CHARACTERIZATION PROPERTIES OF MULTIVARIATE MARSHALL-OLKIN GUMBEL DISTRIBUTION

BY

S.M.UMAR

Department of Mathematical Sciences, Bayero University, P.M.B. 3011

Kano-Nigeria

***Email Address*: surajoumar@yahoo.com**

ABSTRACT

This paper considered the characterization properties of multivariate marshall and Olkin gumbel distribution of independently distributed random vectors based on an integrated lack of memory equations.

Keywords: Multivariate Gumbel distribution; Lack of memory equation; Borel measure; Increasing failure rate.

1. INTRODUCTION

According to Fedilis [1], Gumbel distribution is named after a German Mathematician who founded the Mathematical field of extreme value theory (Emil Julius Gumbel (1891-1966)). It is a special case of the Fisher-Tippett (1902-1985). It is useful in finding extreme number of samples of various distributions. It is applicable in finding the maximum level of a river in a particular year if the list of maximum values for some past years is known. It can also be applied in predicting the chances of extreme events like earthquake, flood or other natural disaster.

Standard Gumbel distribution can be defined as the probability density function in the form of $f(x)=\exp(-x)\exp\{-\exp(-x)\}, \quad -\infty<x<\infty.$ And its corresponding cumulative distribution function is given by $F(x)=\exp\{-\exp(-x)\}$.

Obretenov [2] presented characterization of multivariate Marshall-Olkin exponential distribution using integrated lack of memory equations. This paper intends to apply the same procedure of integrated lack of memory equations to characterize independent multivariate Marshall-Olkin gumbel distributions.

2. PRELIMINARY

Let $\underline{X}=X_1,X_2,...,X_n$ be non-negative and independently distributed random variables with $\overline{F}(x_1,x_2,...,x_n),\ x_k\geq 0,\ (k=1,2,...,n)$, as their survival density function. That is,

$\overline{F}(x_1,x_2,...,x_n)\equiv P(X_1>x_1,X_2>x_2,...,X_n>x_n)$. From Galambos and Kotz [3], we can see that if $\overline{F}(x_1,x_2,...,x_n)$ has the lack of memory property, denoted by (LMP), of the type

$$\overline{F}(x_1+t,x_2+t,...,x_n+t)=\overline{F}(t,t,...,t)\overline{F}(x_1,x_2,...,x_n) \qquad \ldots(2.1)$$

and if all marginal distributions of $\overline{F}(x_1,x_2,...,x_n)$ satisfy equation (2.1), then $\overline{F}(x_1,x_2,...,x_n)$ is a survival distribution function of the multivariate Marshall-Olkin distribution. The condition that the marginal distributions of $\overline{F}(x_1,x_2,...,x_n)$ should satisfy equation (2.1), can be replaced by the following: all marginal distributions of $\overline{F}(x_1,x_2,...,x_n)$ are Marshall-Olkin distributed.

Let the vectors e and x be denoted by $e=(1,1,...,1)$ and $x=(x_1,x_2,...,x_n)$ respectively. From this, equation (2.1) can be rewritten as

$$\overline{F}(x+te)=\overline{F}(te)\overline{F}(x). \qquad \ldots(2.2)$$

Whenever survival distribution function satisfies equation (2.2), we say that it has the lack of memory property.

Next, let us give rather a weak definition of the lack of memory property.

Let $\alpha_i = 0$ or 1, $i = 1, 2, ..., n,$ and $a = \{a_1, a_2, ...a_n\}$. Denote by E the set

$E = \{\alpha : only\ one\ \alpha_i\ is\ 0,\ the\ other\ (n-1)\ \alpha_i\ are\ 1\}.$

DEFINITION 2.1: The survival distribution function $\overline{F}(x_1, x_2, ..., x_n)$ of independent random vector $x = (x_1, x_2, ..., x_n)$ has a weak lack of memory property (WLMP) if

$$\overline{F}(te + x \circ a) = \overline{F}(te)\overline{F}(x \circ a) \qquad \ldots(2.3)$$

for all $t \geq 0$ and $a \in E,$ where $x \circ a \equiv \{x_1 a_1, x_2 a_2, ..., x_n a_n\}.$

The difference between equations (2.3) and (2.2) is that in equation (2.3) at least one of the coordinates of $\underline{x} \circ \underline{a}$ is zero, whereas in equation (2.2) all coordinates of $\underline{x}$ can be different from zero. From the results we have so far, equation (2.3) follows from equation (2.2). On the contrary, the following lemma shows that equation (2.2) is a special case of equation (2.3) with additional assumption for the survival distribution function $\overline{F}(x_1, x_2, ..., x_n)$ of independent random variables $x_1, x_2, ..., x_n$.

