

Odoma Distribution and Its Application

Abstract – In this study, a new continuous one parameter lifetime distribution is proposed. Its mathematical properties such as, moments, order statistics, entropy, survival function, hazard rate function and mean residual life function are derived. The new distribution is applied to real life data from engineering and the method of maximum likelihood is used to estimate the parameter. The goodness of fit of the new distribution is compared with that of some competing distributions and the results show its better fit to the data than the competing distributions.

Keywords: Moment Generating Function, Rényi Entropy, Order Statistics, Reliability Analysis, Maximum Likelihood Estimation.

1 Introduction

The study of lifetime data is considered important in research areas such as engineering, biomedical sciences, environmental sciences and actuarial science. Statistical analyses of data sets depend heavily on the assumed distribution of the data for quality output. The actual distributions of different data sets vary in their shape, nature of hazard rate function, mean residual life etc. and as such efforts are constantly being made to develop models that give the best fit to data sets of interest. Parametric lifetime models such as exponential, gamma, Lindley and Weibull have often been used in literature to model lifetime data, with exponential and Lindley being popular for modeling engineering and biomedical lifetime data. However, studies have shown that the goodness of fit of the classical distributions to available data sets, in comparison with newer distributions, lend more credence to the search for new distributions with better fit.

Ghitany et al. [1] studied the mathematical properties of the Lindley distribution and found them to be more flexible than those of the exponential distribution. In fitting both distributions to the waiting times before service of 100 bank customers, the authors found the Lindley distribution to yield better fit than the exponential distribution. Generalization and extension of the Lindley distribution have been carried out by various researchers including Deniz and Ojeda [2], Nadarajah et al [3], Elbatal et al [4], Ghitany et al [5], Liyanage and Pararai [6] and Abouammoh et al [7]. Shanker [8] proposed a

one parameter distribution, Akash distribution, which yielded a better fit than the exponential and Lindley distributions on the basis of two real life data sets. Other proposed lifetime distributions in literature include the inverse Lindley distribution [9], extended Power Lindley distribution [10], weighted Akash distribution [11], Poisson-Akash distribution [12], Aradhana distribution [13], Sujatha distribution [14], Ishita distribution [15] and Pranav distribution [16].

By modifying and generalizing the well-known distributions, researchers have derived new probability distributions with better goodness of fit to some available data sets. Although new distributions derived from extending and modifying the classical distributions yield superior performances in some data sets, there is also a trade-off in using such distributions. New parameters are usually introduced in the course of deriving these new distributions and estimating the parameters of a distribution with many parameters is typically troublesome. Also, the method of maximum likelihood estimation may encounter problems which may result in numerical difficulties and lead to large standard errors [17].

Consequently, researchers have continued to search for the more mathematically tractable one parameter distributions that can yield good fit to available data sets. Some new one-parameter distributions have been derived and proposed by mixing exponential and gamma distributions using some mixing proportions. It is well known that Lindley distribution is a mixture of exponential (θ) and gamma (2, θ) with mixing proportions $\frac{\theta}{\theta+1}$ and $\frac{1}{\theta+1}$. Other one-parameter distributions have been

derived using this method and shown to perform better than the Lindley distribution in some data sets.

The Akash distribution is a mixture of exponential (θ) and gamma (3, θ) with mixing proportions

$\frac{\theta^2}{\theta^2+2}$ and $\frac{2}{\theta^2+2}$. Ishita distribution is a mixture of exponential (θ) and gamma (3, θ) with mixing

proportions $\frac{\theta^3}{\theta^3+2}$ and $\frac{2}{\theta^3+2}$. Sujatha distribution is a three-component mixture of exponential (θ),

gamma (2, θ) and gamma (3, θ) with mixing proportions $\frac{\theta^2}{\theta^2+\theta+2}$, $\frac{\theta}{\theta^2+\theta+2}$ and $\frac{2}{\theta^2+\theta+2}$. The

recently proposed Pranav distribution is a mixture of exponential (θ) and gamma (4, θ) with mixing

proportions $\frac{\theta^4}{\theta^4+6}$ and $\frac{6}{\theta^4+6}$. All four distributions namely; Akash, Ishita, Sujatha and Pranav have

been shown to outperform the exponential and Lindley distributions in some data sets.

