Journal of King Saud University – Science 31 (2019) 684–691

Contents lists available at ScienceDirect

Journal of King Saud University – Science

journal homepage: www.sciencedirect.com

Original article

The Marshall-Olkin Kappa distribution: Properties and applications

Maria Javed *, Tahir Nawaz, Muhammad Irfan

Department of Statistics, Government College University Faisalabad, Pakistan

ARTICLE INFO

Article history:
Received 14 December 2017
Accepted 9 January 2018
Available online 13 January 2018

Keywords:
Kappa distribution
Marshall-Olkin distribution
Moments
Reliability
Entropy
Estimation

ABSTRACT

This study proposes a new competitive model, the Marshall-Olkin Kappa distribution and presents its various properties. These properties include the derivation of probability density function, quantiles, rth moment, mode, mean deviation and survival function. Results for the various inequality indices are obtained. Expression regarding order statistics is given. The unknown parameters of the proposed distribution are estimated using the maximum likelihood estimation method. Empirical illustration is provided using two real life data sets.

1. Introduction

Kappa distribution is a good choice for model fitting among its competitors in presence of extreme values. The study of extreme events like floods, cyclone and heavy rains is necessary in the planning of water relevant setups, cultivation of crops in agriculture, climatic conditions, overseeing environmental changes and floods basic control systems.

Mielke (1973) and Mielke and Johnson (1973) introduced a class of asymmetric positively skewed distribution, used for explaining and examining rainfall data and weather modifications which obtained attention from the hydrologist. This class of distribution was named as the three parameter kappa distribution. Conventionally, the log normal and the gamma distributions are fitted to precipitation data but these distributions have their own limitations due to non-existence of closed forms of cdfs and quantile functions. Closed algebraic expressions can be analyzed with the help of the class of kappa distribution.A random variable X has the kappa three (KAP-III) distribution with scale parameter $\beta > 0$ and shape parameters $\alpha \& \theta > 0$, if its cumulative distribution function (cdf) is given by

$$F(x) = \left[\frac{\left(\frac{x}{\beta}\right)^{\alpha\theta}}{\alpha + \left(\frac{x}{\beta}\right)^{\alpha\theta}} \right]^{\frac{1}{\alpha}} \tag{1.1}$$

The corresponding probability density function is

$$f(x) = \frac{\alpha\theta}{\beta} \left(\frac{x}{\beta}\right)^{\theta-1} \left[\alpha + \left(\frac{x}{\beta}\right)^{\alpha\theta} \right]^{-\frac{(\alpha+1)}{\alpha}} \tag{1.2}$$

Hussain (2015) proposed three extended forms of kappa distribution namely Exponentiated generalized kappa, Kumaraswamy generalized kappa and McDonald generalized kappa distribution. He explored different statistical properties, survival properties, inequality indices and entropies of new extended forms. He used different methods to estimate the unknown parameters of new forms. He also checked the efficiency of the proposed models of kappa distribution with some real life data sets.

The study in hands is extended as follows. In Section 2, we give a brief introduction of parameter induction in probability distributions and one of the generated family of distribution i.e. Marshall-Olkin (MO) distribution. In Section 3, we introduced a new generalization of the kappa distribution, namely, the Marshall-Olkin Kappa (MOK) distribution. We derived the statistical properties like quantile function, median, mode, moments, moment generating function (mgf), characteristic function (cf), mean deviation from mean and from median and reliability properties i.e. survival function, hazard rate function (hrf), reversed hazard rate and mean residual life function of MOK distribution. We give expressions of inequality measures including Lorenz curve, Bonferroni curve,

* Corresponding author.
E-mail address: maria_mavi786@yahoo.com (M. Javed).

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.jksus.2018.01.001
1018-3647/

Zenga index, Atkinson index, Pietra index and generalized entropy. Most famous and useful method i.e. method of maximum likelihood is used to estimate the unknown parameters of MOK distribution. In Section 4, proposed MOK distribution is applied on two real data sets. Different model selection criterions are used to compare proposed and existing distributions on the same data sets. This compatibility is also shown with the help of graphical representation.

2. Parameter induction in continuous univariate distribution

There is an accelerating trend of inducting shape parameter(s) to the basic distribution in advancing generated families. Doubtlessly the popularity and the utilization of incomplete functions in generated distributions have attracted the thought of statisticians, mathematicians, researchers, engineers, financial specialists, demographers and other connected scientists. Diverse explanations for this thought may be

- The computational and diagnostic facilities available through programming softwares like R (packages), ox5, Python, Matlab, Maple and Mathematica. With the assistance of these softwares scientists can easily deal with the issues in calculating incomplete functions.
- To analyze the tail properties of distributions more extensively one can introduce additional shape parameters to the underlying distribution.
- The goodness of fit can be enhanced with addition of shape parameters.
- Distributions with additional shape parameters performs better in case of skewed data as compared to the conventional distributions (Pescim et al. 2012).

2.1. General theory of Marshall-Olkin distribution

Marshall and Olkin (1997) presented an elastic semi-parametric family of distributions. They defined a new survival function $\bar{F}^{MO}(t)$ by inducting an extra parameter $\delta > 0$ where δ was a tilt parameter. They took δ in the sense of the performance of the hrf of $\bar{F}^{MO}(t)$ and the cdf of the baseline distribution. δ was reintroduced by Nanda and Das (2012) due to the reason that the hrf of the newly introduced distribution is lifted beneath $(\delta 1)$ or overhead $(0 < \delta 1)$ the hrf of the original distribution.

The survival function $\bar{F}^{MO}(t)$ of MO distribution is

$$\bar{F}^{MO}(t) = \frac{\delta \bar{G}(t)}{1 - \bar{\delta}\bar{G}(t)} = \frac{\delta \bar{G}(t)}{G(t) + \delta \bar{G}(t)} \text{or} \frac{\delta[1 - G(t)]}{\delta + \bar{\delta}G(t)}$$

where $g(t), G(t)$ and $\bar{G}(t)$ are the pdf, cdf and the survival function of the existing distribution respectively with condition $-\infty < t < \infty$, $\delta > 0$ and $\bar{\delta} = 1 - \delta$.

The associated cdf and pdf of MO distribution are

$$F^{MO}(t) = \frac{G(t)}{1 - \bar{\delta}\bar{G}(t)} = \frac{G(t)}{G(t) + \delta \bar{G}(t)} \text{or} \frac{G(t)}{\delta + \bar{\delta}G(t)} \tag{2.1}$$

and

$$f^{MO}(t) = \frac{\delta g(t)}{[1 - \bar{\delta}\bar{G}(t)]^2} \text{or} \frac{\delta g(t)}{[\delta + \bar{\delta}G(t)]^2} \tag{2.2}$$

Note that for $\delta = 1, \bar{F}^{MO}(t) = \bar{G}(t) \text{and} F^{MO}(t) = G(t)$.

