The Bivariate Generalized Rayleigh Distribution

Ammar M. Sarhan[1]

[1]*Department of Mathematics & Statistica, Dalhouise University, Canada*

Article Info

Abstract

This paper introduces a new bivariate distribution named the bivariate generalized Rayleigh distribution (BVGR). The proposed distribution is of type of Marshall-Olkin (MO) distribution. The BVGR distribution has generalized Rayleigh marginal distributions. The joint cumulative distribution function, the joint survival function, the joint probability density function and the joint hazard rate function of the proposed distribution are obtained in closed forms. Statistical properties of the BVGR distribution are investigated. The maximum likelihood and Bayes methods are applied to estimate the unknown parameters. Both maximum likelihood and Bayes estimates are not obtained analytically. Therefore, numerical algorithms are required to report on the model parameters and its reliability characteristics. Markov Chain Monte Carlo (MCMC) algorithm is applied for the Bayesian method. A real data set is analyzed using the proposed distribution and compared it with existing distributions. It is observed that the BVGR model fits this dataset better than the MO and the bivariate generalized exponential (BVGE) distributions.

1. Introduction

Surles and Padgett [15] introduced the two parameter Burr Type X distribution and also named as the generalized Rayleigh (GR) distribution. The GR distribution is a particular member of the exponentiated Weibull distribution, originally proposed by Mudholkar and Srivastava [11]. Surles and Padgett [16] and Al-khedhari et al. [1] discussed the parameters' estimations of this distribution using different techniques.

A new bivariate distribution is proposed by Sarhan and Balakrishnan [14], now known as Sarhan-Balakrishnan bivariate (SBBV) distribution, using the GE distribution and exponential distribution and they derived several interesting properties of this new distribution. Although the GE and exponential distributions are used to define the SBBV, the marginal distributions of SBBV distribution are not in known forms. Kundu et al. [5] modified the SBBV to include a scale parameter and discussed the parameters' estimation using maximum likelihood method. Kundu and Gupta [6] followed the same idea using the GE distribution to provide a new bivariate distribution called the bivariate generalized exponential (BVGE) distribution so that the marginal distributions are GE distributions. None of the marginal distributions of the SBBV and BVGE distributions accommodates a bathtub shaped of the hazard rate function. The lack of the bathtub shaped property limits the application of the SBBV and BVGE distributions. Kundu and Gupta [6] derived several interesting properties of the BVGE distribution and discussed the maximum likelihood estimates (MLEs) of the unknown parameters of the distribution. Kundu and Gupta [6] used the BVGE distribution to re-analyze a real data set that was originally analyzed by Meintanis [10] using the bivariate Marshal-Olkin (MO) distribution and concluded that the BVGE distribution provides a better fit than the MO distribution.

The main aim of this paper is to use the similar idea as of Sarhan and Balakrishnan [14] to introduce a new bivariate generalized Rayleigh (BVGR) distribution, using the GR distributions, so that the marginal distributions are GR distributions. The hazard rate functions of marginals of the BVGR distribution can be either increasing or decreasing or constant or of bathtub shaped. This property enriches the application of the BVGR distribution comparing to both the MO and the BVGE distributions. The proposed distribution has four parameters. The BVGR distribution can be interpreted as the joint distribution of the lifetimes of the two components of a reliability system that consists of two non-independent and non-identical components each follows a univariate GR lifetime distribution. The joint cumulative distribution function (jcdf), the joint survival function (jsf), the joint probability density function (jpdf) and the the joint hazard rate function (jhrf) of the BVGR distribution are derived in closed forms. The maximum likelihood and Bayesian methods are used to estimate the four unknown parameters of the BVGR distribution and some of its reliability measures. None of these estimates is derived in a closed form. Therefore, numerical methods are required to calculate them. For Bayesian method, we apply the Markov Chain Monte Carlo.

Email addresses and ORCID numbers: ammar.sarhan@dal, 0000-0002-7750-6910 (A. M. Sarhan)

The article is organized as follows. The BVGR distribution and discusses some of its statistical properties are presented and discussed in Sections 2 and 3. Parameters' estimations of the BVGR distribution are discussed in Section 4. Section 5 analyses a real data set and compares the ability of the BVGR with the MO and BVGE distributions to fit that data set. Finally, Section 6 concludes the paper.

2. The bivariate generalized Rayleigh distribution

The cumulative distribution function of the univariate GR distribution is, Surles and Padgett [15],

$$F_{\text{GR}}(x;\lambda,\alpha) = \left(1 - e^{-(\lambda x)^2}\right)^{\alpha}, \ x \geq 0; \ \alpha,\lambda > 0. \tag{2.1}$$

The corresponding probability density function (pdf) is

$$f_{\text{GR}}(x;\lambda,\alpha) = 2\alpha\lambda^2 x e^{-(\lambda x)^2}\left(1 - e^{-(\lambda x)^2}\right)^{\alpha-1}, \ x \geq 0; \ \alpha,\lambda > 0. \tag{2.2}$$

Here α and λ are the shape and scale parameters, respectively. We will use $\text{GR}(\alpha,\lambda)$ to denoted the GR distribution with parameters α and β. The $\text{GR}(\alpha,\lambda)$ generalizes the Rayleigh distribution.

Now, suppose that there three mutually independent random variables, say U_j, $j = 1,2,3$. The random variable U_j follows $\text{GR}(\alpha_j,\lambda)$ distribution, $j = 1,2,3$. Define $X_i = \max\{U_i, U_3\}$, $i = 1,2$. The bivariate vector (X_1,X_2) follows the bivariate generalized Rayleigh distribution with the shape parameters α_1, α_2, α_3 and scale parameter λ. We will denote it by $\text{BVGR}(\lambda,\alpha_1,\alpha_2,\alpha_3)$. To simplify notation, we write $\alpha_{123} = \alpha_1 + \alpha_2 + \alpha_3$ and $\alpha_{i3} = \alpha_i + \alpha_3$ for $i = 1,2$.

Before providing the details of the BVGR distribution, we first present some practical applications of this distribution to show how it may occur in practice.

Maintenance Model: Suppose a two-component system, each component has been maintained independently and there is an overall common maintenance for the system. Due to component maintenance, suppose the lifetime of the ith component is increased by a random amount U_j, $j = 1,2$, and because of the overall common maintenance, the lifetime of each component is increased by a random amount U_3. Thus, the lifetimes of the component a and 2 are $X_1 = \max\{U_1,U_3\}$ and $X_2 = \max\{U_2,U_3\}$, respectively.

Stress Model: Suppose a system has two components. Component i is subject to individual independent stress say U_i, $i = 1,2$. The system has an overall stress U_3 which has been transmitted to both the components equally, independent of their individual stresses. Therefore, the observed stress at component i is $X_i = \max\{U_i,U_3\}$, $i = 1,2$.

The following results provide the joint cdf, joint pdf, joint sf, joint hazard rate function and conditional pdf.

Theorem 2.1. *If (X_1,X_2) follows BVGR$(\lambda,\alpha_1,\alpha_2,\alpha_3)$, then the joint cdf of (X_1,X_2) for $x_1 > 0, x_2 > 0$, takes the form*

$$F_{X_1,X_2}(x_1,x_2) = \prod_{i=1}^{3}\left\{1 - e^{-(\lambda x_i)^2}\right\}^{\alpha_i} \tag{2.3}$$

where $x_3 = \min\{x_1,x_2\}$.

Proof. This result is a direct consequence of the definition of X_1 and X_2.