In this study, we propose a new one-parameter distribution with its mathematical properties and show its superior goodness of fit to some real life data sets in comparison with some competing distributions.

2 The new distribution

The new one-parameter distribution is defined by

$$f(x; \theta) = \frac{\theta^5}{2(\theta^5 + \theta^3 + 24)} (2x^4 + \theta x^2 + 2\theta) e^{-\theta x}; \quad x > 0, \theta > 0 \quad (1)$$

The probability density function (p.d.f.) in (1) is a three-component mixture of exponential distribution with scale parameter θ , gamma distribution with shape parameter 3 and scale parameter θ and gamma distribution with shape parameter 5 and scale parameter θ . The mixing proportions are $\frac{\theta^5}{\theta^5 + \theta^3 + 6}$

and $\frac{\theta^3}{\theta^5 + \theta^3 + 1}$ such that

$$f(x; \theta) = p_1 g_1(x; \theta) + p_2 g_2(x; 3, \theta) + [1 - (p_1 + p_2)] g_3(x; 5, \theta)$$

Where $p_1 = \frac{\theta^5}{\theta^5 + \theta^3 + 6}$, $p_2 = \frac{\theta^3}{\theta^5 + \theta^3 + 1}$, $g_1 = \exp(\theta) = \theta e^{-\theta x}$, $g_2 = \text{gamma}(3, \theta) = \frac{1}{2} \theta^3 x^2 e^{-\theta x}$

and $g_3 = \text{gamma}(5, \theta) = \frac{1}{24} \theta^5 x^4 e^{-\theta x}$. We would call the new one-parameter distribution with p.d.f. as in (1) “Odoma distribution”.

The cumulative distribution function (c.d.f.) corresponding to the p.d.f. in (1) is given by

$$F(x; \theta) = 1 - \left[1 + \frac{\theta^2 x^2 (\theta^2 x^2 + 4\theta x + 12)}{\theta^5 + \theta^3 + 24} + \frac{\theta x (\theta^4 x + 2\theta^3 + 48)}{2(\theta^5 + \theta^3 + 24)} \right] e^{-\theta x}; \quad x > 0, \theta > 0 \quad (2)$$

The graphs of the p.d.f. and c.d.f. of Odoma distribution for varying values of the parameter are shown in Figures 1 and 2.

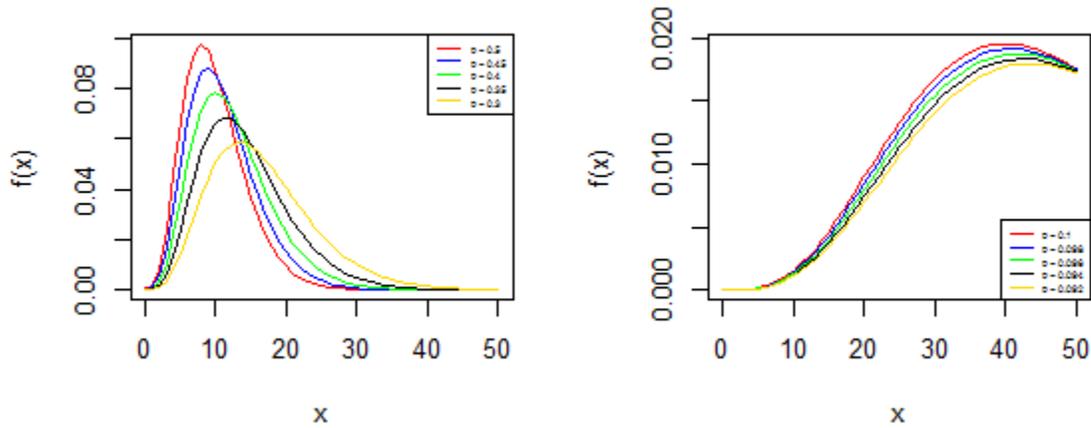


Figure 1. Graphs of the p.d.f. of Odoma distribution for varying values of the parameter