Marshall-Olkin generalizations of different distributions exist in the literature. Some of them are mentioned in our manuscript. Cordeiro and Lemonte (2013) studied the mathematical properties and applications of the Marshall-Olkin extended (MOE) weibull distribution. Other examples include MOE pareto distribution (Alice and Jose, 2003), MOE gamma distribution (Ristić et al., 2007), MOE lomax distribution (Ghitany et al., 2007a), MOE weibull distribution and its application to censored data (Ghitany et al., 2007b), MOE normal distribution (Garcia et al., 2010), MOE lindley distribution (Ghitany et al., 2012), MOE fréchet distribution (Krishna et al., 2013), and MOE birnbaum–saunders distribution (Lemonte, 2013). General properties of the MOE family of distributions were studied recently by Barreto-Souza et al. (2013) and Cordeiro et al. (2014).

3. Marshall-Olkin Kappa (MOK) distribution

Consider the cdf and pdf of kappa distribution from (1.1) and (1.2) then applying the (2.1) and (2.2), we have the Marshall-Olkin Kappa distribution with cdf and pdf given in (3.1) and (3.2)

$$F_{MOK}(x) = \left[\delta\left\{\frac{\alpha}{\left(\frac{x}{\beta}\right)^{\alpha\theta}} + 1\right\}^{\frac{1}{\alpha}} + (1 - \delta)\right]^{-1} \tag{3.1}$$

$$f_{MOK}(x) = \frac{\delta \frac{\alpha\theta}{\beta}\left(\frac{x}{\beta}\right)^{\theta-1}\left(\alpha + \left(\frac{x}{\beta}\right)^{\alpha\theta}\right)^{-\frac{(\alpha+1)}{\alpha}}}{\left[\delta + (1 - \delta)\left(\frac{\left(\frac{x}{\beta}\right)^{\alpha\theta}}{\alpha + \left(\frac{x}{\beta}\right)^{\alpha\theta}}\right)^{\frac{1}{\alpha}}\right]^2} \tag{3.2}$$

with $\alpha, \beta, \theta, \delta > 0$, where β is scale parameter and $\alpha, \theta \text{and} \delta$ are shape parameters. So, we can refer it as MOK $(\alpha, \beta, \theta, \delta)$.

An important use fulid entity is

$$(x + a)^{-n} = \sum_{k=0}^{\infty}(-1)^k \binom{n + k - 1}{k} x^k a^{-n-k} \tag{3.3}$$

(3.2) can be more simplified with the help of identity in (3.3) in this form

$$f_{MOK}(x) = \sum_{k=0}^{\infty} W_k \left(\frac{\alpha\theta}{\beta}\right)\left(\frac{x}{\beta}\right)^{\theta-1}\left[\alpha + \left(\frac{x}{\beta}\right)^{\alpha\theta}\right]^{-\frac{\alpha+1}{\alpha}}\left[\frac{\left(\frac{x}{\beta}\right)^{\alpha\theta}}{\alpha + \left(\frac{x}{\beta}\right)^{\alpha\theta}}\right]^{\frac{k}{\alpha}} \tag{3.4}$$

3.1. Sub models of MOK distribution

(i) In case $\delta = 1$ in (3.2), the MOK distribution reduces to KAP-III distribution.
(ii) If $\theta = \delta = 1$ in (3.2), the MOK distribution reduces to KAP-II distribution.

Graphical representation of pdf and cdf for various values of parameters chosen arbitrary are provided in Figs. 3.1 and 3.2 respectively.

Some important findings from graphs in Fig. 3.1 are

- The density function of MOK distribution tends to normal distribution as the value of δ increases, while the values of α, β and θ are kept constant. Similar pattern is seen with increase in value of β for fixed values of α, θ, δ.
- α being the shape parameter changes the shape of density function for fixed values of β, θ, δ. Increase in α makes the density function more leptokurtic. Similar behavior is found for increase in θ, by fixing the values of $\alpha, \beta \text{and} \delta$.

From Fig. 3.2, graphs of the cdf of MOK distribution satisfy the following properties.

- $F_{MOK}(x)$ goes to 0 as x gets smaller

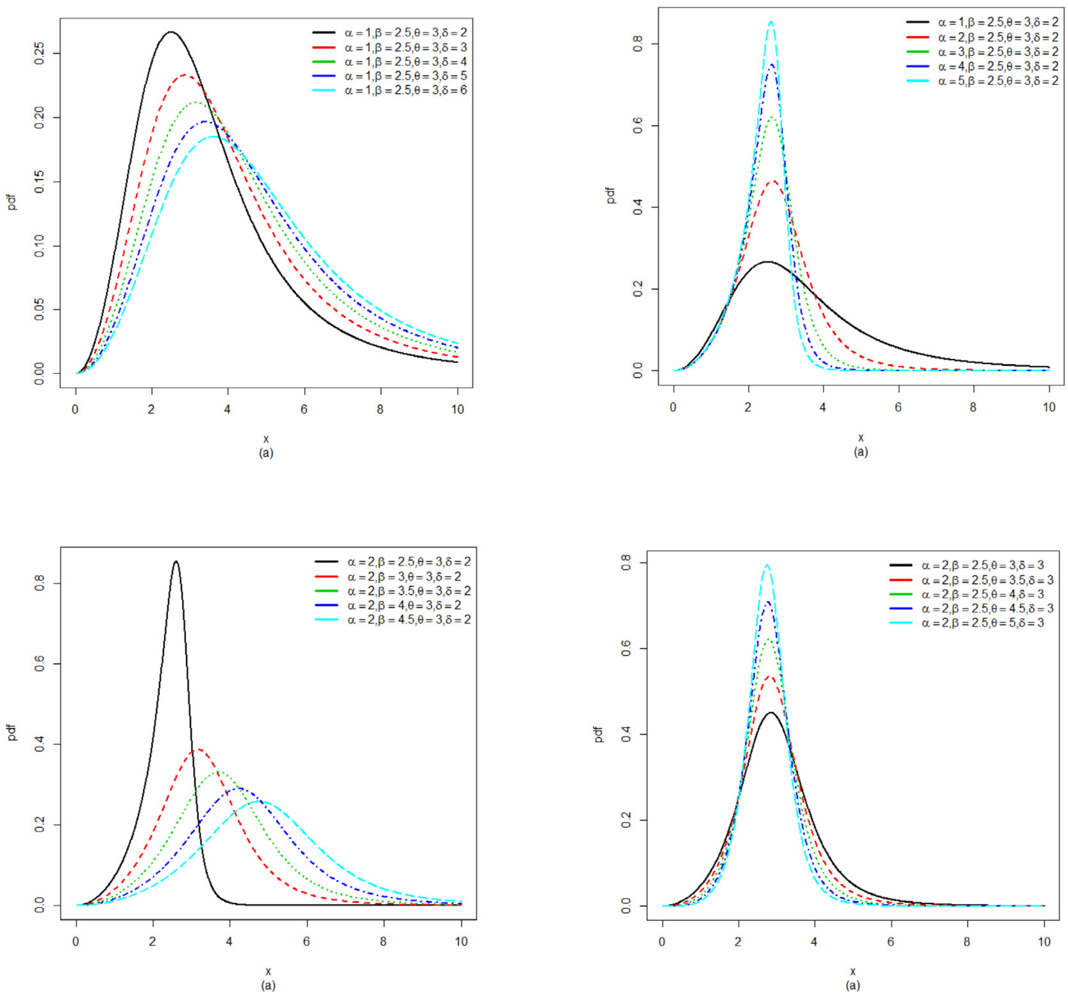

Fig. 3.1. Graphs of the pdf of MOK distribution.

