FUDMA Journal of Sciences (FJS)
ISSN online: 2616-1370
ISSN print: 2645 - 2944
Vol. 8 No. 2, April, 2024, pp 301 - 305
DOI: https://doi.org/10.33003/fjs-2024-0802-2364

ARCSINE RAYLIEGH PARETO DISTRIBUTION: PROPERTIES AND APPLICATION TO CARBON FIBERS DATA SETS

***[1,3]Usman Abubakar, [1]Abdulhameed Ado Osi, [1]Iliyasu Abubakar Salisu, [2]Hassan Muhammad, [3]Yusuf Ibrahim Muhammad and [3]Abbas Abubakar**

[1]Department of Statistics, Aliko Dangote University of Science and Technology Wudil, Kano state. Nigeria.
[2]Department of Mathematics, Federal university of Technology Babura, jigawa state. Nigeria.
[3]Department of Statistics, Jigawa State Polytechnic Dutse, Jigawa State. Nigeria.

*Corresponding authors' email: Usmanabubakar@Jigpoly.edu.ng

ABSTRACT

In this paper, we introduce a new modified distribution called arcsine Rayliegh Pareto (ASRP) Distribution. We derived its mathematical and statistical properties, including survival function, hazard function, entropy, moment, moment generating function, and order statistics. We also used maximum likelihood estimation for estimating the parameters of the distribution. The plots of the cdf, pdf, hazard rate function, and survival function were illustrated with right skewed probability density function, cumulative distribution function with monotone increasing function, which converge at one. The value of some goodness of fit measure (i.e. AIC, AICc, and BIC) were computed, as well as the KS, A, and W statistic. Finally, we suggested that the new modified model outperform better than the other standard distribution using Carbon fibers data sets.

Keywords: Arcsine Rayliegh Pareto distribution, Properties, Maximum likelihood estimation, Application

INTRODUCTION

In general, modeling data from financial, biomedical, engineering, and lifelong testing studies is a significant area of study (Ahmad et al., 2022). (Abubakar et al., 2024) claims that since there are still many scenarios in which none of the traditional or classical probability models can sufficiently describe the actual data, it is imperative to propose new models that can more faithfully capture the real-life phenomena found in a given dataset. In many situations, some researchers focus on modifying an existing probability model in order to increase its performance, versatility, and flexibility to handle the data that an existing distribution cannot fit well. This happens, especially when the data is skewed either to the right or to the left.

The proposed Arcsine rayliegh pareto (ASRP) distribution is a modification to the Rayliegh pareto distribution introduced by (Al-Kadim and Mohammed, 2018), and as similar methods in many researches, including Composite Rayliegh Pareto Distribution by (Benatmane et al., 2021), Arcsine Kumaraswamy generalized family by (Emam and Tashkandy, 2022), Sine topp-leone generalized family of distribution by (Al-Babtain et al., 2020), cosine Topp–Leone family of distributions by (Nanga et al., 2023), Tangent Topp-Leone G family of distributions by (Nanga et al., 2022), among others. (Al-Kadim and Mohammed, 2018) provide several mathematical properties, quantile, reliability function, and parameter estimation using maximum likelihood methods. The cumulative distribution function and probability density function of the Rayliegh Pareto distributions were given by (AlKadim and Mohammed, 2018) as;

$$F(y) = 1 - exp\left(\frac{-\left(\frac{y}{c}\right)^a}{2b^2}\right) \quad (1)$$

The above distribution can be modified as ASRP distribution, and Such modification was used by (Emam and Tashkandy, 2022). The cumulative distribution of Arcsine Rayliegh Pareto Distribution is;

$$G(z, a, b, c) = \frac{2}{\pi} arcsine\left(1 - exp\left(\frac{-\left(\frac{z}{c}\right)^a}{2b^2}\right)\right) \quad (2)$$

After differentiating equation (2) and some further simplification, the probability density function is derived as;

$$g(z, a, b, c) = \frac{aZ^{a-1} exp\left(\frac{-\left(\frac{z}{c}\right)^a}{2b^2}\right)}{\pi b^2 c^a \sqrt{1 - \left(1 - exp\left(\frac{-\left(\frac{z}{c}\right)^a}{2b^2}\right)\right)^2}} \quad (3)$$

The plot of the cdf and pdf of ASRP Distribution for some selected parameters is illustrated in the figures below.

