



## Research Article

# Statistical Properties and Application of Lindley–Kumaraswamy Distribution

<sup>1</sup>Baba Bukar Alhaji and <sup>2</sup>Umaru Haruna

<sup>1</sup>Department of Mathematical Sciences, Nigerian Defence Academy, Kaduna, Nigeria

<sup>2</sup>Department of Mathematical Sciences, Kaduna State University, Kaduna

\*Corresponding author: [bbukar@nda.edu.ng](mailto:bbukar@nda.edu.ng), [doi.org/10.55639/607.838281](https://doi.org/10.55639/607.838281)

### ARTICLE INFO:

#### Keywords:

Distribution,  
Lindley,  
Kumaraswamy,  
Maximum  
Likelihood  
G family

### ABSTRACT

In this study, we proposed an extension of the Kumaraswamy distribution by adding a parameter to the Kumaraswamy distribution using the Lindley-G family of distributions proposed by Cakmakyapan and Ozel (2016). The model parameters were estimated using Maximum Likelihood Estimation (MLE). The performance of the new Lindley-Kumaraswamy distribution was compared with other family of Kumaraswamy distribution. The performance was tested by an application to a real-life data and compared with other family of Kumaraswamy distributions using the Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC) which are two types of information criterion. Our new distribution was found to provide a better fit than the existing distributions when modeling positively skewed data.

**Corresponding author:** Baba Bukar Alhaji **Email:** [bbukar@nda.edu.ng](mailto:bbukar@nda.edu.ng)  
Department of Mathematical Sciences, Nigerian Defence Academy, Kaduna, Nigeria

### INTRODUCTION

Mathematical flexibility and simplicity of probability distributions are factors to consider in the choice of a statistical model otherwise known as distribution. Experts in probability distribution theory have connected these important features

in their choice of distribution for statistical research. This connection had opened another investigation where many researchers put some choosing standard theoretical distributions that best fits real life phenomena.

Lindley (1958) was the first to ever conceptualize the idea of the Lindley distribution. A Lindley distribution is one way to portray the duration of a process or device. Since its formulation the distribution has been used in a wide variety of fields, including engineering, medicine and biology. Ghitany et al. (2008) stated its usefulness especially for modeling in mortality studies. Other

$$G(x) = 1 - \left[ 1 + \frac{\theta x}{\theta + 1} \right] e^{-\theta x} \quad (1)$$

$$g(x) = \frac{\theta^2}{\theta + 1} (1 + x) e^{-\theta x} \quad (2)$$

for interval  $[0, 1]$ , for  $x > 0$ ,  $\theta > 0$  where  $\theta$  represent scale parameter of the Lindley distribution.

The Kumaraswamy distribution describes a distribution within which outcomes are restricted to a selected interval, the characteristics of probability density function inside this interval depends upon the shape parameters. There is a commonality between Kumaraswamy distribution and another distribution called Beta distribution (Handique and Chakraborty, 2016). But Kumaraswamy is probably easier to use due to its easier analytical expressions both for its probability density function and accumulative distribution function.

The Kumaraswamy distribution was originally conceptualized by Poondi Kumaraswamy in 1980 for random processes that are doubly bounded for applications that are hydrological. The capacity to simulate monotone and non-

$$G(x) = 1 - (1 - x^\alpha)^\beta \quad (3)$$

Also supplied is the matching probability density function pdf.:

works which provide more details of Lindley distribution seem conspicuous in the work of Bakouch et al. (2011.), Bhati et al. (2015), Shanker et al. (2015), Hamed and Alzaghal (2021) amongst others. The cumulative distribution function (cdf) and the probability density function (pdf) respectively offer a mathematical definition of the Lindley distribution as:

monotone failure rate functions also lies in this distribution whose commonality in life-time data analysis and consistency appear quite often. Recently, several researches have been carried on different families of Kumaraswamy distributions. For example, Marcelino *et al.* (2011) introduced and studied the Kumaraswamy generalized gamma distribution which will be used in to model bathtub shaped hazard rate functions. For more applications on Kumaraswamy family of distribution, see for example, Bourguignon *et al.* (2012), Santana *et al.* (2012), Mansour *et al.* (2018), Varghese and Jose (2022), among others.