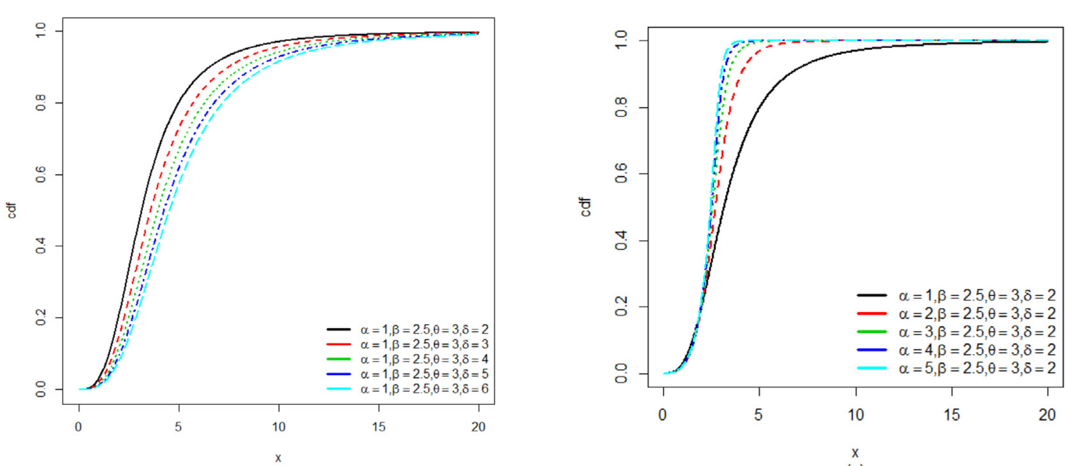

Fig. 3.2. Graphs of the cdf of MOK distribution.

- Converselylim$_{x\to\infty}F_{MOK}(x) = 1$
- $F_{MOK}(x)$ is non-decreasing.

Theorem 3.1. *The mode of the MOK distribution is the solution of the equation*

$$\left[\delta + (1-\delta)y^{\frac{1}{\alpha}}\right]\left[(\theta-1)(\alpha+y)y^{(1-\theta)} - \theta(\alpha+1)y^{\frac{1-\theta}{\alpha\theta}}\right] - 2\theta(1-\delta) = 0 \tag{3.5}$$

with $y = \left(\frac{x}{\beta}\right)^{\alpha\theta}$

Proof. Taking first derivative of the logarithm of (3.2), we have

$$\frac{\partial}{\partial x}\log f(x) = \frac{\theta-1}{x} - \left(\frac{\alpha+1}{\alpha}\right)\frac{\alpha\theta\left(\frac{x}{\beta}\right)^{\alpha\theta-1}\frac{1}{\beta}}{\alpha+\left(\frac{x}{\beta}\right)^{\alpha\theta}} - 2\frac{(1-\delta)\frac{1}{\alpha}\left(\frac{\alpha\beta^{\alpha\theta}}{x^{\alpha\theta}}+1\right)^{-\frac{1}{\alpha}-1}\frac{\alpha\beta^{\alpha\theta}\alpha\theta}{x^{\alpha\theta+1}}}{\delta+(1-\delta)\left(\frac{\left(\frac{x}{\beta}\right)^{\alpha\theta}}{\alpha+\left(\frac{x}{\beta}\right)^{\alpha\theta}}\right)^{\frac{1}{\alpha}}}$$

Equating to zero and simplifying, we get

$$\left[\delta+(1-\delta)y^{\frac{1}{\alpha}}\right]\left[(\theta-1)(\alpha+y)y^{(1-\theta)}-\theta(\alpha+1)y^{\frac{1-\theta}{\alpha\theta}}\right]-2\theta(1-\delta) = 0 \text{ with } y=\left(\frac{x}{\beta}\right)^{\alpha\theta}$$

Hence, the solution of above equation provides the mode value(s) of the MOK distribution.

It is important to mention that for $\delta = 1$ the above equation provides the mode of the KAP-III distribution. □

3.2. Quantiles and moments

The qth quantile of the MOK distribution is

$$x_q = \left[\frac{\alpha\beta^{\alpha\theta}}{\left[\frac{1}{\delta}(p^{-1}-1+\delta)\right]^{\alpha}-1}\right]^{\frac{1}{\alpha\theta}} \quad \text{where } 0<p<1 \tag{3.6}$$

The median, first and third quartiles of MOK distribution are

$$\begin{aligned}\text{Median }(x) &= \beta\alpha^{\frac{1}{\alpha\theta}}\left[\left(1+\frac{1}{\delta}\right)^{\alpha}-1\right]^{-\frac{1}{\alpha\theta}}\\ Q_1 &= \beta\alpha^{\frac{1}{\alpha\theta}}\left[\left(1+\frac{0.33}{\delta}\right)^{\alpha}-1\right]^{-\frac{1}{\alpha\theta}}\\ Q_3 &= \beta\alpha^{\frac{1}{\alpha\theta}}\left[\left(1+\frac{3}{\delta}\right)^{\alpha}-1\right]^{-\frac{1}{\alpha\theta}}\end{aligned} \tag{3.7}$$

The rth moment of MOK random variable X with pdf given in (3.4) is

$$\mu'_r = \sum_{k=0}^{\infty} W_k\beta^r\alpha^{\frac{r}{\alpha\theta}-1}\int_0^1 U^{\frac{k+1}{\alpha}+\frac{r}{\alpha\theta}-1}[1-U]^{1-\frac{r}{\alpha\theta}-1}dU$$

Using beta function,

$$\mu'_r = \sum_{k=0}^{\infty} W_k\beta^r\alpha^{\frac{r}{\alpha\theta}-1}B\left(\frac{k+1}{\alpha}+\frac{r}{\alpha\theta},1-\frac{r}{\alpha\theta}\right);\ r=1,2,3,4 \tag{3.8}$$

where $U = 1-\left(1+\frac{Z}{\alpha}\right)^{-1}, Z=\left(\frac{x}{\beta}\right)^{\alpha\theta}$ and $B(a,b)=\int_0^1 Y^{a-1}[1-Y]^{b-1}dY$ is the beta function. From (3.8), the first four moments, mean, variance and standard deviation can be obtained.

