widehat{\alpha}_3^{(j+1)}(\theta)=-\frac{n}{C_3},\quad \widehat{\alpha}_4^{(j+1)}(\theta)=-\frac{n}{C_4},
$$

when

$$
\begin{aligned}
C_1 &= \sum_{i\in I_0}\ln F_B(x_i(1);\theta)+\sum_{i\in I_{10}}\ln F_B(x_{1i};\theta)+\sum_{i\in I_{20}}\ln F_B(x_{20i}(1);\theta)+\sum_{i\in I_{30}}\ln F_B(x_{30i}(1);\theta)\\
&+\sum_{i\in I_{231}\cup I_{213}\cup I_{321}\cup I_{312}}\ln F_B(x_{1i};\theta)+B_1\sum_{i\in I_{123}\cup I_{132}}\ln F_B(x_{1i};\theta)\\
&+A_1\left(\sum_{i\in I_{123}}\ln F_B(x_{123i}(1);\theta)+\sum_{i\in I_{132}}\ln F_B(x_{132i}(1);\theta)\right),
\end{aligned}
$$

$$
\begin{aligned}
C_2 &= \sum_{i\in I_0}\ln F_B(x_i(2);\theta)+\sum_{i\in I_{10}}\ln F_B(x_{10i}(2);\theta)+\sum_{i\in I_{20}}\ln F_B(x_{2i};\theta)+\sum_{i\in I_{30}}\ln F_B(x_{30i}(2);\theta)\\
&+\sum_{i\in I_{123}\cup I_{132}\cup I_{321}\cup I_{312}}\ln F_B(x_{2i};\theta)+B_2\sum_{i\in I_{213}\cup I_{231}}\ln F_B(x_{2i};\theta)\\
&+A_2\left(\sum_{i\in I_{231}}\ln F_B(x_{231i}(2);\theta)+\sum_{i\in I_{213}}\ln F_B(x_{213i}(2);\theta)\right)
\end{aligned}
$$

$$
C_3 = \sum_{i\in I_0}\ln F_B(x_i(3);\theta)+\sum_{i\in I_{10}}\ln F_B(x_{10i}(3);\theta)+\sum_{i\in I_{20}}\ln F_B(x_{20i}(3);\theta)+\sum_{i\in I_{30}}\ln F_B(x_{3i};\theta)
$$

$$+\sum_{i\in I_{123}\cup I_{132}\cup I_{231}\cup I_{213}} \ln F_B(x_{3i};\theta) + B_3 \sum_{i\in I_{312}\cup I_{321}} \ln F_B(x_{3i};\theta)$$

$$+A_3\left(\sum_{i\in I_{321}} \ln F_B(x_{321i}(3);\theta) + \sum_{i\in I_{312}} \ln F_B(x_{312i}(3);\theta)\right)$$

$$\begin{aligned} C_4 \;=\; & \sum_{i\in I_0} \ln F_B(x_i;\theta) + \sum_{i\in I_{10}} \ln F_B(x_{10i};\theta) + \sum_{i\in I_{20}} \ln F_B(x_{20i};\theta) + \sum_{i\in I_{30}} \ln F_B(x_{30i};\theta) \\ & + \left(A_1 \sum_{i\in I_{123}} \ln F_B(x_{1i};\theta) + B_1 \sum_{i\in I_{123}} \ln F_B(x_{123i}(4);\theta)\right) \\ & + \left(A_1 \sum_{i\in I_{132}} \ln F_B(x_{1i};\theta) + B_1 \sum_{i\in I_{132}} \ln F_B(x_{132i}(4);\theta)\right) \\ & + \left(A_2 \sum_{i\in I_{213}} \ln F_B(x_{2i};\theta) + B_2 \sum_{i\in I_{213}} \ln F_B(x_{213i}(4);\theta)\right) \\ & + \left(A_2 \sum_{i\in I_{231}} \ln F_B(x_{2i};\theta) + B_2 \sum_{i\in I_{213}} \ln F_B(x_{213i}(4);\theta)\right) \\ & + \left(A_3 \sum_{i\in I_{312}} \ln F_B(x_{3i};\theta) + B_3 \sum_{i\in I_{312}} \ln F_B(x_{312i}(4);\theta)\right) \\ & + \left(A_3 \sum_{i\in I_{321}} \ln F_B(x_{3i};\theta) + B_3 \sum_{i\in I_{321}} \ln F_B(x_{321i}(4);\theta)\right) \end{aligned}$$

Note that $\theta^{(j+1)}$ can be obtained by maximizing the profile 'pseudo' log-likelihood function with respect to θ. The profile 'pseudo' log-likelihood function is obtained by adding all the 'pseudo' log-likelihood contributions from different sets as given in this appendix, and replacing $\widehat{\alpha}_k^{(j+1)}(\theta)$ for $k = 1, \cdots, 4$. The iteration process continues until converge takes place.

References

[1] Block, H.W., Savits, T. and Singh, H. (1998), "The reversed hazard rate function", *Probability in the Engineering and Informational Sciences*, vol. 12, 69 - 90.