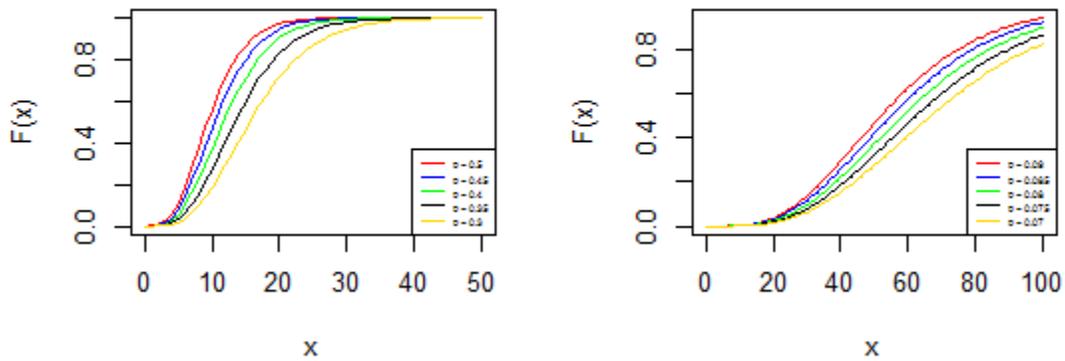


Figure 2. Graphs of the c.d.f. of Odoma distribution for varying values of the parameter.

3 Moment generating function

The moment generating function of Odoma distribution is

$$\begin{aligned}
 M_X(t) &= E(e^{tx}) = \int_0^{\infty} e^{tx} f(x; \theta) dx \\
 &= \frac{\theta^5}{2(\theta^5 + \theta^3 + 24)} \int_0^{\infty} (2x^4 + \theta x^2 + 2\theta) e^{-x(\theta-t)} dx \\
 &= \frac{\theta^5}{2(\theta^5 + \theta^3 + 24)} \left\{ \frac{24}{(\theta-t)^5} + \frac{\theta}{(\theta-t)^3} + \frac{\theta}{(\theta-t)} \right\}
 \end{aligned} \tag{3}$$

$$\begin{aligned}
&= \frac{\theta^5}{2(\theta^5 + \theta^3 + 24)} \left\{ \frac{24}{\theta^5} \sum_{k=0}^{\infty} \binom{k+4}{k} \left(\frac{t}{\theta}\right)^k + \frac{1}{\theta^2} \sum_{k=0}^{\infty} \binom{k+2}{k} \left(\frac{t}{\theta}\right)^k + \sum_{k=0}^{\infty} \left(\frac{t}{\theta}\right)^k \right\} \\
&= \sum_{k=0}^{\infty} \left(\frac{2(k+4)! + \theta^3(k+2)! + 2\theta^5 k!}{2\theta^k (\theta^5 + \theta^3 + 24)} \right) \frac{t^k}{k!}
\end{aligned} \tag{4}$$

Therefore, the r^{th} moment about origin, μ_r' , of Odoma distribution, which is the coefficient of $\frac{t^r}{r!}$ in