MOK distribution has the following moment generating function and characteristic function

$$\begin{aligned}M(t) = E(e^{tx}) &= \sum_{r=0}^{\infty}\frac{t^r}{r!}\mu'_r\\ &= \sum_{r,k=0}^{\infty}\frac{t^r}{r!}W_k\beta^r\alpha^{\frac{r}{\alpha\theta}-1}B\left(\frac{k+1}{\alpha}+\frac{r}{\alpha\theta},1-\frac{r}{\alpha\theta}\right)\end{aligned} \tag{3.9}$$

and

$$\begin{aligned}\varphi(t) = E(e^{itx}) &= \sum_{r=0}^{\infty}\frac{(it)^r}{r!}\mu'_r\\ &= \sum_{r,k=0}^{\infty}\frac{(it)^r}{r!}W_k\beta^r\alpha^{\frac{r}{\alpha\theta}-1}B\left(\frac{k+1}{\alpha}+\frac{r}{\alpha\theta},1-\frac{r}{\alpha\theta}\right)\end{aligned} \tag{3.10}$$

3.3. Mean deviation

The amount of variability in a distribution can be measured up to some extent with the help of totality of deviations about the mean and about the median.

Lemma. If $X \sim MOK(\alpha,\beta,\theta,\delta)$ then $\int_0^x xf(x)dx$ is given as

$$\int_0^x xf(x)dx = \beta\alpha^{\frac{1}{\alpha\theta}-1}\sum_{k=0}^{\infty}W_kB_U\left(\frac{k+1}{\alpha}+\frac{1}{\alpha\theta},1-\frac{1}{\alpha\theta}\right) \tag{3.11}$$

where $B_U(a,b)=\int_0^U U^{a-1}[1-U]^{b-1}dU$ is the incomplete beta function.

Proof.

$$\int_0^x xf(x)dx = \sum_{k=0}^{\infty}W_k\frac{\alpha\theta}{\beta} \times \int_0^x x\left(\frac{x}{\beta}\right)^{\theta-1}\left[\alpha+\left(\frac{x}{\beta}\right)^{\alpha\theta}\right]^{-\frac{(\alpha+1)}{\alpha}}\left[\frac{\left(\frac{x}{\beta}\right)^{\alpha\theta}}{\alpha+\left(\frac{x}{\beta}\right)^{\alpha\theta}}\right]^{\frac{k}{\alpha}}dx$$

Substituting $Z=\left(\frac{x}{\beta}\right)^{\alpha\theta}$ and simplifying, we get

$$\int_0^x xf(x)dx = \sum_{k=0}^{\infty}W_k\frac{\beta}{\alpha^{\frac{\alpha+1+k}{\alpha}}} \times \int_0^Z Z^{\frac{\theta+k\theta+1}{\alpha\theta}-1}\left[1-\left\{1-\left(1+\frac{Z}{\alpha}\right)^{-1}\right\}\right]^{\frac{\alpha+1+k}{\alpha}}dZ$$

$$\int_0^x xf(x)dx = \beta\alpha^{\frac{1}{\alpha\theta}-1}\sum_{k=0}^{\infty}W_kB_U\left(\frac{k+1}{\alpha}+\frac{1}{\alpha\theta},1-\frac{1}{\alpha\theta}\right) \quad \text{where } U = 1-\left(1+\frac{Z}{\alpha}\right)^{-1}$$

The Mean deviation about mean and Mean deviation about median can be obtained by

$$\delta_1(x) = 2\left[\mu F(\mu)-\int_0^{\mu}xf(x)dx\right] \text{ and } \delta_2(x) = E(x)+2MF(M)-M-2\int_0^M xf(x)dx$$

respectively, where μ is the mean of MOK obtained from (3.8) by putting $r=1$, M is the median can be taken from (3.7) and use above lemma to solve $\int_0^{\mu}xf(x)dx$ and $\int_0^M xf(x)dx$

Hence, we get

$$\delta_1(x) = 2\left[\mu F(\mu)-\beta\alpha^{\frac{1}{\alpha\theta}-1}\sum_{k=0}^{\infty}W_kB_U\left(\frac{k+1}{\alpha}+\frac{1}{\alpha\theta},1-\frac{1}{\alpha\theta}\right)\right] \tag{3.12}$$

and

$$\delta_2(x) = E(x)+2MF(M)-M - 2\beta\alpha^{\frac{1}{\alpha\theta}-1}\sum_{k=0}^{\infty}W_kB_{U_1}\left(\frac{k+1}{\alpha}+\frac{1}{\alpha\theta},1-\frac{1}{\alpha\theta}\right) \tag{3.13}$$

□

3.4. Reliability properties

Let t be a life time random variable follows MOK $(\alpha,\beta,\theta,\delta)$ distribution. The hazard rate function of MOK distribution, expressed below, is upside-down shaped. Fig. 3.3 shows the graphical representation of the hrf.

$$h(t) = \frac{\delta \frac{\alpha\theta}{\beta}\left(\frac{t}{\beta}\right)^{\theta-1}\left(\alpha + \left(\frac{t}{\beta}\right)^{\alpha\theta}\right)^{-\frac{(\alpha+1)}{\alpha}}}{\left[1 - \left[\delta\left\{\frac{\alpha}{\left(\frac{t}{\beta}\right)^{\alpha\theta}} + 1\right\}^{\frac{1}{\alpha}} + (1-\delta)\right]^{-1}\right]\left[\delta + (1-\delta)\left(\frac{\left(\frac{t}{\beta}\right)^{\alpha\theta}}{\alpha + \left(\frac{t}{\beta}\right)^{\alpha\theta}}\right)^{\frac{1}{\alpha}}\right]^2}$$

The survival function *S(t)*, cumulative hazard rate function $H(t)$ and reversed hazard rate function $r(t)$ are, respectively, specified as

$$S(t) = \left[1 + \left[\delta\left(\frac{\alpha}{\left(\frac{t}{\beta}\right)^{\alpha\theta}} + 1\right)^{\frac{1}{\alpha}} - \delta\right]^{-1}\right]^{-1}$$

$$H(t) = \int_0^t h(t)dt = -lnS(t)$$

$$= -ln\left[1 - \left[\delta\left\{\frac{\alpha}{\left(\frac{t}{\beta}\right)^{\alpha\theta}} + 1\right\}^{\frac{1}{\alpha}} + (1-\delta)\right]^{-1}\right]$$

and

$$r(t) = \frac{f(t)}{F(t)}$$

$$= \frac{\delta \frac{\alpha\theta}{\beta}\left(\frac{t}{\beta}\right)^{\theta-1}\left(\alpha + \left(\frac{t}{\beta}\right)^{\alpha\theta}\right)^{-\frac{(\alpha+1)}{\alpha}}}{\left[\delta + (1-\delta)\left(\frac{\left(\frac{t}{\beta}\right)^{\alpha\theta}}{\alpha + \left(\frac{t}{\beta}\right)^{\alpha\theta}}\right)^{\frac{1}{\alpha}}\right]^2}\left[\delta\left\{\frac{\alpha}{\left(\frac{t}{\beta}\right)^{\alpha\theta}} + 1\right\}^{\frac{1}{\alpha}} + (1-\delta)\right]$$

