Pak.j.stat.oper.res. Vol.18 No. 1 2020 pp 195-210 DOI: http://dx.doi.org/10.18187/pjsor.v18i1.3096

Pakistan Journal of Statistics and Operation Research

Marshall-Olkin Zubair-G Family of Distributions

Suleman Nasiru[1*] and Abdul Ghaniyyu Abubakari [2]

* Corresponding Author

1. Department of Statistics, School of Mathematical Sciences, C. K. Tedam University of Technology and Applied Sciences, Navrongo, Upper East Region, Ghana, sulemanstat@gmail.com/ snasiru@cktutas.edu.gh

2. Department of Statistics, School of Mathematical Sciences, C. K. Tedam University of Technology and Applied Sciences, Navrongo, Upper East Region, Ghana, abughaniyyu@gmail.com/ aaghaniyyu@cktutas.edu.gh

Abstract

A new class of distributions called Marshall-Olkin Zubair-G family is proposed in this study. Some statistical properties of the family are derived and two special distributions namely, Marshall-Olkin Zubair Nadarajah-Haghighi and Marshall-Olkin Zubair Weibull distributions are developed. The plots of the density and hazard rate functions of the special distributions exhibit different shapes for chosen parameter values, making them good candidates for modeling different types of datasets. A real life application using the Marshall-Olkin Zubair Nadarajah-Haghighi distribution revealed that it performs better than other existing extensions of the Nadarajah-Haghighi distribution for the given dataset.

Key Words*:* Marshall-Olkin, Zubair, generalized classes,Weibull, Nadarajah-Haghighi

Mathematical Subject Classification: 60E05, 62E15

1. Introduction

The quest to develop flexible probability distributions have become an issue of interest to myriad of researchers owing to the usefulness of these distributions in modeling datasets and making inference in areas such as engineering, financial and biological modeling among others. This recent development has led to the proposition of several generalized classes of distributions called generators in literature for modifying existing distributions. The generators usually add one or more extra parameter (s) to the existing classical distributions to make them capable of modeling datasets that exhibit different traits such as bimodality, heavy-tail, monotonic and non-monotonic failure rates, symmetric and non-symmetric shapes. Hence, the aim is to develop new distributions that provide reasonable parametric fit to datasets obtained from different fields.

However, it is worth noting that no single probability distribution can provide good fit to all kinds of datasets. Thus, there is the need to develop new probability distributions for modeling datasets. Some common generators that have been developed for modifying existing distributions include: Marshall-Olkin alpha power family (Nassar et al., 2019); extended odd Fréchet-G family (Nasiru, 2018); Marshall-Olkin extended family (Marshall and Olkin, 1997); Zubair-G family (Zubair, 2018); Kumaraswamy-G family (Cordeiro and de Castro, 2011); odd Fréchet-G family (Haq and Elgarhy, 2018); odd Burr-G Poisson family (Nasir et al., 2018); alpha power transformed family (Mahdavi and Kundu, 2017); exponentiated generalized transformed-transformer family (Nasiru et al., 2017), Marshall-Olkin extended generalized Rayleigh (MirMostafaee et al., 2017) and Marshall-Olkin Burr X family (Jamal et al., 2017).

Recently, Zubair (2018) proposed the Zubair-G family of distributions with cumulative distribution function (CDF) as

 DOI: http://dx.doi.org/10.18187/pjsor.v18i1.3096

$$Z(x)=\frac{e^{\alpha G(x;\boldsymbol{\xi})^2}-1}{e^{\alpha}-1},\alpha>0,\boldsymbol{\xi}>0,x\in\mathbb{R}, \qquad (1)$$

where $G(x;\xi)$ is the CDF of the baseline distribution. The objective of this study is to develop another extension of the Zubair-G family called Marshall-Olkin Zubair (MOZ)-G family by adding an extra shape parameter to the Zubair-G family to make it more flexible. The Zubair-G family adds only a single scale parameter $\alpha>0$ to the baseline distribution. Thus, if the baseline distribution has no shape parameter as in the case of the exponential distribution, the resulting distribution will lack shape parameter. But to produce distribution with heavy-tail, and control skewness and kurtosis, a shape parameter is required. It is therefore necessary to add an extra shape parameter to the Zubair-G family.

Suppose $Z(x)$ is a baseline CDF which depends on a parameter vector $\boldsymbol{\xi}=(\xi_1,\xi_2,...,\xi_p)^T$ of dimension p. Then the CDF of the Marshall-Olkin family is defined as

$$F(x)=\frac{Z(x)}{\theta+(1-\theta)Z(x)},\theta>0,x\in\mathbb{R}, \qquad (2)$$

where θ is an extra shape parameter. Substituting equation (1) into equation (2), the CDF of the MOZ-G family of distribution is defined as

$$F(x)=\frac{e^{\alpha G(x;\boldsymbol{\xi})^2}-1}{(e^{\alpha}-1)\left[\theta+(1-\theta)(e^{\alpha}-1)^{-1}(e^{\alpha G(x;\boldsymbol{\xi})^2}-1)\right]},\alpha>0,\theta>0,\boldsymbol{\xi}>0,x\in\mathbb{R}. \qquad (3)$$

The probability density function (PDF) related to the MOZ-G family is given by

$$f(x)=\frac{e^{\alpha G(x;\boldsymbol{\xi})^2}-1}{(e^{\alpha}-1)\left[\theta+(1-\theta)(e^{\alpha}-1)^{-1}(e^{\alpha G(x;\boldsymbol{\xi})^2}-1)\right]^2},x\in\mathbb{R}. \qquad (4)$$

If we set $\theta=1$, we obtain the Zubair-G family. A physical interpretation of the new family is as follows: given N independent components each with probability mass function $P(N=n)=\theta(1-\theta)^{n-1},n=1,2,...$ and $0<\theta<1$ connected in series. Suppose the lifetime of each component $X_1,X_2,...,X_N$ are independent and identically distributed Zubair-G random variables with parameters α and $\boldsymbol{\xi}$. Then a random variable $X_{(1)}=\min(X_1,...,X_N)$ denotes the time to the first failure with distribution function

$$\begin{aligned}F(x)&=1-\sum_{n=1}^{\infty}P\left(\min(X_1,...,X_N)>x\right)\theta(1-\theta)^{n-1}\\&=1-\theta\frac{e^{\alpha}-e^{\alpha G(x;\boldsymbol{\xi})^2}}{e^{\alpha}-1}\sum_{n=1}^{\infty}\left[\frac{e^{\alpha}-e^{\alpha G(x;\boldsymbol{\xi})^2}}{e^{\alpha}-1}(1-\theta)\right]^{n-1}\\&=\frac{e^{\alpha G(x;\boldsymbol{\xi})^2}-1}{(e^{\alpha}-1)\left[\theta+(1-\theta)(e^{\alpha G(x;\boldsymbol{\xi})^2}-1)\right]},x\in\mathbb{R}.\end{aligned}$$

Hence, the CDF of the MOZ-G family is obtained. Alternatively, the CDF in equation (3) can be interpreted as follows: suppose the random variable N with probability mass function

$P(N=n)=\theta^{-1}(1-\theta^{-1})^{n}, n=0,1,...,\theta>1$ denotes independent components of a parallel system. If $X_1,...,X_N$ constitute independent and identically distributed lifetime of the components and are assumed to be Zubair-G random variables with parameters α and ξ. Then the random variable $X_{(n)}=\max(X_1,...,X_N)$ denotes the lifetime of the system and has the distribution function defined in equation (3).