The Kumaraswamy distribution is denoted by  $K(\alpha, \beta)$  and belongs to the family of continuous probability distributions outlined within the interval  $[0, 1]$  with a cumulative distribution function *cdf* given by:

$$g(x) = \alpha\beta x^{\alpha-1} (1-x^\alpha)^{\beta-1} \tag{4}$$

For interval  $0 \leq x \leq 1$ , where  $\alpha$  and  $\beta$  indicate the shape characteristics of the Kumaraswamy distribution.

**Motivation**

Amongst the flaws in statistical analysis is that most accessible data sets do not adhere appropriately many of the current and popular conventional probability distribution(s) and for that cannot provide an adequate modelling. Our effort in this study is meant to propose a different distribution namely Lindley-Kumaraswamy distribution using Lindley-G family of distributions by adding a parameter in order to obtain a flexible compound distribution which is better than the existing ones. A data set has been used to demonstrate the utility of the new distribution. The distribution has both shape and scale parameters which also make it flexible to model tailed distribution. The study of this distribution is significant, has various

uses in areas which involve tailed distribution problems and in engineering, such as in hydrology, and other fields of study. Also, its extension provides a flexible model fitting among its family.

**Organization of the Study**

The work is organized as follows. Section 2 provides materials and methodology of the research. The data analysis is presented in section 3. Finally, section 4 discussion of findings and section 5 brings to conclusion.

**METHODS AND MATERIALS**

**Lindley-Kumaraswamy Distribution**

Cakmakyapan and Ozel (2016) proposed a Lindley-Generalized distribution that provides more flexibility in modeling of real-life data. Now, the *cdf* and *pdf* of the Lindley-G family of distributions according to Cakmakyapan and Ozel (2016) were derived from the transformation given by;

$$F(x) = \int_{-\infty}^{-\log[1-G(x)]} \frac{\theta^2}{\theta+1} (1+t) e^{-\theta t} dt \tag{5}$$

Now, the *cdf* is derived from the transformation by integration by parts as:

$$F(x) = 1 - \left[ 1 - \frac{\theta}{(\theta+1)} [\log(1-G(x))] \right] [1-G(x)]^\theta \tag{6}$$

*pdf* is obtained by differentiating the *cdf* as:

$$f(x) = \frac{\theta^2}{(\theta+1)} g(x) [1 - \log(1-G(x))] [1-G(x)]^{\theta-1} \tag{7}$$

respectively. For  $x > 0$ ,  $\theta > 0$ ,  $G(x)$  and  $g(x)$  represent the cumulative distribution function (cdf) and the probability distribution function (pdf) of the **baseline** or **parent** distribution

respectively and  $\theta$  is the scale parameter of the Lindley-G family of distribution. Substituting equations (3) and (4) in (6) and (7) respectively, the cumulative distribution function (cdf) and probability distribution function (pdf) of

Lindley-Kumaraswamy distribution respectively are obtained thus:

$$\begin{aligned}
 F(x) &= 1 - \left[ 1 - \frac{\theta}{(\theta+1)} \left[ \log(1 - G(x)) \right] \right] [1 - G(x)]^\theta \\
 F(x) &= 1 - \left[ 1 - \frac{\theta}{(\theta+1)} \left[ \log \left( 1 - \left( 1 - (1 - x^\alpha)^\beta \right) \right) \right] \right] \left[ 1 - \left( 1 - (1 - x^\alpha)^\beta \right) \right]^\theta \\
 F(x) &= 1 - \left[ 1 - \frac{\theta}{(\theta+1)} \left[ \log(1 - x^\alpha)^\beta \right] \right] (1 - x^\alpha)^{\theta\beta}
 \end{aligned}
 \tag{8}$$