3.5. Mean residual life function

In many fields like biomedical science, insurance and industrial reliability, mean residual life function is an imperative consideration. It can be obtained through the expression below

$$m(t) = \frac{1}{S(t)}\left\{E(t) - \int_0^t tf(t)dt\right\} - t$$

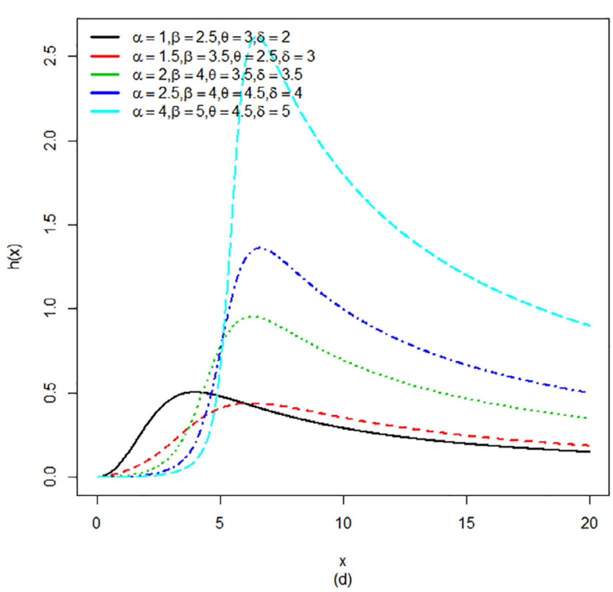

Fig. 3.3. Graph of the hrf of MOK distribution.

Mean residual life function of MOK distribution is given as

$$m(t) = \frac{\sum_{k=0}^{\infty} W_k \beta\alpha^{\frac{1}{\alpha\theta}-1}\left\{B\left(\frac{k+1}{\alpha} + \frac{1}{\alpha\theta}, 1 - \frac{1}{\alpha\theta}\right) - B_U\left(\frac{k+1}{\alpha} + \frac{1}{\alpha\theta}, 1 - \frac{1}{\alpha\theta}\right)\right\}}{1 - \left[\delta\left\{\frac{\alpha}{\left(\frac{t}{\beta}\right)^{\alpha\theta}} + 1\right\}^{\frac{1}{\alpha}} + (1-\delta)\right]^{-1}} - t$$

3.6. Inequality measures

Inequality measures play key role in many fields like economics to study income and poverty, demography, insurance and medicine. Some are discussed in this study.

3.6.1 An American economist Lorenz (1905) developed a graphical diagram of wealth distribution called Lorenz curve. Lorenz index is defined as

$$L(p) = \frac{1}{\mu}\int_0^x xf(x)dx$$

The Lorenz curve is the plot of Lorenz index *L(p)* verses *x*, given below is the Lorenz index for MOK distribution.

$$L(p) = \frac{\beta\alpha^{\frac{1}{\alpha\theta}-1}\sum_{k=0}^{\infty} W_k B_U\left(\frac{k+1}{\alpha} + \frac{1}{\alpha\theta}, 1 - \frac{1}{\alpha\theta}\right)}{\sum_{k=0}^{\infty} W_k \beta\alpha^{\frac{1}{\alpha\theta}-1} B\left(\frac{k+1}{\alpha} + \frac{1}{\alpha\theta}, 1 - \frac{1}{\alpha\theta}\right)}$$

On the diagram perfect equality of wealth distribution is depicted by a straight diagonal line and a line lies under it shows the true wealth distribution. The difference between two above stated lines is actually the inequality of wealth distribution.

3.6.2 A measure of income inequality was projected by Bonferroni (1930), founded on partial means, that is needed when the main source of income inequality is the occurrence of units whose income is much beneath those of others. The Bonferroni index can be determined through the relation

$$BC(p) = \frac{L(p)}{F(x)}$$

The Bonferroni curve is the plot of Bonferroni index *BC(p)* verses *x*, and this index for MOK distribution is as

$$BC(p) = \left[\delta\left\{\frac{\alpha}{\left(\frac{x}{\beta}\right)^{\alpha\theta}} + 1\right\}^{\frac{1}{\alpha}} + (1-\delta)\right]\left[\frac{\beta\alpha^{\frac{1}{\alpha\theta}-1}\sum_{k=0}^{\infty} W_k B_U\left(\frac{k+1}{\alpha} + \frac{1}{\alpha\theta}, 1 - \frac{1}{\alpha\theta}\right)}{\sum_{k=0}^{\infty} W_k \beta\alpha^{\frac{1}{\alpha\theta}-1} B\left(\frac{k+1}{\alpha} + \frac{1}{\alpha\theta}, 1 - \frac{1}{\alpha\theta}\right)}\right]$$

3.6.3 The Zenga index denoted by Z was suggested by Zenga (1984). It measures the disparity between the poorest $p * 100\%$ of the population and the wealthier outstanding $(1-p) * 100\%$ part of the population by looking at the mean salaries of these two disjoint and comprehensive subpopulations. Zenga index for MOK distribution is as follows

$$Z = 1 - \left[\frac{\left[\beta\alpha^{\frac{1}{\alpha\theta}-1}\sum_{k=0}^{\infty} W_k B_U\left(\frac{k+1}{\alpha} + \frac{1}{\alpha\theta}, 1 - \frac{1}{\alpha\theta}\right)\right]\left[\delta\left\{\frac{\alpha}{\left(\frac{x}{\beta}\right)^{\alpha\theta}} + 1\right\}^{\frac{1}{\alpha}} + (1-\delta)\right]}{\left[\mu - \beta\alpha^{\frac{1}{\alpha\theta}-1}\sum_{k=0}^{\infty} W_k B_U\left(\frac{k+1}{\alpha} + \frac{1}{\alpha\theta}, 1 - \frac{1}{\alpha\theta}\right)\right]\left[1 - \left[\delta\left\{\frac{\alpha}{\left(\frac{x}{\beta}\right)^{\alpha\theta}} + 1\right\}^{\frac{1}{\alpha}} + (1-\delta)\right]^{-1}\right]^{-1}}\right]$$

3.6.4 Atkinson (1970) proposed an index ranges from 0 to 1 where 0 means an equal wealth distribution. Mathematically Atkinson index can be defined as

$$A_F(\theta', \epsilon) = 1 - \frac{1}{\mu}\left\{\int_0^{\infty} x^{1-\epsilon} dF(x)\right\}^{\frac{1}{1-\epsilon}}$$

Atkinson index for MOK distribution is given by

$$A_F(\theta',\epsilon) = 1 - \frac{\left\{\sum_{k=0}^{\infty} W_k \beta^{1-\epsilon} \alpha^{\frac{1-\epsilon}{\alpha\theta}-1} B\left(\frac{k+1}{\alpha} + \frac{1-\epsilon}{\alpha\theta}, 1 - \frac{1-\epsilon}{\alpha\theta}\right)\right\}^{\frac{1}{1-\epsilon}}}{\beta\alpha^{\frac{1}{\alpha\theta}-1}\sum_{k=0}^{\infty} W_k B\left(\frac{k+1}{\alpha} + \frac{1}{\alpha\theta}, 1 - \frac{1}{\alpha\theta}\right)}$$

3.6.5 Pietra (1915) offered an index, known as Schutz index or half of the relative mean deviation. Pietra index is defined as

$$P_X = \frac{1}{2\mu}\int_0^{\infty} |x-\mu| dF(x) = \frac{MD_{\bar{x}}}{2\mu}$$