The corresponding hazard rate function of the MOZ-G family is

$$\tau(x)=\frac{2\alpha g(x;\xi)G(x;\xi)e^{\alpha G(x;\xi)^2}}{\left[\theta+(1-\theta)(e^{\alpha}-1)^{-1}(e^{\alpha G(x;\xi)^2}-1)\right]\left[e^{\alpha}-e^{\alpha G(x;\xi)^2}\right]}, x\in \mathbb{R}. \tag{5}$$

For the sake of simplicity, $G(x;\xi)$ can be written as $G(x)$ and a random variable X that follows the MOZ-G family is represented by $X \sim MOZ-G(x;\alpha,\theta,\xi)$. The MOZ-G family has a tractable CDF making it easy to obtain random observations from the family provided the CDF of the baseline distribution is also tractable. The quantile function of the MOZ-G family is given by

$$x_u=Q_G\left[\frac{1}{\alpha}\log\left(\frac{1+u(\theta e^{\alpha}-1)}{1-u(1-\theta)}\right)\right]^{\frac{1}{2}}, u\in[0,1], \tag{6}$$

where $Q_G(\cdot)$ is the quantile function of the baseline distribution. Using some algebraic manipulation, the density function of the MOZ-G family can be expressed in an infinite mixture form as

$$f(x)=2\alpha\sum_{i=0}^{\infty}\sum_{k=0}^{\infty}\sum_{j=0}^{i}\omega_{ijk}\,g(x)G(x)^{2k+1}, \tag{7}$$

where

$$\omega_{ijk}=\frac{(-1)^{i+j}(i+1)\binom{i}{j}\left[\theta(e^{\alpha}-1)\right]^{-(i+1)}(1-\theta)^{i}\left[\alpha(1+i-j)\right]^{k}}{k!}.$$

The mixture representation of the density function is useful when deriving the structural properties of the MOZ-G family of distributions.

2. Statistical Properties

This section presents some useful statistical properties of the MOZ-G family of distributions. The statistical properties derived are the moments, moment generating function (MGF), entropies, stochastic ordering and order statistics.

2.1. Moments and Moment Generating Function

The r^{th} non-central moment is very important when estimating measures of central tendency, dispersion and measures of shapes. The r^{th} non-central moment of the MOZ-G family is given by

$$\begin{aligned}\mu_r' &= \int_{-\infty}^{\infty} x^r f(x)dx \\ &= \int_{-\infty}^{\infty} x^r 2\alpha \sum_{i=0}^{\infty}\sum_{k=0}^{\infty}\sum_{j=0}^{i} \omega_{ijk} g(x)G(x)^{2k+1} dx \\ &= 2\alpha \sum_{i=0}^{\infty}\sum_{k=0}^{\infty}\sum_{j=0}^{i} \omega_{ijk} \int_{-\infty}^{\infty} x^r g(x)G(x)^{2k+1} dx. \end{aligned} \quad (8)$$

The r^{th} non-central moment can be expressed in terms of the quantile function of the baseline distribution. Letting $G(x) = u$, the r^{th} non-central moment can be expressed as

$$\mu_r' = 2\alpha \sum_{i=0}^{\infty}\sum_{k=0}^{\infty}\sum_{j=0}^{i} \omega_{ijk} \int_0^1 Q_G(u)^r u^{2k+1} du,$$

where $Q_G(u)$ is the quantile function of the baseline distribution with CDF $G(x)$. The MGF of a random variable X that follows the MOZ-G family of distributions if it exists is given by

$$\begin{aligned} M_X(t) &= \int_{-\infty}^{\infty} e^{tx} f(x)dx \\ &= \sum_{r=0}^{\infty} \frac{t^r}{r!}\mu_r' \\ &= 2\alpha \sum_{r=0}^{\infty}\sum_{i=0}^{\infty}\sum_{k=0}^{\infty}\sum_{j=0}^{i} \frac{\omega_{ijk} t^r}{r!} \int_{-\infty}^{\infty} x^r g(x)G(x)^{2k+1}. \end{aligned} \quad (9)$$

Alternatively, the MGF can be expressed in terms of the quantile function of the baseline distribution as

$$M_X(t) = 2\alpha \sum_{i=0}^{\infty}\sum_{k=0}^{\infty}\sum_{j=0}^{i} \omega_{ijk} \int_0^1 e^{tQ_G(u)} u^{2k+1} du.$$

2.2 Entropy Measures

Entropies are measures of variation of a random variable. This section presents the Rényi and $\delta-$entropies. The Rényi entropy (Rényi, 1961) of a random variable X with density function $f(x)$ is given by

$$I_R(\delta) = \frac{1}{1-\delta} \log\left[\int_{-\infty}^{\infty} f(x)^{\delta} dx\right], \delta \neq 1, \delta > 0. \quad (10)$$

Expanding $f(x)^{\delta}$, the Rényi entropy of the MOZ-G random variable is

$$I_R(\delta) = \frac{1}{1-\delta} \log\left[(2\alpha)^{\delta} \sum_{i=0}^{\infty}\sum_{k=0}^{\infty}\sum_{j=0}^{i} \varpi_{ijk} \int_{-\infty}^{\infty} g(x)^{\delta} G(x)^{2k+\delta} dx\right], \delta \neq 1, \delta > 0, \quad (11)$$

where

$$\varpi_{ijk} = \frac{(-1)^{i+j}\theta^{-\delta-i}(1-\theta)^{i}(e^{\alpha}-1)^{-\delta-i}\left(\alpha(1+i-j)\right)^{k}}{k!}\binom{i}{j}\binom{2\delta+i-1}{i}.$$

The $\delta-$entropy is defined as

$$I_R(\delta) = \frac{1}{\delta-1}\log\left[1-\int_{-\infty}^{\infty} f(x)^{\delta}\,dx\right], \delta \neq 1, \delta > 0. \qquad (12)$$

Hence, the $\delta-$entropy for the MOZ-G random variable is

$$H(\delta) = \frac{1}{\delta-1}\log\left[1-(2\alpha)^{\delta}\sum_{i=0}^{\infty}\sum_{k=0}^{\infty}\sum_{j=0}^{i}\varpi_{ijk}\int_{-\infty}^{\infty} g(x)^{\delta}G(x)^{2k+\delta}\,dx\right], \delta \neq 1, \delta > 0. \qquad (13)$$

2.3. Stochastic Ordering

A random variable X_1 is said to be stochastically greater than X_2 $\left(X_2 \leq_{st} X_1\right)$ if $F_{X_2}(x) \leq F_{X_1}(x)$ for all x. Also, X_1 is said to be stochastically greater than X_2 in the

i. hazard rate order $(X_2 \leq_{hr} X_1)$ if $\tau_{X_2}(x) \leq \tau_{X_1}(x)$ for all x,

ii. likelihood ratio order $\left(X_2 \leq_{lr} X_1\right)$ if $\frac{f_{X_1}(x)}{f_{X_2}(x)}$ is an increasing function of x.