The joint cdf of the BVGR$(\lambda,\alpha_1,\alpha_2,\alpha_3)$ can also be written as

$$
\begin{aligned}
F_{X_1,X_2}(x_1,x_2) &= \prod_{i=1}^{3} F_{\text{GR}}(x_i;\lambda,\alpha_i) \\
&= \begin{cases}
F_{\text{GR}}(x_1;\lambda,\alpha_{13})F_{\text{GR}}(x_2;\lambda,\alpha_2) & if \quad x_1 < x_2 \\
F_{\text{GR}}(x_1;\lambda,\alpha_1)F_{\text{GR}}(x_2;\lambda,\alpha_{23}) & if \quad x_2 < x_1 \\
F_{\text{GR}}(x;\lambda,\alpha_{123}) & if \quad x_1 = x_2 = x.
\end{cases}
\end{aligned} \tag{2.4}
$$

Theorem 2.2. *Let (X_1,X_2) follow the BVGR$(\lambda,\alpha_1,\alpha_2,\alpha_3)$, then the joint pdf of (X_1,X_2) is*

$$f_{X_1,X_2}(x_1,x_2) = \begin{cases}
f_1(x_1,x_2) & if \ \infty > x_2 > x_1 > 0, \\
f_2(x_1,x_2) & if \ \infty > x_1 > x_2 > 0, \\
f_0(x) & if \ \infty > x_1 = x_2 = x > 0,
\end{cases} \tag{2.5}$$

where

$$
\begin{aligned}
f_1(x_1,x_2) &= f_{GR}(x_1;\lambda,\alpha_{13})f_{GR}(x_2;\lambda,\alpha_2) \\
&= 4\lambda^4\alpha_2\alpha_{13}x_1x_2 e^{-\lambda^2(x_1^2+x_2^2)}\left(1 - e^{-\lambda^2 x_1^2}\right)^{\alpha_{13}-1}\left(1 - e^{-\lambda^2 x_2^2}\right)^{\alpha_2-1} \\
f_2(x_1,x_2) &= f_{GR}(x_1;\lambda,\alpha_1)f_{GR}(x_2;\lambda,\alpha_{23}) \\
&= 4\lambda^4\alpha_1\alpha_{23}x_1x_2 e^{-\lambda^2(x_1^2+x_2^2)}\left(1 - e^{-\lambda^2 x_1^2}\right)^{\alpha_1-1}\left(1 - e^{-\lambda^2 x_2^2}\right)^{\alpha_{23}-1} \\
f_0(x) &= \frac{\alpha_3}{\alpha_{123}}f_{GR}(x;\lambda,\alpha_{123}) \\
&= 2\lambda^2\alpha_3 x e^{-\lambda^2 x^2}\left(1 - e^{-\lambda^2 x^2}\right)^{\alpha_{123}-1}.
\end{aligned}
$$

Proof. *The forms of $f_1(.,.)$ and $f_2(.,.)$ can be directly obtained by differentiating $F_{X_1,X_2}(x_1,x_2)$ in (2.4) with respect to x_1 and x_2 for $x_1 > x_2$ and $x_2 > x_1$. But, $f_0(.)$ can not be derived by using the following identity:*

$$\int_0^\infty\int_0^{x_2} f_1(x_1,x_2)dx_1dx_2 + \int_0^\infty\int_0^{x_1} f_2(x_1,x_2)dx_2dx_1 + \int_0^\infty f_0(x)dx = 1 \tag{2.6}$$

from which we get $f_0(x)$ as given above, which completes the proof of the theorem. □

Comment 2.1: From Theorems 2.1 and 2.2, one can easily verify that if $\alpha_1 + \alpha_3 = \alpha_2 + \alpha_3 = 1$, then both X_1 and X_2 are Rayleigh distributed. Let $\alpha_3 = \alpha$ and $\alpha_1 = 1 - \alpha > 0$ and $\alpha_2 = 1 - \alpha > 0$, then the joint pdf of (X_1, X_2) takes the following form

$$f_{X_1,X_2}(x_1,x_2) = \begin{cases} f_{\mathrm{GR}}(x_1;\lambda,1) f_{\mathrm{GR}}(x_2;\lambda,1-\alpha) & \text{if } \infty > x_2 > x_1 > 0, \\ f_{\mathrm{GR}}(x_1;\lambda,1-\alpha) f_{\mathrm{GR}}(x_2;\lambda,1) & \text{if } \infty > x_1 > x_2 > 0, \\ \frac{\alpha}{2-\alpha} f_{\mathrm{GR}}(x;\lambda,2-\alpha) & \text{if } \infty > x_1 = x_2 = x > 0. \end{cases} \tag{2.7}$$

Therefore, the joint pdf given in (2.7) has Rayleigh marginal distributions.

In the following we show that the BVGR distribution has both a singular part and an absolute continuous part similar to the Marshal-Olkin's bivariate exponential distribution, Sarhan and Balkrishnan bivariate distribution and the bivrariate generalized exponential distribution provided by Kundu and Gupta [6]. The function $f_{X_1,X_2}(.,.)$ may be considered to be a density function for the BVGR distribution if it is understood that the first two terms are densities with respect to two-dimensional Lebesgue measure and the third term is a density function with respect to one dimensional Lebesgue measure, see for example Bemis et al. [3]. It is well known that although in one dimension the practical use of a distribution with this property is usually pathological, but they do arise quite naturally in higher dimension. In case of BVGR distribution, the presence of a singular part means that if X_1 and X_2 are BVGR distribution, then $X_1 = X_2$ has a positive probability. In many practical situations it may happen that X_1 and X_2 both are continuous random variables, but $X_1 = X_2$ has a positive probability, see Marshall and Olkin [9], in this connection. In the following, we provide the explicit forms of the absolute continuous and the singular parts of the BVGR distribution.

Theorem 2.3. *If (X_1, X_2) follows the BVGR$(\lambda, \alpha_1, \alpha_2, \alpha_3)$, then*

$$F_{X_1,X_2}(x_1,x_2) \quad = \quad \frac{\alpha_3}{\alpha_{123}} F_s(x_1,x_2) + \frac{\alpha_{12}}{\alpha_{123}} F_a(x_1,x_2) \tag{2.8}$$

where the singular and the absolute continuous parts $F_s(.,.)$ and $F_a(.,.)$ are, for $x_3 = \min\{x_1,x_2\}$,

$$F_s(x_1,x_2) \quad = \quad \left(1 - e^{-\lambda^2 x_3^2}\right)^{\alpha_{123}}, \tag{2.9}$$

and

$$F_a(x_1,x_2) = \frac{\alpha_{123}}{\alpha_{12}} \prod_{i=1}^{3} \left(1 - e^{-\lambda^2 x_i^2}\right)^{\alpha_i} - \frac{\alpha_3}{\alpha_{12}} \left(1 - e^{-\lambda^2 x_3^2}\right)^{\alpha_{123}}. \tag{2.10}$$

Proof. The joint cdf $F_{X_1,X_2}(x_1,x_2)$ can be written as

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

Let $A = \{U_1 < U_3\} \cap \{U_2 < U_3\} \equiv \{X_1 = X_2\}$, therefore,

$$P(A) = \int_0^\infty 2\alpha_3 \lambda^2 x e^{-\lambda^2 x^2} \left(1 - e^{-\lambda^2 x^2}\right)^{\alpha_{123}-1} dx = \frac{\alpha_3}{\alpha_{123}},$$

and

$$\begin{aligned} F_s(x_1,x_2) \quad &= \quad P(X_1 \leq x_1, X_2 \leq x_2|A) \\ &= \quad \alpha_{123} \int_0^{x_3} 2\lambda^2 x e^{-\lambda^2 x^2} \left(1 - e^{-\lambda^2 x^2}\right)^{\alpha_{123}-1} dx \\ &= \quad \left(1 - e^{-\lambda^2 x_3^2}\right)^{\alpha_{123}}. \end{aligned}$$

Once, $P(A)$ and $F_s(x_1,x_2)$ are obtained, the function $F_a(x_1,x_2)$ can be derived by subtraction as given in (2.10). It can be easily shown that F_s is the singular part as its mixed second partial derivatives is zero when $x_1 \neq x_2$, and F_a is the absolute continuous part as its mixed second partial derivative gives a density function. $\square$

