International Journal of Mathematics Trends and Technology
ISSN: 2231 – 5373 /doi:10.14445/22315373/IJMTT-V66I11P505

Volume 66 Issue 11,67-79, November 2020

On the k – Higher-Order Beta Distribution Function

Benson Ade Afere[#1]

[#]*School of Tecnnology,*

Department of Mathematics and Statistics,

Federal Polytechnic, Idah, Nigeria.

Abstract - *In this work, a new higher-order beta probability distribution function is proposed from the existing beta probability distribution function. The new probability distribution function was derived using the work of [5]. The properties of the two distribution functions were given. A Monte Carlo experiment is performed for two scenarios using small and large sample sizes and it was observed that the proposed distribution has the least mean square error. It was equally observed that for the two scenarios, as the sample size increases, the error decreases which obey the finite sample theory. More importantly, based on the observations, the proposed distribution is efficient even if the data set departs from the standard beta distribution. A real life applications were used to stress further the flexibility of the proposed distribution.*

Keywords: *Probability distribution, classical beta distribution function, higher-order beta distribution function, mean square error, Monte Carlo experiment.*

MSC Subject Classification: 62E10, 62E15

I. INTRODUCTION

A probability distribution can be defined as an assignment of probabilities to the values of a random variable. It can also be viewed as a statistical function that describes all the possible values and likelihoods that a random variable can take within a given range. There are several types of probability distribution such as the uniform distribution, the gamma distribution, the beta distribution, the normal distribution and several others. But the most commonly used distribution is the normal distribution because of its usefulness in finance, inventory, science and technology. However, this does not exclude the development and usage of other distributions.

The basic idea of the word "probability theory" began in the seventeenth century when the two French Mathematicians, Blaise Pascal and Piere de Fermat worked on two problems from game of chance [8]. In recent times, the evolvement of more new methods for constructing simple and robust probability distribution abound in the literature [2, 4, 6, 8]. A survey of these methods can be found in [7] and the references therein.

The beta distribution is a continuous probability distribution with two positive parameters α and β which control its shape and it is defined on a closed interval $x \in [0,1]$. Variants of beta distributions abound in the literature and they have been used in several ways by different authors to model the behaviour of proportions, see for example [4]. Hence mathematically, a random variable X is said to have a beta distribution with shape parameters $\alpha > 0$ and $\beta > 0$ if the probability density function is given by

$$f(x;\alpha,\beta) = \begin{cases} \dfrac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1}, & 0 < x < 1 \\ 0 & , \quad \text{elsewhere} \end{cases} \tag{1}$$

where

$$B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} \tag{2}$$

and

$$\Gamma(\alpha) = \int_0^\infty u^{\alpha-1}e^{-u}du = (\alpha-1)\Gamma(\alpha-1) \tag{3}$$

The main objective of this work is to compare beta distribution proposed by [10] called k – Beta distribution with the higher-order beta distribution proposed in this work. The remaining part of this paper is structured as follows: this section is completed by giving some basic definitions. In Section 2, we review the k-Beta distribution which sets out the framework for developing the proposed higher-order beta distribution discussed in this work. In Section 3, we derive the proposed k-Higher-order beta distribution and give its properties. A Monte-Carlo experiment and real life applications are performed in Section 4 for two data sets. The discussion of results is equally done in the section. Finally Section 5 gives the concluding remark.

A. BASIC DEFINITIONS

We give some definitions here which provide the rudiments for our main results. Definitions 1 – 4 are provided in [11].

1. Let X denote any random variable. The distribution of X, denoted by $F(x)$, is such that $F(x) = P(X \le x)$ for $-\infty < x < \infty$.
2. A random variable X with distribution $F(x)$ is said to be continuous if $F(x)$ is continuous, for $-\infty < x < \infty$.
3. Let $F(x)$ be the distribution function for a continuous random variable X. Then $f(x)$, given by

$$f(x) = \frac{dF(x)}{dx} = F'(x)$$

whenever the derivative exists, it is called the probability density function for the random variable X.

4. If $f(x)$ is a density function for a continuous random variable X, then

i. $f(x) \ge 0 \ \forall\, x, -\infty < x < \infty$

ii. $\int_{-\infty}^{\infty} f(x)dx = 1$.

A survey of the properties of beta distribution as highlighted in [1] is as follows:

1. If $\alpha = \beta$, then the shape of $f(x;\alpha,\beta)$ is symmetric, unimodal and the mode = mean = median =0.5.
2. If $\alpha < \beta$, then the shape of $f(x;\alpha,\beta)$ is right skewed, unimodal and the mean >median> mode <0.5.
3. If $\alpha > \beta$, then the shape of $f(x;\alpha,\beta)$ is left skewed, unimodal and the mean < median < mode >0.5.

II. REVIEW OF *k* – BETA DISTRIBUTION FUNCTION [9]

Let X be a continuous random variable; then it is said to have a k-beta distribution with two parameters α and β if its probability distribution function is defined by

$$f_k(x) = \begin{cases} \dfrac{1}{kB_k(\alpha,\beta)} x^{\alpha/k-1}(1-x)^{\beta/k-1}, & 0 \le x \le 1;\ \alpha,\beta,k > 0, \\ 0, & \text{elsewhere} \end{cases} \tag{4}$$

Rahman, *et al* [9] further defined this distribution as k-beta distribution of the first kind and they labeled it $\beta_{1,k}(m,n)$, (where m = α and n = β in this paper). They also gave its cumulative distribution function as

$$F_k(x) = \begin{cases} 0, & x < 0 \\ \displaystyle\int_0^x \frac{1}{kB_k(\alpha,\beta)} t^{\alpha/k-1}(1-t)^{\beta/k-1}dt, & 0 \le x \le 1; \\ & \alpha,\beta,k > 0 \\ 0, & x > 1 \end{cases} \tag{5}$$

As contained in [10], the k-beta probability distribution function in (4) satisfies the following properties:

i. It is a probability density function. That is, it integrates to 1.

ii. Its mean is $\frac{\alpha}{\alpha+\beta}$

iii. Its variance is $\frac{\alpha\beta k}{(\alpha+\beta)^2(\alpha+\beta+k)}$

Proof: The proof is contained in [10].