If X_1 is MOZ-G random variable and X_2 is Zubair-G random variable, then the likelihood ratio is

$$\frac{f_{X_1}(x)}{f_{X_2}(x)} = \theta\left[\theta+(1-\theta)(e^{\alpha}-1)^{-1}(e^{\alpha G(x)^2}-1)\right]^{-2}. \qquad (14)$$

Finding the first derivative of the likelihood ratio with respect to x yields

$$\frac{d}{dx}\frac{f_{X_1}(x)}{f_{X_2}(x)} = \frac{-4\alpha\theta(1-\theta)(e^{\alpha}-1)^{-1}g(x)G(x)e^{\alpha G(x)^2}}{\left[\theta+(1-\theta)(e^{\alpha}-1)^{-1}(e^{\alpha G(x)^2}-1)\right]^{3}}. \qquad (15)$$

If $\theta < 1$, then $\frac{d}{dx}\frac{f_{X_1}(x)}{f_{X_2}(x)} < 0$. Thus, $X_1 \leq_{lr} X_2$. It is important to note that the likelihood ratio order implies hazard rate order $(X_1 \leq_{hr} X_2)$ and stochastic order $\left(X_1 \leq_{st} X_2\right)$.

2.4. Order Statistics

Order statistics are useful in quality control and reliability analysis. This section presents the density function of the p^{th} order statistic and its r^{th} non-central moment. Let $X_{1:n} \leq X_{2:n} \leq \cdots \leq X_{n:n}$ represent the order statistics of a random sample $X_1, X_2, ..., X_n$ from MOZ-G family of distributions, then the density function of $X_{p:n}$, for $p = 1, 2, ..., n$ is

$$f_{p:n}(x) = \frac{2\alpha n!}{(p-1)!(n-p)!}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty}\sum_{i=0}^{n-p}\sum_{k=0}^{p+i+j-1}\upsilon_{ijkm}\,g(x)G(x)^{2m+1}, \qquad (16)$$

where

$$\upsilon_{ijkm}=\frac{(-1)^{i+j+k}\theta^{-j-p-i}(1-\theta)^{j}(e^{\alpha}-1)^{p+i+j-1}\left(\alpha(p+i+j-k)\right)^{m}}{m!}\binom{n-p}{i}\binom{p+i}{j}\binom{p+i+j-1}{k}.$$

The r^{th} non-central moment of $X_{p:n}$ is given by

$$\mu_r^{'(p:n)}=\frac{2\alpha n!}{(p-1)!(n-p)!}\sum_{j=0}^{\infty}\sum_{m=0}^{\infty}\sum_{i=0}^{n-p}\sum_{k=0}^{p+i+j-1}\upsilon_{ijkm}\int_{-\infty}^{\infty}x^{r}g(x)G(x)^{2m+1}dx. \qquad (17)$$

3. Parameter Estimation

The maximum likelihood technique is employed to estimate the parameters of the MOZ-G family of distributions. Suppose that $x_1,x_2,...,x_n$ is a random sample of size n from the MOZ-G family of distributions. Then, the corresponding log-likelihood function is given by

$$\ell=n\log(2\theta\alpha)-n\log(e^{\alpha}-1)+\sum_{i=1}^{n}\log g(x_i;\boldsymbol{\xi})+\sum_{i=1}^{n}\log G(x_i;\boldsymbol{\xi})+$$
$$\alpha\sum_{i=1}^{n}G(x_i;\boldsymbol{\xi})^2-2\sum_{i=1}^{n}\log\left[\theta+(1-\theta)(e^{\alpha}-1)^{-1}(e^{\alpha G(x_i;\boldsymbol{\xi})^2}-1)\right]. \qquad (18)$$

Finding the first partial derivative of the total log-likelihood function with respect to the parameters, the score functions are given by

$$\frac{\partial\ell}{\partial\theta}=\frac{n}{\theta}-2\sum_{i=1}^{n}\frac{\left[1-(e^{\alpha}-1)^{-1}(e^{\alpha G(x_i;\boldsymbol{\xi})^2}-1)\right]}{\left[\theta+(1-\theta)(e^{\alpha}-1)^{-1}(e^{\alpha G(x_i;\boldsymbol{\xi})^2}-1)\right]}, \qquad (19)$$

$$\frac{\partial\ell}{\partial\alpha}=\frac{n}{\alpha}-\frac{ne^{\alpha}}{e^{\alpha}-1}+\sum_{i=1}^{n}G(x_i;\boldsymbol{\xi})^2-$$
$$2\sum_{i=1}^{n}\frac{(\theta-1)\left[e^{\alpha}(e^{\alpha}-1)^{-2}(e^{\alpha G(x_i;\boldsymbol{\xi})^2}-1)-(e^{\alpha}-1)^{-1}G(x_i;\boldsymbol{\xi})^2e^{\alpha G(x_i;\boldsymbol{\xi})^2}\right]}{\left[\theta+(1-\theta)(e^{\alpha}-1)^{-1}(e^{\alpha G(x_i;\boldsymbol{\xi})^2}-1)\right]}, \qquad (20)$$

and

$$\frac{\partial\ell}{\partial\boldsymbol{\xi}}=\sum_{i=1}^{n}\frac{\partial g(x_i;\boldsymbol{\xi})/\partial\boldsymbol{\xi}}{g(x_i;\boldsymbol{\xi})}+\sum_{i=1}^{n}\frac{\partial G(x_i;\boldsymbol{\xi})/\partial\boldsymbol{\xi}}{G(x_i;\boldsymbol{\xi})}+2\alpha\sum_{i=1}^{n}G(x_i;\boldsymbol{\xi})\partial G(x_i;\boldsymbol{\xi})/\partial\boldsymbol{\xi}-$$
$$4\alpha\sum_{i=1}^{n}\frac{(1-\theta)(e^{\alpha}-1)^{-1}e^{\alpha G(x_i;\boldsymbol{\xi})^2}G(x_i;\boldsymbol{\xi})\partial G(x_i;\boldsymbol{\xi})/\partial\boldsymbol{\xi}}{\left[\theta+(1-\theta)(e^{\alpha}-1)^{-1}(e^{\alpha G(x_i;\boldsymbol{\xi})^2}-1)\right]}. \qquad (21)$$

