Weighted Marshall-Olkin Bivariate Exponential Distribution

Ahad Jamalizadeh[§] & Debasis Kundu[†]

Abstract

Recently Gupta and Kundu [9] introduced a new class of weighted exponential distributions, and it can be used quite effectively to model lifetime data. In this paper, we introduce a new class of weighted Marshall-Olkin bivariate exponential distributions. This new singular distribution has univariate weighted exponential marginals. We study different properties of the proposed model. There are four parameters in this model and the maximum likelihood estimators (MLEs) of the unknown parameters cannot be obtained in explicit forms. We need to solve a four dimensional optimization problem to compute the MLEs. One data set has been analyzed for illustrative purposes and finally we propose some generalization of the proposed model.

Keywords: Joint probability density function; Conditional probability density function; Singular distribution; Maximum likelihood estimators; Fisher information matrix; Asymptotic distribution.

[§]Department of Statistics, Faculty of Mathematics & Computer, Shahid Bahonar University of Kerman, Kerman,Iran,76169-14111.

[†]Department of Mathematics and Statistics, Indian Institute of Technology Kanpur, Pin 208016, India, e-mail: kundu@iitk.ac.in. Visiting Professor at the King Saud University, Saudi Arabia. Corresponding author. Part of this work is supported by a grant from the Department of Science and Technology, Government of India.

1 Introduction

Recently Gupta and Kundu [9] introduced a shape parameter to an exponential distribution using the idea of Azzalini [5] and study its different properties. This new class of distributions turn out to be a weighted exponential distributions, and due to this reason it is named as the weighted exponential (WE) distribution. Several interesting properties of this new WE distribution have been established by Gupta and Kundu [9]. It is observed that the two-parameter WE distribution distribution behaves very similarly, with the other well known two-parameter distributions, like Weibull, gamma or generalized exponential distributions and it has several desirable properties also. The WE distribution can be obtained as a hidden truncation model. Moreover, it has been observed that in certain cases the two-parameter WE distribution may provide a better fit than the two-parameter Weibull, gamma or generalized exponential distributions. Since, its distribution function is in compact form, it can be used very effectively to analyze censored data also. A brief review of the two-parameter WE distribution is presented in section 2, for ready reference.

Recently, Al-Mutairi *et al.* [2] introduced an absolute continuous bivariate distribution with weighted exponential marginals. The main aim of this paper is to introduce a weighted Marshall-Olkin bivariate exponential (WMOBE) distribution, using the similar idea as of Azzalini [5], and it is quite different than the method proposed by Al-Mutairi *et al.* [2]. This new singular WMOBE distribution has four parameters. It can also be obtained as a hidden truncation model as of Arnold and Beaver [3]. Therefore, the interpretation of any multivariate hidden truncation model as it was provided by Arnold and Beaver [4] is valid for this proposed model also. This is one of the basic motivations of the proposed model.

Moreover, it may be mentioned that the most popular singular bivariate distribution is the three-parameter Marshall-Olkin bivariate exponential (MOBE) distribution or the four-parameter Marshall-Olkin bivariate Weibull (MOBW) distribution, see for example Kotz *et al.* [12]. Since it has been observed that the univariate WE distribution may provide a better fit than Weibull or exponential distribution in certain cases, it is expected that the proposed four-parameter WMOBE model may also provide a better fit than the MOBE or MOBW model in certain cases. It is mainly to provide another option to the practitioners to a new bivariate four-parameter model for analyzing singular bivariate data.

Several properties of this new WMOBE distribution have been established. The joint probability density function (PDF) and the joint cumulative distribution function (CDF) can be expressed in explicit forms. The marginals of the WMOBE distribution are univariate WE distributions. The MOBE distribution can be obtained as a limiting distribution of the four-parameter WMOBE distribution. The joint moment generating function (MGF), different moments and the product moments of WMOBE distribution can be obtained in explicit forms. The correlation coefficient between the two variables is always non-negative, and depending on the parameter values it can vary between 0 and 1. The expected Fisher information matrix also can be expressed in compact form. The generation from the WMOBE distribution is quite straight forward, and therefore performing simulation experiments on this particular model becomes quite easy.

As already mentioned, the proposed WMOBE model has four parameters. The maximum likelihood estimators (MLEs) of the unknown parameters can be obtained by solving four nonlinear equations. They do not have explicit solutions. For illustrative purposes we have analyzed one real data set, which was originally analyzed by Csorgo and Welsch [7]. They analyzed the bivariate singular data set using MOBE model and concluded that it does not provide a very good fit. We have re-analyzed the data using the WMOBE model and observe that WMOBE provides a better fit than the MOBW model. It justifies the use of the proposed WMOBE model to analyze certain bivariate singular data sets.

Although, in this paper we have introduced and discussed a singular WMOBE model, several other generalizations are also possible. For example; (a) weighted Marshall-Olkin multivariate exponential model, (b) singular weighted Marshall-Olkin bivariate Weibull model, or (c) singular weighted Marshall-Olkin multivariate Weibull model, can also be obtained along the same lines.

The rest of the paper is organized as follows. In section 2, we briefly review the univariate WE model. In section 3, we introduce WMOBE model, and discuss its different properties. In section 4, the statistical inferences of the unknown parameters are provided. The analysis of a data set is provided in section 5. Finally we conclude the paper and provide some future research in section 6.

2 Weighted Exponential Distribution

Definition: The random variable X is said to have a WE distribution with the shape and scale parameters $\alpha > 0$ and $\lambda > 0$ respectively, if the PDF of X is

$$f_X(x;\alpha,\lambda) = \frac{\alpha+1}{\alpha}\lambda e^{-\lambda x}\left(1-e^{-\alpha\lambda x}\right); \quad x > 0, \tag{1}$$

and 0 otherwise. Form now on a WE distribution with the PDF (1) will be denoted by WE(α, λ).

The PDF of the WE distribution is always unimodal. The CDF and the hazard function (HF) can be expressed in explicit forms. The HF of the WE is always an increasing function, and WE family is a reverse rule of order two (RR_2) family. For different shapes of the PDFs of WE family, the readers are referred to the original work of Gupta and Kundu [9]. It may be mentioned that the shapes of the PDFs of the WE distribution are very similar with the shapes of the PDFs of the well known Weibull, gamma or generalized exponential distributions. Moreover, in this model, λ plays the role of a scale parameter, and α plays the role of a shape parameter.