LEMMA 2.1: If a survival distribution function $\overline{F}(x_1, x_2, ..., x_n)$ of independent random vector $x = (x_1, x_2, ..., x_n)$ has a weak lack of memory property and in addition if $\overline{F}(te) = \exp\{-\exp(-x)\},\ \ x \geq 0,$ then $\overline{F}(x_1, x_2, ..., x_n)$ has a lack of memory property.

PROOF: Let $y = (y_1, y_2, ..., y_n),\ y_k \geq 0,$ and $u \geq 0$ be arbitrary. Suppose that y_i is the smallest coordinate of y, that is $y_i = \min(y_1, y_2, ..., y_n).$ Then, for a fixed y, $\overline{F}(ue + y)$ can be written as

$$\overline{F}(ue+y)=\overline{F}(te+x\circ e), \qquad \ldots(2.4)$$

where we write $t=u+y_i$ for a fixed i, and the vector $x=\{x_1,x_2,...,x_n\}$ has the following coordinates:

$$x_k=\begin{cases} y_k-y_i & for \;\; k\neq i \\ 0 & for \;\; k=i. \end{cases}, \qquad \ldots(2.5)$$

As $\overline{F}(te+x\circ e)=\overline{F}(te+x\circ a)$, where $a\in E$ and the i^{th} coordinate of a is zero, equation (2.4) becomes $\overline{F}(ue+y)=\overline{F}(te+x\circ a)$. This equation together with equation (2.3), gives $\overline{F}(ue+y)=\overline{F}(te)\overline{F}(x\circ a)$. Also considering the assumption $\overline{F}(te)=\exp\{-\exp(-x)\}$,

$$\overline{F}(ue+y)=\overline{F}(ue)\overline{F}(y_i e)\overline{F}(x\circ a). \qquad \ldots(2.6)$$

According to equation (2.3), the last two factors on the right-hand side of equation (2.6) becomes $\overline{F}(y_i e)\overline{F}(x\circ a)=\overline{F}(y_i e+x\circ a)$. Therefore, equation (2.6) can be rewritten as

$$\overline{F}(ue+y)=\overline{F}(ue)\overline{F}(y_i e+x\circ a). \qquad \ldots(2.7)$$

Form equation (2.5), we have $y_i e+x\circ a=y\circ e=y$, and hence equation (2.7) can takes the form of $\overline{F}(ue+y)=\overline{F}(ue)\overline{F}(y)$, therefore by equation (2.2), it is lack of memory property.

3. CHARACTERIZATION PROPERTIES

In section, we present the characterization properties of multivariate Marshall-Olkin Gumbel distribution of independently distributed random variables. The idea of Lack of memory property of independent random variables is used to derive these characterization properties.

The first characterization property of multivariate Marshall-Olkin Gumbel distribution of independently distributed random variables is derived in the following theorem.

THEOREM 3.1: If a survival distribution function, $\overline{F}(x_1, x_2, ..., x_n)$ of independent random variables $x_1, x_2, ..., x_n$, has a weak lack of memory property and $\overline{F}(te)$ is a Gumbel function of t and all marginal distributions of $\overline{F}(x_1, x_2, ..., x_n)$ are Marshall-Olkin distributions, then $\overline{F}(\bullet)$ is also Marshall-Olkin distribution.

PROOF: By lemma (2.1) above**,** $\overline{F}(x_1, x_2, ..., x_n)$ has a lack of memory property. Also, since all marginal distributions of $\overline{F}(x_1, x_2, ..., x_n)$ are Marshall-Olkin ones, therefore $\overline{F}(x_1, x_2, ..., x_n)$ is also Marshall-Olkin distribution.

The next characterization properties of the multivariate Marshall-Olkin Gumbel distribution are based on the result obtained by integrating equation (2) with respect to either t or x.