and

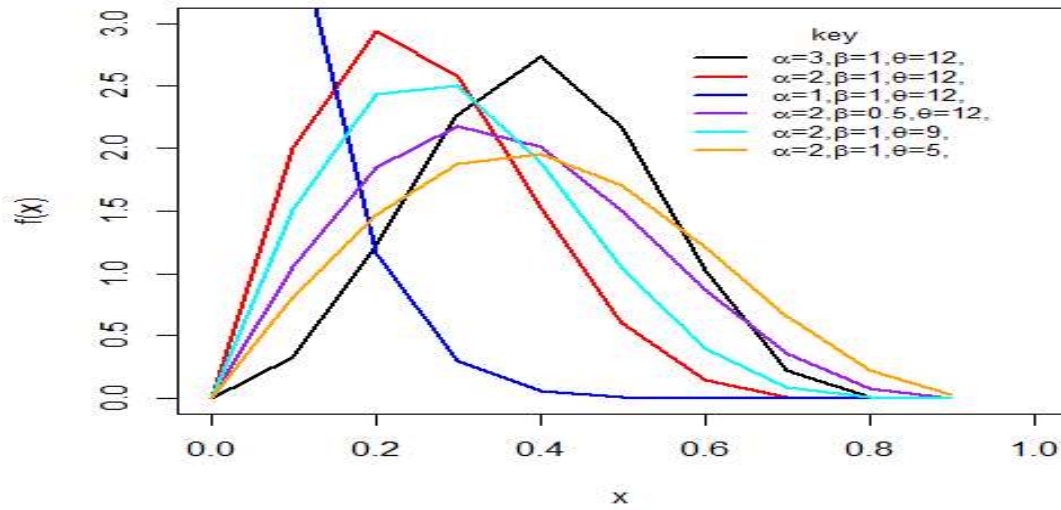
$$\begin{aligned}
 f(x) &= \frac{\theta^2}{(\theta+1)} g(x) [1 - \log(1 - G(x))] [1 - G(x)]^{\theta-1} \\
 f(x) &= \frac{\theta^2}{(\theta+1)} \left( \alpha\beta x^{\alpha-1} (1 - x^\alpha)^{\beta-1} \right) \left[ 1 - \log \left( 1 - \left( 1 - (1 - x^\alpha)^\beta \right) \right) \right] \left[ 1 - \left( 1 - (1 - x^\alpha)^\beta \right) \right]^{\theta-1} \\
 f(x) &= \frac{\alpha\beta\theta^2}{(\theta+1)} x^{\alpha-1} (1 - x^\alpha)^{\beta(\theta-1)} [1 - \log(1 - (1 - (1 - x^\alpha)^\beta))]
 \end{aligned}
 \tag{9}$$

respectively.

For  $0 \leq x \leq 1$ , where  $\alpha, \beta$  are shape parameters and  $\theta$  is a scale parameter. Hence equations (8) and (9) represent the *cdf* and *pdf* of the proposed Lindley-Kumaraswamy Distribution.

**Graphical Presentation of the *pdf* and *cdf***

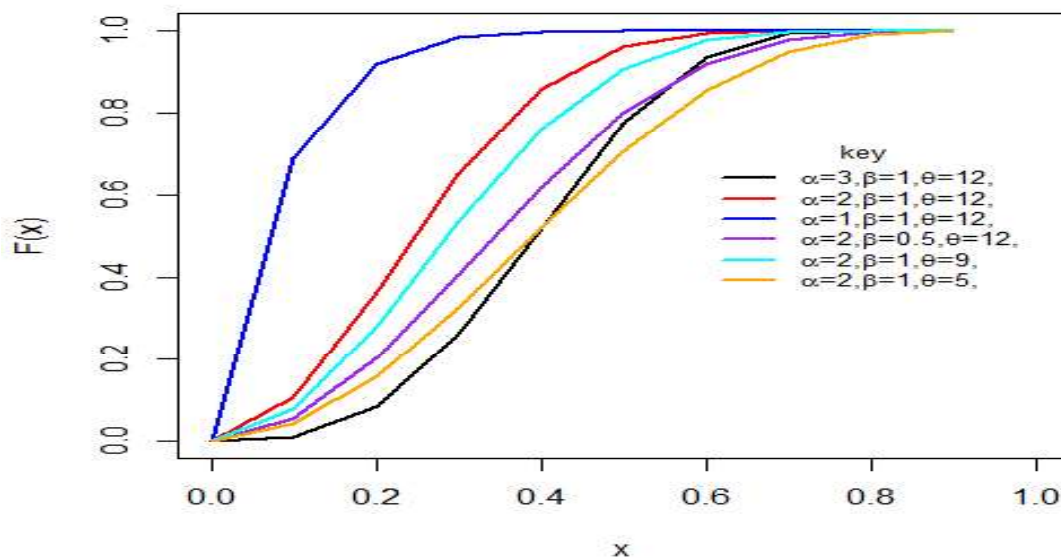
Considering a certain value for parameters  $\alpha, \beta$ , and  $\theta$ , we prepared some potential figures for the probability distribution function (*pdf*) and the cumulative distribution function (*cdf*) of the LinKumD as demonstrated below:



**Figure 1:** Shape of the probability distribution function (pdf) of the projected distribution for various true parameter values.

**Interpretation:** The figure above shows that the growth of the cdf is in direct proportionality with increase in  $x$  and approaches 1 when  $x$  gets bigger. This

means that the distributions are extremely valuable for data sets with many different shapes.



**Figure 2:** Shape of the *cdf* of the proposed distribution for different parameters values

**Interpretation** The figure above indicates, as expected, that the cdf rises when  $x$  rises and approaches 1 as  $x$  rises.

**Estimation of Parameters of the LinKumD**

Let  $X_1, X_2, \dots, X_n$  be independently and identically distributed random variables

$$f(x) = \frac{\alpha\beta\theta^2}{(\theta + 1)} x^{\alpha-1}(1 - x^\alpha)^{\beta(\theta-1)} [1 - \log(1 - (1 - (1 - x^\alpha)^\beta)]$$

The probability function is defined as follows:

$$L(X/\alpha, \beta, \theta) = \left(\frac{\alpha\beta\theta^2}{\theta+1}\right)^n \prod_{i=1}^n \left\{ x_i^{\alpha-1} (1 - x_i^\alpha)^{\beta(\theta-1)} \left[ 1 - \log \left( 1 - (1 - (1 - x_i^\alpha)^\beta) \right) \right] \right\} \tag{10}$$

Let the log-likelihood function,  
 $l = \log L(X/\alpha, \beta, \theta)$ .

Therefore,

$$l = n \log \beta + 2n \log \theta - n \log(\theta + 1) + (\alpha - 1) \sum_{i=1}^n \log x_i + (\theta\beta - 1) \sum_{i=1}^n \log(1 - x_i^\alpha) + \sum_{i=1}^n \log \left[ 1 - \log \left( 1 - (1 - (1 - x_i^\alpha)^\beta) \right) \right] \tag{11}$$

After differentiating  $l$  partially with respect to  $\beta$  and  $\theta$  respectively gives;

$$\frac{\partial l}{\partial \beta} = \frac{n}{\beta} + \theta \sum_{i=1}^n \log(1 - x_i^\alpha) + \sum_{i=1}^n \left\{ \frac{(1 - x_i^\alpha)^\beta \ln(1 - x_i^\alpha)}{\left[ 1 - \log(1 - x_i^\alpha)^\beta \right] (1 - x_i^\alpha)^\beta} \right\} \tag{12}$$

$$\frac{\partial l}{\partial \theta} = \frac{2n}{\theta} - \frac{n}{(\theta+1)} + \beta \sum_{i=1}^n \log(1 - x_i^\alpha) \tag{13}$$

The solution of the non-linear system of equations;  $\frac{\partial l}{\partial \beta} = 0$  and  $\frac{\partial l}{\partial \theta} = 0$  will provide us with the maximum likelihood parameter estimates of  $\beta$  and  $\theta$ . When data sets are available, the solution can only be computed numerically using proper statistical tools such as R, SAS, Maple, and so on.