$M_X(t)$, is

$$\mu_r' = \frac{2(r+4)! + \theta^3(r+2)! + 2\theta^5 r!}{2\theta^r (\theta^5 + \theta^3 + 24)}; \quad r = 1, 2, 3, \dots \tag{5}$$

and the first four moments about origin are obtained as

$$\mu_1' = \frac{\theta^5 + 3\theta^3 + 120}{\theta(\theta^5 + \theta^3 + 24)}, \quad \mu_2' = \frac{2(\theta^5 + 6\theta^3 + 360)}{\theta^2(\theta^5 + \theta^3 + 24)}, \quad \mu_3' = \frac{6(\theta^5 + 10\theta^3 + 840)}{\theta^3(\theta^5 + \theta^3 + 24)}, \quad \mu_4' = \frac{24(\theta^5 + 15\theta^3 + 1680)}{\theta^4(\theta^5 + \theta^3 + 24)}.$$

4 Order statistics

Given a random sample $X_1, X_2, X_3, \dots, X_n$ of size n from Odoma distribution, let the corresponding order statistics be denoted by $Y_1 \leq Y_2 \leq Y_3 \leq \dots \leq Y_n$. Then for $r = 1, 2, 3, \dots, n$, the cumulative distribution function (c.d.f.) and probability density function (p.d.f.) of the r^{th} order statistics are respectively given by

$$F_{Y_r}(y) = \sum_{i=r}^n \binom{n}{i} [F(y; \theta)]^i [1 - F(y; \theta)]^{n-i} \tag{6}$$

and

$$f_{Y_r}(y) = \frac{n!}{(r-1)!(n-r)!} [F(y; \theta)]^{r-1} [1 - F(y; \theta)]^{n-r} f(y; \theta). \tag{7}$$

Substituting (1) and (2) in (6) and (7) accordingly, we obtain the c.d.f. and p.d.f. of the r^{th} order statistics of Odoma distribution as

$$F_{Y_r}(y) = \sum_{i=r}^n \binom{n}{i} \sum_{j=0}^i \binom{i}{j} \sum_{k=0}^{n-i} \binom{n-i}{k} \sum_{l=0}^k \binom{k}{l} (-1)^{j+l} \left[1 + \frac{\theta^2 x^2 (\theta^2 x^2 + 4\theta x + 12)}{\theta^5 + \theta^3 + 24} + \frac{\theta x (\theta^4 x + 2\theta^3 + 48)}{2(\theta^5 + \theta^3 + 24)} \right]^{j+l} e^{-x\theta(j+l)} \tag{8}$$

and

$$f_{Y_r}(y) = \frac{n!\theta^5(2x^4 + \theta x^2 + 2\theta)}{2(\theta^5 + \theta^3 + 24)(r-1)!(n-r)!} \sum_{i=0}^{r-1} \binom{r-1}{i} \sum_{j=0}^{n-r} \binom{n-r}{j} \sum_{k=0}^j \binom{j}{k} (-1)^{i+j+k} \left[1 + \frac{\theta^2 x^2 (\theta^2 x^2 + 4\theta x + 12)}{\theta^5 + \theta^3 + 24} + \frac{\theta x (\theta^4 x + 2\theta^3 + 48)}{2(\theta^5 + \theta^3 + 24)} \right] * e^{-\theta x(i+k+1)} \quad (9)$$

5 Entropy

Entropy is an important property of probability distributions and it measures the uncertainty in a probability distribution. The two popularly known forms of entropy are the Shannon entropy and Rényi entropy. Rényi entropy, which is widely used in literature, is considered in this work.

The Rényi entropy of a random variable, X, following the Odoma distribution is given by

$$\begin{aligned} T_R(\gamma) &= \frac{1}{1-\gamma} \log \left\{ \int_0^\infty f^\gamma(x; \theta) dx \right\} \quad \gamma > 0, \gamma \neq 1 \quad (10) \\ &= \frac{1}{1-\gamma} \log \left\{ \int_0^\infty \frac{\theta^{5\gamma}}{2^\gamma (\theta^5 + \theta^3 + 24)^\gamma} (2x^4 + \theta x^2 + 2\theta)^\gamma e^{-\theta\gamma x} dx \right\} \\ &= \frac{1}{1-\gamma} \log \left\{ \int_0^\infty \frac{(2\theta)^\gamma \theta^{5\gamma}}{2^\gamma (\theta^5 + \theta^3 + 24)^\gamma} \left(1 + \frac{x^4}{\theta} + \frac{x^2}{2} \right)^\gamma e^{-\theta\gamma x} dx \right\} \end{aligned}$$