In 1-dimensional case, we integrate equation (2) with respect to the Borel measure $v(t)$ within the whole real line, and hence we get

$$\int_{-\infty}^{\infty} \overline{F}(t+x)\mu(dt) = \overline{F}(x) \text{ for every } x \geq 0, \text{ with } \mu = \frac{v(t)}{\int_{-\infty}^{\infty} \overline{F}(t)\mu(dt)}. \qquad \ldots(3.1)$$

Naturally, we assume that $\int_{-\infty}^{\infty} \overline{F}(t)\mu(dt) < \infty$. It can be derived from Lau and Rao [4] that if $\mu(t)$ has an infinite support, then the unique solution of equation (3.1) is $\overline{F}(t) = \exp(-\exp(-\alpha t))$, where $\alpha \geq 0$ is obtained from $\int_{-\infty}^{\infty} \exp(-\exp(-\alpha t))\mu(dt) = 1.$

In the case of 2-dimentional function $\overline{F}(x_1, x_2)$, we use the idea presented in Obretenov and Racev [5] for exponential function to derive the one for bivariate Gumbel distribution.

Now, if $\overline{F}(x_1, x_2)$ is survival distribution function with independent Gumbel marginal distributions, if μ is a Borel measure on the whole real line and if

$$\int_{-\infty}^{\infty} \overline{F}\left(t+x_1,t+x_2\right)\mu\left(dt\right)=\overline{F}\left(x_1,x_2\right), \qquad \ldots(3.2)$$

then $\overline{F}\left(x_1,x_2\right)$ is bivariate Marshall-Olkin Gumbel distribution.

Next, we give a theorem that generalized both cases mentioned above to $n-$dimensional case.

THEOREM 3.2: If $\overline{F}\left(x_1,x_2,...,x_n\right)$, $x_k \geq 0$, $k=1,2,...,n,$ is a survival distribution function whose marginal distributions are all of the independent Marshall-Olkin type and $\overline{F}\left(x_1,x_2,...,x_n\right)$ satisfies the equation

$$\int_{-\infty}^{\infty} \overline{F}\left(t+x_1,t+x_2,...,t+x_n\right)\mu\left(dt\right)=\overline{F}\left(x_1,x_2,...,x_n\right) \qquad \ldots(3.3)$$

for some Borel measure $\mu\left(t\right)$ on $\left(-\infty,\infty\right)$ with infinite support, then $\overline{F}\left(x_1,x_2,...,x_n\right)$ is an $n-$dimensional Marshall-Olkin distribution of independent random variables of $x_1,x_2,...,x_n$.

PROOF: Equation (3.3.) can be rewritten in a more compact form as

$$\int_{-\infty}^{\infty} \overline{F}\left(te+x\right)\mu\left(dt\right)=\overline{F}\left(x\right). \qquad \ldots(3.4)$$

If we let $x=0,$ it follows from equation (3.4) that $\int_{-\infty}^{\infty} \overline{F}\left(te\right)\mu\left(dt\right)=\overline{F}\left(0\right)=1.$ On the other hand, if we let $x=xe$, it follows that the function $\overline{F}_1\left(x\right)=\overline{F}\left(xe\right)$ satisfies equation (3.4) in the 1-dimensional case. From what we stated in theorem (3.1), derived from the result of Lau and Rao [3], $\overline{F}_1\left(x\right)=\exp\left(-\exp\left(-\alpha x\right)\right)$, that is $\overline{F}\left(xe\right)=\exp\left(-\exp\left(-\alpha x\right)\right)$.

Assume that $x=\left(x_1,x_2,...,x_n\right)$ and let us fix all $n-1$ coordinates $x_2,...,x_n$ of x. Consider

$$\overline{F}_{n-1}\left(u\right)=\frac{\overline{F}\left(u,u+x_2,...,u+x_n\right)}{\overline{F}\left(0,x_2,...,x_n\right)}=\frac{\overline{F}\left(ue+x\circ\alpha\right)}{\overline{F}\left(x\circ\alpha\right)}, \qquad \ldots(3.5)$$

where $\alpha \in E$ with $\alpha_1 = 0$. Since $\overline{F}(\bullet)$ satisfies equation (3.4), it is easy to check that $\overline{F}_{n-1}(u)$ satisfies equation

$$\int_{-\infty}^{\infty} \overline{F}_{n-1}(u+t)\mu(dt) = \overline{F}_{n-1}(u) \qquad \ldots(3.6)$$

and $\overline{F}_{n-1}(0) = 1$. Therefore, $\overline{F}_{n-1}(t) = \exp(-\exp(-\beta t))$, where $\beta = \beta(x_2, x_3, \ldots x_n)$. However, we have seen that equation (3.6) has a solution $\overline{F}_1(t) = \exp(-\exp(-\alpha t))$ and since this solution is unique, we have

$$\beta(x_2, x_3, \ldots x_n) = \alpha, \quad \overline{F}_{n-1}(t) = \exp(-\exp(-\alpha t)). \qquad \ldots(3.7)$$