**DATA ANALYSIS AND RESULTS**

This section compares the performance of the Lindley-Kumaraswamy Distribution (LinKumD) to that of two other models such as Transmuted

of sample size  $n$  of the LinKumD with previously defined unknowns  $\alpha$ ,  $\beta$ , and  $\theta$ . Equation (10) gives the pdf of the LinKumD as:

Kumaraswamy distribution (TKumD) and Kumaraswamy distribution (KumD) using a real dataset. We used **R statistical package** for all the analysis in

this research work. The density functions of these distributions are depicted below:

**The Lindley-Kumaraswamy Distribution (LinKumD)**

The pdf of the LinKumD obtained from equation (9) is as follows:

:

$$f(x) = \frac{\alpha\beta\theta}{(\theta+1)} x^{\alpha-1} (1-x^\alpha)^{\beta(\theta-1)} [1 - \log(1 - (1 - (1-x^\alpha)^\beta))] \tag{14}$$

**The Transmuted Kumaraswamy Distribution (TKumD)**

Khan *et al.* (2016) defined the *pdf* of the TKumD as follows:

$$f(x; \alpha, \theta, \lambda) = \alpha\theta x^{\alpha-1} (1-x^\alpha)^{\theta-1} \left\{ 1 + \lambda - 2\lambda \left( 1 - (1-x^\alpha)^\theta \right) \right\} \tag{15}$$

**The Kumaraswamy Distribution (KumD)**

According to Kumaraswamy (1980), the *pdf* of the Kumaraswamy distribution is given as:

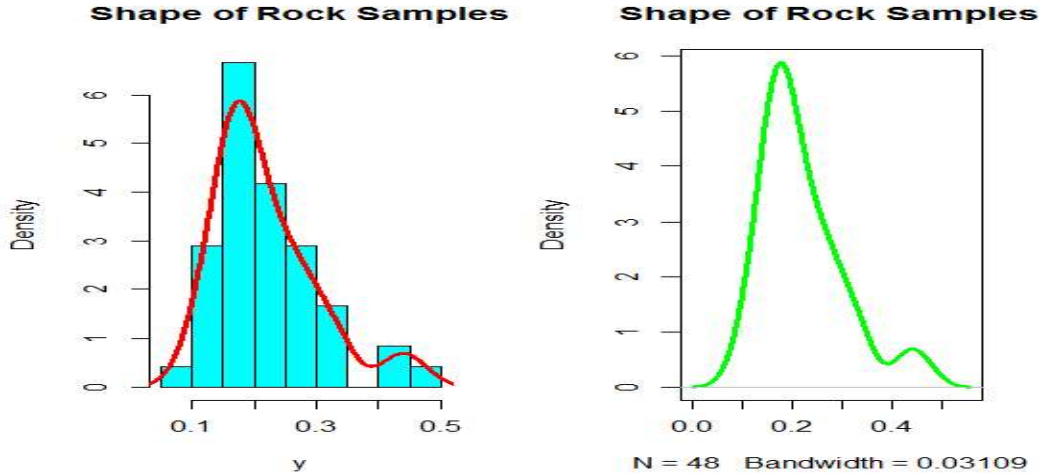
$$f(x; a, b) = ab x^{a-1} (1-x^a)^{b-1} \tag{16}$$

**Table 1:** Shape measurements of 48 rock samples from a petroleum reservoir

|           |          |          |          |          |          |          |
|-----------|----------|----------|----------|----------|----------|----------|
| 0.0903296 | 0.189651 | 0.228595 | 0.200071 | 0.280887 | 0.311646 | 0.176969 |
| 0.464125  | 0.148622 | 0.164127 | 0.231623 | 0.144810 | 0.179455 | 0.276016 |
| 0.438712  | 0.420477 | 0.183312 | 0.203654 | 0.172567 | 0.113852 | 0.191802 |
| 0.19753   | 0.163586 | 0.200744 | 0.117063 | 0.162394 | 0.153481 | 0.291029 |
| 0.133083  | 0.326635 | 0.253832 | 0.262651 | 0.122417 | 0.150944 | 0.204314 |
| 0.240077  | 0.225214 | 0.154192 | 0.328641 | 0.182453 | 0.167045 | 0.148141 |
| 0.262727  | 0.161865 | 0.341273 | 0.276016 | 0.230081 | 0.200447 |          |

This data was extracted from BP research, image analysis by Ronit Katz, University of Oxford and has been used for analysis by Javanshiri *et al.* (2015)

We also provide histogram and density for the data set as shown in the Figure 3 below:



**Figure 3:** A Histogram and Density Plots for the Rock Sample Dataset

From the descriptive statistics, the histogram and the density plot shows the data set as observed is right-skewed and thus appropriate for distributions that are skewed most especially to the right.