To obtain the estimators for the parameters, the score functions are equated to zero and the resulting system of equations are solved numerically. In order to find the interval estimates of the parameters, the observed information matrix can be computed as $I(\vartheta)=\frac{\partial^2\ell}{\partial s\partial z}(\text{for } s,z=\alpha,\theta,\boldsymbol{\xi})$, whose elements are obtained numerically. To determine whether the MOZ-G distributions are superior to the Zubair-G distributions for given datasets, the likelihood ratio (LR) test is carried out using the following hypotheses: $H_0:\theta=1$ versus $H_1:H_0$ is false. The test statistic for the LR test is given by $LR=2\{\ell(\vartheta)-\ell(\bar{\vartheta})\}$, where ϑ is the vector of unrestricted maximum likelihood estimates of parameters under H_1 and $\bar{\vartheta}$ is the vector of restricted parameter estimates under H_0. The LR test statistic asymptotically follows the Chi-square distribution with degrees of freedom equal to the difference between the numbers of parameters of the two models. The null hypothesis is rejected at level ρ when the LR test statistic exceeds the upper $100(1-\rho)\%$ quantile function of the Chi-square distribution.

4. Special Distributions

In this section, two special cases of the MOZ-G family of distributions are discussed.

4.1. Marshall-Olkin Zubair Nadarajah-Haghighi (MOZNH) Distribution

Given that the baseline CDF is that of the Nadarajah-Haghighi (NH) distribution. That is, $G(x)=1-e^{1-(1+\gamma x)^{\beta}}, x>0, \gamma>0, \beta>0$ and the corresponding density function is $g(x)=\beta\gamma(1+\gamma x)^{\beta-1}e^{1-(1+\gamma x)^{\beta}}, x>0$. The PDF of the MOZNH distribution is given by

$$f(x)=\frac{2\alpha\beta\gamma\theta(1+\gamma x)^{\beta-1}e^{1-(1+\gamma x)^{\beta}}(1-e^{1-(1+\gamma x)^{\beta}})e^{\alpha(1-e^{1-(1+\gamma x)^{\beta}})^2}}{(e^{\alpha}-1)\left[\theta+(1-\theta)(e^{\alpha}-1)^{-1}(e^{\alpha(1-e^{1-(1+\gamma x)^{\beta}})^2}-1)\right]^2}, x>0, \qquad (22)$$

where $\alpha>0, \gamma>0$ are scale parameters and $\beta>0, \theta>0$ are shape parameters. The PDF of MOZNH distribution can exhibit different shapes such as symmetric, right skewed and left skewed with different degrees of kurtosis as shown in Figure 1.

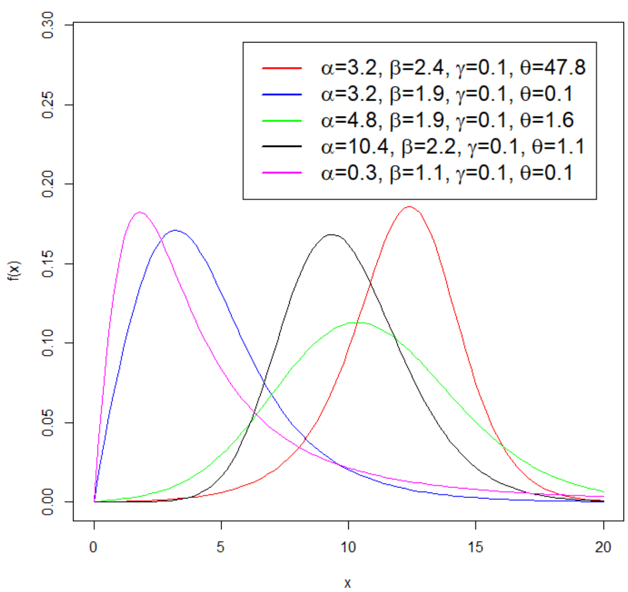

Figure 1: Plot of density function of MOZNH distribution

The hazard rate function of the MOZNH distribution is given by

$$\tau(x)=\frac{2\alpha\beta\gamma(1+\gamma x)^{\beta-1}e^{1-(1+\gamma x)^{\beta}}(1-e^{1-(1+\gamma x)^{\beta}})e^{\alpha(1-e^{1-(1+\gamma x)^{\beta}})^{2}}}{\left[\theta+(1-\theta)(e^{\alpha}-1)^{-1}(e^{\alpha(1-e^{1-(1+\gamma x)^{\beta}})^{2}}-1)\right]\left[e^{\alpha}-e^{\alpha(1-e^{1-(1+\gamma x)^{\beta}})^{2}}\right]},x>0. \qquad (23)$$

The hazard rate function for the MOZNH distribution can exhibit non-monotonic failure rates such as the upside-down bathtub for some given parameter values.

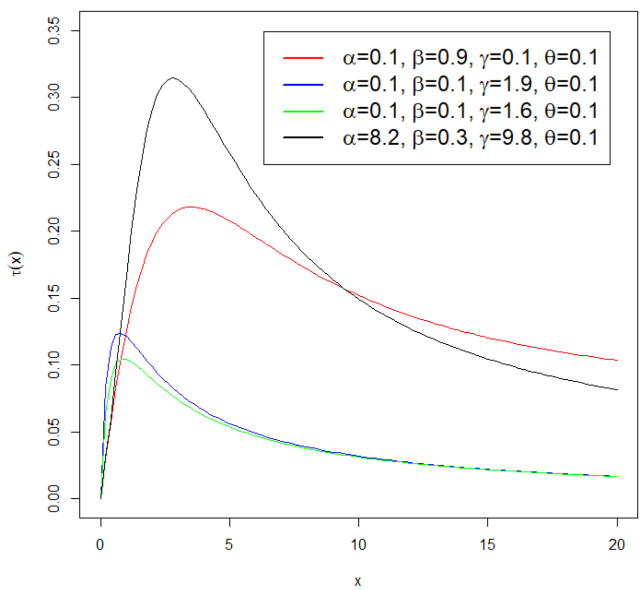

Figure 2: Plot of hazard rate function of MOZNH distribution

The quantile function of the MOZNH distribution is

$$x_{u}=\frac{1}{\gamma}\left\{\left[1-\log\left(1-\left(\frac{1}{\alpha}\log\left(\frac{1+u(\theta e^{\alpha}-1)}{1-u(1-\theta)}\right)\right)^{\frac{1}{2}}\right)\right]^{\frac{1}{\beta}}-1\right\},u\in[0,1]. \qquad (24)$$

