A Class of Bivariate Models with Proportional Reversed Hazard Marginals

Debasis Kundu† and Rameshwar D. Gupta‡

Abstract

Recently the proportional reversed hazard model has received a considerable amount of attention in the statistical literature. The main aim of this paper is to introduce a bivariate proportional reversed hazard model and discuss its different properties. In most of the cases the joint probability distribution function can be expressed in compact forms. The maximum likelihood estimators cannot be expressed in explicit forms in most of the cases. EM algorithm has been proposed to compute the maximum likelihood estimators of the unknown parameters. For illustrative purposes two data sets have been analyzed and the performances are quite satisfactory.

Keywords: Joint probability density function; Conditional probability density function; Maximum likelihood estimators; EM algorithm.

† Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Kanpur, Pin 208016, INDIA. Phone no. 91-512-2597141, Fax No. 91-512-2597500, e-mail: kundu@iitk.ac.in. Corresponding author. Part of his work has been supported by a grant from the Department of Science and Technology, Government of India.

‡ Department of Computer Science and Statistics, The University of New Brunswick at Saint John, New Brunswick, Canada E2L 4L5. Part of his work has been supported by a discovery grant from NSERC, CANADA.

1 Introduction

The hazard rate has been introduced for quite sometime in the statistical literature. The concept of reversed hazard rate is relatively new. The reversed hazard function of a positive random variable can be defined as follows. Suppose X is an absolute continuous positive random variable with the probability density function (p.d.f.) and cumulative distribution function (d.f.) as $g(\cdot)$ and $G(\cdot)$ respectively. Then the reversed hazard rate at the time point t is defined as

$$r(t) = \frac{g(t)}{G(t)}; \qquad t \geq 0. \tag{1}$$

Among different areas, the reversed hazard function has been used quite extensively in forensic studies and some related areas, see for example Block, Savits and Singh (1998), Crescenzo (2000), Gupta and Nanda (2001), Nanda and Gupta (2001), Chandra and Roy (2001), Gupta, Gupta and Sankaran (2004) or Gupta and Gupta (2007) and the references cited therein.

The class of proportional reversed hazard model can be described as follows; Suppose $F_0(\cdot)$ is a distribution function with the support on the positive real axis. Define a class of distribution functions parameterized by $\alpha > 0$ as

$$F(t;\alpha) = (F_0(t))^{\alpha}; \qquad t \geq 0. \tag{2}$$

Clearly, $F(\cdot;\alpha)$ is a distribution function for any $\alpha > 0$ with support on the non-negative real axis. It should be mentioned both (1) and (2) can also be suitably defined for the entire real line. Since we are mainly interested about the lifetime distributions, we restrict our attention to $t \geq 0$. Lehmann (1953) first introduced this class of distribution functions in a testing of hypothesis problem and it is sometime called the Lehmann alternatives, see for example Gupta, Gupta and Gupta (1998). A brief review of proportional reversed hazard models will be presented in section 2.

It may also be observed that if $F_0(\cdot)$ is an absolute continuous distribution function, and has the reversed hazard function $r_0(\cdot)$, then $F(\cdot;\alpha)$ is also absolute continuous, and has the reversed hazard function $\alpha r_0(\cdot)$. Due to this reason, the class of distribution functions defined through (2) is known as the proportional reversed hazard model also. The distribution function $F_0(\cdot)$ is often called as the baseline distribution function.

Recently, proportional reversed hazard model (PRHM) has received considerable attention in the statistical literature. Several PRHMs with various $F_0(\cdot)$ have been introduced and studied quite extensively. For example, the exponentiated Weibull model by Mudholkar *et al.* (1993, 1995), exponentiated Rayleigh by Surles and Padgett (1998, 2001), generalized or exponentiated exponential by Gupta and Kundu (1999), exponentiated Pareto by Shawky and Abu-Zinadah (2008) and exponentiated gamma by Gupta, Gupta and Gupta (1998) have been explored by different authors. In all these cases it has been observed that the various PRHMs can be used quite effectively to analyze lifetime data.

Although extensive work has been done on the PRHMs in the univariate setup, but not much attention attention has been paid to generalize it to the multivariate setup. The authors are familiar with the following work only; Sarhan and Balakrishnan (2007), Kundu and Gupta (2009) and Surles and Padgett (2005), where attempts have been made to introduce bivariate/ multivariate model for some specific PRHMs, for example generalized exponential distribution or generalized Rayleigh distribution.

Our main aim in this paper is to formulate a suitable notion of *bivariate proportional reversed hazard models* (BPRHM) and study its consequences and applicability. We do this by defining BPRHMs in such a way that implies their marginals follow *univariate* proportional reversed hazard (PRHM) distributions. Our proposed model is shown to have a structure that has a singular part, a feature often shared by multivariate distributions (see e.g. Marshall and Olkin (1967)). It is observed that the maximum likelihood estimators of the unknown parameters cannot be obtained in explicit forms as expected. Non-linear equations need to be solved to compute the estimators of the unknown parameters. We recommend to use the EM algorithm for computing the MLEs and discuss how to implement them in different cases.

It may be observed that approaches analogous to those explored here can be used to extend the notion of proportional hazard models to a corresponding bivariate version; a theme that we do not consider further. The rest of the paper is organized as follows. BPRHMs are formally defined and illustrated in section 2. Sections 3 and 4 respectively, explore some of their properties and an EM algorithm for computing the maximum likelihood estimators of the parameters followed by two numerical examples to illustrate the applicability of the models we propose (section 5).

2 Proportional Reversed Hazard Models

2.1 Proportional Reversed Hazard Model

A one-dimensional PRHM (proportional reversed hazard model) is a parametric family of d.f.

$$F_{PRHM}(t;\alpha,\theta) = [F_B(t;\theta)]^{\alpha}; \qquad t \geq 0 \tag{3}$$

with parameters $\alpha > 0$, and θ (may be vector valued) and baseline d.f. $F_B(\cdot;\theta)$. If F_B admits a density f_B, then the PRHM d.f. above has a p.d.f.

$$f_{PRHM}(t;\alpha,\theta) = \alpha[F_B(t;\theta)]^{\alpha-1} f_B(t;\theta), \qquad t \geq 0. \tag{4}$$

This model was originally proposed by Lehmann (1953) in the context of testing of hypotheses. Gupta *et al.* (1998) discussed several interesting properties of the different PRHMs. For some interesting applications of the PRHMs on the breast cancer study see Tsodikov *et al.* (1997). Some of the structural properties of the PRHMs can be easily observed. For example, the distribution of X is always skewed, even though the baseline distribution function $F_0(\cdot)$ is symmetric. It is positively or negatively skewed according as $\alpha > 1$ or $\alpha < 1$ respectively. The median and different moments also can be obtained. Now we briefly provide some descriptions of the different PRHMs which are available in the literature. The readers are referred to Gupta and Gupta (2007) for a recent review article on proportional reversed hazard models.