**Model Selection Criteria**

To test the performance of the above listed distributions, we used the following information criteria: The Akaike Information Criterion (AIC) and the Bayesian Information Criterion (BIC). The statistics are as follows:

$$AIC = -2ll + 2k$$

$$BIC = -2ll + k \log(n)$$

Where  $ll$  denotes the log-likelihood function,  $k$  represents the number of model parameters and  $n$  the sample size.

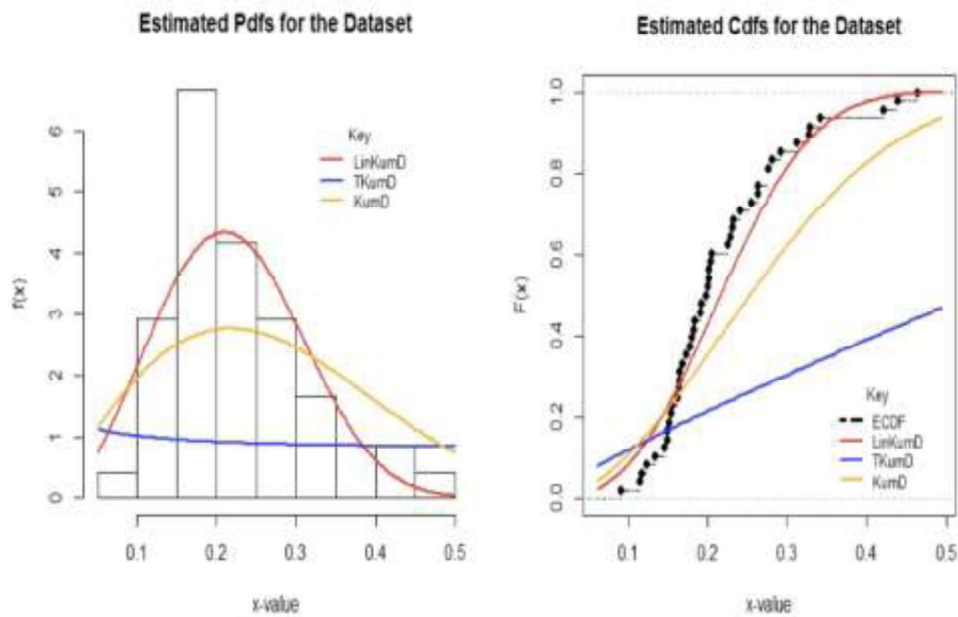
**Bench mark:** The model with the lowest value for these statistics is thought to be the best fit for the data.

Estimates and performance measures of the above distributions are listed in Table 2 for the rock sample data set as follows:

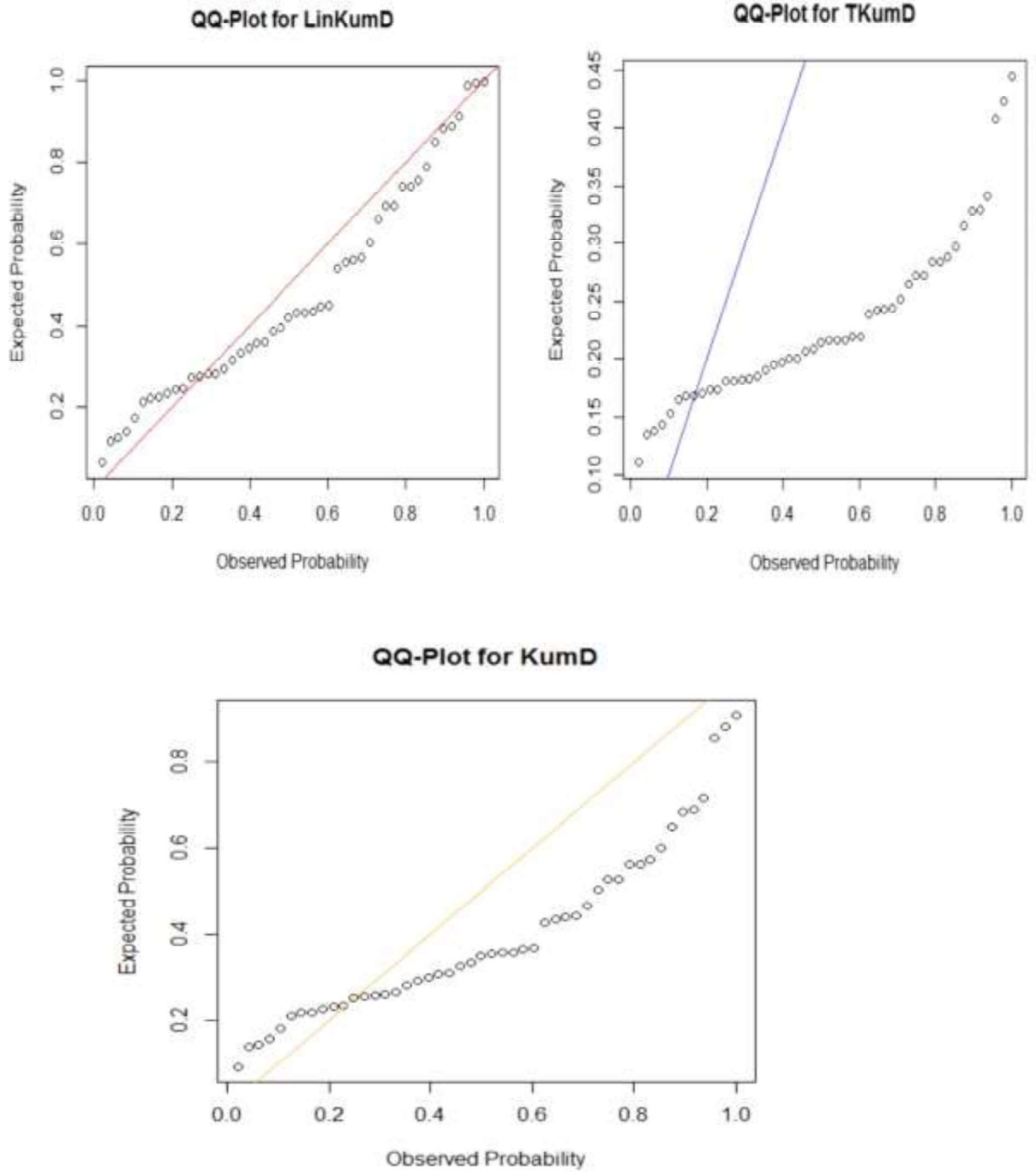


**Table 2:** The performance of the distribution was calculated using R to obtain the AIC and BIC values of the models based on the data set.

| Models  | Estimated Parameter Values   | Log-likelihood Values | Akaike Information Criterion (AIC) | Consistent Akaike Information Criterion (CAIC) | Bayesian Information Criterion (BIC) | Ranks |
|---------|--|-----------------------|------------------------------------|--|--------------------------------------|-------|
| LinKumD | $\alpha = 2.663380$<br>$\beta = 5.009926$<br>$\theta = 9.004390$                       | -52.38064             | -98.7612                           | -98.21583                                      | -93.14768                            | 1     |
| TkumD   | $\hat{\alpha} = 0.8324228$<br>$\hat{\beta} = 0.5579018$<br>$\hat{\lambda} = 0.4570070$ | 4.269415              | 1453883                            | 15.08428                                       | 20.15243                             | 3     |
| KumD    | $\hat{\alpha} = 1.894405$<br>$\hat{\beta} = 9.051699$                                  |                       | -81.4609                           | -81.19424                                      | -77.7185                             | 2     |



**Figure 4:** Histogram shows the estimated density plots, as well as the CDFs and PDFs of the fitted distributions to data set.