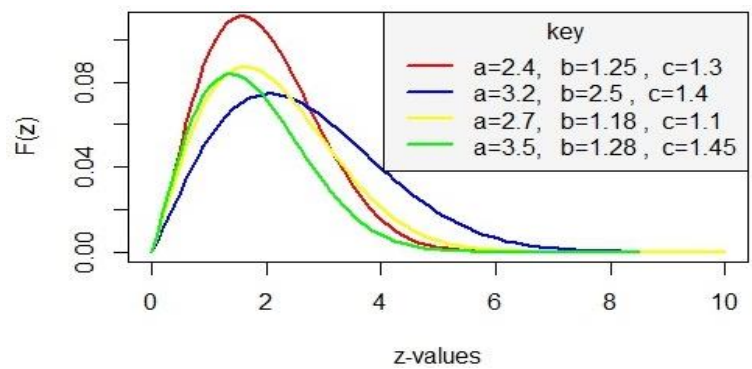

Figure 1: The plot of the pdf of ASRP Distribution for some selected parameters.

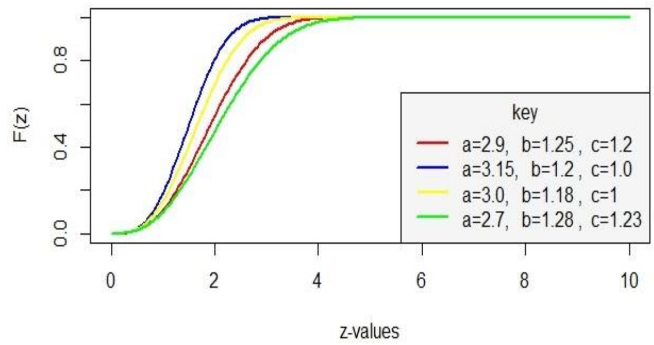

Figure 2: The plot of the cdf of ASRP Distribution for some selected parameters.

Some Statistical Properties of ASRP Distribution

Some properties of ASRP Distribution, which comprises the survival function, Hazard function, Entropy and quantile.

$$S(z) = 1 - \frac{2}{\pi} arcsine\left(1 - exp\left(\frac{-\left(\frac{z}{c}\right)^a}{2b^2}\right)\right) \quad (4)$$

$$H(z) = \frac{aZ^{a-1} exp\left(\frac{-\left(\frac{z}{c}\right)^a}{2b^2}\right)}{\pi b^2 c^a \sqrt{1-\left(1-exp\left(\frac{-\left(\frac{z}{c}\right)^a}{2b^2}\right)\right)^2}\left(1-\frac{2}{\pi} arcsine\left(1-exp\left(\frac{-\left(\frac{z}{c}\right)^a}{2b^2}\right)\right)\right)} \quad (5)$$

$$G^{-1}(m) = C\left(-2b^2 log\left(1 - Sin\frac{\pi m}{2}\right)\right)^{\frac{1}{a}} \quad (6)$$

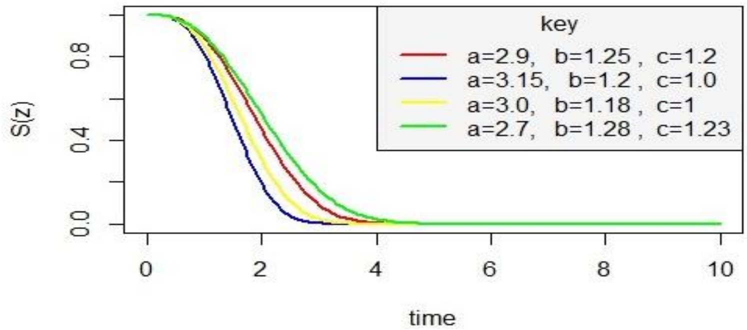

Figure 3: Survival rate function of ASRP distribution.

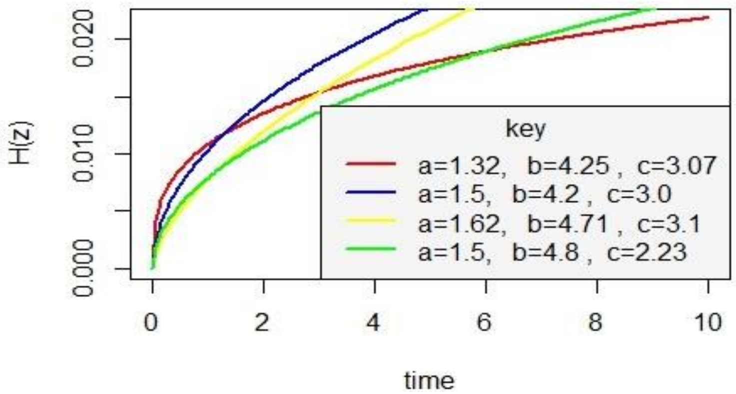

Figure 4: The Hazard rate function of ASRP distribution.