It is observed that the WE(α, λ) can be obtained as a hidden truncation model, as introduced by Arnold and Beaver [3]. For example, if X_1 and X_2 are two independent identically distributed (*i.i.d.*) exponential random variables with mean $\frac{1}{\lambda}$, then

$$X \overset{d}{=} X_1 | \alpha X_1 > X_2, \tag{2}$$

here $\overset{d}{=}$ means equal in distribution. The moment generating function (MGF) of X can be written as

$$M_X(t) = E(e^{tX}) = \left(1 - \frac{t}{\lambda(1+\alpha)}\right)^{-1} \left(1 - \frac{t}{\lambda}\right)^{-1}. \tag{3}$$

From (3) it is immediate that

$$X = U + V, \tag{4}$$

where U and V are independent exponential random variables with means $\frac{1}{\lambda(1+\alpha)}$ and $\frac{1}{\lambda}$ respectively. Using the representation (4) random samples from WE can be easily generated.

The mean and variance of X becomes

$$E(X) = \frac{1}{\lambda}\left(1 + \frac{1}{1+\alpha}\right), \quad \text{and} \quad V(X) = \frac{1}{\lambda^2}\left(1 + \frac{1}{(1+\alpha)^2}\right),$$

respectively. The coefficient of variation (CV) and skewness are both functions of the shape parameter only, as expected. The CV increases from $\frac{1}{\sqrt{2}}$ to 1, and skewness increases from $\sqrt{2}$ to 3. The mean residual lifetime is a decreasing function of time. The convolution of WE distribution can be expressed as a weighted mixture of gamma distributions. For different estimation procedures, different other properties and for comparison with other lifetime distributions, like Weibull, gamma or generalized exponential distributions, the readers are referred to Gupta and Kundu [9].

3 Bivariate Weighted Exponential Distribution

The bivariate random vector (Y_1, Y_2) has the MOBE distribution, if it has the joint PDF

$$f_{Y_1,Y_2}(y_1, y_2) = \begin{cases} g_1(y_1, y_2) & \text{if} \quad 0 < y_1 < y_2 \\ g_2(y_1, y_2) & \text{if} \quad 0 < y_2 < y_1 \\ g_0(y) & \text{if} \quad 0 < y_1 = y_2 = y, \end{cases} \tag{5}$$

where

$$\begin{aligned} g_1(y_1, y_2) &= \lambda_1 e^{-\lambda_1 y_1}(\lambda_2 + \lambda_{12})e^{-(\lambda_2+\lambda_{12})y_2}, \quad y_1 < y_2 \\ g_2(y_1, y_2) &= (\lambda_1 + \lambda_{12})e^{-(\lambda_1+\lambda_{12})y_1}\lambda_2 e^{-\lambda_2 y_2}, \quad y_2 < y_1 \\ g_0(y) &= \lambda_{12}e^{-(\lambda_1+\lambda_2+\lambda_{12})y}, \end{aligned}$$

and it will be denoted by MOBE$(\lambda_1, \lambda_2, \lambda_{12})$. Note that if U_1, U_2 and U_0 are independent exponential random variables with parameters λ_1, λ_2 and λ_{12} respectively, then

$$(Y_1, Y_2) \stackrel{d}{=} (\min\{U_0, U_1\}, \min\{U_0, U_2\}). \tag{6}$$

Now based on the MOBE distribution, we introduce the WMOBE distribution as follows:

Definition: Let $(Y_1, Y_2) \sim$ MOBE$(\lambda_1, \lambda_2, \lambda_{12})$ and $Z \sim \exp(1)$, and they are independently distributed. A random vector (X_1, X_2) is said to have a WMOBE distribution with parameter $\theta = (\alpha, \lambda_1, \lambda_2, \lambda_{12})$, if

$$X_1 \stackrel{d}{=} Y_1 | Z < \alpha \min\{Y_1, Y_2\} \quad \text{and} \quad X_2 \stackrel{d}{=} Y_2 | Z < \alpha \min\{Y_1, Y_2\}. \tag{7}$$

and it will be denoted by WMOBE$(\alpha, \lambda_1, \lambda_2, \lambda_{12})$.

Note that (7) can also be written as

$$X_1 \stackrel{d}{=} Y_1 | W < \min\{Y_1, Y_2\} \quad \text{and} \quad X_2 \stackrel{d}{=} Y_2 | W < \min\{Y_1, Y_2\}. \tag{8}$$

where $W \sim \exp(\alpha)$. It is immediate that as $\alpha \to \infty$, $(X_1, X_2) \xrightarrow{d} (Y_1, Y_2)$, where $\xrightarrow{d}$ means convergence in distribution. Therefore, MOBE is not a member of WMOBE family, but it can be obtained as a limiting distribution of the WMOBE family.

Now based on the above definition, we provide the joint survival function and the joint PDF of (X_1, X_2) in the following theorems.

THEOREM 3.1: Let $(X_1, X_2) \sim \text{WMOBE}(\alpha, \lambda_1, \lambda_2, \lambda_{12})$, then the joint survival function of (X_1, X_2) is

$$S(x_1, x_2) = P(X_1 > x_1, X_2 > x_2) = \begin{cases} S_1(x_1, x_2) & \text{if} \quad 0 < x_1 < x_2 \\ S_2(x_1, x_2) & \text{if} \quad 0 < x_2 < x_1 \\ S_0(x) & \text{if} \quad 0 < x_1 = x_2 = x, \end{cases}$$

where

$$\begin{aligned} S_1(x_1, x_2) &= \frac{\alpha + \lambda}{\alpha} \left[e^{-(\lambda_1 x_1 + \lambda_2 x_2 + \lambda_{12} x_2)} (1 - e^{-x_1 \alpha}) + \frac{\alpha}{\alpha + \lambda_1} e^{-(\lambda_2 x_2 + \lambda_{12} x_2 + \lambda_1 x_1 + x_1 \alpha)} \right. \\ &\quad \left. - \frac{\alpha(\lambda_2 + \lambda_{12})}{(\alpha + \lambda)(\alpha + \lambda_1)} e^{-(\lambda + \alpha) x_2} \right] \qquad (9) \\ S_2(x_1, x_2) &= \frac{\alpha + \lambda}{\alpha} \left[e^{-(\lambda_2 x_2 + \lambda_1 x_1 + \lambda_{12} x_1)} (1 - e^{-x_2 \alpha}) + \frac{\alpha}{\alpha + \lambda_2} e^{-(\lambda_1 x_1 + \lambda_{12} x_1 + \lambda_2 x_2 + x_2 \alpha)} \right. \\ &\quad \left. - \frac{\alpha(\lambda_1 + \lambda_{12})}{(\alpha + \lambda)(\alpha + \lambda_2)} e^{-(\lambda + \alpha) x_1} \right] \qquad (10) \\ S_0(x) &= \frac{\alpha + \lambda}{\alpha} \left[e^{-\lambda x} (1 - e^{-x\alpha}) + \frac{\alpha}{\alpha + \lambda} e^{-(\alpha + \lambda_2) x} \right], \qquad (11) \end{aligned}$$

and $\lambda = \lambda_1 + \lambda_2 + \lambda_{12}$.