Considering equations (3.5) and (3.7), we have

$\overline{F}(te + x\circ\alpha) = \exp(-\exp(-\alpha t))\overline{F}(x\circ\alpha) = \overline{F}_1(t)\overline{F}(x\circ\alpha) = \overline{F}(te)\overline{F}(x\circ\alpha)$, that is equation (2.3) is satisfied for all $\alpha \in E$ with the first coordinate $\alpha_1 = 0$. We can also establish that equation (2.3) is satisfies as well for $\alpha \in E$ with $\alpha_2 = 0$ and in general, for every $\alpha \in E$. Hence, the function $\overline{F}(\bullet)$ has the weak lack of memory property. Also, we have seen that $\overline{F}(te)$ is independent Gumbel distribution with respect to *t*. Thus, by theorem (3.1) above, $\overline{F}(x_1, x_2, \ldots, x_n)$ is an independent $n-$ dimensional Marshall-Olkin distribution. Hence, theorem (3.2) follows.

Next theorem gives yet another characterization property of independent $n-$dimensional Marshall-Olkin Gumbel distribution. This time around using an integrated equation obtained from equation (2.2) after integration with respect to x. In this case, we restrict our studies to the multivariate distributions of the class "Increasing failure rate", denoted by IFR. Precisely, we shall use one of the possible definitions given in Marshall [6] about the definitions with a monotone failure rate. Before the theorem, let us see one definition. This definition can be seen in Obretenov [2].

DEFINITION 3.1: A survival distribution function $\overline{F}(x_1, x_2, ..., x_n),\ x_i \geq 0,\ \ i = 1, 2, ..., n$ belongs to the class "Increasing failure rate", if

$$R(x) = \frac{\overline{F}(te + x)}{\overline{F}(x)} \qquad \ldots(3.8)$$

is decreasing with respect to x for all $t \geq 0,$ and $\overline{F}(x_1, ..., x_n) > 0$ for every $x_i,\ \ i = 1, 2, ..., n.$

This theorem provides us with the next characterization of independent $n-$ dimensional Marshall-Olkin Gumbel distribution.

THEOREM 3.3: Let $\overline{F}(x_1, ..., x_n)$ belong to the class IFR and $\mu(t)$ be a Borel measure on $R_n^+ = \{t : t_k \geq 0,\ k = 1, 2, ..., n\}$ such that

$$\int_{R_n^+(\alpha)} d\mu(t) \neq 0 \qquad \ldots(3.9)$$

for every $\alpha = (\alpha_1, \alpha_2, ..., \alpha_n)$, where $R_n^+(\alpha) = \{t : t_k \geq \alpha_k,\ 1 \leq k \leq n\}$. If $\overline{F}(x_1, ..., x_n)$ satisfies equation

$$\int_{R_n^+} \overline{F}(te + x)\, d\mu(x) = \overline{F}(te), \quad t \geq 0, \qquad \ldots(3.10)$$

and if all marginal distributions of $\overline{F}(x_1, ..., x_n)$ are independent Marshall-Olkin Gumbel distributions, then $\overline{F}(x_1, ..., x_n)$ is multivariate Marshall-Olkin Gumbel distribution.

PROOF: When $t = 0$, it follows from equation (3.10) that

$$\int_{R_n^+(\alpha)} \overline{F}(x)\, d\mu(x) = 1, \qquad \ldots(3.11)$$

therefore

$$I(\alpha) = \int_{R_n^+(0)} \overline{F}(x)\, d\mu(x) = 1 - \int_{R_n^+(\alpha)} \overline{F}(x)\, d\mu(x). \qquad \ldots(3.12)$$

Considering equation (3.9), $I(\alpha) \neq 0$ for every α. Hence equation (3.10) becomes

$$\int_{R_n^+(0)} R(x)\overline{F}(x)\, d\mu(x) + \int_{R_n^+(\alpha)} R(x)\overline{F}(x)\, d\mu(x) = \overline{F}(te). \qquad \ldots(3.13)$$