2.2 Examples

The following are typical examples of univariate PRHM families, known in the literature. In each example below, one may include a *location parameter* to allow flexibility in modeling and data analysis.

(i) *Generalized Exponential* (GE), with an exponential baseline distribution. The PRHM family GE(α, λ) has d.f.

$$F_{GE}(t; \alpha, \lambda) = \left(1 - e^{-\lambda t}\right)^{\alpha}; \quad t \geq 0. \tag{5}$$

See Gupta and Kundu (1999, 2007). It is observed that the GE distribution has an increasing or decreasing hazard rate depending on the shape parameter. Several properties of the GE distribution are very close with the corresponding properties of the Weibull or gamma distributions.

(ii) *Exponentiated Weibull* (EW) model d.f.

$$F_{EW}(t; \alpha, \beta, \lambda) = \left(1 - e^{-\lambda t^{\beta}}\right)^{\alpha}; \quad t \geq 0 \tag{6}$$

with an Weibull baseline distribution and parameters $\alpha > 0$, $\beta > 0$, $\lambda > 0$ (see Mudholkar and Srivastava (1993), Mudholkar *et al.* (1995)). The hazard (failure) rate function of the

EW model can be increasing, decreasing, bathtub, or inverted bathtub depending on the parameter values and thus allow substantial flexibility in modeling.

(iii) *Exponentiated Rayleigh* (ER) This is the EW model with parameters (α, $\beta = 2$, λ), and thus is the two-parameter family of d.f.s

$$F_{ER}(t; \alpha, \lambda) = \left(1 - e^{-\lambda t^2}\right)^{\alpha}; \qquad t \geq 0. \tag{7}$$

See Surles and Padgett (1998, 2001, 2005), Kundu and Raqab (2005).

(iv) *Linear Failure Rate* (LF) model has the d.f.

$$F_{LF}(t; \alpha, \beta, \lambda) = \left(1 - e^{-(\lambda t + \beta \lambda t^2)}\right)^{\alpha}; \quad t \geq 0, \tag{8}$$

with parameters $\alpha > 0$, $\beta > 0$, $\lambda > 0$ considered by Sarhan and Kundu (2009). The shape of its failure rate can be any of the four types as in the EW model. The p.d.f. is unimodal. We further note that for $\alpha \geq 1$, the LF model is in fact *strongly unimodal* (SU) *i.e.* its convolution with any unimodal distribution remains unimodal. This mainly follows from the fact that a necessary and sufficient condition for the SU property is that the corresponding p.d.f. is log-concave, see Ibragimov (1956). For $\alpha \geq 1$, it can be easily verified that the p.d.f. of LF model is log-concave.

The bivariate versions of the univariate proportional reversed hazard models (i) - (iv) above defined using the formulation in section 3 which follows, will be referred to as BGE, BEW, BER and BLF models respectively.

3 Bivariate Proportional Reversed Hazard Model

Suppose $U_1 \sim$ PRHM(α_1, θ), $U_2 \sim$ PRHM(α_2, θ), $U_3 \sim$ PRHM(α_3, θ) and they are mutually independent. Here $\sim$ means distributed as. Define $X_1 = \max\{U_1, U_3\}$ and $X_2 =$

$\max\{U_2, U_3\}$, then we say (X_1, X_2) has a bivariate proportional reversed hazard model (BPRHM) with parameters α_1, α_2, α_3, θ and will be denoted by BPRHM$(\alpha_1, \alpha_2, \alpha_3, \theta)$.

The following result is an easy consequence of the definition above.

THEOREM 3.1: If $(X_1, X_2) \sim$ BPRHM$(\alpha_1, \alpha_2, \alpha_3, \theta)$, then the joint d.f. of (X_1, X_2) for $x_1 > 0$ and $x_2 > 0$ is

$$F_{X_1,X_2}(x_1, x_2) = (F_B(x_1; \theta))^{\alpha_1} (F_B(x_2; \theta))^{\alpha_2} (F_B(z; \theta))^{\alpha_3}, \tag{9}$$

where $z = \min\{x_1, x_2\}$.

Note that (9) can be written as

$$F_{X_1,X_2}(x_1, x_2) = \begin{cases} (F_B(x_1; \theta))^{\alpha_1+\alpha_3} (F_B(x_2; \theta))^{\alpha_2} & \text{if} \quad x_1 < x_2 \\ (F_B(x_1; \theta))^{\alpha_1} (F_B(x_2; \theta))^{\alpha_2+\alpha_3} & \text{if} \quad x_2 < x_1 \\ (F_B(x; \theta))^{\alpha_1+\alpha_2+\alpha_3} & \text{if} \quad x_1 = x_2 = x. \end{cases} \tag{10}$$

THEOREM 3.2: If $(X_1, X_2) \sim$ BPRHM$(\alpha_1, \alpha_2, \alpha_3, \theta)$, then the joint p.d.f. of (X_1, X_2) for $x_1 > 0$ and $x_2 > 0$ is

$$f_{X_1,X_2}(x_1, x_2) = \begin{cases} f_1(x_1, x_2) & \text{if} \quad 0 < x_1 < x_2 < \infty \\ f_2(x_1, x_2) & \text{if} \quad 0 < x_2 < x_1 < \infty \\ f_0(x) & \text{if} \quad 0 < x_1 = x_2 = x \end{cases} \tag{11}$$

where

$$\begin{aligned} f_1(x_1, x_2) &= f_{PRHM}(x_1; \alpha_1 + \alpha_3, \theta) \times f_{PRHM}(x_2; \alpha_2, \theta) \\ f_2(x_1, x_2) &= f_{PRHM}(x_1; \alpha_1, \theta) \times f_{PRHM}(x_2; \alpha_2 + \alpha_3, \theta) \\ f_0(x) &= \frac{\alpha_3}{\alpha_1 + \alpha_2 + \alpha_3} f_{PRHM}(x; \alpha_1 + \alpha_2 + \alpha_3, \theta). \end{aligned}$$