PROOF: We will prove the result for the case $x_1 < x_2$, the other two cases will follow along the same lines. For $x_1 < x_2$, and for U_0, U_1, U_2 same as defined in (6), we have

$$\begin{aligned} S(x_1, x_2) &= P(X_1 > x_1, X_2 > x_2) = P(Y_1 > x_1, Y_2 > x_2 | Z < \alpha \min\{Y_1, Y_2\}) \\ &= \frac{1}{P(Z < \alpha \min\{U_0, U_1, U_2\})} P(U_1 > x_1, U_2 > x_2, U_0 > x_2, Z < \alpha \min\{U_0, U_1, U_2\}). \end{aligned}$$

Now if $A = P(U_1 > x_1, U_2 > x_2, U_0 > x_2, Z < \alpha \min\{U_0, U_1, U_2\})$, then

$$\begin{aligned} A &= \int_0^\infty e^{-y-\lambda_1 \max\{x_1, \frac{y}{\alpha}\} - (\lambda_2+\lambda_{12}) \max\{x_2, \frac{y}{\alpha}\}} dy \\ &= \int_0^{x_1\alpha} e^{-y-\lambda_1 x_1 - (\lambda_2+\lambda_{12})x_2} dy + \int_{x_1\alpha}^{x_2\alpha} e^{-(1+\frac{\lambda_1}{\alpha})y - (\lambda_2+\lambda_{12})x_2} dy + \int_{x_2\alpha}^\infty e^{-(1+\frac{\lambda_1}{\alpha})y} dy. \end{aligned}$$

Now after simplification and using $P(Z < \alpha \min\{U_0, U_1, U_2\}) = \dfrac{\alpha}{\alpha + \lambda}$, the result immediately follows. ∎

Theorem 3.2: Let $(X_1, X_2) \sim$ WMOBE$(\alpha, \lambda_1, \lambda_2, \lambda_{12})$, then the joint PDF of (X_1, X_2) is

$$f(x_1, x_2) = \begin{cases} f_1(x_1, x_2) & \text{if} \quad 0 < x_1 < x_2 \\ f_2(x_1, x_2) & \text{if} \quad 0 < x_2 < x_1 \\ f_0(x) & \text{if} \quad 0 < x_1 = x_2 = x, \end{cases} \tag{12}$$

where

$$f_1(x_1, x_2) = \frac{\alpha + \lambda}{\alpha} \lambda_1 e^{-\lambda_1 x_1} (\lambda_2 + \lambda_{12}) e^{-(\lambda_2+\lambda_{12})x_2} (1 - e^{-x_1\alpha}) \tag{13}$$

$$f_2(x_1, x_2) = \frac{\alpha + \lambda}{\alpha} (\lambda_1 + \lambda_{12}) e^{-(\lambda_1+\lambda_{12})x_1} \lambda_2 e^{-\lambda_2 x_2} (1 - e^{-x_2\alpha}) \tag{14}$$

$$f_0(x) = \frac{\alpha + \lambda}{\alpha} \lambda_{12} e^{-\lambda x} (1 - e^{-x\alpha}). \tag{15}$$

Proof: The expressions of $f_1(\cdot, \cdot)$ and $f_2(\cdot, \cdot)$ can be obtained by simply taking $\dfrac{\partial^2}{\partial x_1 \partial x_2} S(x_1, x_2)$ for $x_1 < x_2$ and $x_1 > x_2$ respectively. But naturally $f_0(\cdot)$ cannot be obtained similarly. Using the similar ideas as of Sarhan and Balakrishnan [15] or Kundu and Gupta [9], and also using the fact

$$\int_0^\infty \int_{x_1}^\infty f_1(u, v) dv du + \int_0^\infty \int_{x_2}^\infty f_2(u, v) du dv + \int_0^\infty f_0(w) dw = 1,$$

the result can be easily obtained. ∎

We provide the surface plot of the absolute continuous part of the joint PDF of (12) in Figure 1 for different values of α for fixed $\lambda_1 = \lambda_2 = \lambda_{12} = 1$. The joint PDF is always unimodal, and it can take various shapes. It is clear that α plays the role of a shape parameter also in this bivariate model.

COMMENTS: Note that the joint PDF of (X_1, X_2) as obtained in Theorem 3.2, can be written as follows:

$$f(x_1, x_2) = \frac{\alpha + \lambda}{\alpha} \left(1 - e^{-\alpha \min\{x_1, x_2\}}\right) f_{Y_1,Y_2}(x_1, x_2), \tag{16}$$

where $f_{Y_1,Y_2}(\cdot,\cdot)$ is same as defined in (5). Therefore, it is clear that the proposed bivariate distribution is a weighted Marshall-Olkin bivariate exponential distribution, where the weight function is $\frac{\alpha + \lambda}{\lambda} \left(1 - e^{-\alpha \min\{x_1, x_2\}}\right)$.

The following corollary provides explicitly the absolute continuous and singular parts explicitly.

COROLLARY: The joint PDF of X_1 and X_2 as provided in Theorem 3.2, can also be expressed in the following form for $x = \max\{x_1, x_2\}$, and for $f_1(\cdot,\cdot)$, $f_2(\cdot,\cdot)$ same as defined in (13) and (14) respectively;

$$f(x_1, x_2) = \frac{\lambda_1 + \lambda_2}{\lambda} f_a(x_1, x_2) + \frac{\lambda_{12}}{\lambda} f_s(x), \tag{17}$$

where

$$f_s(x) = \frac{\alpha + \lambda}{\alpha} \lambda e^{\lambda x}(1 - e^{-\alpha x}),$$

and

$$f_a(x_1, x_2) = \frac{\lambda}{\lambda_1 + \lambda_2} \times \begin{cases} f_1(x_1, x_2) & \text{if} \quad x_1 < x_2 \\ f_2(x_1, x_2) & \text{if} \quad x_2 < x_1. \end{cases}$$

Here clearly, $f_a(\cdot,\cdot)$ is the absolute continuous part and $f_s(\cdot)$ is the singular part of $f(\cdot,\cdot)$.