III. THE *k* - HIGHER-ORDER BETA PROBABILITY DISTRIBUTION FUNCTION

In this section, we propose a new distribution function from the existing one. Since the kernel density is a probability density function, we adapt the method for developing higher-order kernels to construct our proposed new probability distribution function. Using the work of [5] and the *k*-Beta distribution of [10], a three parameters beta distribution is proposed as:

$$g_{kH}(x;\alpha,\beta,k)=\begin{cases}\frac{1}{2kB(\alpha,\beta)}(\alpha+2k-x(\alpha+\beta+k))x^{\frac{\alpha}{k}-1}(1-x)^{\frac{\beta}{k}-2}, & 0<x<1;\ \alpha,\beta,k>0\\ 0 & ,\quad \text{elsewhere}\end{cases} \tag{6}$$

Equation (6) is thus the proposed *k*-HBDF with the cumulative distribution function given as:

$$G_{kH}(x)=\begin{cases}0, & x<0\\ \frac{1}{2kB_k(\alpha,\beta)}\int_0^x(\alpha+2k-t(\alpha+\beta+k))t^{\alpha/k-1}(1-t)^{\beta/k-2}dt, & 0\le x\le 1;\ \ \alpha,\beta,k>0\\ 0, & x>1\end{cases} \tag{7}$$

Theorem 1
The proposed *k*-HBDF represents a probability density function with the respective mean and variance:

(i) $\mathrm{E}(x)=\frac{\alpha}{\alpha+\beta}$

(ii) $|\mathrm{Var}(x)|=\frac{\alpha^2}{(\alpha+\beta)^2}$

Proof
Using the relation in Definition 4(ii), we observe that

$$\int_0^1 g_{kH}(x;\alpha,\beta)dx=\int_0^1\frac{1}{2kB_k(\alpha,\beta)}(\alpha+2k-x(\alpha+\beta+k))x^{\frac{\alpha}{k}-1}(1-x)^{\frac{\beta}{k}-2}dx$$

$$=\frac{1}{2k\,\mathrm{B}_k(\alpha,\beta)}[2k\,\mathrm{B}_k(\alpha,\beta)=1$$

(i) $$\mathrm{E}(x)=\int_0^1 xg_{kH}(x;\alpha,\beta)dx=\int_0^1 x\frac{1}{2kB_k(\alpha,\beta)}(\alpha+2k-x(\alpha+\beta+k))x^{\frac{\alpha}{k}-1}(1-x)^{\frac{\beta}{k}-2}dx$$

$$=\int_0^1\frac{1}{2kB_k(\alpha,\beta)}(\alpha+2k-x(\alpha+\beta+k))x^{\frac{\alpha+k}{k}-1}(1-x)^{\frac{\beta}{k}-2}dx$$

$$=\frac{B_k(\alpha+k,\beta)}{B_k(\alpha,\beta)}$$

$$=\frac{\Gamma_k(\alpha+k)\Gamma_k(\beta)}{\Gamma_k(\alpha+\beta+k)B_k(\alpha,\beta)} \text{ [applying definition ()]}$$

$$=\frac{\Gamma_k(\alpha+k)\Gamma_k(\beta)\Gamma_k(\alpha+\beta)}{\Gamma_k(\alpha+\beta+k)\Gamma_k(\alpha)\Gamma_k(\beta)}$$

$$=\frac{\alpha}{\alpha+\beta}$$

(ii) $\mathrm{Var}(x)=\mathrm{E}(x^2)-(\mathrm{E}(x))^2$

$$\mathrm{E}(x^2)=\int_0^1 x^2 g_{kH}(x;\alpha,\beta)dx=\int_0^1 x^2\frac{1}{2kB_k(\alpha,\beta)}(\alpha+2k-x(\alpha+\beta+k))x^{\frac{\alpha}{k}-1}(1-x)^{\frac{\beta}{k}-2}dx$$

$$=\int_0^1\frac{1}{2kB_k(\alpha,\beta)}(\alpha+2k-x(\alpha+\beta+k))x^{\frac{\alpha+2k}{k}-1}(1-x)^{\frac{\beta}{k}-2}dx$$

$$=\frac{1}{2kB_k(\alpha,\beta)}([\alpha+2k]2\,\mathrm{k}\,B_k(\alpha+2k,\beta-k)-[\alpha+\beta+k]2\,\mathrm{k}\,B_k(\alpha+3k,\beta-k))$$

$$=\frac{1}{B_k(\alpha,\beta)}([\alpha+2k]B_k(\alpha+2k,\beta-k)-[\alpha+2k]B_k(\alpha+2k,\beta-k))$$

$$=\frac{1}{B_k(\alpha,\beta)}(0)=0$$

But $\mathrm{Var}(x)=\mathrm{E}(x^2)-(\mathrm{E}(x))^2$

$\therefore \quad |\mathrm{Var}(x)| \le |\mathrm{E}(x^2)|+|(\mathrm{E}(x))^2|$

Hence

$$|\mathrm{Var}(x)|=\frac{\alpha^2}{(\alpha+\beta)^2}$$

A. PROPERTIES OF THE PROPOSED k-HBDF

When k = 0.5

1. If $k=1$ and $\alpha>\beta$, then the shape of $g(x;\alpha,\beta)$ is left skewed, unimodal and the median > mean > mode > 0.5.
2. If $k=1$ and $\alpha=\beta$, then the shape of $g(x;\alpha,\beta)$ is symmetric, unimodal and the mode = mean = median > 0.5
3. If $k=1$ and $\alpha<\beta$, then the shape of $g(x;\alpha,\beta)$ is right skewed, unimodal and the mean =median = mode < 0.5.