Expression for the Pietra index P_X for MOK distribution is

$$P_X = \frac{\left[\mu F(\mu) - \beta\alpha^{\frac{1}{\alpha\theta}-1}\sum_{k=0}^{\infty} W_k B_U\left(\frac{k+1}{\alpha} + \frac{1}{\alpha\theta}, 1 - \frac{1}{\alpha\theta}\right)\right]}{\sum_{k=0}^{\infty} W_k \beta\alpha^{\frac{1}{\alpha\theta}-1} B\left(\frac{k+1}{\alpha} + \frac{1}{\alpha\theta}, 1 - \frac{1}{\alpha\theta}\right)}$$

3.7. Generalized entropy (GE)

Variation of the uncertainty can be measured for a random variable through entropy. The generalized entropy (GE) index, suggested by Cowell (1980) and Shorrocks (1980) is

$$GE_F(\theta',\omega) = \frac{1}{\omega(\omega-1)}\left\{\frac{\mu'_\omega}{\mu^\omega} - 1\right\}$$

where μ'_ω is the ωth moment about origin.

The GE for MOK distribution can be obtained as below

$$GE_F(\theta',\omega) = \frac{1}{\omega(\omega-1)}\left\{\frac{\sum_{k=0}^{\infty} W_k \beta^{\omega}\alpha^{\frac{\omega}{\alpha\theta}-1} B\left(\frac{k+1}{\alpha} + \frac{\omega}{\alpha\theta}, 1 - \frac{\omega}{\alpha\theta}\right)}{\left(\sum_{k=0}^{\infty} W_k \beta\alpha^{\frac{1}{\alpha\theta}-1} B\left(\frac{k+1}{\alpha} + \frac{1}{\alpha\theta}, 1 - \frac{1}{\alpha\theta}\right)\right)^{\omega}} - 1\right\}$$

3.8. Order statistics

In many areas of statistical theory and applications, order statistics have significant role. Consider $X_1, \ldots, X_n$ is a random sample from a population with the MOK $(\alpha,\beta,\theta,\delta)$ distribution. Let $X_{i,n}$ denote the *ith* order statistics then its pdf is given as

$$f_{i,n}(x) = \frac{n!}{(i-1)!(n-i)!} f(x) F(x)^{i-1}\{1 - F(x)\}^{n-i}$$

Inserting pdf and cdf of Kappa distribution and simplifying we get

$$f_{i,n}(x) = \frac{n!}{(i-1)!(n-i)!}\sum_{j=0}^{n-i}\sum_{k=0}^{\infty}(-1)^j\binom{n-i}{j} W_k\left(\frac{\alpha\theta}{\beta}\right)\left(\frac{x}{\beta}\right)^{\theta-1}$$
$$\left[\alpha + \left(\frac{x}{\beta}\right)^{\alpha\theta}\right]^{-\frac{\alpha+1}{\alpha}} \times \left[\frac{\left(\frac{x}{\beta}\right)^{\alpha\theta}}{\alpha + \left(\frac{x}{\beta}\right)^{\alpha\theta}}\right]^{\frac{k}{\alpha}}\left[\delta\left\{\frac{\alpha}{\left(\frac{x}{\beta}\right)^{\alpha\theta}} + 1\right\}^{\frac{1}{\alpha}} + (1-\delta)\right]^{(1-j-i)}$$

3.9. Estimation of parameters

Section given below considers the famous and most useful method for estimating the parameters α,β,θ and δ of MOK distribution. Let $X \sim \mathrm{MOK}(\alpha,\beta,\theta,\delta)$ distribution. We estimate these parameters with the help of method of maximum likelihood. Consider the log-likelihood function $\ell = L(\Theta)$

$$\ell = nlog(\delta\alpha\theta) - nlog\beta + (\theta-1)\sum_{i=1}^{n} log\left(\frac{x_i}{\beta}\right)$$
$$-\left(\frac{\alpha+1}{\alpha}\right)\sum_{i=1}^{n} log\left[\alpha + \left(\frac{x_i}{\beta}\right)^{\alpha\theta}\right]$$
$$-2\sum_{i=1}^{n} log\left\{\delta + (1-\delta)\left[\alpha\left(\frac{x_i}{\beta}\right)^{-\alpha\theta} + 1\right]^{-\frac{1}{\alpha}}\right\}$$

Partially differentiating ℓ with respect to $\alpha, \beta, \theta and \delta$ and then equate to zero,

We have four equations as

$$\frac{\partial\ell}{\partial\alpha} = \frac{n}{\alpha} - \left(\frac{\alpha+1}{\alpha}\right)\sum_{i=1}^{n}\left[\left[\alpha + \left(\frac{x_i}{\beta}\right)^{\alpha\theta}\right]^{-1}\left[1 + \left(\frac{x_i}{\beta}\right)^{\alpha\theta}\right]\log\left(\frac{x_i}{\beta}\right)\right]$$
$$+\frac{1}{\alpha^2}\sum_{i=1}^{n} log\left[\alpha + \left(\frac{x_i}{\beta}\right)^{\alpha\theta}\right]$$
$$-2\sum_{i=1}^{n}\frac{(1-\delta)\left[\alpha + \left(\frac{x_i}{\beta}\right)^{\alpha\theta}\right] log\left[\alpha\left(\frac{x_i}{\beta}\right)^{-\alpha\theta} + 1\right] + \log\left(\frac{x_i}{\beta}\right)\alpha^2\theta - \alpha}{\left[\delta + (1-\delta)\left[\alpha\left(\frac{x_i}{\beta}\right)^{-\alpha\theta} + 1\right]^{-\frac{1}{\alpha}}\right]\left[\alpha\left(\frac{x_i}{\beta}\right)^{-\alpha\theta} + 1\right]^{\frac{1}{\alpha}}\alpha^2\left[\alpha + \left(\frac{x_i}{\beta}\right)^{\alpha\theta}\right]}$$
$$= 0$$

$$\frac{\partial\ell}{\partial\beta} = -\frac{n}{\beta} - \left(\frac{\theta-1}{\beta}\right) + \frac{(\alpha+1)\theta}{\beta}\sum_{i=1}^{n}\left[\left[\alpha + \left(\frac{x_i}{\beta}\right)^{\alpha\theta}\right]^{-1}\left(\frac{x_i}{\beta}\right)^{\alpha\theta}\right]$$
$$+\frac{2(1-\delta)\theta}{\beta}\sum_{i=1}^{n}\left[\left\{\delta + (1-\delta)\left[\alpha\left(\frac{x_i}{\beta}\right)^{-\alpha\theta} + 1\right]^{-\frac{1}{\alpha}}\right\}^{-1}\left[\alpha\left[\frac{x_i}{\beta}\right]^{-\alpha\theta}\right.\right.$$
$$\left.\left. +1\right]^{-\frac{1}{\alpha}-1}\left(\frac{x_i}{\beta}\right)^{-\alpha\theta}\right] = 0$$