Corollary 2.4. *If (X_1, X_2) follows the BVGR$(\lambda, \alpha_1, \alpha_2, \alpha_3)$, then the joint pdf of (X_1, X_2) can be expressed as a mixture of the singular and absolute parts as*

$$f_{X_1,X_2}(x_1,x_2) = \frac{\alpha_3}{\alpha_{123}} f_s(x_3) + \frac{\alpha_{12}}{\alpha_{123}} f_a(x_1,x_2) \tag{2.11}$$

where

$$f_s(x_3) \quad = \quad f_{\mathrm{GR}}(x_3; \alpha_{123}, \lambda) \ \text{if } x_1 = x_2 = x_3 = \min\{x_1,x_2\}$$
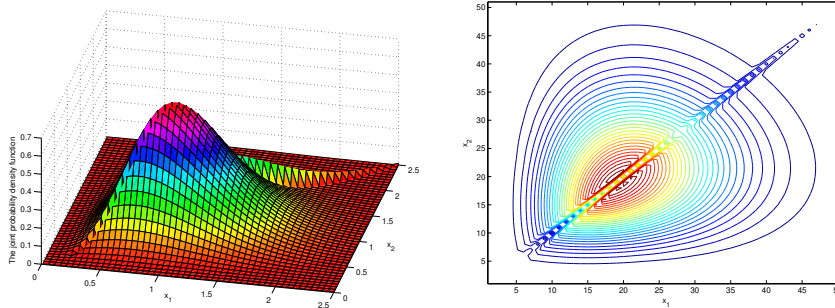
and

$$f_a(x_3) = \frac{\alpha_{123}}{\alpha_{12}} \begin{cases} f_{\mathrm{GR}}(x_1,\alpha_{13},\lambda) f_{\mathrm{GR}}(x_2;\alpha_2,\lambda) & \text{if } x_1 < x_2 \\ f_{\mathrm{GR}}(x_1,\alpha_1,\lambda) f_{\mathrm{GR}}(x_2;\alpha_{23},\lambda) & \text{if } x_1 > x_2. \end{cases}$$

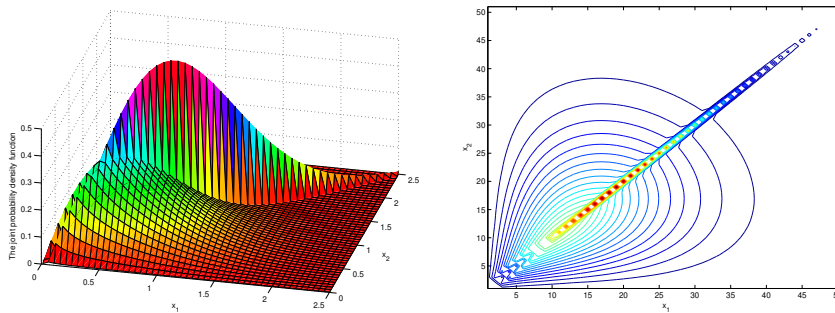*Here $f_s(.)$ and $f_a(.,.)$ are the singular and absolute continuous parts, respectively.*

Figures 2.1 and 2.2 show different shapes of the joint pdf and the jhrf of the BVGR distribution along with their contours for different sets of values of parameters.
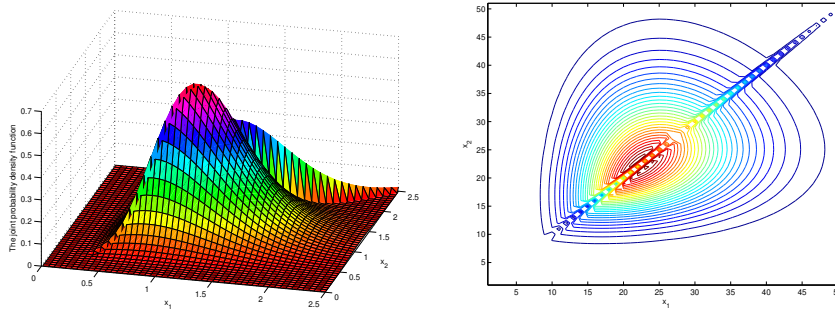
Set (1):$(\alpha_1, \alpha_2, \alpha_3, \lambda) = (1.2, 1.2, 0.2, 1)$

Set (2): $(\alpha_1, \alpha_2, \alpha_3, \lambda) = (1.2, 1.2, 2.1, 1)$

Set (3): $(\alpha_1, \alpha_2, \alpha_3, \lambda) = (1.2, 1.2, 1.0, 1)$

Set (4): $(\alpha_1, \alpha_2, \alpha_3, \lambda) = (1.2, 0.2, 2.1, 1)$

Figure 2.1: The joint probability density function of BVGR distribution.

Set (1): $(\alpha_1, \alpha_2, \alpha_3, \lambda) = (1.2, 1.2, 0.2, 1)$

Set (2): $(\alpha_1, \alpha_2, \alpha_3, \lambda) = (1.2, 1.2, 2.1, 1)$

Set (3): $(\alpha_1, \alpha_2, \alpha_3, \lambda) = (1.2, 1.2, 1.0, 1)$

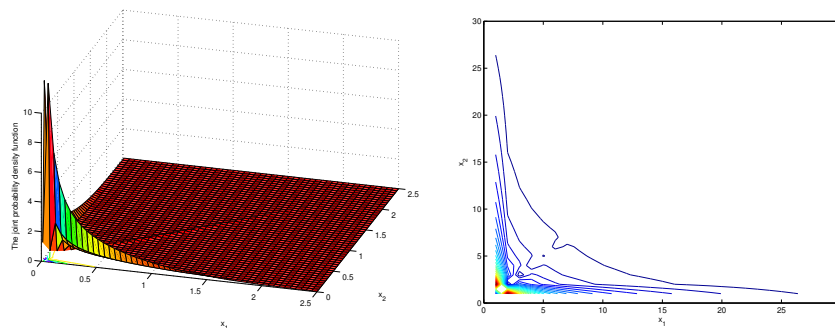Set (4): $(\alpha_1, \alpha_2, \alpha_3, \lambda) = (1.2, 0.2, 2.1, 1)$
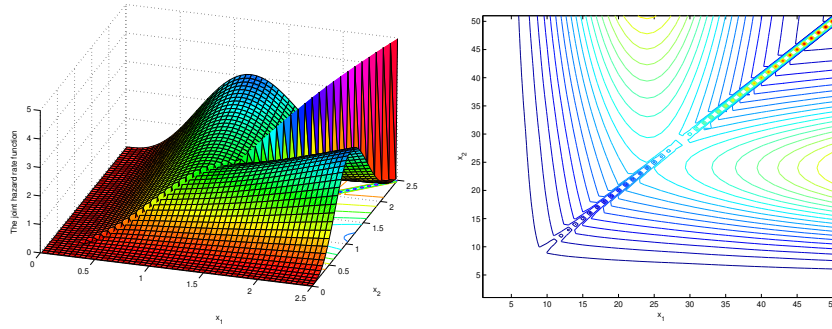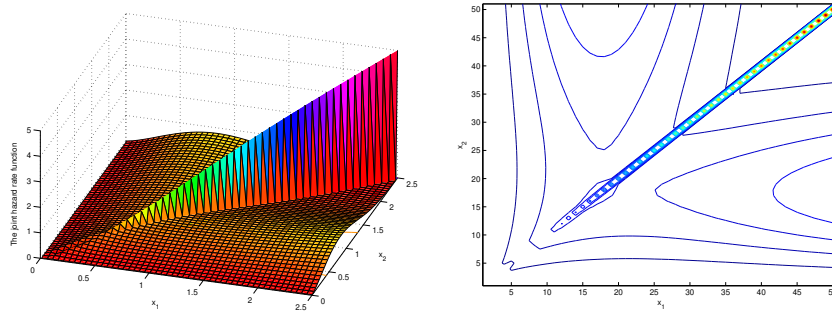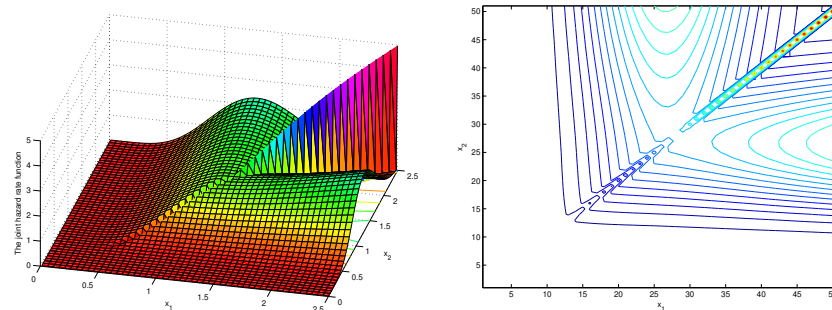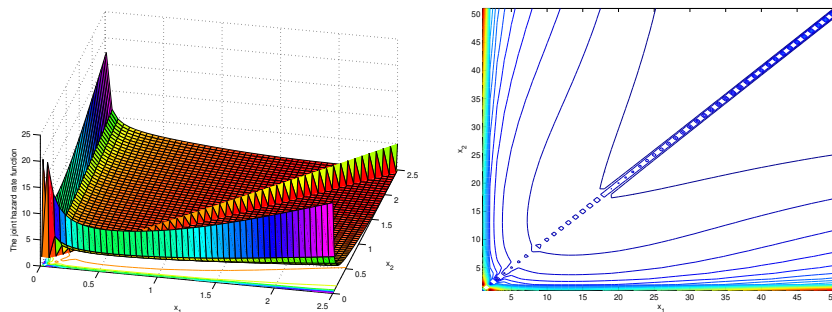

Figure 2.2: The joint hazard rate function of the BVGR distribution.