When k = 1

1. If k > 1 and $\alpha>\beta$, then the shape of $g(x;\alpha,\beta)$ is left skewed.
2. If k > 1 and $\alpha=\beta$, then the shape of $g(x;\alpha,\beta)$ is symmetric.
3. If k > 1 and $\alpha<\beta$, then the shape of $g(x;\alpha,\beta)$ is strictly decreasing.

When k = 1.5

1. If k < 1 and $\alpha>\beta$, then the shape of $g(x;\alpha,\beta)$ is symmetric and the mean = median = mode = 0.8.
2. If k < 1 and $\alpha=\beta$, then the shape of $g(x;\alpha,\beta)$ is symmetric and the mean = median = mode = 0.5.
3. If k < 1 and $\alpha<\beta$, then the shape of $g(x;\alpha,\beta)$ is symmetric and the mean = median = mode = 0.2.

B. SERIES REPRESENTATION OF k – HBDF

$$g_{kH}=\frac{1}{2kB_k(a,b)}[(a+2k-x(a+b+k))(1-x)^{\frac{b}{k}-2}]x^{\frac{a}{k}-1}$$

$$=\frac{1}{kB_k(a,b)}[(\sum_{\rho=0}^{\infty}(-1)^{\rho}2^{-1}\begin{pmatrix}\frac{b}{k}-1\\ \rho\end{pmatrix}(a+(2+\rho)k)]x^{\frac{a}{k}+\rho-1}$$

Now, $B_k(a,b)=\frac{1}{k}\int_0^1(1-x)^{\frac{b}{k}-1}x^{\frac{a}{k}-1}dx$

$$=\frac{1}{k}\sum_{\upsilon=0}^{\infty}(-1)^{\upsilon}\begin{pmatrix}\frac{b}{k}-1\\ \upsilon\end{pmatrix}x^{\upsilon}\int_0^1 x^{\frac{a}{k}-1}dx$$

$$=\sum_{\upsilon=0}^{\infty}(-1)^{\upsilon}\begin{pmatrix}\frac{b}{k}-1\\ \upsilon\end{pmatrix}\frac{1}{a+\upsilon k}$$

Hence,

$$g_{kH}(x)=\frac{\frac{1}{2kB_k(a,b)}[(\sum_{\rho=0}^{\infty}(-1)^{\rho}\begin{pmatrix}\frac{b}{k}-1\\ \rho\end{pmatrix}(a+(2+\rho)k)]x^{\frac{a}{k}+\rho-1}}{\sum_{\upsilon=0}^{\infty}(-1)^{\upsilon}\begin{pmatrix}\frac{b}{k}-1\\ \upsilon\end{pmatrix}\frac{1}{a+\upsilon k}}$$

$$\therefore \quad g_{kH}(x)=(\sum_{\rho=0}^{\infty}\sum_{\upsilon=0}^{\infty}\tau_{\rho\upsilon}x^{\frac{a}{k}+\rho-1} \tag{8}$$

where

$$\tau_{\rho\upsilon}=(-1)^{\rho-\upsilon}\begin{pmatrix}\frac{b}{k}-1\\ \rho\end{pmatrix}\begin{pmatrix}\frac{b}{k}-1\\ \upsilon\end{pmatrix}^{-1}\frac{2^{-1}(a+(2+\rho)k)}{a+\upsilon k} \tag{9}$$

And thus, the corresponding cdf in (7) can be re-written in series form as

$$G_{kH}(x)=(\sum_{\rho=0}^{\infty}\sum_{\upsilon=0}^{\infty}(-1)^{\rho-\upsilon}\begin{pmatrix}\frac{b}{k}-1\\ \rho\end{pmatrix}\begin{pmatrix}\frac{b}{k}-1\\ \upsilon\end{pmatrix}^{-1}\frac{(a+(2+\rho)k)}{a+\upsilon k}\int_0^x t^{\frac{a}{k}+\rho-1}dt$$

$$=(\sum_{\rho=0}^{\infty}\sum_{\upsilon=0}^{\infty}\frac{\tau_{\rho\upsilon}}{(a+\rho k)}x^{\frac{a}{k}+\rho} \tag{10}$$

where $\tau_{\rho\upsilon}$ is as defined in (9).

Theorem 2

If the shape parameters α, β and k are strictly greater than zero, then the higher-order moments of k-HBDF are given by:

$$|\operatorname{E}(x^r)|\leq|2-r|\frac{a(a+k)(a+2k)\cdots(a+(r-1)k)}{(a+b)(a+b+k)(a+b+2k)\cdots(a+b+(r-1)k)} \tag{11}$$

Proof

Consider

$$\operatorname{E}(x^r)=\int_0^1 x^r\frac{1}{2kB_k(\alpha,\beta)}(\alpha+2k-x(\alpha+\beta+k))x^{\frac{\alpha}{k}-1}(1-x)^{\frac{\beta}{k}-2}dx$$

$$= \frac{1}{2kB_k(\alpha,\beta)}[(\alpha+2k)\int_0^1 x^r (x^{\frac{\alpha}{k}-1}(1-x)^{\frac{\beta}{k}-2}dx - (\alpha+\beta+k)\int_0^1 x^r x^{\frac{\alpha}{k}}(1-x)^{\frac{\beta}{k}-2}dx]$$

$$= \frac{1}{2kB_k(\alpha,\beta)}[(\alpha+2k)\int_0^1 x^r (x^{\frac{\alpha+rk}{k}-1}(1-x)^{\frac{\beta}{k}-2}dx - (\alpha+\beta+k)\int_0^1 x^r x^{\frac{\alpha+(r+1)k}{k}-1}(1-x)^{\frac{\beta}{k}-2}dx]$$

$$= \frac{1}{2kB_k(\alpha,\beta)}[(\alpha+2k)B_k(\alpha+rk,\beta) = (\alpha+\beta+k)B_k(\alpha+(r+1)k,\beta)]$$