4.2. Marshall-Olkin Zubair Weibull (MOZW) Distribution

Suppose the baseline CDF is that of Weibull distribution. That is $G(x)=1-e^{-\gamma x^{\beta}},x>0,\gamma>0,\beta>0$ and the corresponding density function is $g(x)=\beta\gamma x^{\beta-1}e^{-\gamma x^{\beta}},x>0.$ The PDF of the MOZW distribution is given by

$$f(x)=\frac{2\alpha\beta\gamma\theta x^{\beta-1}e^{-\gamma x^{\beta}}(1-e^{-\gamma x^{\beta}})e^{\alpha(1-e^{-\gamma x^{\beta}})^{2}}}{(e^{\alpha}-1)\left[\theta+(1-\theta)(e^{\alpha}-1)(e^{\alpha(1-e^{-\gamma x^{\beta}})^{2}}-1)\right]^{2}},x>0, \qquad (25)$$

where $\alpha > 0, \gamma > 0$ are scale parameters and $\beta > 0, \theta > 0$ are shape parameters. The density function of the MOZW distribution can exhibit symmetric and right skewed shapes with different degrees of kurtosis as shown in Figure 3 for some selected parameter values.

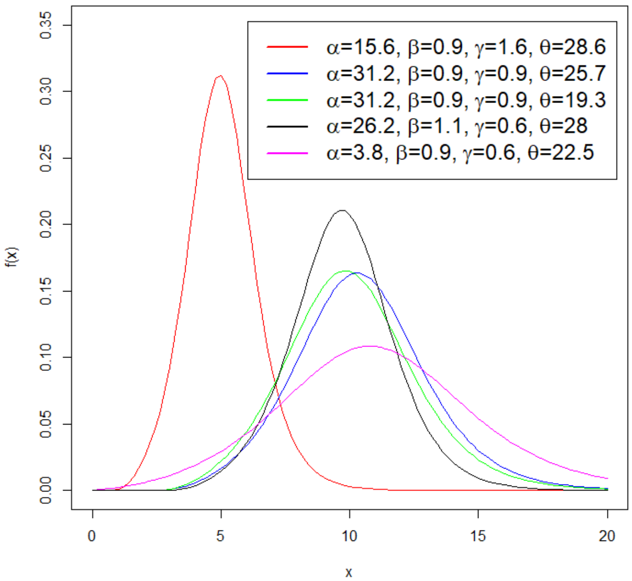

Figure 3: Plot of density function of MOZW distribution

The hazard rate function of the MOZW distribution is given by

$$\tau(x) = \frac{2\alpha\beta\gamma x^{\beta-1} e^{-\gamma x^{\beta}} (1-e^{-\gamma x^{\beta}}) e^{\alpha(1-e^{-\gamma x^{\beta}})^2}}{\left[\theta + (1-\theta)(e^{\alpha}-1)(e^{\alpha(1-e^{-\gamma x^{\beta}})^2}-1)\right]\left[e^{\alpha} - e^{\alpha(1-e^{-\gamma x^{\beta}})^2}\right]}, x > 0. \quad (26)$$

Figure 4 displays the hazard rate function of the MOZW distribution for some given parameter values. It is obvious from Figure 4 that the hazard rate function of the MOZW distribution can exhibit decreasing and upside-down bathtub failure rates.

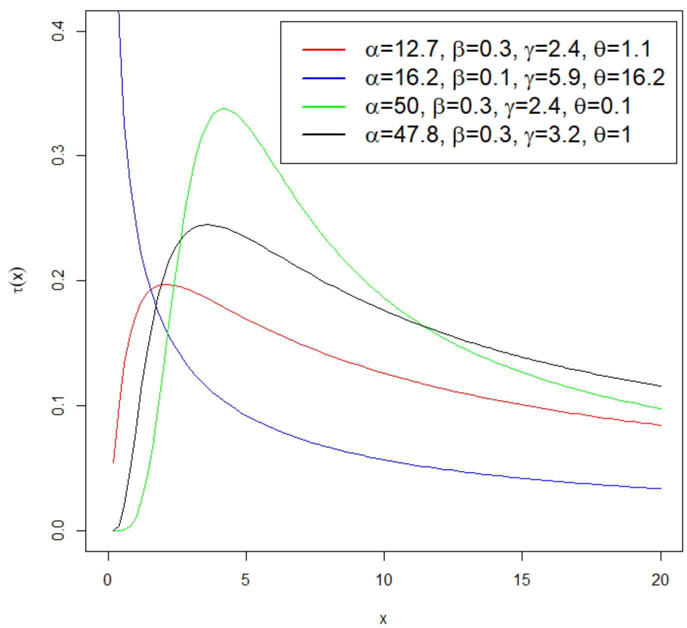

Figure 4: Plot of hazard rate function of MOZW distribution

The quantile function of the MOZW distribution is

$$x_u = \left[\frac{1}{\gamma} \log\left(1 - \left(\frac{1}{\alpha} \log\left(\frac{1+u(\theta e^{\alpha} - 1)}{1-u(1-\theta)} \right) \right) \right) \right]^{\frac{1}{\beta}}, u \in [0,1]. \qquad (27)$$

5. Simulation Studies

This section presents the Monte Carlo simulation results used to assess the performance of the estimators of the parameters. For illustration purpose the MOZNH distribution was used for the simulation. The experiment was repeated $10,000$ times with sample sizes $n = 25, 50, 75, 100$ and 125. The root mean square error (RMSE) for the parameters as shown in Table 1 decays to zero as the sample size increases. The coverage probabilities (CP) for the 95% confidence interval for the parameters in some cases were quite close to the nominal level of 0.95.