3. Statistical properties

3.1. Marginal distributions

We can easily verify that the marginal distribution of X_i, $i = 1, 2$, follows $GR(\lambda, \alpha_{i3})$. That is, the marginal cdf of X_i is

$$F_{X_i}(x_i) = \left(1 - e^{-(\lambda x_i)^2}\right)^{\alpha_{i3}}, \; x > 0. \tag{3.1}$$

Therefore, the marginal survival function of X_i is

$$S_{X_i}(x) = 1 - F_{X_i}(x) = 1 - \left(1 - e^{-(\lambda x)^2}\right)^{\alpha_{i3}}, \; x > 0. \tag{3.2}$$

The relation (3.2) gives the survival function of a parallel system with two independent units, say i and 3. The lifetime of unit i is U_i which follows $GR(\lambda, \alpha_i)$ and the lifetime of unit 3 is U_3 which follows $GR(\lambda, \alpha_3)$. Thus, the system lifetime $X_i = \max\{U_i, U_3\}$ follows $GR(\lambda, \alpha_i + \alpha_3)$.

Using (3.1), the marginal pdf of X_i is

$$f_{X_i}(x) = 2\lambda^2 \alpha_{i3} x e^{-(\lambda x)^2} \left(1 - e^{-(\lambda x)^2}\right)^{\alpha_{i3}-1}, \tag{3.3}$$

and the marginal hrf of X_i is

$$h_{X_i}(x) = \frac{2\lambda^2 \alpha_{i3} x e^{-(\lambda x)^2} \left(1 - e^{-(\lambda x)^2}\right)^{\alpha_{i3}-1}}{1 - \left(1 - e^{-(\lambda x)^2}\right)^{\alpha_{i3}}}. \tag{3.4}$$

It is observed by Raqab and Kundu (2006) that: (1) if $\alpha_{i3} \leq \frac{1}{2}$, the pdf is decreasing and the hrf takes a bathtub shape, (2) if $\alpha_{i3} > \frac{1}{2}$, the pdf is right-skewed and unimodal and the hrf is increasing. Shapes of the pdf and hrf of X_i for different values of λ and α_{i3} are provided in Figure 3.1. Surles and Padgett [15] showed that the GR distribution can be used quite effectively in modeling strength univariate data and also modeling general univariate lifetime data. Hence, we expect that the BVGR distribution can be used quite effectively in modeling strength bivariate data and also modeling general bivariate lifetime data comparing to the MO and the BVGE distributions.
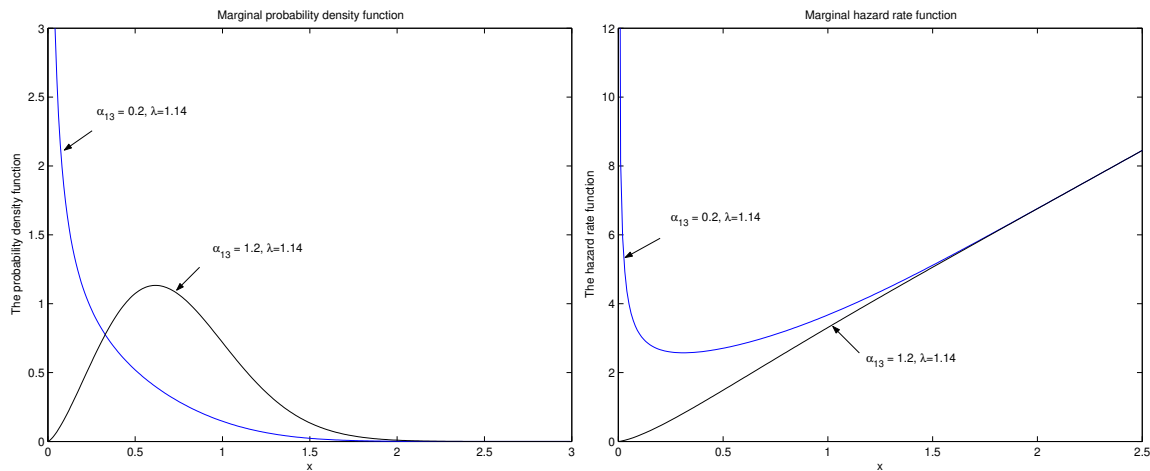

Figure 3.1: The probability density [left panel] and hazard rate [right panel] functions of the marginal distribution of X_1.

3.2. Conditional distributions

The following theorem provides the conditional pdf of X_1 given $X_2 = x_2$, say $f_{X_1|X_2}(x_1|x_2)$, which is not continuous at $X_1 = x_2$.

Theorem 3.1. *If* $(X_1, X_2) \sim BVGR(\lambda, \alpha_1, \alpha_2, \alpha_3)$, *then conditional pdf of X_1 given $X_2 = x_2$ is*

$$f_{X_1|X_2}(x_1|x_2) = \frac{\alpha_3}{\alpha_{123}} f_{1|2}^{(s)}(x_1|x_2) + \frac{\alpha_{12}}{\alpha_{123}} f_{1|2}^{(a)}(x_1|x_2) \tag{3.5}$$

where

$$f_{1|2}^{(s)}(x_1|x_2) = \frac{\alpha_{123}}{\alpha_{23}} F_{GR}(x_2; \lambda, \alpha_1) \; if \; x_1 = x_2$$

and

$$f_{1|2}^{(a)}(x_1|x_2) = \frac{\alpha_{123}}{\alpha_{12}} \begin{cases} \frac{\alpha_2}{\alpha_{23}} \frac{f_{GR}(x_1; \lambda, \alpha_{13})}{F_{GR}(x_2; \lambda, \alpha_3)} & if \; x_1 < x_2, \\ f_{GR}(x_1; \lambda, \alpha_1) & if \; x_1 > x_2. \end{cases}$$

Proof. *These results are obtained using the definition of conditional probability and the results of Corollary 2.4 and equation (3.3).*

The conditional pdf of X_1 given $X_2 = x_2$ can be rewritten as

$$f_{X_1|X_2}(x_1|x_2) = \frac{\alpha_3}{\alpha_{23}} F_{GR}(x_1; \lambda, \alpha_1) I_{x_1=x_2} + f_{GR}(x_1; \lambda, \alpha_1) I_{x_1>x_2}$$
$$+ \frac{\alpha_2}{\alpha_{23}} \frac{f_{GR}(x_1; \lambda, \alpha_{13})}{F_{GR}(x_2; \lambda, \alpha_3)} I_{x_1<x_2}$$

where $I_A = 1$ if A is true and 0 otherwise.

Equivalently, $f_{X_1|X_2}(x_1|x_2)$ can be rewritten as in the following form

$$f_{X_1|X_2}(x_1|x_2) = \begin{cases} \dfrac{2\alpha_2\alpha_{13}\lambda^2 x_1 e^{-\lambda^2 x_1^2}\left(1-e^{-\lambda^2 x_1^2}\right)^{\alpha_{13}-1}}{\alpha_{23}\left(1-e^{-\lambda^2 x_2^2}\right)^{\alpha_3-1}} & x_1 < x_2, \\ 2\lambda^2\alpha_1 x_1 e^{-\lambda^2 x_1^2}\left(1-e^{-\lambda^2 x_1^2}\right)^{\alpha_1-1} & x_1 > x_2, \\ \dfrac{\alpha_3}{\alpha_{23}}\left(1-e^{-\lambda^2 x^2}\right)^{\alpha_1} & x_1 = x_2. \end{cases}$$

Obviously, the $f_{X_1|X_2}(x_1|x_2)$ is not continuous when $X_1 = x_2$. Plots in Figure 3.2 give different patterns of the conditional pdf's of X_1 given $X_2 = x_2$ plotted at different values of x_2 ($x_2 = 1.5, 0.5$) and different sets of parameters.