The Renyi entropy, moment, moment generating function, and Order Statistics of ASRP Distribution are;

$$I_x(z) = \frac{1}{1-\psi} log \int_{-\infty}^{\infty} g(z)^{\psi}\, dz \quad (7)$$

$$g(z)^{\psi} = (\mu\eta)^{\psi} \quad (8)$$

Where $\eta = \frac{aZ^{a-1} exp\left(\frac{-\left(\frac{z}{c}\right)^a}{2b^2}\right)}{\pi b^2 c^a \sqrt{1-\left(1-exp\left(\frac{-\left(\frac{z}{c}\right)^a}{2b^2}\right)\right)^2}}$

And $\mu = \frac{a}{\pi b^2 c^a}$

Implies that, the entropy of ASRP Distribution is

$$I_x(z) = \frac{1}{1-\psi} log \int_0^\infty (\mu\eta)^\psi \, dz \quad (9)$$

$$= \frac{1}{1-\psi}\left[\psi log\mu + log \int_0^\infty \eta^\psi \, dz\right] \quad (10)$$

The rth moment of a continuous distribution is given by;

$$\mu_1^r = E(z^r) = \int_{-\infty}^\infty z^r g(z, a, b, c)\, dz \quad (11)$$

$$\frac{1}{\sqrt{1-m^2}} = \sum_{u=0}^\infty \frac{1x3x5x\ldots x(2u-1)}{2^u xU!} m^{2u} \quad (12)$$

$$\mu_1^r = \frac{a}{\pi b^2 c^a} \sum_{u=0}^\infty \frac{1x3x5x\ldots x(2u-1)}{2^u xU!} \theta_{r,a,b,c} \quad (13)$$

Where

$$\theta_{r,a,b,c} = \int_0^\infty \frac{aZ^{a+r-1} exp\left(\frac{-\left(\frac{z}{c}\right)^a}{2b^2}\right)}{\pi b^2 c^a \sqrt{1-\left(1-exp\left(\frac{-\left(\frac{z}{c}\right)^a}{2b^2}\right)\right)^2}} dz \quad (14)$$

The moment generating function $M_y(t)$ is,

$$M_z(t) = E(e^{tz}) = \int_{-\infty}^\infty e^{tz} g(z, a, b, c)\, dz \quad (15)$$

$$\frac{at^a}{\pi b^2 c^a u!} \sum_{u=0}^\infty \sum_{v=0}^\infty \frac{1x3x5x\ldots x(2u-1)}{2^u xU!} k_{r,a,b,c} \quad (16)$$

Where

$$k_{r,a,b,c} = \int_0^\infty \frac{aZ^{a+r-1} exp\left(\frac{-\left(\frac{z}{c}\right)^a}{2b^2}\right)}{\pi b^2 c^a \sqrt{1-\left(1-exp\left(\frac{-\left(\frac{z}{c}\right)^a}{2b^2}\right)\right)^2}} dz$$

For Order Statistics, Let $z_1, z_2 \ldots, z_n$ be a random sample from the ASRP distribution and let z(1), . . . , z(n) be the corresponding order statistics. The nth order statistic's pdf is written as;

$$g_{(i,n)}(z) = \frac{n!(-1)^j}{(i-1)(n-i-j)!j!} \sum_{j=0}^{n-i} g(z)[G(z)]^{i+j-1} \quad (17)$$

$$g_{(i,n)}(z) = \sum_{j=0}^{n-i} \frac{n!(-1)^j aZ^{a+r-1} exp\left(\frac{-\left(\frac{z}{c}\right)^a}{2b^2}\right)\left(\frac{2}{\pi} arcsine\left(1-exp\left(\frac{-\left(\frac{z}{c}\right)^a}{2b^2}\right)\right)\right)^{i+j-1}}{(i-1)(n-i-j)!j!\pi b^2 c^a \sqrt{1-\left(1-exp\left(\frac{-\left(\frac{z}{c}\right)^a}{2b^2}\right)\right)^2}} \quad (18)$$

When i=1,

$$g_{(1,n)}(z) = \sum_{j=0}^{n-i} \frac{aZ^{a+r-1} exp\left(\frac{-\left(\frac{z}{c}\right)^a}{2b^2}\right)\left(\frac{2}{\pi} arcsine\left(1-exp\left(\frac{-\left(\frac{z}{c}\right)^a}{2b^2}\right)\right)\right)^{j}}{\pi b^2 c^a \sqrt{1-\left(1-exp\left(\frac{-\left(\frac{z}{c}\right)^a}{2b^2}\right)\right)^2}} \quad (19)$$