PROOF OF THEOREM 3.2: The expressions of $f_1(\cdot, \cdot)$ and $f_2(\cdot, \cdot)$ can be obtained by taking $\frac{\partial^2}{\partial x_1 \partial x_2} F_{X_1,X_2}(x_1, x_2)$ for $x_1 < x_2$ and $x_2 < x_1$ respectively. To derive the p.d.f. $f_0(x)$ in the case when $x_1 = x_2 = x$; use

$$\int_0^\infty \int_0^{x_2} f_1(x_1, x_2) dx_1 dx_2 + \int_0^\infty \int_0^{x_1} f_2(x_1, x_2) dx_2 dx_1 + \int_0^\infty f_0(x) dx = 1, \tag{12}$$

$$\int_0^\infty \int_0^{x_2} f_1(x_1, x_2) dx_1 dx_2 = \alpha_2 \int_0^\infty (F_B(x;\theta))^{\alpha_1+\alpha_2+\alpha_3-1} f_B(x;\theta) dx = \frac{\alpha_2}{\alpha_1+\alpha_2+\alpha_3} \tag{13}$$

and similarly,

$$\int_0^\infty \int_0^{x_1} f_2(x_1, x_2) dx_2 dx_1 = \frac{\alpha_1}{\alpha_1+\alpha_2+\alpha_3}. \tag{14}$$

Note that

$$\int_0^\infty f_0(x) dx = \alpha_3 \int_0^\infty (F_B(x;\theta))^{\alpha_1+\alpha_2+\alpha_3-1} f_B(x;\theta) dx = \frac{\alpha_3}{\alpha_1+\alpha_2+\alpha_3}, \tag{15}$$

therefore the result follows. ∎

It should be mentioned that the BPRHM has both an absolute continuous part and a singular part similar to the Marshall-Olkin bivariate exponential or bivariate Weibull model. Therefore, if there are ties in the data then similar to the Marshall-Olkin bivariate exponential or Weibull model, BPRHM model can be used quite successfully because of its flexibility. In Theorem 3.2, the function $f_{X_1,X_2}(\cdot,\cdot)$ is considered to be the joint p.d.f. of the BPRHM, if it is understood that the first two terms are densities with respect to the two dimensional Lebesgue measure and the third term is a density function with respect to the one dimensional Lebesgue measure, see for example Bemis, Bain and Higgins (1972) for a discussion on it relating to the Marshall-Olkin bivariate exponential.

THEOREM 3.3: If $(X_1, X_2) \sim \text{BPRHM}(\alpha_1, \alpha_2, \alpha_3, \theta)$ with an absolute continuous baseline d.f.; then

$$F_{X_1,X_2}(x_1, x_2) = \frac{\alpha_1+\alpha_2}{\alpha_1+\alpha_2+\alpha_3} F_a(x_1, x_2) + \frac{\alpha_3}{\alpha_1+\alpha_2+\alpha_3} F_s(x_1, x_2), \tag{16}$$

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

$$F_s(x_1, x_2) = (F_B(z;\theta))^{\alpha_1+\alpha_2+\alpha_3} \tag{17}$$

and

$$\begin{aligned} F_a(x_1, x_2) &= \frac{\alpha_1+\alpha_2+\alpha_3}{\alpha_1+\alpha_2}(F_B(x_1;\theta))^{\alpha_1} \quad (F_B(x_2;\theta))^{\alpha_2} \quad (F_B(z;\theta))^{\alpha_3} - \\ & \frac{\alpha_3}{\alpha_1+\alpha_2}(F_B(z;\theta))^{\alpha_1+\alpha_2+\alpha_3}. \end{aligned} \tag{18}$$

Here $F_a(\cdot,\cdot)$ and $F_s(\cdot,\cdot)$ are the absolute continuous part and singular part respectively.

PROOF OF THEOREM 3.3: Suppose A is the following event $A = \{U_1 < U_3\} \cap \{U_2 < U_3\}$, then $P(A) = \dfrac{\alpha_3}{\alpha_1+\alpha_2+\alpha_3}$. Therefore,

$$F_{X_1,X_2}(x_1, x_2) = P(X_1 \le x_1, X_2 \le x_2|A)P(A) + P(X_1 \le x_1, X_2 \le x_2|A')P(A').$$

For $z = \min\{x_1, x_2\}$, note that

$$\begin{aligned} P(X_1 \le x_1, X_2 \le x_2|A) &= [P(A)]^{-1}P(U_1 \le U_3,\ U_2 \le U_3,\ U_3 \le z) \\ &= [P(A)]^{-1}\int_0^z \alpha_3 f_B(t;\theta)(F_B(t;\theta))^{\alpha_1+\alpha_2+\alpha_3-1}dt \\ &= (F_B(z;\theta))^{\alpha_1+\alpha_2+\alpha_3}, \end{aligned}$$

and $P(X_1 \le x_1, X_2 \le x_2|A')$ can be obtained by subtraction. It is immediate that $(F_B(x;\theta))^{\alpha_1+\alpha_2+\alpha_3}$ is the singular part as its mixed second partial derivatives is zero when $x_1 \ne x_2$, and $P(X_1 \le x_1, X_2 \le x_2|A')$ is the absolute continuous part as its mixed second partial derivatives is a bivariate density function. ∎

Using Theorem 3.3, we immediately obtain the joint p.d.f. of (X_1, X_2) also in the following form for $z = \min\{x_1, x_2\}$;

$$f_{X_1,X_2}(x_1, x_2) = \frac{\alpha_1+\alpha_2}{\alpha_1+\alpha_2+\alpha_3}f_a(x_1, x_2) + \frac{\alpha_3}{\alpha_1+\alpha_2+\alpha_3}f_s(z), \tag{19}$$

where

$$f_a(x_1, x_2) = \frac{\alpha_1 + \alpha_2 + \alpha_3}{\alpha_1 + \alpha_2} \times \begin{cases} f_{PRHM}(x_1; \alpha_1 + \alpha_3, \theta) \times f_{PRHM}(x_2; \alpha_2, \theta) & \text{if} \quad x_1 < x_2 \\ f_{PRHM}(x_1; \alpha_1, \theta) \times f_{PRHM}(x_2; \alpha_2 + \alpha_3, \theta) & \text{if} \quad x_1 > x_2 \end{cases} \tag{20}$$

and

$$f_s(z) = f_{PRHM}(z; \alpha_1 + \alpha_2 + \alpha_3, \theta).$$

COMMENTS: It is clear from Theorem 3.3 that for fixed α_1 and α_2, for $\alpha_3 = 0$ (as $\alpha_3 \to 0^+$ respectively), X_1 and X_2 are independent (approach statistically independence, respectively); and as $\alpha_3 \to \infty$, then $P(X_1 = X_2) \to 1$, *i.e.* X_1 and X_2 are asymptotically almost surely equal.