Similarly, the conditional pdf of X_2 given $X_1 = x_1$, which is not continuous at $X_2 = x_1$, can be derived in a similar form as above.

Theorem 3.2. *If $(X_1, X_2) \sim BVGR(\lambda, \alpha_1, \alpha_2, \alpha_3)$, then the conditional cumulative distribution function of X_1 given $X_2 \leq x_2$, say $F_{X_1|X_2\leq x_2}(x_1)$, is an absolute continuous function which is given by*

$$F_{X_1|X_2\leq x_2}(x_1) = P(X_1 \leq x_1 | X_2 \leq x_2)$$
$$= \frac{F_{GR}(x_1; \alpha_1) F_{GR}(x_3; \alpha_3)}{F_{GR}(x_2; \alpha_3)}$$
$$= \begin{cases} \left(1-e^{-\lambda^2 x_1^2}\right)^{\alpha_{13}}\left(1-e^{-\lambda^2 x_2^2}\right)^{-\alpha_3} & \text{if } x_1 \leq x_2 \\ \left(1-e^{-\lambda^2 x_1^2}\right)^{\alpha_1} & \text{if } x_1 > x_2 \end{cases}$$

Proof. *These results are obtained using the definition of conditional probability and the results of Theorems 2.3 and 3.1.*

Comment: Using Theorem 3.2, different moments of X_1, X_2, and conditional moments of $X_2|X_1 = x_1$ or $X_1|X_2 = x_2$ can be obtained in terms of infinite series.

Remarks If $(X_1, X_2) \sim \text{BVGR}(\lambda, \alpha_1, \alpha_2, \alpha_3)$ distribution, then

1. The cdf of $\max\{X_1, X_2\}$ is

$$P(\max\{X_1, X_2\} \leq x) = \left(1-e^{-\lambda^2 x^2}\right)^{\alpha_1+\alpha_2+\alpha_3}. \tag{3.6}$$

   That is, the $\max\{X_1, X_2\}$ has $\text{GR}(\lambda, \alpha_1 + \alpha_2 + \alpha_3)$. This result means that the reliability function of a parallel system with three independent units such that the lifetime of unit i follows $\text{GR}(\lambda, \alpha_i)$, $i = 1, 2, 3$, is

$$R(x) = 1 - \left(1-e^{-\lambda^2 x^2}\right)^{\alpha_1+\alpha_2+\alpha_3}. \tag{3.7}$$

2. For all $0 < x_1, x_2 < \infty$,

$$F_{X_1,X_2}(x_1, x_2) \geq F_{X_1}(x_1) F_{X_2}(x_2). \tag{3.8}$$

   One can easily prove (3.8), as follows. Since

$$F_{X_1,X_2}(x_1, x_2) = \bar{F}_{U_1}(x_1) F_{U_2}(x_2) F_{U_3}(x_3), \quad x_3 = \min(x_1, x_2) \tag{3.9}$$

   and

$$F_{X_1}(x_1) F_{X_2}(x_2) = F_{U_1}(x_1) F_{U_3}(x_1) F_{U_2}(x_2) F_{U_3}(x_2). \tag{3.10}$$

   From (3.9) and (3.10), we get

$$F_{X_1,X_2}(x_1, x_2) - F_{X_1}(x_1) F_{X_2}(x_2) = F_{U_1}(x_1) F_{U_2}(x_2)$$
$$\times \begin{cases} F_{U_3}(x_1)\bar{F}_{U_3}(x_2) & \text{if } x_1 \leq x_2 \\ \bar{F}_{U_3}(x_1) F_{U_3}(x_2) & \text{if } x_1 > x_2 \end{cases}$$
$$\geq 0.$$

   which completes the proof of (3.8).

3. For all $0 < x_1, x_2 < \infty$,

$$\bar{F}_{X_1,X_2}(x_1, x_2) \geq \bar{F}_{X_1}(x_1)\bar{F}_{X_2}(x_2). \tag{3.11}$$

4. From (3.8), one can say that X_1 and X_2 are positive quadratic dependent. That is, for every pair of increasing functions $g_1(.)$ and $g_2(.)$,
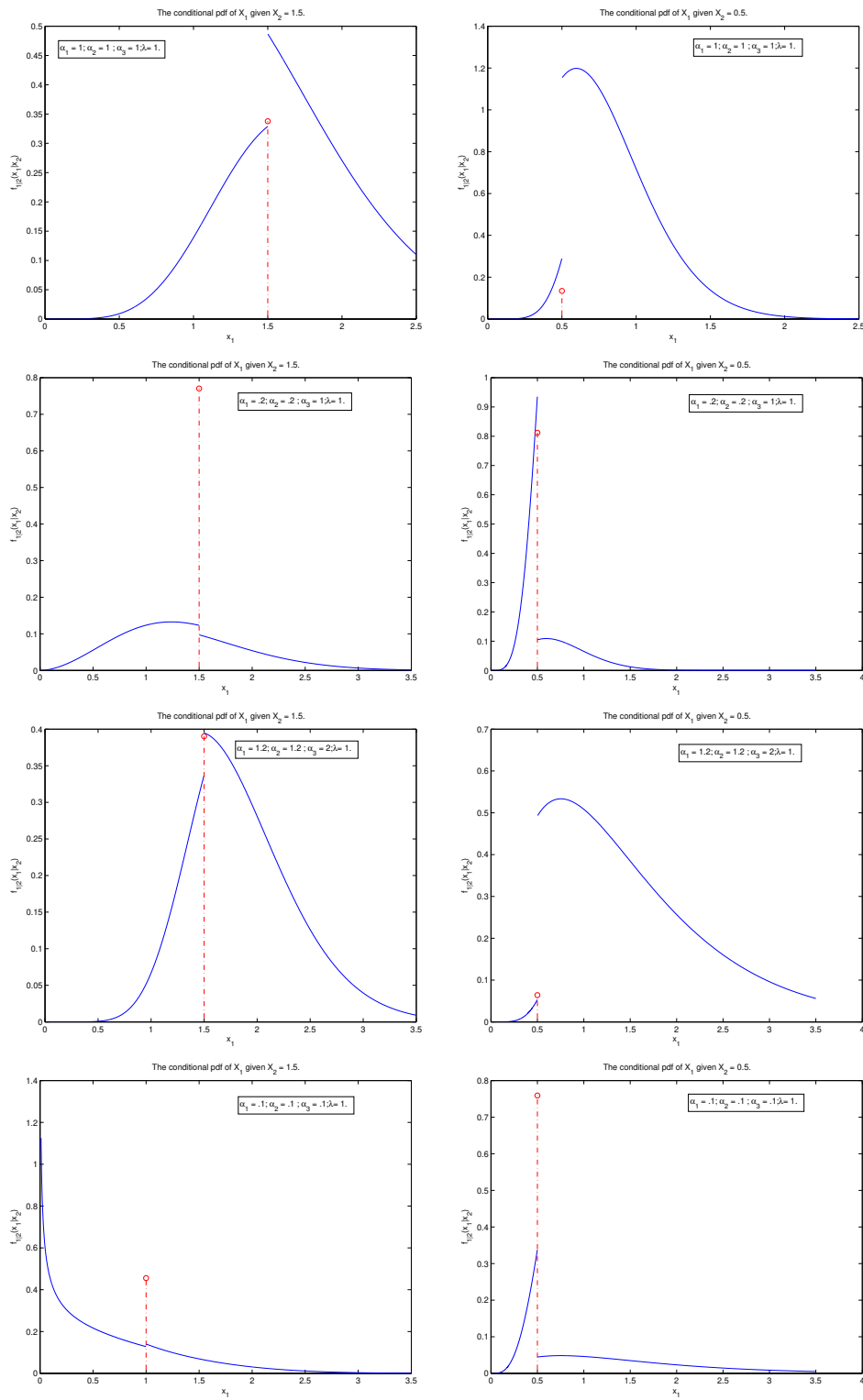
$$Cov(g_1(X_1), g_2(X_2)) \geq 0.$$

Figure 3.2: The conditional probability density function of X_1, given X_2 at different sets of the parameters.