THEOREM 3.3: Let $(X_1, X_2) \sim \text{WMOBE}(\alpha, \lambda_1, \lambda_2, \lambda_{12})$, then

(a) $X_1 \sim \text{WE}\left(\frac{\alpha + \lambda_2}{\lambda_1 + \lambda_{12}}, \lambda_1 + \lambda_{12}\right)$

(b) $X_2 \sim \text{WE}\left(\frac{\alpha + \lambda_1}{\lambda_2 + \lambda_{12}}, \lambda_2 + \lambda_{12}\right)$

(c) $\min\{X_1, X_2\} \sim \text{WE}\left(\frac{\alpha}{\lambda}, \lambda\right)$.

Proof: To prove (a), note that from the joint survival function, it can be observed

$$P(X_1 > x) = \frac{\alpha+\lambda}{\alpha+\lambda_2} e^{-(\lambda_1+\lambda_{12})x} - \frac{\lambda_1+\lambda_{12}}{\alpha+\lambda_2} e^{-(\alpha+\lambda)x}.$$

Then the result can be easily obtained by observing the fact $f_{X_1}(x) = -\frac{d}{dx} P(X_1 > x)$. (b) follows along the same line. The proof of (c) is trivial. ∎

The moment generating function of the WMOBE distribution can be expressed in explicit form. It is provided in the following theorem.

Theorem 3.4: Let $(X_1, X_2) \sim \text{WMOBE}(\alpha, \lambda_1, \lambda_2, \lambda_{12})$, then the moment generating function of (X_1, X_2) is

$$M(t_1, t_2; \theta) = \frac{(\alpha+\lambda)}{(\lambda - t_1 - t_2)(\lambda + \alpha - t_1 - t_2)} \left\{ \frac{\lambda_2 (\lambda_1 + \lambda_{12})}{\lambda_1 + \lambda_{12} - t_1} + \frac{\lambda_1 (\lambda_2 + \lambda_{12})}{\lambda_2 + \lambda_{12} - t_2} + \lambda_{12} \right\}. \quad (18)$$

Proof: It is straight forward, therefore it is avoided. ∎

Note that the moment generating function of (X_1, X_2) can be written as

$$\begin{aligned} M(t_1, t_2; \theta) &= \frac{\lambda_2}{\lambda}\left(1 - \frac{t_1+t_2}{\lambda}\right)^{-1}\left(1 - \frac{t_1+t_2}{\alpha+\lambda}\right)^{-1}\left(1 - \frac{t_1}{\lambda_1+\lambda_{12}}\right)^{-1} + \quad (19)\\ &\quad \frac{\lambda_1}{\lambda}\left(1 - \frac{t_1+t_2}{\lambda}\right)^{-1}\left(1 - \frac{t_1+t_2}{\alpha+\lambda}\right)^{-1}\left(1 - \frac{t_2}{\lambda_2+\lambda_{12}}\right)^{-1} + \quad (20)\\ &\quad \frac{\lambda_{12}}{\lambda}\left(1 - \frac{t_1+t_2}{\lambda}\right)^{-1}\left(1 - \frac{t_1+t_2}{\alpha+\lambda}\right)^{-1}. \quad (21) \end{aligned}$$

Suppose, $V_1 \sim \exp(\lambda)$, $V_2 \sim \exp(\alpha+\lambda)$, $V_3 \sim \exp(\lambda_1+\lambda_{12})$, $V_4 \sim \exp(\lambda_2+\lambda_{12})$, and they are independently distributed. If we define $V = V_1 + V_2$, have the following corollary of Theorem 3.4;

Corollary:

$$(X_1, X_2) \overset{d}{=} \begin{cases} (V + V_3, V) & \text{with probability } \frac{\lambda_2}{\lambda} \\ (V, V + V_4) & \text{with probability } \frac{\lambda_1}{\lambda} \\ (V, V) & \text{with probability } \frac{\lambda_{12}}{\lambda}. \end{cases} \quad (22)$$

Note that (22) can be used quite effectively to generate samples from the WMOBE distribution. Moreover, the product moments of X_1 and X_2 can also be easily computed using the representation of (22) and exponential moments, which has not attempted here.

COROLLARY: From the cumulant generating function (logarithm of the moment generating function), it can be easily seen that the correlation between X_1 and X_2, say ρ_{X_1,X_2} can be written as

$$\rho_{X_1,X_2} = (\lambda_1+\lambda_{12})(\lambda_2+\lambda_{12}) \times \frac{\left[\frac{1}{\lambda^2}+\frac{1}{(\alpha+\lambda)^2}-\frac{\lambda_1\lambda_2}{(\lambda_1+\lambda_{12})(\lambda_2+\lambda_{12})\lambda^2}\right]}{\sqrt{1+\left(\frac{\lambda_1+\lambda_{12}}{\alpha+\lambda}\right)^2}\times\sqrt{1+\left(\frac{\lambda_2+\lambda_{12}}{\alpha+\lambda}\right)^2}}. \tag{23}$$

From (23) it is clear that for all $\alpha > 0, \lambda_1 > 0, \lambda_2 > 0, \lambda_{12} > 0$, $\rho_{X_1,X_2} > 0$. It can be easily seen that as $\alpha \to \infty$ and $\lambda_{12} \to 0$ then $\rho_{X_1,X_2} \to 0$. Moreover, when $\alpha \to \infty$ and $\lambda_{12} \to \infty$ so that $\lambda_{12}/\alpha \to 0$, then $\rho_{X_1,X_2} \to 1$.

THEOREM 3.5: Let $(X_1, X_2) \sim \text{WMOBE}(\alpha, \lambda_1, \lambda_2, \lambda_{12})$, then if $\lambda_1 = \lambda_2$, the absolute continuous part of (X_1, X_2) has a total positivity of order two (TP_2) property.