The following properties (Theorems 3.4 and 3.5) of BPRHMs are easily proved.

THEOREM 3.4: If $(X_1, X_2) \sim$ BPRHM$(\alpha_1, \alpha_2, \alpha_3, \theta)$, then

(a) $X_1 \sim$ PRHM$(\alpha_1 + \alpha_3,\, \theta)$ and $X_2 \sim$ PRHM$(\alpha_2 + \alpha_3,\, \theta)$

(b) $P(X_1 < X_2) = \dfrac{\alpha_1}{\alpha_1 + \alpha_2 + \alpha_3}$, $P(X_2 < X_1) = \dfrac{\alpha_2}{\alpha_1 + \alpha_2 + \alpha_3}$, $P(X_1 = X_2) = \dfrac{\alpha_3}{\alpha_1 + \alpha_2 + \alpha_3}$.

(c) $\max\{X_1, X_2\} \sim$ PRHM$(\alpha_1 + \alpha_2 + \alpha_3; \theta)$.

THEOREM 3.5: Let $(X_1, X_2) \sim$ BPRHM$(\alpha_1, \alpha_2, \alpha_3, \theta)$. Suppose the baseline d.f. F_B is absolutely continuous. Then

(a) The conditional distribution of X_1 given $X_2 \leq x_2$, say $F_{X_1|X_2 \leq x_2}(x_1)$ is an absolute continuous distribution function as follows;

$$F_{X_1|X_2 \leq x_2}(x_1) = P(X_1 \leq x_1 | X_2 \leq x_2) = \begin{cases} (F_B(x_1; \theta))^{\alpha_1 + \alpha_3} (F_B(x_2; \theta))^{-\alpha_3} & \text{if} \quad x_1 < x_2 \\ (F_B(x_1; \theta))^{\alpha_1} & \text{if} \quad x_1 > x_2. \end{cases}$$

(b) The conditional d.f. in (a) above has a representation

$$F_{X_1|X_2\leq x_2}(x_1) = p_2 G_2(x_1) + (1-p_2)H_2(x_1),$$

as a mixture of an absolute continuous d.f. and a degenerate distribution, where

$$\begin{aligned} p_2 &:= 1 - \frac{\alpha_3}{\alpha_2+\alpha_3}(F_B(x_2;\theta))^{\alpha_1} \\ H_2(x) &:= 0 \text{ or, } 1 \text{ if } x < x_2 \text{ or } x > x_2 \text{ respectively,} \end{aligned}$$

with $H_2(x_2)$ defined via right continuity, and

$$G_2(x_1) = p_2^{-1} \times \begin{cases} (F_B(x_2;\theta))^{-\alpha_3} \times (F_B(x_1;\theta))^{\alpha_1+\alpha_3} & \text{if} \quad x_1 < x_2 \\ (F_B(x_1;\theta))^{\alpha_1} - \frac{\alpha_3}{\alpha_2+\alpha_3}(F_B(x_2;\theta)^{\alpha_1} & \text{if} \quad x_1 > x_2. \end{cases}$$

∎

Moreover, it is interesting to observe the following properties of (X_1, X_2), if $(X_1, X_2) \sim$ BPRHM$(\alpha_1, \alpha_2, \alpha_3, \theta)$,

(i) Since $F_{X_1,X_2}(x_1, x_2) \geq F_{X_1}(x_1)F_{X_2}(x_2)$ for all x_1, x_2, therefore, they will be positive quadrant dependent, i.e, for every pair of increasing functions $h_1(\cdot)$ and $h_2(\cdot)$,

$$\text{Cov}(h_1(X_1), h_2(X_2)) > 0.$$

(ii) From part (a) of Theorem 3.5 it easily follows that for every x_1, $P(X_1 \leq x_1 | X_2 \leq x_2)$ is a decreasing function of x_2, therefore X_1 is left tail decreasing in X_2. By symmetry it follows that X_2 is left tail decreasing in X_1.

(iii) From part (b) of Theorem 3.5 it easily follows that for every x_1, $P(X_1 \leq x_1 | X_2 = x_2)$ is a decreasing function of x_2, therefore X_2 is positive regression dependent of X_1. By symmetry it follows that X_1 is positive regression dependent of X_2.

4 Maximum Likelihood Estimation

In this section we address the problem of computing the maximum likelihood estimators of the unknown parameters of the BPRHMs. Suppose $\{(x_{11}, x_{21}), \cdots, (x_{1n}, x_{2n})\}$ is a random sample from BPRHM$(\alpha_1, \alpha_2, \alpha_3, \theta)$, then we want to compute the MLEs of the unknown parameters. We consider two cases separately, (i) θ is known, (ii) All the parameters are unknown. Let us use the following notations;

$$I_1 = \{i; x_{1i} < x_{2i}\}, \quad I_2 = \{i; x_{1i} > x_{2i}\}, \quad I_0 = \{i; x_{1i} = x_{2i} = x_i\}, \quad I = I_0 \cup I_1 \cup I_1,$$

$$|I_0| = n_0, \quad |I_1| = n_1, \quad |I_2| = n_2, \quad n = n_0 + n_1 + n_2.$$

Based on the above notation, the log-likelihood function can be written as

$$\begin{aligned} l(\alpha_1, \alpha_2, \alpha_3, \theta|data) &= n_1 \ln(\alpha_1 + \alpha_3) + (\alpha_1 + \alpha_3 - 1) \sum_{i \in I_1} \ln F_B(x_{1i}; \theta) + \sum_{i \in I_1} \ln f_B(x_{1i}; \theta) + \\ & n_1 \ln \alpha_2 + (\alpha_2 - 1) \sum_{i \in I_1} \ln F_B(x_{2i}; \theta) + \sum_{i \in I_1} \ln f_B(x_{2i}; \theta) + n_2 \ln \alpha_1 + \\ & (\alpha_1 - 1) \sum_{i \in I_2} \ln F_B(x_{1i}; \theta) + \sum_{i \in I_2} \ln f_B(x_{1i}; \theta) + n_2 \ln(\alpha_2 + \alpha_3) + \\ & (\alpha_2 + \alpha_3 - 1) \sum_{i \in I_2} \ln F_B(x_{2i}; \theta) + \sum_{i \in I_2} \ln f_B(x_{2i}; \theta) + n_0 \ln \alpha_3 + \\ & (\alpha_1 + \alpha_2 + \alpha_3 - 1) \sum_{i \in I_0} \ln F_B(x_i; \theta) + \sum_{i \in I_0} \ln f_B(x_i; \theta). \end{aligned} \tag{21}$$

Now we consider two cases separately.