[2] Blomqvist, N. (1950), "On a measure of dependence between two random variables", *Annals of Mathematical Statistics*, vol. 21, 593 - 600.

[3] Boland, P.J., Hollander, M., Joag-Dev, K. and Kochar, S. (1996), "Bivariate dependence properties of order statistics", *Journal of Multivariate Analysis*, vol. 56, 75 - 89.

[4] Brindley, E.J.Jr. and Thompson, W.A.Jr. (1972), "Dependence and aging aspects of multivariate survival", *Journal of the American Statistical Association*, vol. 67, 822 - 830.

[5] Colangelo, A., Scarsini, M. and Shaked, M. (2005), "Some notions of multivariate positive dependence", *Insurance Mathematical Economics*, vol. 37, 13 - 26.

[6] Colangelo, A., Hu, T. and Shaked. M. (2008), "Conditional ordering and positive dependence", *Journal of Multivariate Analsis*, vol. 99, 358 - 371.

[7] Crescenzo, A.D. (2000), "Some results on proportional reversed hazard model", *Statistics and Probability Letters*, vol. 50, 313 - 321.

[8] Dinse, G.E. (1982), "Non-parametric estimation of partially incomplete time and type of failure data", *Biometrics*, vol. 38, 417 - 431.

[9] Domma, F. (2010), "Some properties of the bivariate Burr type III distribution", *Statistics*, vol. 44, 203 - 215.

[10] Domma, F. (2011), "Bivariate Reversed Hazard Rate, Notions, and Measures of Dependence and their Relationships", *Communications in Statistics - Theory and Methods*, vol. 40, 989 - 999.

[11] Franco, M., Kundu, D. and Vivo, J.M. (2011), "Multivariate extension of the modified Sarhan-Balakrishnan bivariate distribution", *Journal of Statistical Planning and Inference*, vol. 141, 3400 - 3412.

[12] Gupta, R.C. and Gupta, R.D. (2007), "Proportional reversed hazard model and its applications", *Journal of Statistical Planning and Inference*, vol. 137, 3525 - 3536.

[13] Gupta, R.D. and Kundu, D. (1999), "Generalized exponential distribution", *Australian & New Zealand Journal of Statistics*, vol. 41, 173 - 188.

[14] Gupta, R.D. and Kundu, D. (2007), "Generalized exponential distribution: existing results and some recent developments", *Journal of Statistical Planning and Inference*, vol. 137, 3525 - 3536.

[15] Joe, H. (1997) *Multivariate model and dependence concepts*, Chapman and Hall, London.

[16] Johnson, N.L. and Kotz, S. (1975), "A vector multivariate hazard rate", *Journal of Multivariate Analysis*, vol. 5, 53 - 66.

[17] Karlis, D. (2003), "ML estimation for multivariate shock models via an EM algorithm", *Annals of the Institute of Statistical Mathematics*, vol. 55, 817 - 830.

[18] Kundu, D. (2004), "Parameter estimation of partially incomplete time and type of failure data", *Biometrical Journal*, vol. 46, 165 - 179.

[19] Kundu, D. and Dey, A.K. (2009), "Estimating the parameters of the Marshall-Olkin bivariate Weibull distribution by EM algorithm", *Computational Statistics and Data Analysis*, vol. 53, 956 - 965.

[20] Kundu, D. and Gupta, R.D. (2004), "Characterization of the (reversed) hazard rate distributions", *Communications in Statistics - Theory and Methods*, vol. 33, 3095-3102.

[21] Kundu, D. and Gupta, R.D. (2009), "Bivariate generalized exponential distribution", *Journal of Multivariate Analysis*, vol. 100, 581 - 593.

[22] Kundu, D. and Gupta, R.D. (2010), "A class of bivariate models with proportional reversed hazard marginals", *Sankhya, B*, vol. 72, 236 - 253.

[23] Lai, C.D. and Xie, M. (2006), *Stochastic Ageing and Dependence for Reliability*, Springer, New York.

[24] Lin, J. and Li, X. (2012), "Multivariate generalized Marshall-Olkin distribution and copulas", *Methodology and Computing in Applied Probability*, DOI 10.1007/s11009-012-9297-4.

[25] Marshall, A.W. and Olkin, I. (1967), "A mutivariate exponential model", *Journal of the American Statistical Association*, vol. 62, 30 - 44.

[26] Meeker, W.Q. and Escobar, L.A. (1998), *Statistical methods for reliability data*, Wiley, New York.

[27] Nelsen, R.B. (2006), *An introduction to copulas*, 2nd-edition, Springer, New York.

[28] Roy, D. (2002), "A characterization of model approach for generating bivariate life distributions using reversed hazard rates", *Journal of the Japan Statistical Society*, vol. 32, 239 - 245.

[29] Sarhan, A. and Kundu, D. (2009), "Generalized linear failure rate distribution", *Communications in Statistics - Theory and Methods*, vol. 38, 642 - 660.