PROOF: Note that an absolute continuous bivariate random vector, say (T_1, T_2) has TP_2 property, if and only if for any $t_{11}, t_{12}, t_{21}, t_{22}$, whenever $t_{11} < t_{12}$ and $t_{21} < t_{22}$, we have

$$f_{T_1,T_2}(t_{11}, t_{21}) f_{T_1,T_2}(t_{12}, t_{22}) - f_{T_1,T_2}(t_{12}, t_{21}) f_{T_1,T_2}(t_{11}, t_{22}) \geq 0, \tag{24}$$

where $f_{T_1,T_2}(\cdot,\cdot)$ is the joint PDF of (T_1, T_2). Note that by taking different ordered $t_{11}, t_{12}, t_{21}, t_{22}$, such that $t_{11} < t_{12}$ and $t_{21} < t_{22}$, the result can be proved very easily. ∎

4 INFERENCE

4.1 MAXIMUM LIKELIHOOD ESTIMATION:

In this section we propose to compute the MLEs of the unknown parameters $\alpha, \lambda_1, \lambda_2, \lambda_{12}$ based on a random sample of size n, $\{(x_{1i}, x_{2i}), i = 1, \ldots, n\}$. We will be using the following notations.

$I = \{1, \cdots, n\}$, $I_0 = \{i \in I; x_{1i} = x_{2i} = x_i\}$, $I_1 = \{i \in I; x_{1i} < x_{2i}\}$, $I_2 = \{i \in I; x_{1i} > x_{2i}\}$,

and n_0, n_1, n_2 denote the number of elements in I_0, I_1 and I_2 respectively.

Based on the observed sample, the log-likelihood contribution becomes;

$$
\begin{aligned}
l(\alpha, \lambda_1, \lambda_2, \lambda_{12}) &= \sum_{I_0} \ln f_0(x_i, x_i) + \sum_{I_1} \ln f_1(x_{1i}, x_{2i}) + \sum_{I_2} \ln f_2(x_{1i}, x_{2i}) \\
&= \sum_{i \in I_0} \ln\left(1 - e^{-\alpha x_i}\right) + \sum_{i \in I_1} \ln\left(1 - e^{-\alpha x_{1i}}\right) + \sum_{i \in I_2} \ln\left(1 - e^{-\alpha x_{2i}}\right) \\
&\quad + n_0 \ln \lambda_{12} + n_1 \ln \lambda_1 + n_2 \ln \lambda_2 + n_1 \ln(\lambda_2 + \lambda_{12}) + n_2 \ln(\lambda_1 + \lambda_{12}) \\
&\quad - \lambda_0 \sum_{i \in I_0} x_i - \lambda_1 \left(\sum_{i \in I_0} x_i + \sum_{i \in I_1 \cup I_2} x_{1i}\right) - \lambda_2 \left(\sum_{i \in I_0} x_i + \sum_{i \in I_1 \cup I_2} x_{2i}\right) \\
&\quad - \lambda_{12} \left(\sum_{i \in I_0} x_i + \sum_{i \in I_2} x_{1i} + \sum_{i \in I_1} x_{2i}\right) + n \ln(\alpha + \lambda) - n \ln \alpha. \qquad (25)
\end{aligned}
$$

The MLEs of the unknown parameters can be obtained by solving the following four normal equations simultaneously;

$$
\frac{\partial l}{\partial \lambda_1} = \frac{n_1}{\lambda_1} + \frac{n_2}{\lambda_1 + \lambda_{12}} - T_1 + \frac{n}{\alpha + \lambda} = 0, \qquad (26)
$$

$$
\frac{\partial l}{\partial \lambda_2} = \frac{n_2}{\lambda_2} + \frac{n_1}{\lambda_2 + \lambda_{12}} - T_2 + \frac{n}{\alpha + \lambda} = 0, \qquad (27)
$$

$$
\frac{\partial l}{\partial \lambda_{12}} = \frac{n_0}{\lambda_{12}} + \frac{n_1}{\lambda_2 + \lambda_{12}} + \frac{n_2}{\lambda_1 + \lambda_{12}} - T_{12} + \frac{n}{\alpha + \lambda} = 0, \qquad (28)
$$

$$
\frac{\partial l}{\partial \alpha} = \sum_{i \in I_0} \frac{x_i e^{-\alpha x_i}}{(1 - e^{-\alpha x_i})} + \sum_{i \in I_1} \frac{x_{1i} e^{-\alpha x_{1i}}}{(1 - e^{-\alpha x_{1i}})} + \sum_{i \in I_2} \frac{x_{2i} e^{-\alpha x_{2i}}}{(1 - e^{-\alpha x_{2i}})} + \frac{n}{\alpha + \lambda} - \frac{n}{\alpha} = 0, \qquad (29)
$$

where

$$
\begin{aligned}
T_1 &= \left(\sum_{i \in I_0} x_i + \sum_{i \in I_1 \cup I_2} x_{1i}\right) \\
T_2 &= \left(\sum_{i \in I_0} x_i + \sum_{i \in I_1 \cup I_2} x_{2i}\right) \\
T_{12} &= \left(\sum_{i \in I_0} x_i + \sum_{i \in I_2} x_{1i} + \sum_{i \in I_1} x_{2i}\right).
\end{aligned}
$$

The MLEs of the unknown parameters cannot be obtained explicitly. They have to be obtained by solving some numerical methods, like Newton-Raphson or Gauss-Newton methods or their variants. Alternatively, some optimization algorithm for example genetic algorithm, simulated annealing or down hill simplex method can be used directly to maximize the function (25).

Initial Guess Values

To solve the above normal equations, we need to use some iterative methods and for that we need to provide some initial guesses of the unknown parameters. Theorem 3.3 can be used quite effectively to obtain these initial guesses. For example first we can fit WE to X_1, to get initial guesses of $\alpha + \lambda_2$ and $\lambda_1 + \lambda_{12}$. Similarly by fitting WE to X_2 and to $\min\{X_1, X_2\}$, we can obtain initial guesses of $\alpha + \lambda_1$, $\lambda_2 + \lambda_{12}$, α and λ. From these initial guesses we can obtain initial guesses of all the unknown parameters easily.

Asymptotic Properties

Although, the proposed WMOBE distribution is a singular distribution, but it can be shown along the same line as the MOBE model that the asymptotic distribution of the MLEs is multivariate normal. The result is given below without proof.

Theorem 4.1 As $n \to \infty$, the MLEs of θ, say $\widehat{\theta} = (\widehat{\alpha}, \widehat{\lambda}_1, \widehat{\lambda}_2, \widehat{\lambda}_{12})$ has the following asymptotic property;

$$\sqrt{n}\left\{(\widehat{\alpha}, \widehat{\lambda}_1, \widehat{\lambda}_2, \widehat{\lambda}_{12}) - (\alpha, \lambda_1, \lambda_2, \lambda_{12})\right\} \longrightarrow N_4(0, I^{-1}),$$

here I is the expected Fisher information matrix, and the exact expression of I is provided in Appendix B.