Case (i): θ is known: In this case the baseline distribution is completely known. The only unknown parameters are α_1, α_2 and α_3. In this case, the three normal equations are

$$\frac{n_1}{\alpha_1 + \alpha_3} + \frac{n_2}{\alpha_1} = -\sum_{I_1 \cup I_2} \ln F_B(x_{1i}; \theta) - \sum_{I_0} \ln F_B(x_i; \theta), \tag{22}$$

$$\frac{n_2}{\alpha_2 + \alpha_3} + \frac{n_1}{\alpha_2} = -\sum_{I_1 \cup I_2} \ln F_B(x_{2i}; \theta) - \sum_{I_0} \ln F_B(x_i; \theta), \tag{23}$$

$$\frac{n_1}{\alpha_1 + \alpha_3} + \frac{n_0}{\alpha_3} + \frac{n_2}{\alpha_2 + \alpha_3} = -\sum_{I_1} \ln F_B(x_{1i}; \theta) - \sum_{I_2} \ln F_B(x_{2i}; \theta) - \sum_{I_0} \ln F_B(x_i; \theta). \tag{24}$$

It may be easily seen that for fixed $\alpha_3 > 0$, the MLEs of $\widehat{\alpha}_1(\alpha_3)$ and $\widehat{\alpha}_2(\alpha_3)$ become

$$\begin{aligned}\widehat{\alpha}_1(\alpha_3) &= -\frac{(-k_1\alpha_3 + n_1 + n_2) + \sqrt{(-k_1\alpha_3 + n_1 + n_2)^2 + 4k_1 n_2 \alpha_3}}{2k_1} \\ \widehat{\alpha}_2(\alpha_3) &= -\frac{(-k_2\alpha_3 + n_1 + n_2) + \sqrt{(-k_2\alpha_3 + n_1 + n_2)^2 + 4k_2 n_1 \alpha_3}}{2k_2},\end{aligned}$$

where

$$\begin{aligned}k_1 &= -\sum_{I_1 \cup I_2} \ln F_B(x_{1i};\theta) - \sum_{I_0} \ln F_B(x_i;\theta) \\ k_2 &= -\sum_{I_1 \cup I_2} \ln F_B(x_{2i};\theta) - \sum_{I_0} \ln F_B(x_i;\theta).\end{aligned}$$

Now the MLE of α_3 can be obtained from (24) as a solution of the following fixed point equation

$$g(\alpha_3) = \alpha_3, \tag{25}$$

where

$$g(\alpha_3) = -n_0 \left[\sum_{I_1} \ln F_B(x_{1i};\theta) + \sum_{I_2} \ln F_B(x_{2i};\theta) + \sum_{I_0} \ln F_B(x_i;\theta) + \frac{n_1}{\widehat{\alpha}_1(\alpha_3) + \alpha_3} + \frac{n_2}{\widehat{\alpha}_2(\alpha_3) + \alpha_3} \right]^{-1}.$$

To solve (25) iteratively, one may use any suitable numerical procedure.

Case (ii): θ is unknown: In this case it is difficult to compute the MLEs of the unknown parameters directly. We propose to use EM algorithm, which is computationally more tractable. We treat this as a missing value problem as follows.

Assume that for a bivariate random vector (X_1, X_2), there is an associated random vector (Δ_1, Δ_2), where $\Delta = 1$ or 3 if $U_1 > U_3$ or $U_1 < U_3$ respectively. Similarly, $\Delta_2 = 2$ or 3, if $U_1 > U_3$ or $U_2 < U_3$ respectively. Clearly, for given (X_1, X_2), the associated (Δ_1, Δ_2) may not be known always. If $X_1 = X_2$, then $\Delta_1 = \Delta_2 = 3$, therefore, in this case (Δ_1, Δ_2) is completely known. On the other hand of $X_1 < X_2$ or $X_1 > X_2$, the associated (Δ_1, Δ_2) is missing. If $(x_1, x_2) \in I_1$, then the possible values of (Δ_1, Δ_2) are (1, 2) or (3, 2). Similarly, if $(x_1, x_2) \in I_2$, then the possible values of (Δ_1, Δ_2) are (1, 3) or (1, 2) with non-zero probabilities.

Now we provide the 'E'-step and 'M'-step of the EM algorithm. In the 'E'-step, we treat the observations belong to I_0 as the complete observations. If (x_1, x_2) belongs to either I_1 or I_2, we treat the observation as a missing observation. If $(x_1, x_2) \in I_1$, we form the 'pseudo observation' similarly as in Dinse (1982), see also Kundu and Gupta (2009), by fractioning (x_1, x_2) to two partially complete observation of the form $(x_1, x_2, u_1(\gamma))$ and $(x_1, x_2, u_2(\gamma))$. Here $\gamma = (\alpha_1, \alpha_2, \alpha_3, \theta)$, and the fractional mass $u_1(\gamma)$ and $u_2(\gamma)$ assigned to the 'pseudo observation' (x_1, x_2) is the conditional probability that the random vector (Δ_1, Δ_2) takes values (1, 2) and (3, 2) respectively, given that $X_1 < X_2$. Similarly, for $(x_1, x_2) \in I_2$ we form the 'pseudo observation' of the form $(x_1, x_2, w_1(\gamma))$ and $(x_1, x_2, w_2(\gamma))$. Here the fractional mass $w_1(\gamma)$ and $w_2(\gamma)$ assigned to the 'pseudo observation' is the conditional probability that the random vector (Δ_1, Δ_2) takes the values (1, 2) and (1, 3) respectively, given that $X_1 > X_2$. Since

$$P(U_3 < U_1 < U_2) = \frac{\alpha_1 \alpha_2}{(\alpha_1 + \alpha_3)(\alpha_1 + \alpha_2 + \alpha_3)}, \quad P(U_1 < U_3 < U_2) = \frac{\alpha_2 \alpha_3}{(\alpha_1 + \alpha_3)(\alpha_1 + \alpha_2 + \alpha_3)},$$

therefore

$$u_1(\gamma) = \frac{\alpha_1}{\alpha_1 + \alpha_3} \quad \text{and} \quad u_2(\gamma) = \frac{\alpha_3}{\alpha_1 + \alpha_3}.$$

Similarly,

$$w_1(\gamma) = \frac{\alpha_2}{\alpha_2 + \alpha_3} \quad \text{and} \quad w_2(\gamma) = \frac{\alpha_3}{\alpha_2 + \alpha_3}.$$