3.3. Joint hazard rate function

Using the forms (2.3) and (3.1) together with the relations between the joint sf and the joint cdf of (X_1, X_2) given by

$$S_{X_1,X_2}(x_1,x_2) = 1 - F_{X_1}(x_1) - F_{X_2}(x_2) + F_{X_1,X_2}(x_1,x_2),$$

we can derive the joint sf of (X_1, X_2) as

$$S_{X_1,X_2}(x_1,x_2) = 1 - \sum_{j=1}^{2} \left\{ 1 - e^{-(\lambda x_j)^2} \right\}^{\alpha_{j3}} + \prod_{i=1}^{3} \left\{ 1 - e^{-(\lambda x_j)^2} \right\}^{\alpha_i}. \tag{3.12}$$

Using the relation between the joint pdf and the jsf of (X_1, X_2), one can derive the joint hazard rate function (jhrf) of (X_1, X_2) according to the following relation

$$h_{X_1,X_2}(x_1,x_2) = \frac{f_{X_1,X_2}(x_1,x_2)}{S_{X_1,X_2}(x_1,x_2)}.$$

Different shapes of the joint hrf and corresponding contour are provided in Figures 4 for different sets of values of the parameters.

4. Parameters' estimations

In this section we discuss the problem of computing the maximum likelihood estimates of the four unknown parameters of the BVRD. Let us assume that (x_{11}, x_{12}), (x_{21}, x_{22}), $\cdots$, (x_{n1}, x_{n2}) be a simple random sample "data" from the BVRD$(\alpha_1, \alpha_2, \alpha_3, \lambda)$. The likelihood function for this sample is

$$l(data; \theta) = \prod_{i=1}^{n} [f_1(x_{1i}, x_{2i})]^{I(x_{1i} < x_{2i})} [f_2(x_{1i}, x_{2i})]^{I(x_{2i} < x_{1i})} [f_0(x_{1i})]^{I(x_{1i} = x_{2i})}, \tag{4.1}$$

where $I(A)$ is an indicator function that is equal to 1 if A is true and 0 otherwise. Substituting (2.11) into (4.1) and take the natural logarithm, we get the log-likelihood function as

$$
\begin{aligned}
\mathscr{L} \;=\; & \sum_{i=1}^{n} \Bigg\{ I(x_{1i} < x_{2i}) \Big[\ln\left(4\lambda^4 \alpha_2 \alpha_{13}\right) + \ln(x_{1i} x_{2i}) - \lambda^2 \left(x_{1i}^2 + x_{2i}^2\right) + \\
& \qquad\qquad (\alpha_{13} - 1)\ln\left(1 - e^{-\lambda^2 x_{1i}^2}\right) + (\alpha_2 - 1)\ln\left(1 - e^{-\lambda^2 x_{2i}^2}\right) \Big] \\
& + I(x_{2i} < x_{1i}) \Big[\ln\left(4\lambda^4 \alpha_1 \alpha_{23}\right) + \ln(x_{1i} x_{2i}) - \lambda^2 \left(x_{1i}^2 + x_{2i}^2\right) + \\
& \qquad\qquad (\alpha_{23} - 1)\ln\left(1 - e^{-\lambda^2 x_{2i}^2}\right) + (\alpha_1 - 1)\ln\left(1 - e^{-\lambda^2 x_{1i}^2}\right) \Big] \\
& + I(x_{2i} = x_{1i}) \Big[\ln\left(2\lambda^2 \alpha_3\right) + \ln(x_{1i}) - \lambda^2 x_{1i}^2 + (\alpha_{123} - 1)\ln\left(1 - e^{-\lambda^2 x_{1i}^2}\right) \Big] \Bigg\}.
\end{aligned}
\tag{4.2}
$$

4.1. Maximum likelihood method

Maximum likelihood point estimates (MLE) of the vector of the four unknown parameters is the value of that vector that maximizes the log-likelihood function (4.2). That is the MLE of $\theta = (\alpha_1, \alpha_2, \alpha_3, \lambda)$ is the solution of the following system of four non-linear equations with respect to $\alpha_1, \alpha_2, \alpha_3$, and λ.

$$
\begin{aligned}
0 \;=\; & \frac{n_1}{\alpha_1 + \alpha_3} + \frac{n_2}{\alpha_1} + \sum_{i=1}^{n} \ln\left(1 - e^{-\lambda^2 x_{1i}^2}\right), \\[4pt]
0 \;=\; & \frac{n_2}{\alpha_2 + \alpha_3} + \frac{n_1}{\alpha_2} + \sum_{i=1}^{n} \ln\left(1 - e^{-\lambda^2 x_{2i}^2}\right), \\[4pt]
0 \;=\; & \frac{n_0}{\alpha_3} + \frac{n_1}{\alpha_1 + \alpha_3} + \frac{n_2}{\alpha_2 + \alpha_3} + \sum_{i=1}^{n} \Big[I(x_{1i} < x_{2i}) \ln\left(1 - e^{-\lambda^2 x_{1i}^2}\right) \\
& + \{ I(x_{1i} > x_{2i}) + I(x_{1i} = x_{2i}) \} \ln\left(1 - e^{-\lambda^2 x_{2i}^2}\right) \Big], \\[4pt]
0 \;=\; & 2\lambda \sum_{i=1}^{n} \Bigg\{ I(x_{1i} < x_{2i}) \left[\frac{(\alpha_2 - 1)x_{2i}^2}{e^{\lambda^2 x_{2i}^2} - 1} + \frac{(\alpha_{13} - 1)x_{1i}^2}{e^{\lambda^2 x_{1i}^2} - 1} - x_{1i}^2 - x_{2i}^2 \right] + \\
& \qquad I(x_{1i} > x_{2i}) \left[\frac{(\alpha_1 - 1)x_{1i}^2}{e^{\lambda^2 x_{1i}^2} - 1} + \frac{(\alpha_{23} - 1)x_{2i}^2}{e^{\lambda^2 x_{2i}^2} - 1} - x_{1i}^2 - x_{2i}^2 \right] + \\
& \qquad I(x_{1i} = x_{2i}) \left[\frac{(\alpha_{123} - 1)x_{1i}^2}{e^{\lambda^2 x_{1i}^2} - 1} - x_{1i}^2 \right] \Bigg\} + \frac{4}{\lambda}\left(n_1 + n_2 + \frac{1}{2}n_0\right),
\end{aligned}
\tag{4.3}
$$

where $n_1 = \sum_{i=1}^{n} I(x_{1i} < x_{2i})$, $n_2 = \sum_{i=1}^{n} I(x_{1i} > x_{2i})$, and $n_0 = \sum_{i=1}^{n} I(x_{1i} = x_{2i})$. The likelihood equations (4.3) do not have an explicit solution. Therefore, the distribution of the MLE of the parameters cannot be derived in an explicit form. Hence, we could not obtain the explicit confidence intervals for the parameters. We use R to get the MLEs and the corresponding Fisher information matrix. Using normality property of MLEs, we can construct the asymptotic confidence interval for each parameter.

4.2. Bayes estimation

Let us assume that the four parameters are independent random variables with gamma priori distributed. That is, the joint prior pdf of θ is

$$g_0(\boldsymbol{\theta}) \propto \alpha_1^{a_{11}-1} \alpha_2^{a_{21}-1} \alpha_3^{a_{31}-1} \lambda^{a_{41}-1} e^{-a_{12}\alpha_1 - a_{22}\alpha_2 - a_{32}\alpha_3 - a_{42}\lambda}, \ \alpha_1, \alpha_2, \alpha_3, \lambda > 0, \tag{4.4}$$

where the hyperparameters a_{ij}, $i = 1, 2, 3, 4$ and $j = 1, 2$ are all positive and reflect the prior knowledge about the parameters. The log-prior density function is

$$\mathscr{G}_0(\boldsymbol{\theta}) \propto (a_{41}-1)\ln(\lambda) - a_{42}\lambda + \sum_{i=1}^{3} [(a_{i1}-1)\ln(\alpha_i) - a_{i2}\alpha_i]. \tag{4.5}$$