When i=n

$$g_{(n,n)}(z) = \frac{naZ^{a+r-1} exp\left(\frac{-\left(\frac{z}{c}\right)^a}{2b^2}\right)\left(\frac{2}{\pi} arcsine\left(1-exp\left(\frac{-\left(\frac{z}{c}\right)^a}{2b^2}\right)\right)\right)^{n-1}}{\pi b^2 c^a \sqrt{1-\left(1-exp\left(\frac{-\left(\frac{z}{c}\right)^a}{2b^2}\right)\right)^2}} \quad (20)$$

Parameter Estimation Using Maximum Likelihood

To estimate the parameters of the ASRP distribution using a complete sample, let z1,z2,...,zn be a random sample from the ASRP distribution. The log-likelihood of the parameter vector is written as;

$$logl(a, b, c) = log \prod_{i=1}^n g(z, a, b, c) \quad (21)$$

$$logl(a, b, c) = nloga + (a-1)logz + \frac{1}{2b^2} \sum_{i=0}^\infty \left(\frac{z}{c}\right)^a - nlog\pi - 2nlogb - anlogc + \frac{1}{2} \sum_{i=0}^\infty log\left(1-\left(1-exp\left(\frac{-\left(\frac{z}{c}\right)^a}{2b^2}\right)\right)^2\right) \quad (22)$$

$$\frac{dl}{dc} = \frac{an}{c} \quad (23)$$

$$\frac{dl}{da} = \frac{n}{a} + \sum_{i=0}^\infty logz - \frac{1}{2b^2 c^a} \sum_{i=0}^\infty z^a log\left(\frac{z}{c}\right) - nlogc +$$

$$\sum_{i=0}^{\infty} \frac{z^{a}\left(1-exp\left(\frac{-\left(\frac{z}{c}\right)^{a}}{2b^{2}}\right)\right)log\left(\frac{z}{c}\right)exp\left(\frac{-\left(\frac{z}{c}\right)^{a}}{2b^{2}}\right)}{c^{a}\left(1-\left(1-exp\left(\frac{-\left(\frac{z}{c}\right)^{a}}{2b^{2}}\right)\right)^{2}\right)} \quad (24)$$

$$\frac{dl}{db} = \frac{1}{b^{3}c^{a}}\sum_{i=0}^{\infty} z^{a} - \frac{2n}{b} - \sum_{i=0}^{\infty} \frac{z^{a}\left(1-exp\left(\frac{-\left(\frac{z}{c}\right)^{a}}{2b^{2}}\right)\right)exp\left(\frac{-\left(\frac{z}{c}\right)^{a}}{2b^{2}}\right)}{b^{3}c^{a}\left(1-\left(1-exp\left(\frac{-\left(\frac{z}{c}\right)^{a}}{2b^{2}}\right)\right)^{2}\right)} \quad (25)$$

Application to Carbon Fibers Data Sets

The real data set was used by (Selim and Badr, 2016). The data represent the strength data measured in GPA, for single carbon fibers were tested under tension at gauge lengths of 1, 10, 20 and 50 mm. For illustrative purpose, we consider only the data set consisting the single fibers of 20 mm, with a sample of size 63. The data are;

1.901, 2.132, 2.203, 2.228, 2.257, 2.350, 2.361, 2.396, 2.397, 2.445, 2.454, 2.474, 2.518, 2.522, 2.525, 2.532, 2.575, 2.614, 2.616, 2.618, 2.624, 2.659, 2.675, 2.738, 2.740, 2.856, 2.917, 2.928, 2.937, 2.937, 2.977, 2.996, 3.030, 3.125, 3.139, 3.145, 3.220, 3.223, 3.235, 3.243, 3.264, 3.272, 3.294, 3.332, 3.346, 5 3.377, 3.408, 3.435, 3.493, 3.501, 3.537, 3.554, 3.562, 3.628, 3.852, 3.871, 3.886, 3.971, 4.024, 4.027, 4.225, 4.395, 5.020.