Recall the binomial expansion, $(1+a)^n = \sum_{i=0}^n \binom{n}{i} a^i$. Applying the binomial expansion and simplifying, we obtain

$$T_R(\gamma) = \frac{1}{1-\gamma} \log \left\{ \sum_{i=0}^{\infty} \binom{\gamma}{i} \sum_{j=0}^{\infty} \binom{i}{j} \left(\frac{\theta^{6\gamma-j} 2^{j-i}}{(\theta^5 + \theta^3 + 24)^\gamma} \int_0^\infty x^{2(i+j)} e^{-\theta\gamma x} dx \right) \right\}.$$

Recall that $\int_0^\infty x^n e^{-ax} dx = \frac{\Gamma(n+1)}{a^{n+1}} = \frac{n!}{a^{n+1}}$. Therefore

$$\begin{aligned} T_R(\gamma) &= \frac{1}{1-\gamma} \log \left\{ \sum_{i=0}^{\infty} \binom{\gamma}{i} \sum_{j=0}^{\infty} \binom{i}{j} \left(\frac{\theta^{6\gamma-j} 2^{j-i}}{(\theta^5 + \theta^3 + 24)^\gamma} \right) \left(\frac{(2i+2j)!}{(\theta\gamma)^{2i+2j+1}} \right) \right\} \\ T_R(\gamma) &= \frac{1}{1-\gamma} \log \left\{ \sum_{i=0}^{\infty} \binom{\gamma}{i} \sum_{j=0}^{\infty} \binom{i}{j} \left(\frac{\theta^{6\gamma-2i-3j-1} 2^{j-i}}{(\theta^5 + \theta^3 + 24)^\gamma} \right) \left(\frac{(2i+2j)!}{\gamma^{2i+2j+1}} \right) \right\} \quad (11) \end{aligned}$$

6 Reliability

In this section, we present the survival function, the hazard rate function and the mean residual life function of the new distribution which are important functions used in reliability analysis.

A. Survival function

This is the probability that a system (such as mechanical system) or an organism survives beyond a specified time. It is usually computed as the complement of the cumulative distribution function of the probability density function in question. The survival function of Odoma distribution is given by

$$S(x; \theta) = 1 - F(x; \theta) \quad (12)$$

$$S(x; \theta) = 1 - \left[1 - \left\{ 1 + \frac{\theta^2 x^2 (\theta^2 x^2 + 4\theta x + 12)}{\theta^5 + \theta^3 + 24} + \frac{\theta x (\theta^4 x + 2\theta^3 + 48)}{2(\theta^5 + \theta^3 + 24)} \right\} e^{-\theta x} \right]$$

$$S(x; \theta) = \left\{ 1 + \frac{\theta^2 x^2 (\theta^2 x^2 + 4\theta x + 12)}{\theta^5 + \theta^3 + 24} + \frac{\theta x (\theta^4 x + 2\theta^3 + 48)}{2(\theta^5 + \theta^3 + 24)} \right\} e^{-\theta x} \quad (13)$$

B. Hazard rate function

This is the instantaneous rate of failure of a system or an organism at a specified time given that it has survived up to that time. The hazard rate function of Odoma distribution is given by

$$h(x; \theta) = \frac{f(x; \theta)}{1 - F(x; \theta)} = \frac{f(x; \theta)}{S(x; \theta)} \quad (14)$$

$$h(x; \theta) = \frac{\theta^5 (2x^4 + \theta x^2 + 2\theta) e^{-\theta x}}{2(\theta^5 + \theta^3 + 24)}$$

$$h(x; \theta) = \frac{\theta^5 (2x^4 + \theta x^2 + 2\theta)}{2\theta^2 x^2 (\theta^2 x^2 + 4\theta x + 12) + \theta x (\theta^4 x + 2\theta^3 + 48) + 2(\theta^5 + \theta^3 + 24)} \quad (15)$$

It can be observed that $h(0) = f(0) = \frac{\theta^6}{\theta^5 + \theta^3 + 24}$. The graph of the hazard rate function of Odoma distribution is presented in figure 3.