For brevity, we write $u_1(\gamma)$, $u_2(\gamma)$, $w_1(\gamma)$, $w_2(\gamma)$ as u_1, u_2, w_1 and w_2 respectively. Now based on the above notations, the 'E'-step or the log-likelihood function of the 'pseudo data' can be written as

$$\begin{aligned} l_{pseudo}(\alpha_1, \alpha_2, \alpha_3, \theta) \;\; = \;\; & n_0 \ln \alpha_3 + (\alpha_1 + \alpha_2 + \alpha_3 - 1) \sum_{i \in I_0} \ln F_B(x_i; \theta) + \sum_{i \in I_0} \ln f_B(x_i; \theta) + \\ & u_1 \left[n_1 \ln \alpha_1 + (\alpha_1 + \alpha_3 - 1) \sum_{i \in I_1} \ln F_B(x_{1i}; \theta) + \sum_{i \in I_1} \ln f_B(x_{1i}; \theta) \right] + \end{aligned}$$

$$u_2 \left[n_1 \ln \alpha_3 + (\alpha_1 + \alpha_3 - 1) \sum_{i \in I_1} \ln F_B(x_{1i};\theta) + \sum_{i \in I_1} \ln f_B(x_{1i};\theta) \right] +$$
$$n_1 \ln \alpha_2 + (\alpha_2 - 1) \sum_{i \in I_1} \ln F_B(x_{2i};\theta) + \sum_{i \in I_1} \ln f_B(x_{2i};\theta) +$$
$$w_1 \left[n_2 \ln \alpha_2 + (\alpha_2 + \alpha_3 - 1) \sum_{i \in I_2} \ln F_B(x_{2i};\theta) + \sum_{i \in I_2} \ln f_B(x_{2i};\theta) \right] +$$
$$w_2 \left[n_2 \ln \alpha_3 + (\alpha_2 + \alpha_3 - 1) \sum_{i \in I_2} \ln F_B(x_{2i};\theta) + \sum_{i \in I_2} \ln f_B(x_{2i};\theta) \right] +$$
$$n_2 \ln \alpha_1 + (\alpha_1 - 1) \sum_{i \in I_2} \ln F_B(x_{1i};\theta) + \sum_{i \in I_2} \ln f_B(x_{1i};\theta).$$

Now the 'M'-step involves maximizing $l_{pseudo}(\alpha_1, \alpha_2, \alpha_3, \theta)$ with respect to α_1, α_2, α_3 and θ at each step. For fixed θ, the maximization of the 'pseudo log-likelihood' function occurs at

$$\begin{aligned}
\widehat{\alpha}_1(\theta) &= -\frac{n_1 u_1 + n_2}{\sum_{i \in I_0} \ln F_B(x_i;\theta) + \sum_{i \in I_1 \cup I_2} \ln F_B(x_{1i};\theta)} \\
\widehat{\alpha}_2(\theta) &= -\frac{n_1 + w_1 n_2}{\sum_{i \in I_0} \ln F_B(x_i;\theta) + \sum_{i \in I_1 \cup I_2} \ln F_B(x_{2i};\theta)} \\
\widehat{\alpha}_3(\theta) &= -\frac{n_0 + n_1 u_2 + n_2 w_2}{\sum_{i \in I_0} \ln F_B(x_i;\theta) + \sum_{i \in I_1} \ln F_B(x_{1i};\theta) + \sum_{i \in I_2} \ln F_B(x_{2i};\theta)}.
\end{aligned}$$

Finally $\widehat{\theta}$ can be obtained by maximizing $l_{pseudo}(\widehat{\alpha}_1(\theta), \widehat{\alpha}_2(\theta), \widehat{\alpha}_3(\theta), \theta)$ with respect to θ.

Now we can formally describe how to obtain the $(i+1)$-th step from the i-th step of the EM algorithm. Suppose at the i-th step the estimates of α_1, α_2, α_3 and θ are $\alpha_1^{(i)}$, $\alpha_2^{(i)}$, $\alpha_3^{(i)}$ and $\theta^{(i)}$ respectively.

Algorithm

- Step 1: Compute u_1, u_2, w_1 and w_2 using $\alpha_1^{(i)}$, $\alpha_2^{(i)}$, $\alpha_3^{(i)}$.
- Step 2: Compute $\theta^{(i+1)}$, by maximizing $l_{pseudo}(\widehat{\alpha}_1(\theta), \widehat{\alpha}_2(\theta), \widehat{\alpha}_3(\theta), \theta)$.
- Step 3: Once $\theta^{(i+1)}$ is obtained compute $\alpha_1^{(i+1)} = \widehat{\alpha}_1(\theta^{(i+1)})$, $\alpha_2^{(i+1)} = \widehat{\alpha}_2(\theta^{(i+1)})$ and $\alpha_3^{(i+1)} = \widehat{\alpha}_3(\theta^{(i+1)})$.

The process should be continued until the convergence criterion is met. It should be mentioned that this version of EM algorithm is popularly known as ECM (expectation-conditional maximization) algorithm.

5 Data Analysis

For illustrative purposes, we have analyzed two data sets. We have fitted four different bivariate proportional reversed hazard models namely (i) Bivariate exponentiated Weibull (BEW), (ii) Bivariate exponentiated Rayleigh (BER), (iii) Bivariate generalized exponential (BGE) and (iv) Bivariate linear failure rate (BLF). Note that both BER and BGE are four-parameter models and they are particular cases of the five-parameter BEW model. BLF is also a five-parameter model, but it cannot be obtained from BEW or vice versa.

In this section our main aim is to see, how the different bivariate proportional reversed hazard models and the proposed EM algorithm works in practice.

Data Set 1: This data set represents football (soccer) data of the UEFA Champion's League data for the year 2004-2005 and 2005-2006, and it has been obtained from Meintanis (2007). The data set is presented in Table 1, and it represents those games, where at least one goal has been scored by the home team and one goal directly from a *kick* (penalty kick, foul kick or any other direct free kick) by any team. Here X_1 represents the time in minutes of the first *kick* goal scored by any team and X_2 represents the first goal of any type scored by the home team.