$$\frac{\partial\ell}{\partial\theta} = \frac{n}{\theta} + \sum_{i=1}^{n} log\left(\frac{x_i}{\beta}\right) - (\alpha+1)\sum_{i=1}^{n}\left[\frac{\left(\frac{x_i}{\beta}\right)^{\alpha\theta} log\left(\frac{x_i}{\beta}\right)}{\alpha + \left(\frac{x_i}{\beta}\right)^{\alpha\theta}}\right] + 2\alpha(1-\delta)$$
$$\times\sum_{i=1}^{n}\left[\left\{\delta + (1-\delta)\left[\alpha\left(\frac{x_i}{\beta}\right)^{-\alpha\theta} + 1\right]^{-\frac{1}{\alpha}}\right\}^{-1}\left[\alpha\left(\frac{x_i}{\beta}\right)^{-\alpha\theta} + 1\right]^{-\frac{1}{\alpha}-1}\right.$$
$$\left.\times\left(\frac{x_i}{\beta}\right)^{-\alpha\theta} log\left(\frac{\alpha x_i}{\beta}\right)\right] = 0$$

$$\frac{\partial\ell}{\partial\delta} = \frac{n}{\delta} - 2\sum_{i=1}^{n}\left[\left\{\delta + (1-\delta)\left[\alpha\left(\frac{x_i}{\beta}\right)^{-\alpha\theta} + 1\right]^{-\frac{1}{\alpha}}\right\}^{-1}\right.$$
$$\left.\times\left[1 - \left\{\alpha\left(\frac{x_i}{\beta}\right)^{-\alpha\theta} + 1\right\}^{-\frac{1}{\alpha}}\right]\right] = 0$$

Hence, the MLEs of $\alpha, \beta, \theta and\delta$ for MOK distribution can be obtained by solving above equations. This task can be successfully done with any statistical package for example R.

4. Application of the MOK distribution

A comparison of proposed MOK distribution has been made with the Exponentiated generalized kappa distribution, Kumaraswamy generalized kappa distribution, McDonald generalized kappa distribution, two parameters kappa and three parameters kappa distribution with the help of data sets given in Sections 4.1 and 4.2.

Goodness of fit is generally decided utilizing a likelihood ratio approach. We used two goodness of fit criterions Cramer-Von Mises W^* statistic and Anderson–Darling A^* statistic along with minimum value of the log likelihood function $\ell(\cdot)$.

Table 4.1
MLEs for stream flow amounts.

Distribution	α	β	θ	δ	a	b
K2	10.7420	312.1797	–	–	–	–
	(5.9393)	(46.5958)				
K3	0.0457	161.5442	56.7357	–	–	–
	(0.2369)	(15.9801)	(287.0343)			
McGK	0.1951	24.1709	4.3609	0.0024	42.7565	7.5042
	(0.5998)	(53.9753)	(13.2095)	(23.2242)	(99.1328)	(20.1226)
KGK	0.1678	17.3661	7.2191	–	37.6739	4.8300
	(1.0243)	(219.7929)	(44.7869)		(563.6182)	(9.7572)
EGK	0.0264	42.6137	34.2247	–	3.2326	79.3117
	(0.0118)	(39.2628)	(145.7587)		(19.3128)	(102.2513)
MOK	0.0626	129.8663	50.1215	2.9310	–	–
	(0.1406)	(34.1181)	(111.7535)	(3.7595)		

(Standard errors in parenthesis).

Table 4.2
MLEs for failure times of mechanical components.

Distribution	α	β	θ	δ	a	b
K2	17868.2697	4.66187	–	–	–	–
	(13831.0322)	(0.0015)				
K3	17060.1568	4.6624	1.3066	–	–	–
	(4510.9833)	(0.0012)	(0.1427)			
McGK	29.1913	4.6808	34.1854	0.0274	0.0473	1.3184
	(176.2242)	(0.0823)	(78.5917)	(0.1189)	(0.1089)	(0.2988)
KGK	23.9728	4.6815	34.7390	–	0.0453	1.3333
	(125.6004)	(0.0806)	(112.8132)		(0.1473)	(0.3046)
EGK	6.0757	5.9738	4.9868	–	11.5468	0.3326
	(1.0824)	(0.3541)	(0.7825)		(2.8778)	(0.0788)
MOK	6.0909	1.4164	0.7576	15.4506	–	–
	(5.4648)	(0.7806)	(0.7299)	(30.1753)		

(Standard errors in parenthesis).

Table 4.3
Performance indices.

Data	Stream flow amounts			Failure times of mechanical components		
Distribution	$\ell(\cdot)$	A^*	W^*	$\ell(\cdot)$	A^*	W^*
K2	214.1449	1.0876	0.1802	129.3613	1.4302	0.2574
K3	207.0037	0.4919	0.0799	126.2338	1.3418	0.2397
McGK	206.3862	1.4356	0.2222	125.2638	1.5101	0.1786
KGK	206.4671	0.4555	0.0788	125.2886	1.0720	0.1864
EGK	206.8389	0.4761	0.0782	127.3661	0.7816	0.1191
MOK	206.6291	0.4583	0.0774	129.8588	0.6006	0.0802

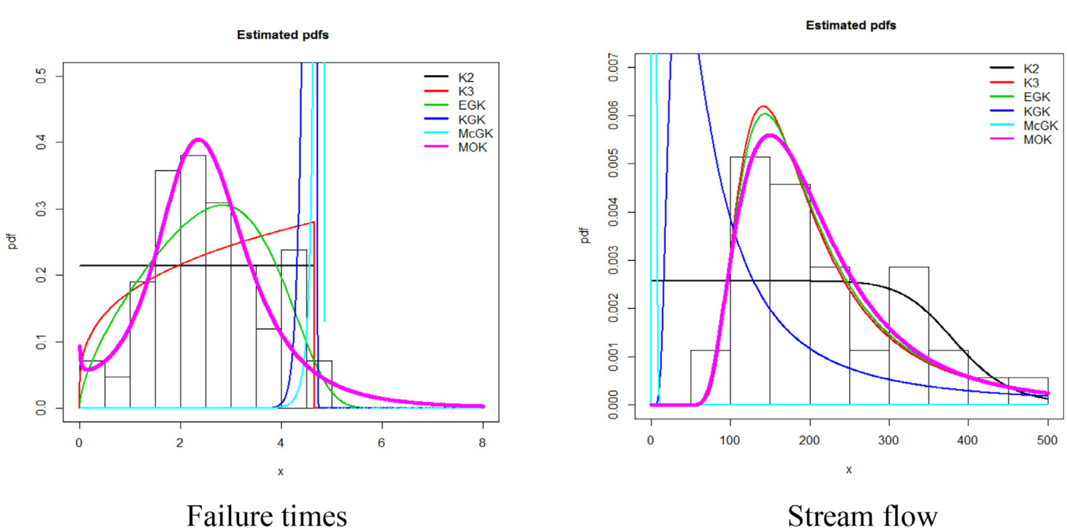

Fig. 4.1. Graphs of the estimated pdfs for G-Kappa distributions.