C. Mean residual life function

The mean residual life function of a continuous random variable with p.d.f. and c.d.f. given by $f(x)$ and $F(x)$ respectively is defined as

$$m(x) = E[X - x | X > x] = \frac{1}{1 - F(x)} \int_x^{\infty} [1 - F(t)] dt \quad (16)$$

Therefore the corresponding mean residual life function of Odoma distribution is given by

$$m(x) = \frac{1}{\left[2\theta^4 x^4 + 8\theta^3 x^3 + (24\theta^2 + \theta^5)x^2 + (2\theta^4 + 48\theta)x + 2(\theta^5 + \theta^3 + 24) \right] e^{-\theta x}} \int_x^{\infty} \left[2\theta^4 t^4 + 8\theta^3 t^3 + (24\theta^2 + \theta^5)t^2 + (2\theta^4 + 48\theta)t + 2(\theta^5 + \theta^3 + 24) \right] e^{-\theta t} dt$$

$$m(x) = \frac{2\theta^4 x^4 + 16\theta^3 x^3 + (72\theta^2 + \theta^5)x^2 + (192\theta + 26\theta^4)x + 2(\theta^5 + 3\theta^3 + 120)}{\theta \left[2\theta^4 x^4 + 8\theta^3 x^3 + (24\theta^2 + \theta^5)x^2 + (2\theta^4 + 48\theta)x + 2(\theta^5 + \theta^3 + 24) \right]} \quad (17)$$

It can also be observed that $m(0) = \mu_1' = \frac{\theta^5 + 3\theta^3 + 120}{\theta(\theta^5 + \theta^3 + 24)}$. The graph of the mean residual life function of Odoma distribution is presented in figure 4.

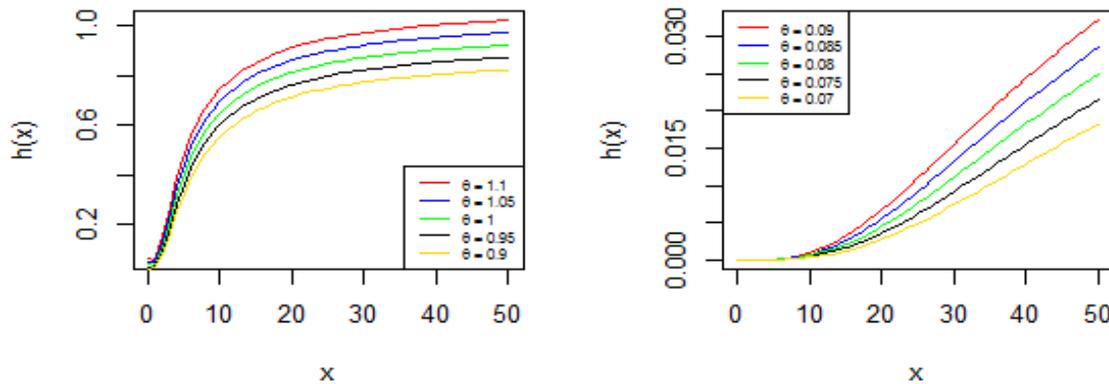


Figure 3. Graphs of the hazard rate function of Odoma distribution for varying values of the parameter