Applying Bayes' theorem and using the likelihood function of the available data and the joint prior distribution of θ, the joint posterior probability density function of θ, given data, is

$$g(\boldsymbol{\theta}|\text{data}) = \frac{1}{K} \exp\{\mathscr{L} + \mathscr{G}_0(\boldsymbol{\theta})\}, \tag{4.6}$$

where K is the normalizing constant. Obviously, the joint posterior distribution of θ is analytically intractable. Therefore, under the squared-error loss, Bayes estimators of the parameters and/or of any parametric function of θ, say $w(\theta)$, involve ratio of two multidimensional integrals as

$$\hat{w}(\text{data}) = \frac{\int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty w(\alpha_1, \alpha_2, \alpha_3, \lambda) \exp\{\mathscr{L} + \mathscr{G}_0(\boldsymbol{\theta})\} d\alpha_1 \, d\alpha_2 \, d\alpha_3 \, d\lambda}{\int_0^\infty \int_0^\infty \int_0^\infty \int_0^\infty \exp\{\mathscr{L} + \mathscr{G}_0(\boldsymbol{\theta})\} d\alpha_1 \, d\alpha_2 \, d\alpha_3 \, d\lambda}. \tag{4.7}$$

The integrals in (4.7) do not have analytical solution. Thus, some approximation methods should be used to solve these integrals an calculate the ratio of the integrals. Lindley [7] and Tierney and Kadane [13] discussed some methods of approximations that work well for low dimension. In this paper we will use Markov Chain Monte Carlo method that works well with higher dimensions cases, see Metropolis et al. [8] and Hastings [4]. One of the main advantages of the MCMC is it does not require to calculate the integrals that are needed for the normalizing constant K and/or included in the ratio that gives the Bayes estimators.

The MCMC method generates random draws from the joint posterior distribution by generating draws from an an arbitrary distribution that easy to simulate from then apply an accept-reject method. The arbitrary distribution is named a proposal distribution that satisfies two conditions: (1) it mimics the posterior distribution, (2) easy to simulate from. In this article, we use multivariate normal distribution as a proposal. We follow steps below to generate random draws from the joint posterior distribution (4.6) without computing the normalizing constant K:

1. Set the size of the random draws we wish to generate, say M.
2. Choose an initial guess of θ, say $\theta^{(0)}$.
3. $i = 1, 2, \cdots, M$, perform the following steps:
   (a) Generate θ^* from the multivariate normal with mean $\theta^{(i-1)}$ and variance-covariance Σ.
   (b) Compute the ratio $\kappa = \min\left\{1, \dfrac{g(\theta^*|\text{data})}{g(\theta^{(i-1)}|\text{data})}\right\}$.
   (c) Generate a random value u from uniform distribution on $(0,1)$.
   (d) If $\kappa \geq u$ set $\theta^{(i)} = \theta^*$, otherwise set $\theta^{(i)} = \theta^{(i-1)}$.

We discard the early M_0 number of burn-in draws and use the remaining $M - M_0$, "$\theta^{(M_0+1)}, \theta^{(M_0+2)}, \cdots, \theta^{(M)}$", as the desired draws from the joint posterior distribution. Thus Bayes estimate of θ_j is

$$\hat{\theta}_j = \frac{\sum_{i=M_0+1}^{M-M_0} \theta_j^{(i)}}{M - M_0}, \ j = 1, 2, 3, 4.$$

Furthermore, the lower and upper bounds of the $100(1-\vartheta)\%$, $0 < \vartheta < 1$, Bayesian probability interval (BPI) of θ_j are the $\frac{\vartheta}{2} 100$th and $(1-\frac{\vartheta}{2}) 100$ th percentiles of the sequence of the $M - M_0$ draws $\theta_j^{(M_0+1)}, \theta_j^{(M_0+2)}, \cdots, \theta_j^{(M)}$, respectively.

5. Application

In order to discuss how the proposed distribution can be implemented in real life, we re-analyze the UEFA Champion's League Data. This data set was originally analyzed in Meintanis [10] using the Marshall-Olkin bivariate exponential model (MO) with three parameters $\lambda_1, \lambda_2, \lambda_3$, then by Kundu and Gupta [6] using the bivariate generalized exponential (BVGE) distribution with four parameters $\alpha_1, \alpha_2, \alpha_3$ and β. Kundu and Gupta [6] reported that the BVGE model fits the data better than MO model. In the following, we use the BVGR distribution to reanalyze this data and compare it with the MO and BVGE models.

Table 1 shows the MLEs of the unknown parameters of the proposed distribution along with the values of the log-likelihhod values and the Akaike information criterion (AIC; see [2]). The AIC suggests that the BVGR distribution provides a better fit than both the MO and BVGE distributions.

For Bayesian computations, we assumed that: (1) the four parameters follow gamma prior distributions with all hyperparameters equal and equal to 0,001, (2) the proposal distribution is multinormal with variance covariance matrix, which is the same as the Fisher information matrix,

$$\Sigma = \begin{pmatrix} 0.016461 & 9.574e-04 & 0.0011642 & -1.135e-04 \\ 0.000957 & 4.188e-03 & -0.0010077 & -2.663e-05 \\ 0.001164 & -1.008e-03 & 0.0086362 & -6.134e-05 \\ -0.000114 & -2.663e-05 & -0.0000613 & 3.688e-06 \end{pmatrix}.$$

Distribution	MLE	$\mathscr{L}$	AIC
MO	$\hat{\lambda}_1 = 0.012, \hat{\lambda}_2 = 0.014, \hat{\lambda}_3 = 0.022$	-339.006	684.012
MVGE	$\hat{\alpha}_1 = 1.351, \hat{\alpha}_2 = 0.465, \hat{\alpha}_3 = 1.153, \hat{\beta} = 0.039$	-296.935	601.870
MVGR	$\hat{\alpha}_1 = 0.492, \hat{\alpha}_2 = 0.166, \hat{\lambda}_3 = 0.410, \hat{\lambda} = 0.020$	-293.357	594.714

Table 1: The MLEs of the parameters, the log-likelihood values and AIC values.
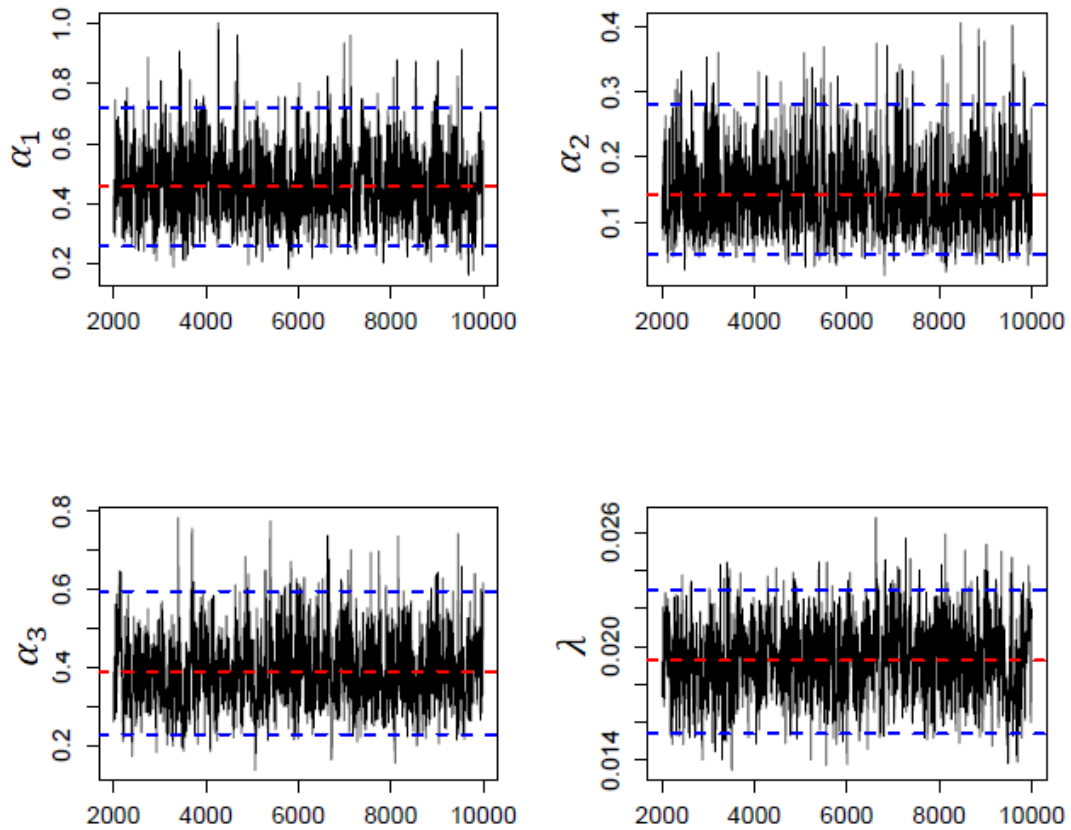

Figure 5.1: The trace plot of the random draws from the joint posterior distribution along with the posterior mean and the limits of the 95% BPI after discarding the early 2000 draws.