Clearly all possibilities are open, for example $X_1 < X_2$, or $X_1 > X_2$, or $X_1 = X_2$. Meintanis (2007) analyzed this data set using Marshall-Olkin bivariate exponential model and the authors, see Kundu and Gupta (2009), analyzed this data set using BGE model. Although, Marshall-Olkin bivariate exponential model is not a sub-model of the BGE model,

2005-2006	X_1	X_2	2004-2005	X1	X2
Lyon-Real Madrid	26	20	Internazionale-Bremen	34	34
Milan-Fenerbahce	63	18	Real Madrid-Roma	53	39
Chelsea-Anderlecht	19	19	Man. United-Fenerbahce	54	7
Club Brugge-Juventus	66	85	Bayern-Ajax	51	28
Fenerbahce-PSV	40	40	Moscow-PSG	76	64
Internazionale-Rangers	49	49	Barcelona-Shakhtar	64	15
Panathinaikos-Bremen	8	8	Leverkusen-Roma	26	48
Ajax-Arsenal	69	71	Arsenal-Panathinaikos	16	16
Man. United-Benfica	39	39	Dynamo Kyiv-Real Madrid	44	13
Real Madrid-Rosenborg	82	48	Man. United-Sparta	25	14
Villarreal-Benfica	72	72	Bayern-M. TelAviv	55	11
Juventus-Bayern	66	62	Bremen-Internazionale	49	49
Club Brugge-Rapid	25	9	Anderlecht-Valencia	24	24
Olympiacos-Lyon	41	3	Panathinaikos-PSV	44	30
Internazionale-Porto	16	75	Arsenal-Rosenborg	42	3
Schalke-PSV	18	18	Liverpool-Olympiacos	27	47
Barcelona-Bremen	22	14	M. Tel-Aviv-Juventus	28	28
Milan-Schalke	42	42	Bremen-Panathinaikos	2	2
Rapid-Juventus	36	52			

Table 1: UEFA Champion's League data

but using AIC and BIC, it is observed that BGE provides a better fit than the Marshall-Olkin bivariate exponential model. Moreover, it is also observed by the authors, Kundu and Gupta (2009), using the TTT plot of Aarset (1987), that both X_1 and X_2 have increasing hazard functions and that also explains why BGE provides a better fit than the Marshall-Olkin bivariate exponential model.

Now we provide in Table 2 the MLEs, the associate 95% confidence intervals (within brackets below) and the log-likelihood (LL) values of the four bivariate proportional hazard models mentioned above. It should be noted that all the data points are divided by 100 just for computational purposes. In all the cases we have used the EM algorithm to compute the MLEs of the unknown parameters and we have used the idea of Louis (1982) to compute the corresponding confidence intervals from the EM algorithm.

Models	λ	β	α_1	α_2	α_3	LL
BEW	3.8526	1.1123	1.2071	0.3387	0.8434	-17.23
	(2.716, 4.989)	(0.715, 1.510)	(0.561, 1.853)	(0.088, 0.589)	(0.418,1.269)	
BER	4.0263	—	0.4921	0.1659	0.4101	-22.73
	(2.514, 5.538)		(0.170, 0.734)	(0.064, 0.268)	(0.246, 0.574)	
BGE	3.8991	—	1.4552	0.4686	1.1703	-20.59
	(2.800, 4.499)		(0.657, 2.233)	(0.167, 0.769)	(0.651, 1.689)	
BGL	3.8631	0.0125	1.4361	0.4655	1.1636	-18.25
	(2.772, 4.955)	(0.007, 0.018)	(0.653,2.220)	(0.166,0.765)	(0.648,1.680)	

Table 2: The MLEs, associates 95% confidence intervals and the corresponding log-likelihood (LL) values for four different bivariate proportional reversed hazard models.

From Table 2 it is clear based on the log-likelihood values that BGE fits better than BER (both are four-parameter models) and BEW fits better than BGL ((both are five-parameter models). Now if we want to test the following hypothesis: H_0 : BGE, vs. H_1 : BEW, since $p > 0.25$, H_0 cannot be rejected. Even based on AIC and BIC, BGE is preferable than BEW.

Data Set 2: This data set is obtained from the American Football (National Football League) League from the matches on three consecutive weekend in 1986. The data were first published in 'Washington Post', and they are available in Csorgo and Welsch (1989).

In this bivariate data set (X_1, X_2), the variable X_1 represents the game time to the first points scored by kicking the ball between goal posts and X_2 represents the 'game time' by moving the ball into the end zone.

The data in average time in minutes and seconds, are represented in Table 3. The variables X_1 and X_2 have the following structure: (i) $X_1 < X_2$ means that the first score is a field goal, (ii) $X_1 = X_2$ means the first score is a converted touchdown, (iii) $X_1 > X_2$ means the first score is an unconverted touchdown or safety. In this case the ties are exact because no 'game time' elapses between a touchdown and a point-after conversion attempt. Therefore, here ties occur quite naturally and they cannot be ignored.

X_1	X_2	X_1	X_2	X_1	X_2
2:03	3:59	5:47	25:59	10:24	14:15
9:03	9:03	13:48	49:45	2:59	2:59
0:51	0:51	7:15	7:15	3:53	6:26
3:26	3:26	4:15	4:15	0:45	0:45
7:47	7:47	1:39	1:39	11:38	17:22
10:34	14:17	6:25	15:05	1:23	1:23
7:03	7:03	4:13	9:29	10:21	10:21
2:35	2:35	15:32	15:32	12:08	12:08
7:14	9:41	2:54	2:54	14:35	14:35
6:51	34:35	7:01	7:01	11:49	11:49
32:27	42:21	6:25	6:25	5:31	11:16
8:32	14:34	8:59	8:59	19:39	10:42
31:08	49:53	10:09	10:09	17:50	17:50
14:35	20:34	8:52	8:52	10:51	38:04

Table 3: American Football League (NFL) data

The data set was analyzed by Csorgo and Welsch (1989) by using the Marshall-Olkin bivariate exponential model. Csorgo and Welsch (1989) proposed a test procedure, where the null hypothesis is that the data are coming from Marshall-Olkin bivariate exponential. The test rejects the null hypothesis. They claimed that X_1 may be exponential but X_2 is not exponential. No further investigations were made.

We analyze the data set using the four proposed bivariate reversed hazard models. First the seconds in the data have been converted to the decimal minutes similarly as in Csorgo and Welsh (1989), *i.e.* 2:03 has been converted to 2.05, 3:59 to 3.98 and so on. It should further be noted that the possible scoring times are restricted by the duration of the game but it has been ignored similarly as in Csorgo and Welsh (1989). Here also all the data points are divided by 100 just for computational purposes.