On simplification, we have

$$\mathrm{E}(x^r) = \frac{1}{2kB_k(\alpha,\beta)} \frac{2[\alpha(\alpha+k)(\alpha+2k)\cdots(\alpha+(r-1)k][(\alpha+2k)-(\alpha+rk)]\Gamma_k(\alpha)\Gamma_k(\beta)}{(\alpha+\beta)(\alpha+\beta+k)(\alpha+\beta+2k)\cdots(\alpha+\beta+(r-1)k)\Gamma_k(\alpha+\beta)}$$

$$= \frac{1}{kB_k(\alpha,\beta)} \frac{[\alpha(\alpha+k)(\alpha+2k)\cdots(\alpha+(r-1)k][k(2-r)]}{(\alpha+\beta)(\alpha+\beta+k)(\alpha+\beta+2k)\cdots(\alpha+\beta+(r-1)k)} B_k(\alpha,\beta)$$

$$= \frac{(2-r)\alpha(\alpha+k)(\alpha+2k)\cdots(\alpha+(r-1)k)}{(\alpha+\beta)(\alpha+\beta+k)(\alpha+\beta+2k)\cdots(\alpha+\beta+(r-1)k)}$$

Thus, by applying the triangle inequality in the above equation, the desired result is achieved. If $r = 1$, $\mathrm{E}(x^1) = \mu = \alpha/(\alpha+\beta)$. If $r = 2$, then $\mathrm{E}(x^2) = 0$ and thus, $\mathrm{Var}(x) = \mathrm{E}(x^2) - \mathrm{E}^2(x) = 0 - [\alpha/(\alpha+\beta)]^2$ and hence, we have the desired results in Theorem 2