4.1. On stream flow amounts (1000 acre-feet)

Following data set (Mielke and Johnson, 1973) consists of stream flow amounts (1,000 acre-feet) for 35 years (1936–70) at the U.S. Geological Survey (USGS) gaging station number 9–3425 for 1st April to 31st August of each year.

192.48, 303.91, 301.26, 135.87, 126.52, 474.25, 297.17, 196.47, 327.64, 261.34, 96.26, 160.52, 314.60, 346.30, 154.44, 111.16, 389.92, 157.93, 126.46, 128.58, 155.62, 400.93, 248.57, 91.27, 238.71, 140.76, 228.28, 104.75, 125.29, 366.22, 192.01, 149.74, 224.58, 242.19, 151.25.

4.2. On failure times of mechanical components

Following data represents the failure times of mechanical components obtained from Silva et al. (2015).

0.040, 1.866, 2.385, 3.443, 0.301, 1.876, 2.481, 3.467, 0.309, 1.899, 2.610, 3.478, 0.557, 1.911, 2.625, 3.578, 0.943, 1.912, 2.632, 3.595, 1.070, 1.914, 2.646, 3.699, 1.124, 1.981, 2.661, 3.779, 1.248, 2.010, 2.688, 3.924, 1.281, 2.038, 2.823, 4.035, 1.281, 2.085, 2.890, 4.121, 1.303, 2.089, 2.902, 4.167, 1.432, 2.097, 2.934, 4.240, 1.480, 2.135, 2.962, 4.255, 1.505, 2.154, 2.964, 4.278, 1.506, 2.190, 3.000, 4.305, 1.568, 2.194, 3.103, 4.376, 1.615, 2.223, 3.114, 4.449, 1.619, 2.224, 3.117, 4.485, 1.652, 2.229, 3.166, 4.570, 1.652, 2.300, 3.344, 4.602, 1.757, 2.324, 3.376, 4.663

Tables 4.1 and 4.2 present the maximum likelihood estimates of the parameters of the distributions and their standard errors in parenthesis. Table 4.3 shows the minimum values of the log likelihood function $\ell(\cdot)$ and performance indices W^* and A^* for all the distributions under study on both data sets. It is worth pointing that MOK distribution has smallest values of W^* and A^* as compared to the rest distributions. Fig. 4.1 depicts the graphical representation of fitness of MOK and other forms of kappa distribution. It is evident from performance indices and graph that MOK is the good fit distribution for both data sets so it can be considered that the proposed model is a good competitive model.

5. Concluding remarks

In this manuscript, we proposed a new generalization of kappa distribution named as Marshall-Olkin Kappa (MOK) distribution. Statistical properties i.e. pdf, cdf, mode, median, quantiles, moments and mean deviations of the new model are derived. Various reliability properties, inequality measures and generalized entropy are also studied. Maximum likelihood method is applied to estimate the unknown parameters of the new model. Two real life data sets are used to show the competitiveness of proposed model and finally conclusion has been made that new model may serve better than other competing models.

References

Alice, T., Jose, K.K., 2003. Marshall-Olkin Pareto process. Far East J. Theor. Stat. 9, 117–132.

Atkinson, A.B., 1970. On the measurement of the inequality. J. Econ. Theory. 2, 244–263.

Barreto-Souza, W., Lemonte, A.J., Cordeiro, G.M., 2013. General results for the Marshall and Olkin's family of distributions. Ann. Braz. Acad. Sci. 85, 3–21.

Bonferroni, C., 1930. Elementi di statistica generale. Libreria Seber, Firenze.

Cordeiro, G.M., Lemonte, A.J., 2013. On the Marshall-Olkin extended Weibull distribution. Stat. Pap. 54, 333–353.

Cordeiro, G.M., Lemonte, A.J., Ortega, E.M.M., 2014. The Marshall-Olkin family of distributions: mathematical properties and new models. J. Stat. Theor. Pract. 8, 343–366.

Cowel, E.A., 1980. On the structure of additive inequality measures. Rev. Econ. Stud. 47, 521–531.

Garcia, V.J., Gómez-Deniz, E., Vázquez Polo, F.J., 2010. A new skew generalization of the normal distribution: properties and applications. Comput. Stat. Data. An. 54, 2021–2034.

Ghitany, M.E., Al-Awadhi, F.A., Alkhalfan, L.A., 2007a. Marshall-Olkin extended Lomax distribution and its application to censored data. Commun. Stat-Theor. M. 36, 1855–1866.

Ghitany, M.E., Al-Hussaini, E.K., Al-Jarallah, R.A., 2007b. Marshall-Olkin extended Weibull distribution and its application to censored data. J. Appl. Stat. 32, 1025–1036.

Ghitany, M.E., Al-Mutairi, D.K., Al-Awadhi, F.A., Al-Burais, M.M., 2012. Marshall-Olkin extended Lindley distribution and its application. Int. J. Ap. Mat. 25, 709–721.

Hussain, S., 2015. Extension, Properties and Application of Kappa Distribution. M. Phil. Thesis, Submitted to Government College University, Faisalabad, Pakistan. (Unpublished).

Krishna, E., Jose, K.K., Alice, T., Ristić, M.M., 2013. The Marshall-Olkin Fréchet distribution. Commun. Stat-Theor. M. 42, 4091–4107.

Lemonte, A.J., 2013. A new extension of the Birnbaum-Saunders distribution. Braz. J. Probab. Stat. 27, 133–149.

Lorenz, M.O., 1905. Method of measuring the concentration of wealth. J. Am. Stat. Assoc. 9, 209–219.

Marshall, A.W., 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.

Mielke, P.W.J., 1973. Another family of distributions for describing and analyzing precipitation data. J. Appl. Meteorol. 12 (2), 275–280.

Mielke, P.W.J., Johnson, E.S., 1973. Three parameter kappa distribution maximum likelihood estimations and likelihood ratio tests. Mon. Weather Rev. 101, 701–707.

Nanda, A.K., Das, S., 2012. Stochastic orders of the Marshall-Olkin extended distribution. Stat. Probabil. Lett. 82, 295–302.

Pescim, R.R., Cordeiro, G.M., Demétrio, C.G.B., Ortega, E.M., Nadarajah, S., 2012. The new class of Kummer beta generalized distributions. SORT-Stat. Oper. Res. T. 36, 153–180.

Pietra, G., 1915. Delle relazioni fra indici di variabilita, note I e II. Atti del Reale Istituto Veneto di scienze. Lettere ed Arti. 74, 775–804.

Ristić, M.M., Jose, K.K., Ancy, J., 2007. A Marshall-Olkin gamma distribution and minification process. STARS: Stress and Anxiety Research. Society. 11, 107–117.

Silva, R.V., Silva, F.G., Ramos, M.W.A., Cordeiro, G.M., 2015. A new extended gamma generalized model. Int. J. Pure Appl. Math.. 100 (2), 309–335.

Shorrocks, A.F., 1980. The Class of additively decomposable inequality measures. Econometrica. 48, 613–625.

Zenga, M., 1984. Proposta per un indice di concentrazione basto sui rapporti fra quantili di populizone e quantili ti reddito. Giornale degli economisti e Annali di Economia 48, 301–326.