4.2 Testing of Hypothesis

It is already mentioned that the MOBE distribution is a limiting member of the class of WMOBE distributions. In this subsection only let us denote $\beta = \frac{1}{\alpha}$. With the abuse of notation, let us denote the new parameter space also as $\theta = (\beta, \lambda_1, \lambda_2, \lambda_{12})$, where $\beta > 0, \lambda_1 > 0, \lambda_2 > 0, \lambda_{12} \geq 0$. When $\beta = 0$, it corresponds to the MOBE model. So one of the natural testing of hypothesis problem will be to test the following:

$$H_0 : \beta = 0, \quad \text{vs} \quad H_1 : \beta > 0. \tag{30}$$

In this case since β is in the boundary under the null hypothesis, the standard results do not work. But using Theorem 3 of Self and Liang [16], it follows that

$$2(l_{WMOBE}(\widehat{\beta}, \widehat{\lambda}_1, \widehat{\lambda}_2, \widehat{\lambda}_{12}) - l_{MOBE}(\widehat{\lambda}_1, \widehat{\lambda}_2, \widehat{\lambda}_{12})) \longrightarrow \frac{1}{2} + \frac{1}{2}\chi^2_1.$$

Here $l_{WMOBE}(\cdot)$ and $l_{MOBE}(\cdot)$ denote the log-likelihood functions at the maximum likelihood values of WMOBE and MOBE models respectively.

5 Data Analysis

American Football Data Set: This data set is obtained from the American Football (National Football League) League from the matches on three consecutive weekends in 1986. The data were first published in 'Washington Post', and they are available in Csorgo and Welsch [7]. The 'seconds' in the data have been converted to the decimal points, as it has been done by Csorgo and Welsch [7], and they are presented in Table 1. 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.

X_1	X_2	X_1	X_2	X_1	X_2
2.05	3.98	5.78	25.98	10.40	14.25
9.05	9.05	13.80	49.75	2.98	2.98
0.85	0.85	7.25	7.25	3.88	6.43
3.43	3.43	4:25	4.25	0.75	0.75
7.78	7.78	1.65	1.65	11.63	17.37
10.57	14.28	6.42	15.08	1.38	1.38
7.05	7.05	4.22	9.48	10.35	10.35
2.58	2.58	15.53	15.53	12.13	12.13
7.23	9.68	2.90	2.90	14.58	14.58
6.85	34.58	7.02	7.02	11.82	11.82
32.45	42.35	6.42	6.42	5.52	11.27
8.53	14.57	8.98	8.98	19.65	10.70
31.13	49.88	10.15	10.15	17.83	17.83
14.58	20.57	8.87	8.87	10.85	38.07

Table 1: American Football League (NFL) data

The data represent the time in minutes and seconds. 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.

The data set was analyzed by Csorgo and Welsch [7] by using the Marshall-Olkin bivariate exponential model. Csorgo and Welsch [7] proposed a test procedure, where the null hypothesis is that the data are coming from the Marshall-Olkin bivariate exponential model. 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. Preliminary data analysis from the scaled TTT transform of Aarset [1] indicate that the hazard functions of X_1 and X_2 cannot be constant. They are more likely to be increasing functions. It indicates that WMOBE distribution may be used to analyze this data set.

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 [7]. Here all the data points are divided by 10 just for computational purposes. It should not make any difference in the statistical inference.

The MLEs of α, λ_1, λ_2 and λ_{12} are obtained by maximizing the log-likelihood function (25) with respect to the four unknown parameters and they are as follows $\widehat{\lambda}_1 = 0.5996$, $\widehat{\lambda}_2 = 0.0346$, $\widehat{\lambda}_{12} = 0.8639$ and $\widehat{\alpha} = 2.5302$. The corresponding log-likelihood value is -85.4447. The corresponding 95% confidence intervals are (0.3655, 0.8337), (0.0221,0.0471), (0.4652,1.2626) and (1.4054,3.6550) respectively.

Now the natural question is how good is the fit. Unfortunately, we do not have any proper bivariate goodness of fit test for general models like the univariate case. We examine the marginals and the minimum of the marginals, definitely they provide some indication about the goodness of fit of the proposed WMOB to the given data set. We fit WE(1.7525,1.4635), WE(3.4834,0.8985) and WE(1.6889, 1.4981) to X_1, X_2 and $\min\{X_1, X_2\}$ respectively. The parameters of the corresponding WE model are obtained from Theorem 3.3, by replacing the true values with their estimates.

The Kolmogorov-Smirnov (KS) distance between the empirical distribution function and the fitted distribution function and the associated p values reported in brackets in three cases are 0.0958 (0.8351), 0.1040 (0.7536) and 0.1027 (0.7671) respectively. From the p values, we cannot reject the hypotheses that X_1, X_2 and $\min\{X_1, X_2\}$ follow WE. We have fitted four-parameter MOBW model also to this data set. Using the same notation as in Kundu and Dey [10], we obtain the MLEs as $\widehat{\alpha} = 1.2889$, $\widehat{\lambda}_0 = 0.5761$, $\widehat{\lambda}_1 = 0.4297$, $\widehat{\lambda}_2 = 0.0244$. The corresponding log-likelihood value is -90.4169. Based on the log-likelihood value it is clear that WMOBE provides a better fit to this data set than MOBW model.

Now we want to perform the testing of hypothesis problem as defined in (30). We have fitted three-parameter MOBE model to this data set, and obtain the estimates of λ_0, λ_1 and λ_2 (using the same notations as in Kundu and Dey [10]) as 0.7145, 0.4562 and 0.0298 respectively. The log-likelihood value is -93.3058, and the p value of the test statistic is less than 0.0001. Therefore, we reject the null hypothesis that the data are coming from a MOBE distribution.

6 Conclusions

In this paper we have proposed a new singular bivariate distribution whose marginals are weighted exponential distribution recently proposed by Gupta and Kundu [9]. This singular distribution has been obtained as a hidden truncation model as of Azzalini [5]. This new four-parameter bivariate exponential distribution has explicit probability density function and the distribution function. We have derived several interesting properties of this distribution. Marshall-Olkin bivariate exponential distribution can be obtained as a limiting distribution of this model. It is observed that in some cases it may provide a better fit than the existing well known MOBE or MOBW models.