Table 1: Monte Carlo simulation results

Parameter	n	RMSE					CP			
$(\alpha,\beta,\gamma,\theta)=(0.4,0.5,0.2,4.8)$		α	β	γ	θ		α	β	γ	θ
	25	0.233	0.120	0.116	0.870		0.960	0.973	0.963	0.930
	50	0.221	0.092	0.104	0.778		0.956	0.969	0.951	0.964
	75	0.217	0.075	0.093	0.666		0.953	0.901	0.923	0.954
	100	0.209	0.072	0.087	0.703		0.962	0.998	0.978	0.960
	125	0.122	0.058	0.077	0.629		0.964	0.999	0.945	0.916
$(\alpha,\beta,\gamma,\theta)=(0.5,0.5,0.5,0.5)$	25	2.211	0.218	3.463	3.183		0.919	0.970	0.865	0.950
	50	1.659	0.167	2.145	1.584		0.941	0.956	0.956	0.935
	75	1.041	0.133	1.351	1.260		0.945	0.941	0.979	0.830
	100	1.011	0.128	0.214	0.817		0.963	0.954	0.997	0.959
	125	0.834	0.109	0.269	0.672		0.915	0.931	0.986	0.850
$(\alpha,\beta,\gamma,\theta)=(0.4,0.5,0.1,0.2)$	25	1.473	0.294	3.770	5.272		0.574	0.968	0.913	0.909
	50	1.399	0.178	0.942	1.109		0.643	0.943	0.904	0.975
	75	1.142	0.146	0.793	0.794		0.671	0.940	0.908	0.965
	100	0.059	0.123	0.192	0.442		0.665	0.906	0.923	0.958
	125	0.038	0.110	0.085	0.384		0.624	0.900	0.925	0.981

6. Empirical illustration

Here, we demonstrated the application of the MOZNH distribution using dataset. The dataset comprises the remission time of 128 bladder cancer patients presented in Lee and Wang (2003) and are:

0.08	9.22	2.62	15.96	5.49	5.85	12.07	3.52	25.82	7.39	1.19	17.36	2.02
2.09	13.8	3.82	36.66	7.66	8.26	21.73	4.98	0.51	10.34	2.75	1.4	3.31
3.48	25.74	5.32	1.05	11.25	11.98	2.07	6.97	2.54	14.83	4.26	3.02	4.51
4.87	0.5	7.32	2.69	17.14	19.13	3.36	9.02	3.7	34.26	5.41	4.34	6.54
6.94	2.46	10.06	4.23	79.05	1.76	6.93	13.29	5.17	0.9	7.63	5.71	8.53
8.66	3.64	14.77	5.41	1.35	3.25	8.65	0.4	7.28	2.69	17.12	7.93	12.03
13.11	5.09	32.15	7.62	2.87	4.5	12.63	2.26	9.74	4.18	46.12	11.79	20.28
23.63	7.26	2.64	10.75	5.62	6.25	22.69	3.57	14.76	5.34	1.26	18.1	2.02
0.2	9.47	3.88	16.62	7.87	8.37	3.36	5.06	26.31	7.59	2.83	1.46	
2.23	14.24	5.32	43.01	11.64	12.02	6.76	7.09	0.81	10.66	4.33	4.4	

We compared the performance of the MOZNH distribution with that of the Zubair NH (ZNH), exponentiated NH (ENH) (Abdul-Moniem, 2015) and Kumaraswamy NH (KNH) (Lima, 2015) distributions using the Akaike information criterion (AIC), consistent Akaike information criterion (CAIC) and Bayesian information criterion (BIC). The Kolmogorov-Smirnov (K-S), Cramér-von Mises (CM) and Anderson-Darling (AD) statistics were used to investigate the goodness-of –fit of the models. The density functions of the ENH and KNH distributions are respectively given by:

$$f(x)=\beta\gamma\theta(1+\gamma x)^{\beta-1}e^{1-(1+\gamma x)^{\beta}}\left(1-e^{1-(1+\gamma x)^{\beta}}\right)^{\theta-1},x>0, \qquad (28)$$

and

$$f(x)=ab\beta\gamma(1+\gamma x)^{\beta-1}e^{1-(1+\gamma x)^{\beta}}\left(1-e^{1-(1+\gamma x)^{\beta}}\right)^{a-1}\left(1-\left(1-e^{1-(1+\gamma x)^{\beta}}\right)^{a}\right)^{b-1},x>0. \qquad (29)$$

Table 2 presents the maximum likelihood estimates of parameters of the fitted distributions with their corresponding standard errors and 95% confidence intervals (CI). The maximum likelihood estimates of the parameters were obtained by maximizing the log-likelihood functions of the fitted distributions via the subroutine *mle2* using the *bbmle* package in R (Bolker, 2014). The optimizations were carried out using the BFGS technique and the initial values for the optimization were obtained using the *GenSA* package in R. The estimates of the parameters were all significant at the 5% level.

Table 2: Estimates, standard errors and CI

Distribution	**Estimates**	**Standard error**	**CI**
MOZNH	$\alpha = 5.2352$	1.6678	$[1.9663, 8.5041]$
	$\beta = 0.2524$	5.8951×10^{-3}	$[0.2409, 0.2640]$
	$\gamma = 299.5666$	1.5910×10^{-2}	$[299.5354, 299.5978]$
	$\theta = 28.8791$	2.2813×10^{-1}	$[28.4320, 29.3262]$
ZNH	$\alpha = 11.6732$	1.6874	$[8.3659, 14.9805]$
	$\beta = 0.2372$	0.0060	$[0.2254, 0.2490]$
	$\gamma = 99.5888$	0.0593	$[99.4726, 99.7050]$
ENH	$\beta = 0.2021$	0.0054	$[0.1905, 0.2137]$
	$\gamma = 387.1700$	0.1185	$[386.9377, 387.4023$
	$\theta = 27.5698$	4.6654	$[18.4256, 36.7140]$
KNH	$a = 10.5978$	1.0370	$[8.5653, 12.6303]$
	$\hat{b} = 48.4529$	2.1583×10^{-2}	$[48.4106, 48.4952]$
	$\beta = 0.0997$	4.6959×10^{-3}	$[0.0905, 0.1089]$
	$\gamma = 285.8821$	7.5794×10^{-3}	$[285.8672, 285.8970]$

Table 3 displays the model selection criteria and the goodness-of-fit statistics for investigating how well the distributions fit the given dataset. The results indicate that the MOZNH distribution provides a better fit to the datasets than the other candidate distributions because it has the least values for thse model selection criteria and the goodness-of-fit statistics.

Table 3: Information criteria and goodness-of-fit statistics

Distribution	**AIC**	**BIC**	**CAIC**	**AD**	**CM**	**K-S**
MOZNH	827.1233	838.5314	827.4485	0.1406	0.0234	0.0370
ZNH	839.7497	851.1578	840.0749	0.9492	0.1415	0.0672
ENH	851.6863	860.2424	851.8798	1.8426	0.2855	0.0953
KNH	831.3724	842.7805	831.6976	0.3929	0.0573	0.0496

The LR test was conducted to compare the performance of the MOZNH distribution with the ZNH distribution. The test yielded a test statistic of 12.6260 with corresponding p-value of 0.0004. This is an indication that the MOZNH distribution gives a better fit to the dataset than the ZNH distribution. Figure 5 displays the density plots and the distribution function plots of the fitted distributions. The plots revealed that the MOZNH distribution fits the dataset well.