We replace the integrand $R(x)$ in first integral of the left of equation (3.13) with $R(0) \geq R(x) \quad for \;\; x \in R_n^+(\alpha)$ and in the second integral with $R(\alpha) \geq R(x) \quad for \;\; x \in R_n^+(\alpha)$. From these, we obtain

$$R(0)\int_{R_n^+(\alpha)} \overline{F}(x)d\mu(x) + R(\alpha)\int_{R_n^+(\alpha)} \overline{F}(x)d\mu(x) \geq \overline{F}(te). \qquad \ldots(3.14)$$

Using equations (3.12) and (3.14),we have $R(0)I(\alpha)+R(\alpha)\left[1-I(\alpha)\right] \geq \overline{F}(te) = R(0)$, from where we get $R(\alpha) \geq R(0)$ for arbitrary α. But $R(x)$ is decreasing and, hence, $R(\alpha) = R(0)$. The last equation is equivalent to $\overline{F}(te+\alpha) = \overline{F}(te)\overline{F}(\alpha)$, which shows that $\overline{F}(\bullet)$ has a lack of memory property. By our assumptions, all independent marginal distributions of $\overline{F}(\bullet)$ are Marshall-Olkin distributed, therefore, $\overline{F}(x_1, x_2, ..., x_n)$ is also independent multivariate Marshall-Olkin distribution. Hence, theorem (3.3) is completely proved.

Before we go to the next theorem, which give additional characterization property of independent multivariate Gumbel distribution, let us consider the following definition of Gumbel minima.

DEFINITION 3.2: Independent random variables $X_1, X_2, ..., X_n$ are said to have a joint Gumbel minima distribution whenever $P\left(\min_{i \in I} X_i > t\right) = \exp\left(-\exp\left(-\lambda_i t\right)\right), \;\; \lambda_i > 0,$ for every non-empty subset $I \subset \{1, 2, ..., n\}$.

THEOREM 3.4: Let, for every fixed $x \geq 0,$ the quotient

$$\varphi(te) = \frac{\overline{F}(te+xe)}{\overline{F}(te)} \qquad \ldots(3.15)$$

do not increase with respect to $t \geq 0$, and all independent marginal distributions of $\overline{F}(x_1, x_2, ..., x_n)$ be of the "Gumbel minima" type. If for some Borel measure μ *on* $(-\infty, \infty)$ with infinite support, we have

$$\int_{-\infty}^{\infty} \overline{F}(xe + te)\mu(dt) = \overline{F}(xe) \qquad \ldots(3.16)$$

for every $x \geq 0$, then $\overline{F}(x_1, x_2, ..., x_n)$ is "Gumbel minima" distributed.

PROOF: Similar to the proof of theorem (3.3), we use here the fact that $\varphi(te)$ does not increase in t and obtain $\varphi(\delta e) = \varphi(0)$ for every $\delta \geq 0$. The last equation is

$$\overline{F}(\delta e + xe) = \overline{F}(\delta e)\overline{F}(xe). \qquad \ldots(3.17)$$

From Galamgos and Kotz [3], it follows using equation (3.17) and assumption about the "Gumbel minima" type of the independent marginal distributions that the joint distribution $\overline{F}(x_1, x_2, ..., x_n)$ is also of the "Gumbel minima" type. Hence this theorem follows.

4. CONCLUSION

From what we have presented, we can conclude that independent multivariate Gumbel distribution can be extended to multivariate Marshall-Olkin Gumbel distributions. The characterization properties of these type of distributions can be derived using lack of memory property. Multivariate "Gumbel minima" type distributions can be derived and its characterization property can also be stated and proved.

ACKNOWLEDGEMENTS: The Author is highly grateful to all those that have contributed toward improved version of this paper.

REFERENCES

[1] Fedilis, I.U (2009): Intro. To Prob. And distribution theory, Imilvo publishers, Nsukka, Nig.

[2] Obretenov, A (1985): Characterizations of the multivariate Marshall-Olkin exponential distribution. Journal of prob. and Math. Statistics. 6: 51-56.

[3] Galambos, J and Kotz, S (1978): Characterizations of probability Distributions, Lecture notes in Math. 675.

[4] Lau, K.S and Rao, C.R (1989): Integrated Cauchy functional equations and characterizations of exponential law. Tech reports , Inst. Of stat. and Applied prob. University of Pittsburgh.

[5] Obretenov, A and Racev, S (1983): Characterizations of the bivariate exponential distribution and Marshall-Olkin distribution and stability. Lecture notes in Math, Stab. of stoch. models. 982: 136-149.

[6] Marshall, A. W (1975): Multivariate distributions with monotone hazard rate, Reliability and Fault Tree Analysis, Philadelphia: 259-284.