It may be mentioned that although in this paper we have proposed bivariate weighted exponential model, but similarly, weighted Marshall-Olkin multivariate exponential distribution also can be defined. Moreover, along the same line the weighted Marshall-Olkin bivariate or multivariate Weibull distribution also can be introduced. More work is needed in these direction.

Acknowledgements:

The authors would like to thank the two referees for their valuable suggestions which has helped them to improve the manuscript significantly. Part of this work has been supported by a grant from the Department of Science and Technology, Government of India.

References

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

[2] Al-Mutairi, D.K., Ghitany, M.E. and Kundu, D. (2011), "A new bivariate distribution with weighted exponential marginals and its multivariate generalization", *Statistical Papers*, to appear.

[3] Arnold, B. C. and Beaver, R. J. (2000), "Hidden truncation model", *Sankhya*, Ser. A, vol. 62, 23 - 35.

[4] Arnold, B.C. and Beaver, R. J. (2002), "Selective multivariate models related to hidden truncation and/ or selective reporting", *Test*, vol. 11, 7 - 54.

[5] Azzalini, A. (1985), "A class of distributions which includes the normal ones", *Scandinavian Journal of Statistics*, vol. 12, 171 - 185.

[6] 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.

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

[8] Genton, M. G. (ed.) (2004), *Skew-Elliptical Distributions and their Applications: A Journey Beyond Normality*, Chapman and Hall/CRS, New York.

[9] Gupta, R. D. and Kundu, D. (2009), "A new class of weighted exponential distributions", *Statistics*, to appear, DOI: 10.1080/02331880802605346.

[10] 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, no. 4, 956 - 965.

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

[12] Kotz, S., Johnson, N.L. and Balakrishnan, N. (2000), *Continuous Multivariate Distributions, vol. 1, Models and Applications*, 2nd-edition, John Wiley and Sons, New York.

226 - 233.

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

[14] Press, W.H., Teukolsky, S.A., Vellerling, W.T. and Flannery, B.P. (1992), *Numerical Recipes in FORTRAN, The Art of Scientific Computing*, 2nd. ed., Cambridge University, Cambridge, U.K.

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

[16] Self, S.G. and Liang, K-L., (1987), "Asymptotic properties of the maximum likelihood estimators and likelihood ratio test under non-standard conditions", *Journal of the American Statistical Association*, vol. 82, 605 - 610.

Appendix: Expected Fisher Information Matrices

Let the Fisher information matrix be;

$$I = -E\begin{bmatrix} \frac{\partial^2 l}{\partial\alpha^2} & \frac{\partial^2 l}{\partial\alpha\partial\lambda_1} & \frac{\partial^2 l}{\partial\alpha\partial\lambda_2} & \frac{\partial^2 l}{\partial\alpha\partial\lambda_{12}} \\ \frac{\partial^2 l}{\partial\lambda_1\partial\alpha} & \frac{\partial^2 l}{\partial\lambda_1^2} & \frac{\partial^2 l}{\partial\lambda_1\partial\lambda_2} & \frac{\partial^2 l}{\partial\lambda_1\partial\lambda_{12}} \\ \frac{\partial^2 l}{\partial\lambda_2\partial\alpha} & \frac{\partial^2 l}{\partial\lambda_2\lambda_1} & \frac{\partial^2 l}{\partial\lambda_2^2} & \frac{\partial^2 l}{\partial\lambda_2\partial\lambda_{12}} \\ \frac{\partial^2 l}{\partial\lambda_{12}\partial\alpha} & \frac{\partial^2 l}{\partial\lambda_{12}\lambda_1} & \frac{\partial^2 l}{\partial\lambda_{12}\lambda_2} & \frac{\partial^2 l}{\partial\lambda_{12}^2} \end{bmatrix}. \tag{31}$$

Now we provide all the elements of the Fisher information matrix, and for that we need the following results;

$$P(X_1 < X_2) = \frac{\lambda_1}{\lambda}, \quad P(X_1 > X_2) = \frac{\lambda_2}{\lambda}, \quad P(X_1 = X_2) = \frac{\lambda_{12}}{\lambda},$$

moreover, we will be using the following notation;

$$\psi(\alpha,\lambda) = \frac{\alpha+\lambda}{\lambda}\int_0^\infty \frac{x^2 e^{-(\alpha+\lambda)x}}{(1-e^{-\alpha x})}dx.$$

$$\begin{aligned}
E\left[\frac{\partial^2 l}{\partial\alpha^2}\right] &= -E\left[\frac{n}{(\alpha+\lambda)^2} - \frac{n}{\alpha^2} + \sum_{i\in I_1}\frac{X_{1i}^2 e^{-\lambda X_{1i}}}{(1-e^{-\alpha X_{1i}})^2} + \sum_{i\in I_2}\frac{X_{2i}^2 e^{-\lambda X_{2i}}}{(1-e^{-\alpha X_{2i}})^2} + \sum_{i\in I_0}\frac{X_i^2 e^{-\lambda X_i}}{(1-e^{-\alpha X_i})^2}\right] \\
&= -n\left[\frac{1}{(\alpha+\lambda)^2} - \frac{1}{\alpha^2} + \frac{\psi(\alpha,\lambda)}{\lambda}\left(\lambda_1^2+\lambda_2^2+\lambda_{12}^2\right)\right] \\
E\left[\frac{\partial^2 l}{\partial\lambda_1^2}\right] &= -E\left[\frac{n_1+n_2}{\lambda_1^2} + \frac{n}{(\alpha+\lambda)^2} + \frac{n_2}{(\lambda_1+\lambda_{12})^2}\right] \\
&= -n\left[\frac{\lambda_1+\lambda_2}{\lambda\lambda_1^2} + \frac{1}{(\alpha+\lambda)^2} + \frac{\lambda_2}{\lambda(\lambda_1+\lambda_{12})^2}\right] \\
E\left[\frac{\partial^2 l}{\partial\lambda_2^2}\right] &= -E\left[\frac{n_1+n_2}{\lambda_2^2} + \frac{n}{(\alpha+\lambda)^2} + \frac{n_1}{(\lambda_1+\lambda_{12})^2}\right] \\
&= -n\left[\frac{\lambda_1+\lambda_2}{\lambda\lambda_1^2} + \frac{1}{(\alpha+\lambda)^2} + \frac{\lambda_2}{\lambda(\lambda_1+\lambda_{12})^2}\right] \\
E\left[\frac{\partial^2 l}{\partial\lambda_{12}^2}\right] &= -E\left[\frac{n_1}{(\lambda_2+\lambda_{12})^2} + \frac{n_2}{(\lambda_1+\lambda_{12})^2} + \frac{n}{(\alpha+\lambda)^2} + \frac{n_0}{\lambda_{12}^2}\right] \\
&= -n\left[\frac{\lambda_1}{\lambda(\lambda_2+\lambda_{12})^2} + \frac{\lambda_2}{\lambda(\lambda_1+\lambda_{12})^2} + \frac{1}{(\alpha+\lambda)^2} + \frac{1}{\lambda\lambda_{12}}\right]
\end{aligned}$$