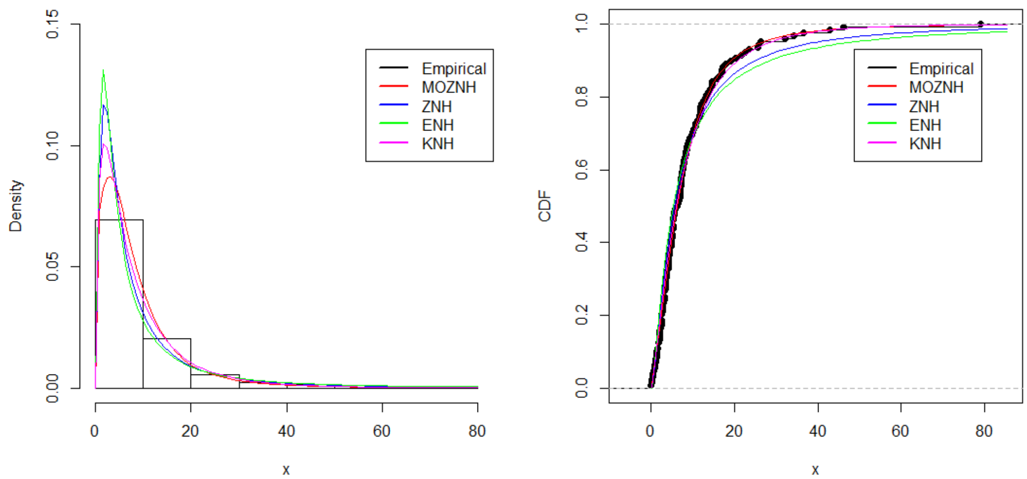

Figure 5: Density and distribution functions plots of the fitted distributions

The probability-probability plots of the MOZNH, ZNH, ENH and KNH distributions for the dataset are presented in Figure 6. Figure 6 revealed that the MOZNH distribution fitted the dataset well.

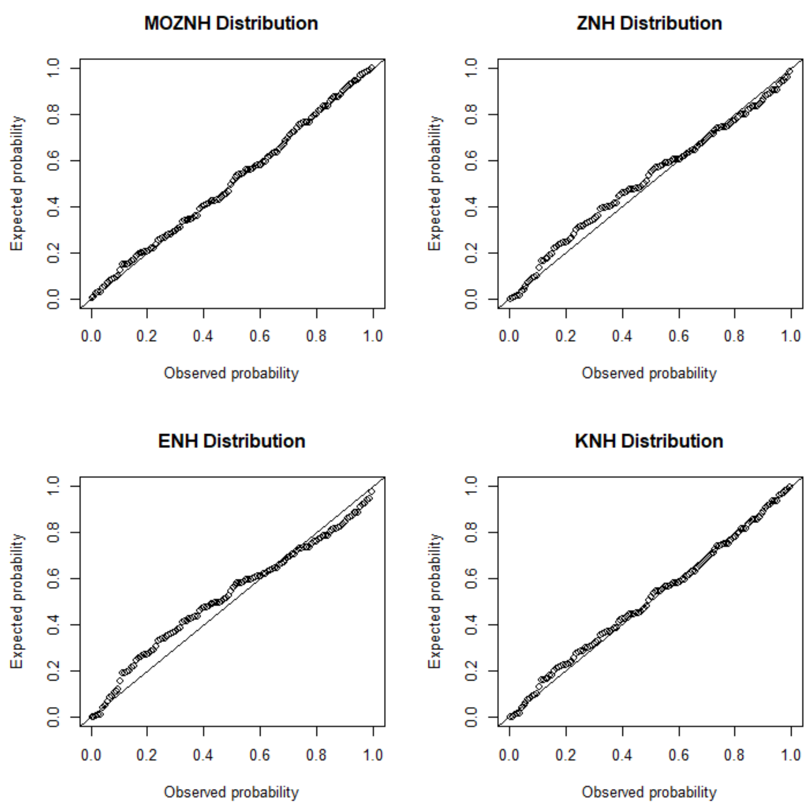

Figure 6: Probability-Probability plots of fitted distributions

7. Conclusion

The MOZ-G family of distributions was developed in this study. The proposed generator was used to develop the MOZNH and MOZW distributions. The density and hazard rate functions of the MOZNH and MOZW distributions exhibit different type of shapes making them suitable for analyzing datasets with either monotonic or non-monotonic failure rates. The application of the MOZNH distribution was illustrated using datasets and the findings indicated that the distribution fitted the dataset well.

References

1. Abdul-Moniem, I. B. (2015). Exponentiated Nadarajah and Haghighi's exponential distribution. International Journal of Mathematical Analysis and Applications, **2**(5): 68-73.
2. Bolker, B. (2014). Tools for general maximum likelihood estimation. R development core team.
3. Cordeiro, G. M. and de Castro, M. (2011). A new family of generalized distributions. Journal of Statistical Computation and Simulation, **81**(7): 883-898.
4. Haq, M. A. and Elgarhy, M. (2018). The odd Fréchet-G family of probability distributions. Journal of Statistics Applications and Probability, **7**(1): 189-203.
5. Jamal, F., Tahir, M. H., Alizadeh, M. and Nasir, M. A. (2017). On Marshall-Olkin Burr X family of distributions. Tbilisi Mathematical Journal, **10**(4): 175-199.
6. Lee, E. T. and Wang, J. (2003). Statistical methods for survival data analysis. Volume 476. John Wiley and Sons.
7. Lima, S. R. L. (2015). The half normal generalized family and Kumaraswamy Nadarajah-Haghighi distribution. Master thesis, Universidade Federal de Pernumbuco.
8. Mahdavi, A. and Kundu, D. (2017). A new method for generating distributions with an application to exponential distribution. Communications in Statistics-Theory and Methods, **46**(13): 6543-6557.
9. Marshall, A. W. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. Biometrika, **84**:641-652.
10. MirMostafaee, S. M. T. K., Mahdizadeh, M. and Lemonte, A. J. (2017). The Marshall-Olkin extended generalized Rayleigh distribution: Properties and applications. Communications in Statistics: Theory and Methods, **46**: 653-671.
11. Nasir, M. A., Jamal, F., Silva, G. O. and Tahir, M. H. (2018). Odd Burr-G Poisson family of distributions. Journal of Statistics Applications and Probability, **7**(1): 9-28.
12. Nasiru, S. (2018). Extended odd Fréchet-G family of distributions. Journal of Probability and Statistics, **2018**: 1-12.
13. Nasiru, S., Mwita, P. N. and Ngesa, O. (2017). Exponentiated generalized transformed-transformer family of distributions. Journal of Statistical and Econometric Methods, **6**(4): 1-17.
14. Nassar, M. Kumar, D., Dey, S., Cordeiro, G. M. and Afify, A. Z. (2019). The Marshall-Olkin alpha power family of distributions with applications. Journal of Computational and Applied Mathematics, **351**: 41-53.
15. Rényi, A. (1961). On measures of entropy and information. In proceedings of the 4^{th} Berkeley Symposium on Mathematical Statistics and Probability. Pages 547-561, University of California Press.
16. Zubair, A. (2018). The Zubair-G family of distributions: properties and applications. Annals of Data Science. http://doi.org/10.1007/s40745-018-0169-9.