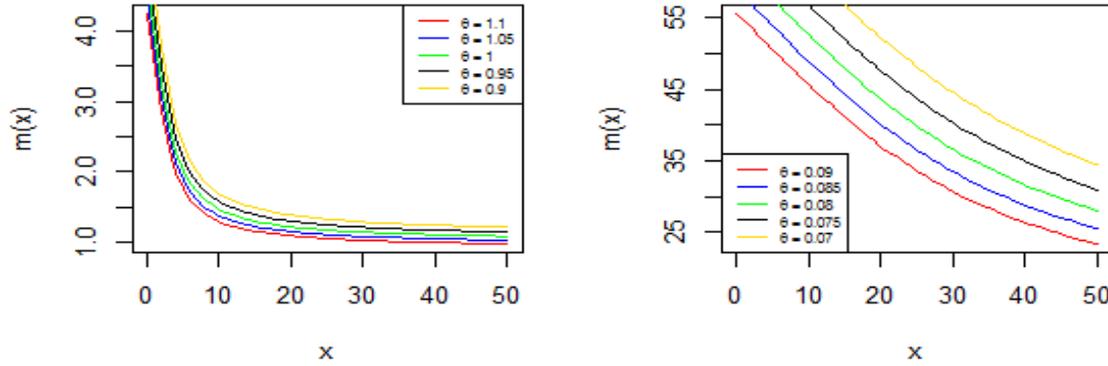


Figure 4. Graphs of the mean residual life function of Odoma distribution for varying values of the parameter.

7. Parameter estimation

Let $X_1, X_2, X_3, \dots, X_n$ be a random sample of size n from Odoma distribution with p.d.f. as in (1).

Therefore the likelihood function, L , of Odoma distribution is obtained as

$$L = \left(\frac{\theta^5}{2(\theta^5 + \theta^3 + 24)} \right)^n \prod_{i=1}^n (2x_i^4 + \theta x_i^2 + 2\theta) e^{-\theta \sum_{i=1}^n x_i} \quad (18)$$

and the log likelihood function is

$$\ln L = 5n \ln \theta - n \ln 2 - n \ln (\theta^5 + \theta^3 + 24) + \sum_{i=1}^n \ln (2x_i^4 + \theta x_i^2 + 2\theta) - \theta \sum_{i=1}^n x_i \quad (19)$$

It follows, therefore, that the maximum likelihood estimate (MLE), $\hat{\theta}$ of θ is the solution of the nonlinear equation

$$\frac{\partial \ln L}{\partial \theta} = \frac{5n}{\theta} - \frac{n(5\theta^4 + 3\theta^2)}{\theta^5 + \theta^3 + 24} + \sum_{i=1}^n \left(\frac{x_i^2 + 2}{2x_i^4 + \theta x_i^2 + 2\theta} \right) - \sum_{i=1}^n x_i = 0. \quad (20)$$

The value of the parameter estimate was computed using R-Software.

8. Application

In this section, we demonstrate the usefulness of Odoma distribution distribution by application to real life data. Using the widely used goodness of fit indices such as Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and $-2 \ln L$, the distribution is compared with the following distributions: Pranav, Sujatha, Aradhana, Akash, Lindley and exponential distributions. The maximum

likelihood estimate of the parameter and the goodness of fit indices are computed for the strength data of glass of the aircraft window reported by Fuller et al. (1994).

Data Set: Table 1 presents the strength data of glass of the aircraft window as reported by Fuller et al. (1994)

Table 1: Strength data of glass of the aircraft window reported by Fuller et al. (1994)

18.83	20.8	21.657	23.03	23.23	24.05
24.321	25.5	25.52	25.8	26.69	26.77
26.78	27.05	27.67	29.9	31.11	33.2
33.73	33.76	33.89	34.76	35.75	35.91
36.98	37.08	37.09	39.58	44.045	45.29
45.381					

Table 2: MLE's, $-2\ln L$, AIC and BIC of the fitted distributions

Model	Parameter estimate	$-2\ln L$	AIC	BIC
Odoma	0.162264	227.26	229.26	230.69
Pranav	0.129818	232.77	234.77	236.68
Sujatha	0.095613	241.50	243.50	244.94
Aradhana	0.094319	242.22	244.22	245.64
Akash	0.097062	240.68	242.68	244.11
Lindley	0.062988	253.99	255.99	257.42
Exponential	0.032455	274.53	276.53	277.96

It can be seen from Table 2 that Odoma distribution, having the lowest values across the goodness of fit indices, gave a better fit to the data than the competing distributions.