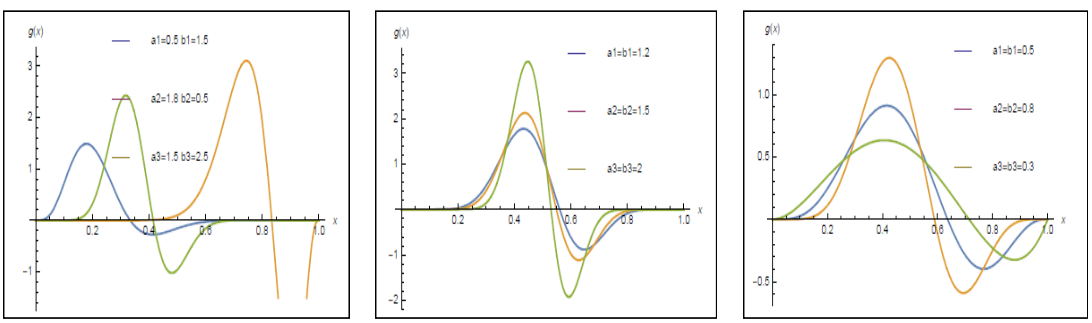

Fig. 1: Graph of the pdf of *k* – HBDF when *k* << 1

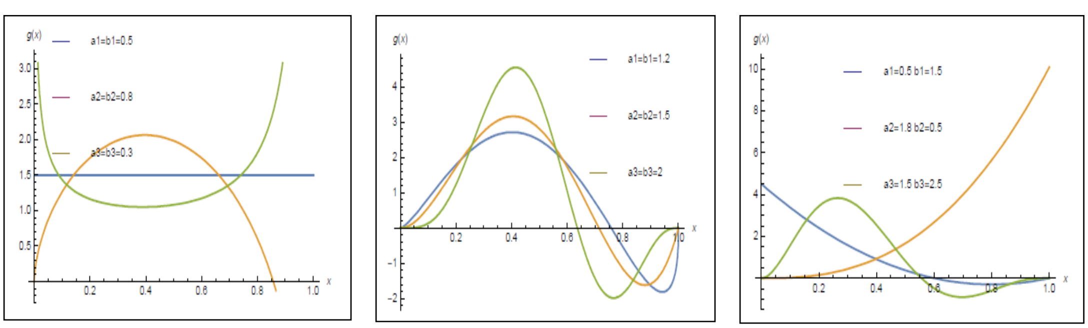

Fig. 2: Graph of the pdf of *k* – HBDF when *k* < 1

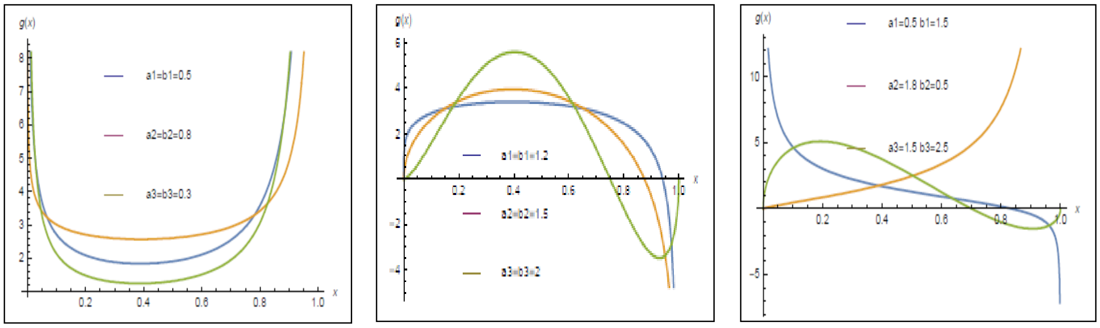

Fig. 3: Graph of the pdf of *k* – HBDF when *k* = 1

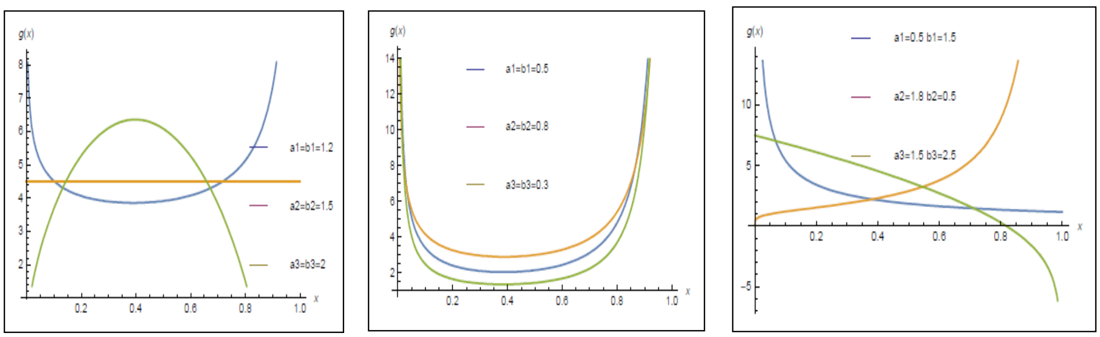

Fig. 4: Graph of the pdf of *k* – HBDF when *k* > 1

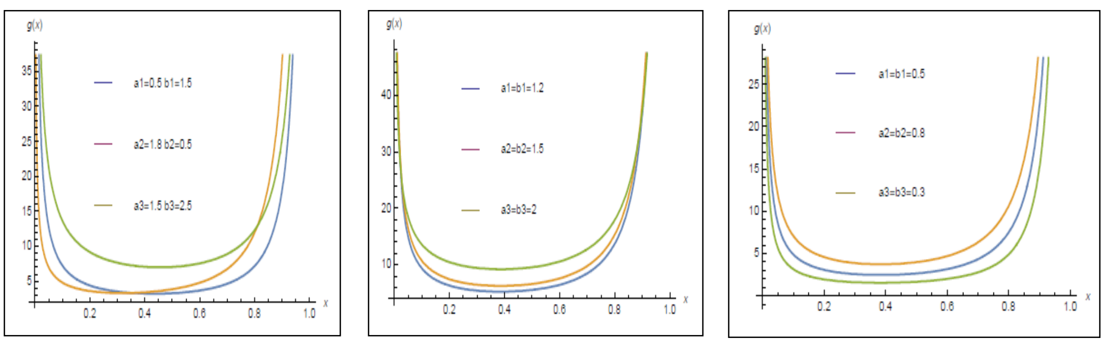

Fig. 5: Graph of the pdf of *k* – HBDF when *k* > > 1

Figure 1 shows the density functions for some selected values of parameters α, β and k. These plots indicate that $k-$HBDF are all symmetric, unimodal and ununimodal irrespective of the when $\alpha=\beta<1$, $\alpha=\beta>1$ and when $\alpha>\beta$ or $\alpha<\beta$ for the value of k when it is far less than one. Figure 2(a,b,c) shows the density of $k-$HBDF for some selected values of parameters α and β when k is less than one. The plots in Figure 2(a) shows that the $k-$HBDF is U-shaped, uniform and symmetric when $\alpha=\beta=0.5$, $\alpha=\beta=0.8$ and $\alpha=\beta=0.3$. Figure 2(b) plots take the same shape as Figure 1(a,b,c). The plots in Figure 2(c) show that $k-$HBDF could be increasing, decreasing or symmetric depending on the values of α and β

Figure 3(a,b,c) shows the density of $k-$HBDF for some selected values of parameters α and β when k is equal to one. The plots in Figure 3(a) indicate that $k-$HBDF are U-shaped when $\alpha=\beta<1$. Those in Figure 3(b) are all symmetric but as the value of α and β increases, the shape becomes flattened. Figure 3(c) plots show that $k-$HBDF could be increasing, decreasing or skewed depending on the values of α and β.

Figure 4(a,b,c) shows the density of $k-$HBDF for some selected values of parameters α and β when k is greater than one. The plots in Figure 4(a) take the same shape as those in Figure 2(a). Equally, the shape of the plots in Figure 4(b) takes the same shape as those in Figure 3(a). However, plots in Figure 4(c) indicate that the plots are either decreasing or increasing. But, when $\alpha>\beta$ or $\alpha<\beta$, the curves are asymptotic to the x-axis.

Figure 3(a,b,c) shows the density of $k-$HBDF for some selected values of parameters α and β when k is far greater than one. The plots in Figure 5(a,b,c) are all U-shaped irrespective of when $\alpha=\beta<1$, $\alpha=\beta>1$ and when $\alpha>\beta$ or $\alpha<\beta$.

IV. NUMERICAL EXPERIMENT

In this section, we highlight the performance of the proposed distribution by comparing it with the existing probability distribution function studied by [10]. This is done by visualising by a Monte Carlo experiment. Thereafter, two real life data sets are used to show its flexibility.

A. MONTE – CARLO EXPERIMENT

To study the performance of the higher-order beta probability distribution function, Monte Carlo Simulation experiments are conducted for two different scenarios. The first experiment is performed using small and large sample sizes for three different combinations of values of α and β. That is, when $\alpha<\beta$, $\alpha=\beta$ and $\alpha>\beta$. In the second instance, a Monte Carlo experiment is performed using small and large sample sizes for two different mixtures beta densities. These are $A=(3/10)B(1,1)+(7/10)B(2,2)$ and $B=(2/3)B(1,2)+(1/3)B(7/4,7/4)$. Given a random sample $x_1,x_2,\ldots,x_n$ on a unit interval $x\in(0,1)$, the simulation is performed for $r=1000$ runs such that mean squared error (MSE) is given as:

$$MSE=\frac{1}{n}\sum_{i=1}^{n}[f(x_i)-\hat{f}(x_i)]^2 \tag{12}$$

where $f(.)$ is the pdf of a distribution and $\hat{f}(.)$ is its estimate. Equation (12) is then computed for four different standard beta densities and two different mixture beta densities for small and large sample sizes. The results are presented in Figures 6 through 11 below:

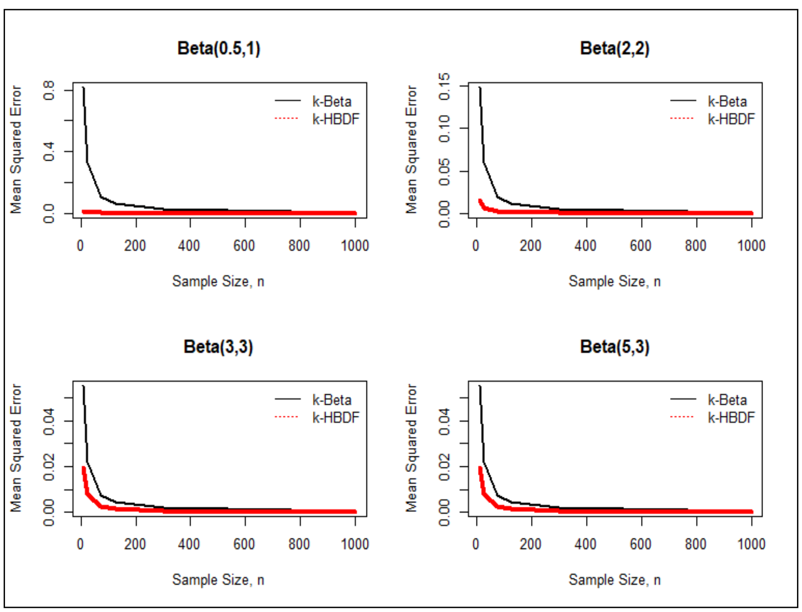

Fig. 6: Graph of the MSE of *k* – Higher-order Beta Distribution and *k* – Beta Distribution Functions for the four standard densities when *k* = 0.5

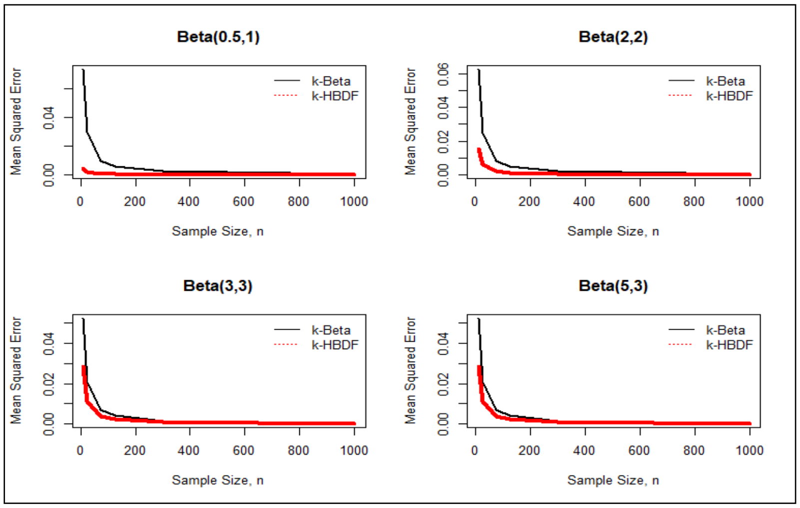

Fig. 7: Graph of the MSE of k – Higher-order Beta Distribution and k – Beta Distribution Functions for the four standard densities when k = 1

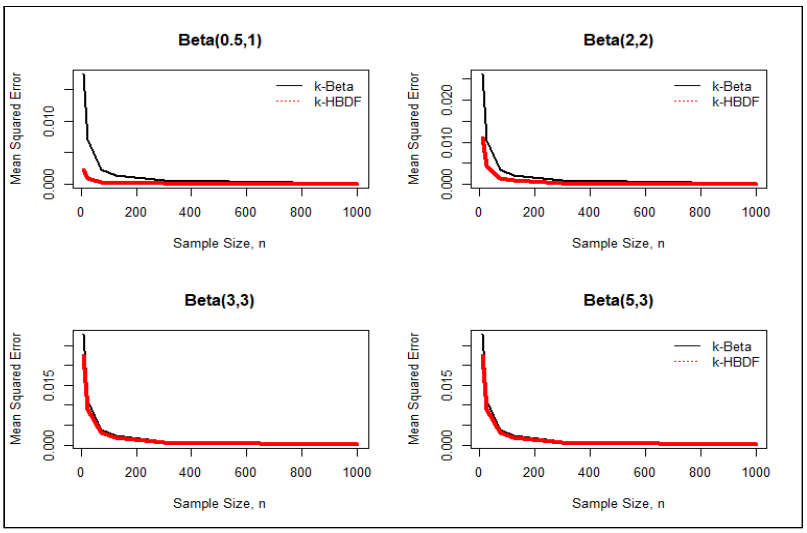

Fig. 8: Graph of the MSE of k – Higher-order Beta Distribution and k – Beta Distribution Functions for the four standard densities when k = 1.5

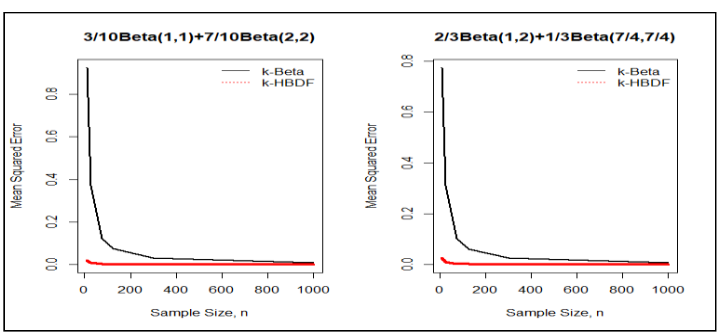

Fig. 9: Graph of the MSE of k – Higher-order Beta Distribution and k – Beta Distribution Functions for the two mixture densities when k = 0.5

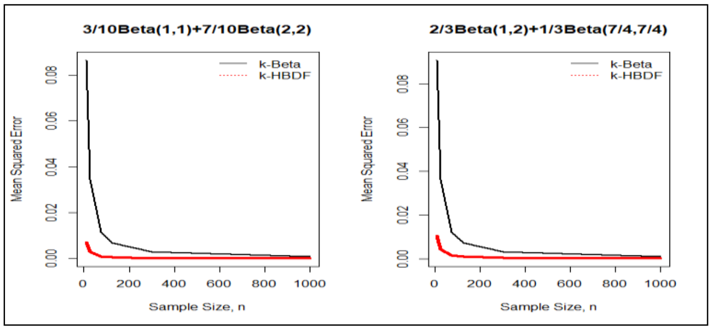

Fig. 10: Graph of the MSE of $k-$Higher-order Beta Distribution and $k-$Beta Distribution Functions for the two mixture densities when $k = 1$