 DOI: http://dx.doi.org/10.18187/pjsor.v18i1.3096

Appendix

```
################## MOZNH PDF ########################
MOZNH_PDF<-function(x,alpha,beta,gamma,theta){
  A<-2*theta*alpha*gamma*beta*((1+gamma*x)^(beta-1))
  B<-exp(1-((1+gamma*x)^(beta)))
  num<-A*B*(1-B)*exp(alpha*((1-B)^2))
  deno<-(exp(alpha)-1)*((theta+(1-theta)*((exp(alpha)-1)^(-1))*(exp(alpha*((1-
B)^2))-1))^2)
  PDF<-num/deno
  return(PDF)
}
########################## Hazard of MOZNH #####################
MOZNH_H<-function(x,alpha,beta,gamma,theta){
  A<-2*alpha*gamma*beta*((1+gamma*x)^(beta-1))
  B<-exp(1-((1+gamma*x)^(beta)))
  num<-A*B*(1-B)*exp(alpha*((1-B)^2))
  deno<-((theta+(1-theta)*((exp(alpha)-1)^(-1))*(exp(alpha*((1-B)^2))-
1)))*(exp(alpha)-exp(alpha*((1-B)^2)))
  H<-num/deno
  return(H)
}
##################### MOZNH Quantile function ###################
quantile<-function(alpha,beta,gamma,theta,u){
  A<-((1/alpha)*log((1+u*(theta*exp(alpha)-1))/(1-u*(1-theta))))^(1/2)
  quant<-(1/gamma)*((1-log(1-A))^(1/beta)-1)
  return(quant)
}
  ######################### NEGATIVE LOG-LIKELIHOOD OF MOZNH
###########################
MOZNH_LL<-function(alpha,beta,gamma,theta){
  A<-2*theta*alpha*gamma*beta*((1+gamma*x)^(beta-1))
  B<-exp(1-((1+gamma*x)^(beta)))
  num<-A*B*(1-B)*exp(alpha*((1-B)^2))
  deno<-(exp(alpha)-1)*((theta+(1-theta)*((exp(alpha)-1)^(-1))*(exp(alpha*((1-
B)^2))-1))^2)
  PDF<-num/deno
  LL<--sum(log(PDF))
  return(LL)
}
########################### MOZW PDF ##################
MOZW_PDF<-function(x,alpha,beta,gamma,theta){
  A<-2*theta*alpha*gamma*beta*(x^(beta-1))*exp(-gamma*(x^beta))
  B<-1-exp(-gamma*(x^beta))
  C<-exp(alpha*((B^2)))
  num<-A*B*C
  deno<-(exp(alpha)-1)*((theta+(1-theta)*((exp(alpha)-1)^(-1))*(C-1))^2)
  PDF<-num/deno
  return(PDF)
}
########################## MOZW Hazard ###############################
MOZW_H<-function(x,alpha,beta,gamma,theta){
  A<-2*alpha*gamma*beta*(x^(beta-1))*exp(-gamma*(x^beta))
  B<-1-exp(-gamma*(x^beta))
  C<-exp(alpha*((B^2)))
  num<-A*B*C
  deno<-((theta+(1-theta)*((exp(alpha)-1)^(-1))*(C-1)))*(exp(alpha)-C)
```

 DOI: http://dx.doi.org/10.18187/pjsor.v18i1.3096

```
####################### MOZW Quantile ###############################
quantile<-function(alpha,beta,gamma,theta,u){
 A<-((1/alpha)*log((1+u*(theta*exp(alpha)-1))/(1+u*(1-theta))) )^(1/2)
 quant<-((-1/gamma)*log(1-A))^(1/beta)
 return(quant)
}
############################# NEGATIVE LOG-LIKELIHOOD OF MOZW
#########################
MOZW_LL<-function(alpha,beta,gamma,theta){
  A<-2*theta*alpha*gamma*beta*(x^(beta-1))*exp(-gamma*(x^beta))
  B<-1-exp(-gamma*(x^beta))
  C<-exp(alpha*((B^2)))
  num<-A*B*C
  deno<-(exp(alpha)-1)*((theta+(1-theta)*((exp(alpha)-1)^(-1))*(C-1))^2)
  PDF<-num/deno
  LL<--sum(log(PDF))
  return(LL)
}

################ APPLICATION OF MOZNH ###################
Install.packages("GenSA")
Install.packages("bbmle")
library(GenSA)
library(bbmle)
########################### Generating Initial Values ############
fit.sa2<-function(data,density){
  minusllike<-function(x)-sum(log(density(data,x[1],x[2],x[3],x[3])))
  lower<-c(0.001,0.001,0.001,0.001)
  upper<-c(1000,1000,1000,1000)
  out<-
GenSA(lower=lower,upper=upper,fn=minusllike,control=list(verbose=TRUE,max.tim
e=2))
  return(out[c("value","par","counts")])
}

############  Fitting the Data Using MOZNH ###########################
x<-c(0.08, 2.09, 3.48, 4.87, 6.94, 8.66, 13.11, 23.63, 0.20, 2.23, 3.52,
4.98, 6.97, 9.02, 13.29, 0.40, 2.26, 3.57, 5.06, 7.09, 9.22, 13.80, 25.74,
0.50, 2.46, 3.64, 5.09, 7.26, 9.47, 14.24, 25.82, 0.51, 2.54, 3.70, 5.17,
7.28, 9.74, 14.76, 26.31, 0.81, 2.62, 3.82, 5.32, 7.32, 10.06, 14.77, 32.15,
2.64, 3.88, 5.32, 7.39, 10.34, 14.83, 34.26, 0.90, 2.69, 4.18, 5.34, 7.59,
10.66, 15.96, 36.66, 1.05, 2.69, 4.23, 5.41, 7.62, 10.75, 16.62, 43.01, 1.19,
2.75, 4.26, 5.41, 7.63, 17.12, 46.12, 1.26, 2.83, 4.33, 5.49, 7.66, 11.25,
17.14, 79.05, 1.35, 2.87, 5.62, 7.87, 11.64, 17.36, 1.40, 3.02, 4.34, 5.71,
7.93, 11.79, 18.10, 1.46, 4.40, 5.85, 8.26, 11.98, 19.13, 1.76, 3.25, 4.50,
6.25, 8.37, 12.02, 2.02, 3.31, 4.51, 6.54, 8.53, 12.03, 20.28, 2.02, 3.36,
```