$$
\begin{aligned}
E\left[\frac{\partial^2 l}{\partial\alpha\partial\lambda_1}\right] &= E\left[\frac{\partial^2 l}{\partial\alpha\partial\lambda_2}\right] = E\left[\frac{\partial^2 l}{\partial\alpha\partial\lambda_{12}}\right] = E\left[\frac{\partial^2 l}{\partial\lambda_1\partial\lambda_2}\right] = -\frac{n}{(\alpha+\lambda)^2} \\
E\left[\frac{\partial^2 l}{\partial\lambda_1\partial\lambda_{12}}\right] &= -E\left[\frac{n_2}{(\lambda_1+\lambda_{12})^2} + \frac{n}{(\alpha+\lambda)^2}\right] = -n\left[\frac{\lambda_2}{\lambda(\lambda_1+\lambda_{12})^2} + \frac{1}{(\alpha+\lambda)^2}\right] \\
E\left[\frac{\partial^2 l}{\partial\lambda_2\partial\lambda_{12}}\right] &= -E\left[\frac{n_1}{(\lambda_2+\lambda_{12})^2} + \frac{n}{(\alpha+\lambda)^2}\right] = -n\left[\frac{\lambda_1}{\lambda(\lambda_2+\lambda_{12})^2} + \frac{1}{(\alpha+\lambda)^2}\right]
\end{aligned}
$$

Note that the above elements can be obtained using routine calculation as in Kundu and Gupta [11].

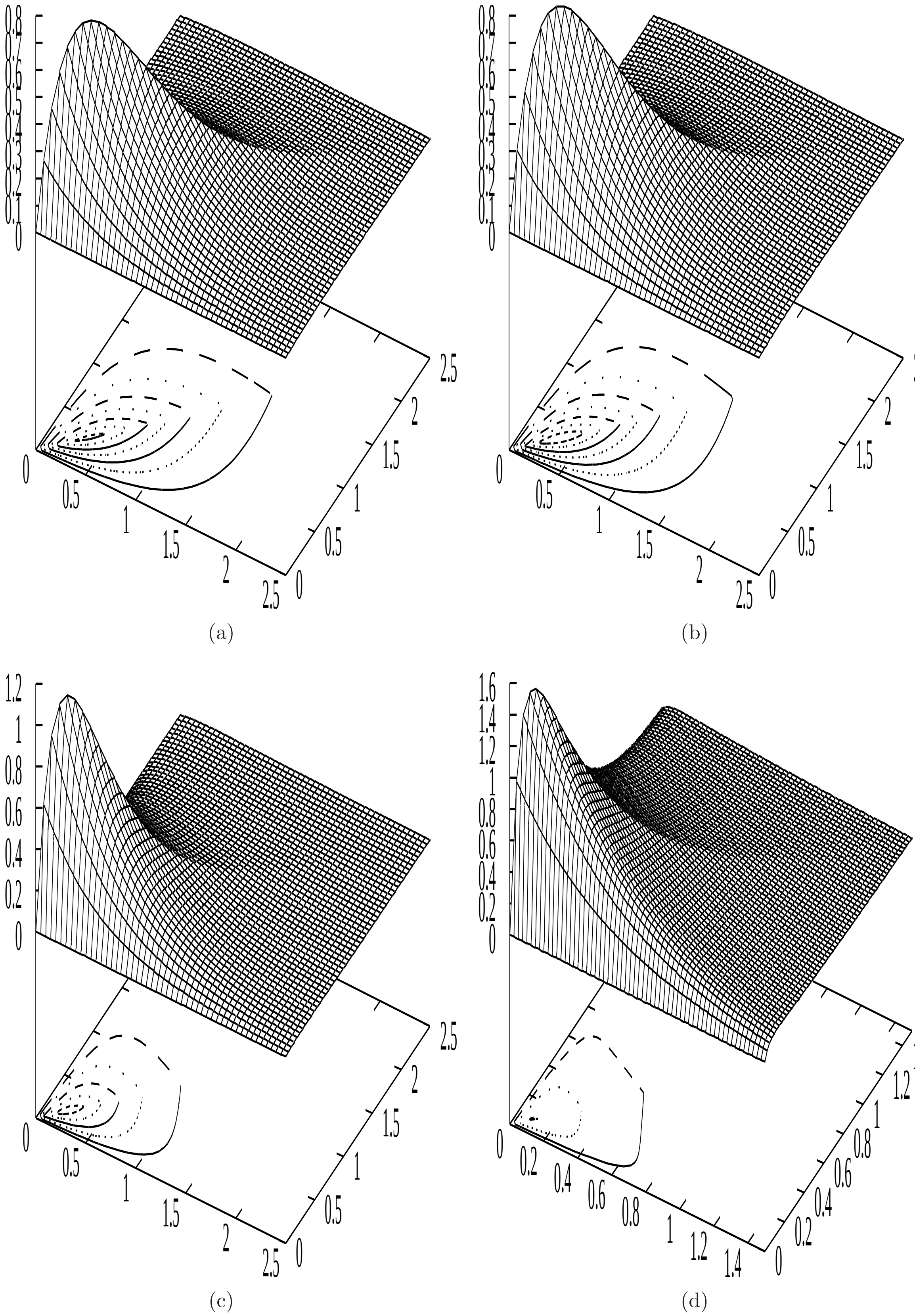

Figure 1: The surface plot of the absolute continuous part of the joint PDF of (12) when $\lambda_1 = \lambda_2 = \lambda_{12} = 1$ and (a) $\alpha = 0.001$, (b) $\alpha = 0.5$, (c) $\alpha = 5$, and (d) $\alpha = 25$.