**Figure 5:** Probability plots for the fit of the LinKumD, TKumD and KumD based on the data set

## DISCUSSIONS

The MLEs of the parameters for the fitted models for the data set are provided in Table 2. The AIC and BIC statistics for the fitted distributions are also included in the rock samples data set table. Based on the data sets and model selection values, it is obvious that the proposed LinKumD with three parameters outperforms the TKumD and the standard KumD. Furthermore, the estimated pdfs and cdfs in figure 2 and the Q-Q plots in figure 3 clearly support and confirm the results in Tables 2 and 3.

## CONCLUSION

This research work proposed a new extension of the Kumaraswamy distribution called Lindley-Kumaraswamy distribution. We estimated the parameters by using the method of Maximum Likelihood Estimation (MLE). We demonstrated the performance of the new distribution using real life data sets. Results of the model indicated that Lindley-Kumaraswamy distribution outperforms Transmuted Kumaraswamy and Kumaraswamy distributions.

## REFERENCES

- Bakouch, H. S., Bander, M. A., Al-Shaomrani, A. A., Marchi, V. A. A., and Louzada, F. (2012). An ex-tended Lindley distribution Bakouch. *Journal of the Korean Statistical Society*, 41, 75-85.
- Bhati, D., Malik M. A., and Vaman H. J. (2015). Lindley-Exponential Distribution: Properties and Applications. *Metron*, 73(3): 335-357.
- Bourguignon, M. R.B., Silva, L. M., and Cordeiro, G. M., (2012). The Kumaraswamy Pareto distribution, *AMS 2000 Subject Classifications*: 62E10, 62E15, 62E99
- Cakmakyapan, S., and Ozel G. (2016). The Lindley Family of Distributions: Properties and Applications. *Hacetatepe Journal of Mathematics and Statistics*, 46, 1-27, doi: 10.15672/hjms.201611615850.
- Ghitany M. E., Atieh B., and Nadarajah, S. (2008). Lindley Distribution and Its Application. *Mathematics and Computers in Simulation*, 78 (4): 493-506.
- Hamed D. and Alzaghal A. (2021), New class of Lindley distributions: properties and applications. *Journal of Statistical Distributions and Applications* 8:11, doi.org/10.1186/s40488-021-00127-y.
- Handique, L., and Chakraborty, S. (2016). Beta generated Kumaraswamy-G and other new families of distributions. *Mathematics and Statistics Theory*, arXiv:1603.00634.
- Javanshiri, Z., Habibi, R. A. and Arghami, N. R. (2015). Exp-Kumaraswamy distributions: Some properties and Applications. *Journal of Sciences, Islamic Republic of Iran*, 26(1): 57-69.
- Khan, M. S., King, R. and Hudson, I. L. (2016). Transmuted kumaraswamy distribution. *Statistics in Transition*, 17(2): 183-210.
- Kumaraswamy, P. (1980). Generalized probability density-function for double-bounded random-processes. *Journal of Hydrology*, 46: 79-88.
- Lindley, D.V. (1958). Fiducial distributions and Bayes' theorem. *Journal of Royal*

- Statistics*, Soc. Series B, 20, 102-107.
- Mansour M. M., Aryal G., Afify, A. Z. and Ahmad M. (2018). The Kumaraswamy exponentiated Frechet distribution. *Pakistan Journal of Statistics*, 34(3), 177-193.
- Marcelino, A.R., Edwin, M.M., and Cordeiro, G. M. (2011). The Kumaraswamy generalized gamma distribution with application in survival analysis. *Statistical Methodology*, 8: 411–433.
- Santana, T. V., Ortega, E. M. M., Cordeiro, G. M. and Silva, G. (2012). The Kumaraswamy Log-logistic distribution. *Journal of Statistical Theory and Applications*, 11(3): 265-291.
- Shanker R., Fesshayee H., and Selvaraj, S. (2015). On Modeling of Lifetimes Data Using Exponential and Lindley Distributions. *Biometrics & Biostatistics International Journal*, 2 (5): 00042. DOI: 10.15406/bbij.2015.02.00042.
- Varghese J, and Jose KK. (2022). Kumaraswamy harris generalized kumaraswamy distribution and its Application in Survival Analysis. *Biom Biostat Int J*. 11(1):28–34. DOI: 10.15406/bbij.2022.11.00351