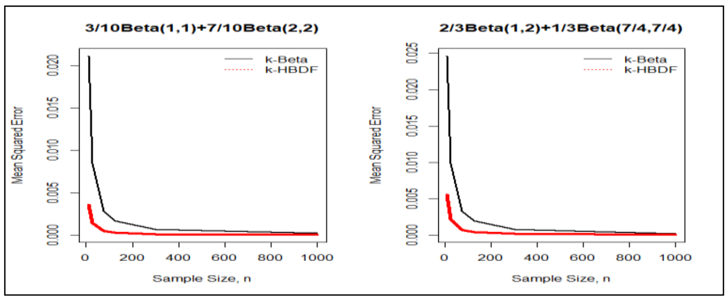

Fig. 11: Graph of the MSE of $k-$Higher-order Beta Distribution and $k-$Beta Distribution Functions for the two mixture densities when $k = 1.5$

B. REAL LIFE APPLICATIONS

In this subsection, a demonstration of the real life applications of $k-\text{HBDF}$ **is** carried out. Equation (12) is used for the two data sets. The first data set is the first 58 observations recorded from the failure times of Kevlar 49/epoxy strands when the pressure is at 90% stress level [3]. The $k-\text{HBDF}$ and $k-\text{Beta}$ distributions are used to fit the data set. The estimates of the parameters α and β when $k = 0.5$ were obtained by using the modified R-Code of [9]. The result of this fit is presented in Table 1 and Figure 12 below.

Table 1: Method of moments fit of the failure times data

Distribution	$k-\text{Beta}$	$k-\text{HBDF}$
Parameter estimates	α =1.57939 β =1.30348	α =2.49065 β =2.76633
MSE	1.32829	1.58894e-35
RMSE	1.15252	3.98616e-18
MAD	0.13078	0.00326178

From Table 1, it can be seen that $k-\text{HBDF}$ and $k-\text{Beta}$ distributions gave adequate fit for the data set. However, $k-\text{HBDF}$ distribution provides the best fit. This is buttressed by its lower MSE as compared with $k-\text{Beta}$ distribution.

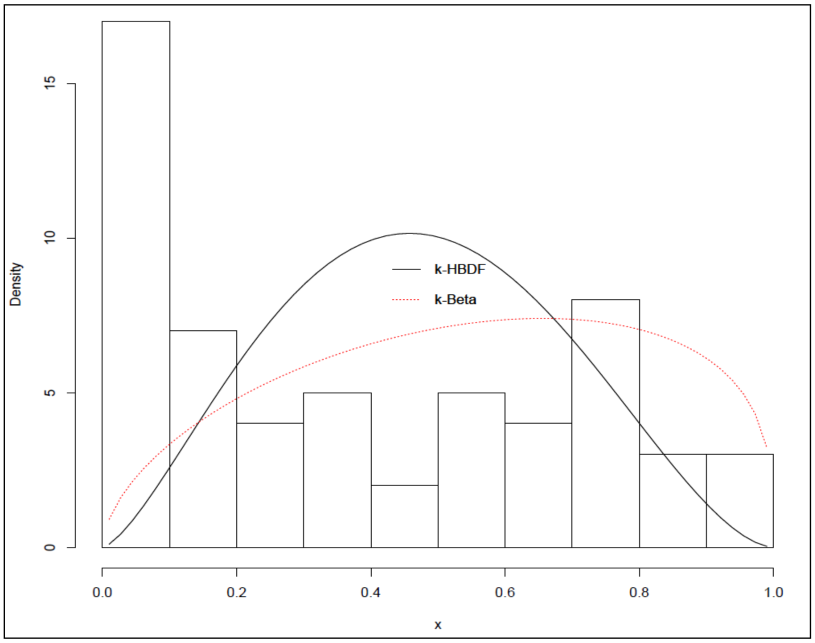

Fig. 12: Histogram and fitted densities of the failure times data

The second data set is the set of 272 observations of the old faithful gersey eruption data extracted from Old Faithful Geyser Data – CMU Statistics available at https:www.stat.cmu.edu>~larry. This data was converted to cumulative relative and was used for the fiting in Table 2 and Figure 13 below.

Table 2: Method of moments fits of the cumulative relative version of the old faithful geyser eruption data

Distribution	$k-$Beta	$k-$HBDF
Parameter estimates	α =1.49375 β =1.93338	α =2.26151 β =3.32982
MSE	0.69705	1.06661e-34
RMSE	0.83489	1.03277e-17
MAD	0.05144	0.0057433

The results from Table 2 signify the flexibility and superiority of $k-$HBDF distribution over the $k-$Beta distribution since it posses lower MSE. This is clearly exhibit in Figure 13.