The acceptance rate is 37.07% which is very high for a four parameter case. The trace plots of the draws are plotted in Figure 5.1 after discarding the early 2000 draws (burn-in period). The trace plots show a good mix of the simulated draws. As a further diagnostic test for the draws, we provide the autocorrelations (ACFs) of the simulated draws from the joint posterior distribution of the four parameters after discarding the early 20% of the draws as shown in Figure 5.2. From the ACFs plot, we see that the Lag goes to zero very rapidly which indicates the draws converge to the actual distribution very fast.

Using the simulated draws, the posterior mean (Bayes estimates under squared error loss), median (Bayes estimate under absolute error loss) and the bounds of a 95% Bayesian probability intervals for the parameters are calculated as given in Table 2. Furthermore, the marginal posterior density functions are estimated as presented in Figure **??**.

6. Conclusions

A new bivariate lifetime distribution named bivariate generalized Rayleigh (BVGR) distribution function whose marginals are generalized Rayleigh distributions is proposed in this paper. The BVGR distribution is of Marshal-Olkin type. It is observed that the BVGR distribution is a singular distribution and has an absolute continuous part and a singular part. Since the joint distribution function and the joint density

Parameter	Mean	Median	95% Bayesian Probability Interval	
			Lower limit	upper limit
α_1	0.4569	0.4467	0.2568	0.7194
α_2	0.1423	0.1337	0.0507	0.2802
α_3	0.3895	0.3794	0.2288	0.5951
λ	0.0193	0.0192	0.0154	0.0230

Table 2: The basic posterior characteristics for the four parameters of the BVGR distribution.
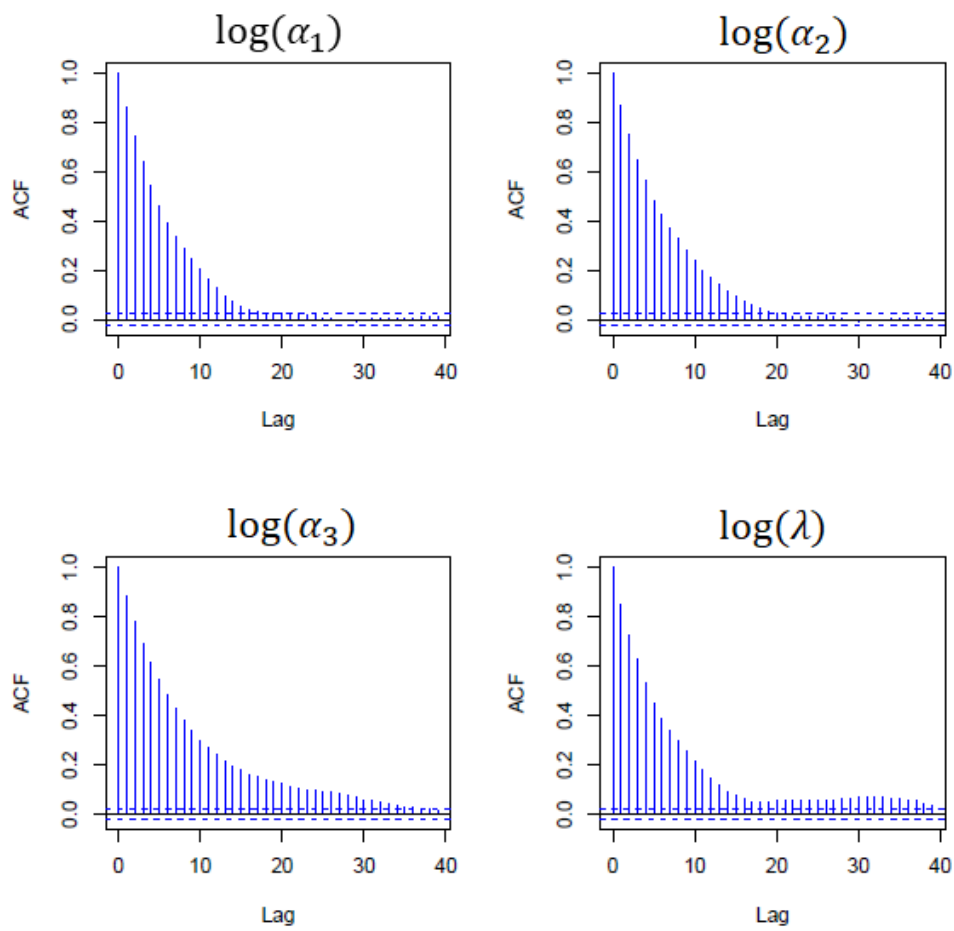
Figure 5.2: The autocorrelations of the simulated draws from the joint posterior distribution of the four parameters after discarding the early 20% of the draws.

function are in closed forms, this distribution can be used in practice for non-negative and positively correlated random variables. We investigated some statistical properties of the proposed distribution. Also, the marginal and conditional distribution of the BVGR distribution are derived in closed forms. The maximum likelihood and Bayesian methods are applied to estimate the four unknown parameters of the BVGR distribution. For Bayesian method, we used the Markov chain Monte Carlo (MCMC) method. One real data set is analyzed using the BVGR distribution which showed a better fit than the MO and BVGE distributions for this data set.

Acknowledgement

References

[1] Al-khedhari A, Sarhan AM and Tadj L. Estimation of the generalized Rayleigh distribution parameters, *International Journal of Reliability and Applications*, (2008), **7**(1), *1-12*.

[2] Akaike H. Fitting autoregressive model for regression. *Annals of the Institute of Statistical Mathematics*, (1969), **21**, 243-247.

[3] Bemis B, Bain LJ and Higgins JJ. Estimation and hypothesis testing for the parameters of a bivariate exponential distribution, *Journal of the American Statistical Association,* (1972), **67**, 927-929.

[4] Hastings WK. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, (1970), **57**, 97-109.

[5] Kundu D, Sarhan AM and Gupta RD. On Sarhan-Balakrishnan bivariate distribution. *Journal of Statistics and Probability*, (2012), **1**, no. 3, 163-170.

[6] Kundu D and Gupta RD. Bivariate Generalized Exponential Distribution, *Journal of Multivariate Analysis*, (2009), **100**, 581-593.

[7] Lindley DV. Approximate Bayesian methods. *Trabajos de Estadistica*, (1980), **31**, 223-245.

[8] Metropolis N, Rosenbluth AW, Rosenbluth MN, Teller AH, Teller E. Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, (1953), **21**, 1087-1092.

[9] Marshall AW and Olkin IA. A multivariate exponential distribution, *Journal of American Statistical Association*, (1967), **62**, 30-44.

[10] Meintanis SG. Test of fit for Marshall-Olkin distributions with applications, *Journal of Statistical Planning and inference*, (2007), **137**, 3954-3963.

[11] Mudholkar GS and Srivastava DK. Exponentiated Weibull family for analyzing bathtub failure-rate data, *IEEE Transactions on Reliability*, (1993), **42**, 299-302.

[12] Raqab MZ, Kundu D. Burr Type X distribution; revisited, *Journal of Probability and Statistical Science,* (2006), **4**(2), 179-193.

[13] Tierney L, Kadane JB. Accurate approximations for posterior moments and marginal densities. *Journal of American Statistical Association*, (1986), **81**, 82-86.

[14] Sarhan AM and Balakrishnan N. A new class of bivariate distributions and its mixture. *Journal of Multivariate Analysis*, (2007), **98**, 1508-1527.

[15] Surles JG and Padgett WJ. Inference for reliability and stress-strength for a scaled Burr Type X distribution, *Lifetime Data Analysis*, (2001), **7**, 187-200.

[16] Surles JG and Padgett WJ. Some properties of a scaled Burr type X distribution, *Journal of Statistical Planning and Inference*, (2005), **128**, 271-280.