The MLEs of the different parameters for four different models are presented in Table 4. In this case between BGE and BER, BER provides a better fit and between BEW and

Models	λ	β	α_1	α_2	α_3	LL
BEW	7.2994 (4.184, 10.415)	0.4124 (0.196, 0.628)	0.3347 (0.186, 0.484)	4.0957 (1.898, 6.294)	8.1678 (6.020,10.316)	-22.56
BER	18.0844 (2.514, 5.538)	—	0.0152 (0.170, 0.734)	0.1880 (0.064, 0.268)	0.3705 (0.246, 0.574)	-36.53
BGE	9.5634 (7.122, 12.005)	—	0.0481 (0.030, 0.066)	0.5959 (0.423, 0.769)	1.1706 (0.673, 1.669)	-38.25
BGL	6.6721 (2.456, 10.888)	1.1267 (0.651, 1.601)	0.0389 (0.027,0.051)	0.4823 (0.224,0.740)	0.9467 (0.489,1.405)	-34.75

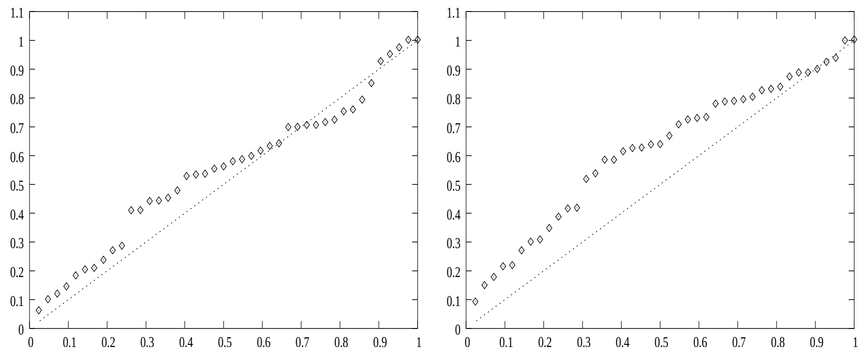
Table 4: The MLEs, associates 95% confidence intervals and the corresponding log-likelihood (LL) values for four different bivariate proportional reversed hazard models.

BGL, BEW provides a better fit. Now consider the following testing of hypothesis problem; H_0 : BER, vs. H_1 : BEW, since $p < 0.001$, H_0 is rejected. Even based on AIC and BIC, BEW is preferable than BER.

Now we would like to see the empirical hazard functions of X_1 and X_2. We have plotted in Figure 1, the scaled TTT transform of X_1 and X_2, as suggested by Aarset (1987). It is clear from Figure 1 that TTT transform of X_2 is concave where the TTT transform of X_1 is first concave and then convex. It indicates that the hazard function of X_2 will be an increasing function and the hazard function of X_1 will be an unimodal (inverted bathtub) shape. Therefore, although Csorgo and Welsh (1989) indicated that the hazard function of X_1 is constant (exponential distribution), but it may not be true. Moreover, it also explains why BEW provides a better fit than BGL and BER provides a better fit than BGE.

Acknowledgements:

The authors would like to thank the referee and the editor Professor Ayanendranath Basu for their valuable comments.



(a) (b)

Figure 1: Scaled TTT transform of (a) X_1 and (b) X_2

References

[1] Aarset, M.V. (1987), "How to identify a bathtub hazard rate?", *IEEE Transactions on Reliability*, vol. 36, 106 -108.

[2] 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, 927-929.

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

[4] Chandra, N.N. and Roy, D. (2001), "Some results on reversed hazard rate", *Probab. in Eng. and Inform. Sci.*, vol. 15, (2001), 95 - 102.

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

[6] Csorgo, S. and Welsh, A.H. (1989), "Testing for exponential and Marshall-Olkin distribution", *Journal of Statistical Planning and Inference*, vol. 23, 287 - 300.

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

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

[9] Gupta, R.D. and Kundu, D. (2007), "Generalized exponential distribution: existing methods and recent developments", *Journal of the Statistical Planning and Inference* vol. 137, 3537 - 3547.

[10] Gupta, R.D. and Nanda, A.K. (2001), "Some results on (reversed) hazard rate ordering", *Communications in Statistics - Theory and Methods*, vol. 30, 2447 - 2458.

[11] Gupta, R.D., Gupta, R.C. and Sankaran, P.G. (2004), "Some characterization results based on the (reversed) hazard rate function", *Communications in Statistics - Theory and Methods*, vol. 33, 3009 - 3031.

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

[13] Gupta, R. C., Gupta, P. L. and Gupta, R. D. (1998), "Modeling failure time data by Lehmann alternatives", *Communications in Statistics - Theory and Methods*, vol. 27, 887 - 904.

[14] Ibragimov, I.A. (1956), 'On the composition unimodal distribution", *Theory of Probability and its Applications*, vol. 1, 255 - 260.

[15] 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.

[16] Kundu, D. and Gupta, R. D. (2009), "Bivariate generalized exponential distribution', *Journal of Multivariate Analysis*, vol. 100, no. 4, 581 - 593.

[17] Kundu, D. and Raqab, M.Z. (2005), " Generalized Rayleigh Distribution: Different Methods of Estimation " *Computational Statistics and Data Analysis*" , vol. 49, 187-200.

[18] Lehmann, E.L. (1953), "The power of rank test", *Annals of Mathematical Statistics*, vol. 24, 23 - 42.

[19] Louis, T. A. (1982), "Finding the observed information matrix when using the EM algorithm", *Journal of the Royal Statistical Society*, Series B 44, 2, 226233.

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

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

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

[23] Mudholkar, G.S., Srivastava, D.K. and Freimer, M. (1995), "The exponentiated Weibull family: a reanalysis of the bus-motor-failure data", *Technometrics*, vol. 37, 436 - 445..

[24] Nanda, A. K. and Gupta, R. D. (2001), "Some properties of reversed hazard function", *Statistical Methods*, vol 3, 108 - 124.

[25] Sarhan, A. and Balakrishnan, N. (2007), "A new class of bivariate distribution and its mixture", *Journal of the Multivariate Analysis*, vol. 98, 1508 - 1527.

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

[27] Shawky, A.I. and Abu-Zinadah, H.H. (2008), "Characterizations of the exponentiated Pareto distribution based on record values", *Applied Mathematical Sciences*, vol. 2, 1283 - 1290.

[28] Surles, J.G. and Padgett, W.J. (1998), "Inference for P(Y $<$ X) in the Burr Type X model", *Journal of Applied Statistical Science*, vol. 7, 225-238.

[29] Surles, J.G. and Padgett, W.J. (2001), "Inference for reliability and stress-strength for a scaled Burr Type X distribution", *Lifetime Data Analysis*, vol. 7, 187-200.

[30] Surles, J.G. and Padgett, W.J. (2005), "Some properties of a scaled Burr type X distribution", *Journal of Statistical Planning and Inference*, vol. 128, 271 - 280. ;

[31] Tsodikov, A.D., Aselain, B.,Yakovlev, A.Y. (1997), "A distribution of tumor size at detection: an application to breast cancer data", *Biometrics* vol. 53, 1495 1502.