9. Conclusion

A new one parameter lifetime distribution is proposed and various statistical properties of the distribution are derived. The method of maximum likelihood is used to estimate the parameter. Finally, the new distribution is fitted to the strength data of glass of the aircraft window and compared with some competing distribution. The results show its better fit over the competing distributions.

References

- [1] Ghitany, M. E., Atieh, B., & Nadarajah, S. (2008). Lindley distribution and its application. *Mathematics and computers in simulation*, 78(4), 493-506.
- [2] Gómez-Déniz, E., & Calderín-Ojeda, E. (2011). The discrete Lindley distribution: properties and applications. *Journal of Statistical Computation and Simulation*, 81(11), 1405-1416.
- [3] Nadarajah, S., Bakouch, H. S., & Tahmasbi, R. (2011). A generalized Lindley distribution. *Sankhya B*, 73(2), 331-359.
- [4] Elbatal, I., & Elgarhy, M. (2013). Transmuted quasi Lindley distribution: A generalization of the quasi Lindley distribution. *Int. J. Pure Appl. Sci. Technol*, 18(2), 59-70.
- [5] Ghitany, M. E., Al-Mutairi, D. K., Balakrishnan, N., & Al-Enezi, L. J. (2013). Power Lindley distribution and associated inference. *Computational Statistics & Data Analysis*, 64, 20-33.
- [6] Warahena-Liyanage, G., & Pararai, M. (2014). A generalized power Lindley distribution with applications. *Asian Journal of Mathematics and Applications*, 2014.
- [7] Abouammoh, A. M., Alshangiti, A. M., & Ragab, I. E. (2015). A new generalized Lindley distribution. *Journal of Statistical Computation and Simulation*, 85(18), 3662-3678.
- [8] Shanker, R. (2015). Akash distribution and its applications. *International Journal of Probability and Statistics*, 4(3), 65-75.
- [9] Sharma, V. K., Singh, S. K., Singh, U., & Agiwal, V. (2015). The inverse Lindley distribution: a stress-strength reliability model with application to head and neck cancer data. *Journal of Industrial and Production Engineering*, 32(3), 162-173.
- [10] Alkarni, S. H. (2015). Extended power Lindley distribution: A new statistical model for non-monotone survival data. *European journal of statistics and probability*, 3(3), 19-34.
- [11] Shanker, R., & Shukla, K. K. (2016). Weighted Akash distribution and its applications to model lifetime data. *International Journal of Statistics*, 39(2), 1138-1147.
- [12] Shanker, Rama. (2016). On Poisson-Akash Distribution and its Applications. *Biometrics & Biostatistics International Journal*. 3. 10.15406/bbij.2016.03.00075.

- [13] Shanker, R. (2016). Aradhana distribution and its Applications. *International Journal of Statistics and Applications*, 6(1), 23-34.
- [14] Shanker, R. (2016). Sujatha distribution and its Applications. *Statistics in Transition new series*, 17(3), 391-410.
- [15] Shanker, R., & Shukla, K. K. (2017). Ishita distribution and its applications. *Biometrics & Biostatistics International Journal*, 5(2), 1-9.
- [16] Shukla K. K. (2018). Pranav distribution with properties and its applications. *Biom Biostat Int J.*, 7(3), 244–254. DOI: 10.15406/bbij.2018.07.00215.
- [17] Pinho, L. G. B. (2017). Building new probability distributions: the composition method and a computer based method.

UNDER PEER REVIEW