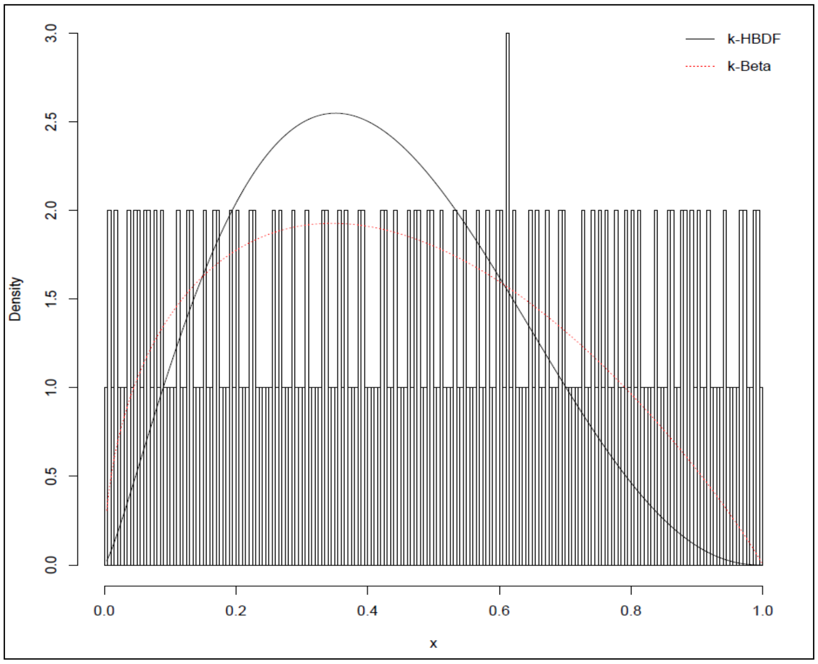

Fig. 13: Histogram and fitted densities of the Old faithful geyser eruption (cumulative relative) data

C. SUMMARY OF RESULTS

This section discusses the summary of results obtained from the Monte-Carlo simulation experiment as well as results arising from this study. Specific outcomes of the Monte-Carlo simulation for the two scenarios and real life applications are presented below.

In the first instance, the *k*-HBDF has relatively low mean squared error for both the small and large sample sizes for the standard beta distribution considered as compared with the *k*-Beta distribution proposed by [10] (see Figures 6 to 8). Another major finding of this work is that the *k*-HBDF has relatively low mean squared error for both the small and large sample sizes for the mixture beta distribution considered as compared with the *k*-Beta distribution proposed by [10] (see Figures 9 to 11).

Another major finding of significance in this work is that the *k*-HBDF has relatively low mean squared error for both the cases k = 0.5, k = 1and k = 1.5 for the standard beta distribution considered as compared with the *k*-Beta distribution proposed by [10] (see Figures 6 to 8). Closely related to the above is that the *k*-HBDF has relatively low mean squared error for both the cases k = 0.5, k = 1 and k = 1.5 for the mixture beta distribution considered as compared with the *k*-Beta distribution proposed by [10] (see Figures 9 to 11).

In addition to the above, it is observed that the *k*-HBDF and *k*-beta distribution proposed by [9] both competes favourably well for a single and mixture density. However, the MSE tends to be consistent at the mixture density than the single density (see Figures 6 through 11).

Another notable result from the experiment (using the real life data - $n = 58$ and $n = 272$) shows that the $k-$HBDF distribution has relatively low mean squared error as compared with the $k-$Beta distribution (see Tables 1 and 2). The importance of the current research is further corroborated by the discovery that the $k-$HBDF and $k-$Beta distributions provided the best fit for the data sets (see Figures 12 and 13).

Apart from the above, another finding of significance from the experiment is that the $k-$HBDF and $k-$Beta distributions both compete favourably well for the simulated data and real life data. However, the MSE tends to be consistent for $k-$HBDF . Furthermore, the *k*-HBDF also satisfies the skewness property since it is symmetric when $\alpha = \beta$, skewed to the right when $\alpha > \beta$ and skewed to the left when $\alpha < \beta$ (see Figures 1 to 5). Lastly, in all, the proposed *k*-HBDF is a better probability estimator compared to the *k*-Beta distribution proposed by [9] in terms of MSE.

V. CONCLUSION

In this work, a new probability distribution known as the *k*-HBDF is proposed. The properties of the proposed *k*-HBDF were derived and efficiency of the *k*-HBDF was compared with the *k*-beta (classical when k =1) distribution function through the Monte-Carlo simulation experiment. Results obtained signify that the proposed *k*-HBDF has smaller MSE compared to *k*-beta and hence, it is more efficient than the *k*-beta and classical beta distribution functions.

REFERENCES

[1] K. M. Aludant(2018), On the Beta Cumulative Distribution Function, *Applied Mathematical Sciences,* Vol. 12, No. 10, 461 – 466.
[2] A. Alzaatreh, C. Lee and F. Famoye(2014),T-normal family of distributions: a new approach to generate the normal distribution, *Journal of Statistical Distribution and Applications*,1:6,1-8, http://www.jsdajournal.com/content/1/1/16
[3] D. F. Andrews and A. M. Herzberg(1985), *Data*: *A Collection of Problems from Many Fields for the Student and Research Worker* (Springer Series in Statistics), Springer, New York.
[4] N. Eugene, C.Lee and F. Famoye(2002), Beta-Normal Distribution and its Applications, *Communication Statistical Theory Method*, Vol. 31(4), 497 – 512.
[5] M. C. Jones and P. J. Foster (1993), Generalized Jackknifing and higher-order kernels, Journal *of Nonparametric Statistics* 3, 81 – 94.
[6] S. Kotz and D. Vicari (2005), Survey of developments in the theory of continuous skewed distributions, *Metron*, 63, 225 – 261.
[7] C. Lee, F. Famoye and A. Y. Alzaatreh (2013), Methods for generating families of univariate continuous distributions in the recent decades, *WIREs Computational Statistics*, 5, 219 – 238.
[8] J. E. Lighter (1991), A Brief Look at the History of Probability and Statistics, *The Mathematics Teacher*, Vol. 84, No. 8, 623 – 630.
[9] C. E. B. Owen (2008), Parameter estimation for the beta distribution, *All Theses and Dissertations*-1614, https://scholararchive-byu.edu/etd/1614.
[10] G. Rahman, S. Mubeen, A. Rehman and M. Naz (2014), On *k*-Gamma and *k*-Beta Distribution and Moments Generating Functions, *Journal of Probability and Statistics*, Vol. 2014, Article ID:982013, 6 pages, http://dx.doi.org/10.1155/2014/982013.
[11] D. D. Wackerly, W. Mendenhall III and R. L.Scheaffer (2008), *Mathematical Statistics with Applications*, The Thomson Corporation, USA.