Table 1: Summary of the carbon fibers data sets

Min	Q(1)	Median	Mean	Q(3)	Max
1.901	2.554	2.996	3.059	3.421	5.020

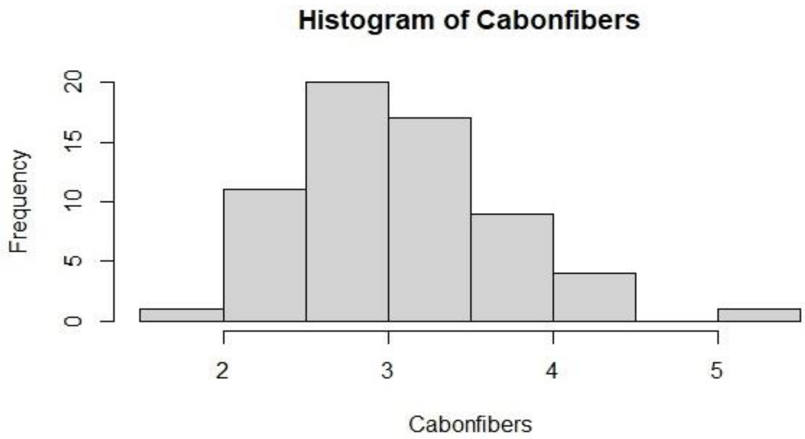

Figure 5: Histogram illustration of right skewed carbon fiber data sets

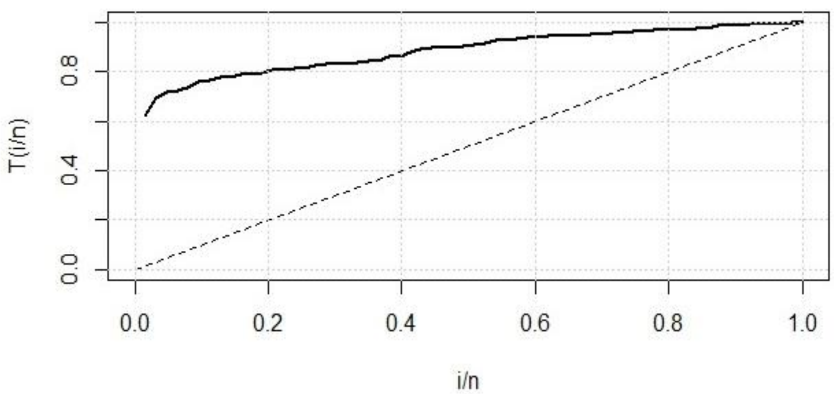

Figure 6: TTT plot illustration for model adequacy.

To compare the fitted model with some standard models using the carbon fibers and bladder cancer Data sets, we used methods of goodness of fit, which comprises the log-likelihood function evaluated at the MLEs, the Akaike information criterion (AIC), the Bayesian information criterion (BIC), the Akaike information corrected criterion (AICc), Anderson-Darling (A*), Cramer-von Mises (W*), and Kolmogorov-Smirnov (K-S*) (Aryal et al., 2017). generally speaking, the smaller the values of these statistics, the better the fit (Al-Kadim and Mohammed, 2018).

$AIC = -2L+2k$ (26)

$BIC = -2L+klog(n)$ (27)

And $CAIC = -2L + \frac{2kn}{n-k-1}$ (28)

Table 2: Competitors model

Distributions.	Author(s)
Rayliegh Pareto	(Al-Kadim and Mohammed, 2018)
Exponentiated Weibull	(Shawky and Abu-Zinadah, 2009)
Exponentiated Pareto	(Nadarajah, 2005)

Table 3: Goodness of fit measure and estimate of the parameters using MLE

Distr.	$\hat{a}$	$\hat{b}$	$\hat{c}$	LL	AIC	AICc	BIC
ASRP	5.330	2.114	1.970	-18.29134	42.58267	42.98945	49.01208
RP	0.002471	4.801721	93.189154	-24.88366	55.76733	56.17411	62.19673
EXPW	0.7878	1.54616	28.96688	-56.31814	118.6363	119.0431	125.0657
EXPP	34.2529	2.7867		-92.27645	188.5529	188.7529	192.8392

Table 4: Normality test for the fitted models

Distribution	KS	A	W	P
ASRP	0.085799	0.87734	0.12667	0.7426
RP	0.97151	0.32514	0.056579	2.2×10^{16}
EXPW	0.079744	0.032222	0.060305	0.8179
EXPP	0.2703	0.32758	0.062385	0.0002009

The summary of carbon fibers data sets in Table 1 and Figure 5 illustrates the right-skewness nature of the data. By considering the values in Table 1, we can observe that the mean is greater than the median of the distribution. Likewise, in Table 3, the likelihood of the modified Arcsine Rayliegh Pareto distribution is higher than for other standard distributions. Also, the goodness of fit measure for AIC, BIC, and AICC for ASRP is smaller than that of other distributions, which shows the better performance of ASRP.

CONCLUSION

In this paper, we introduce a new modified distribution called the Arcsine Rayliegh pareto distribution, which is a modification of the Rayliegh pareto distribution. We examined some of its statistical and mathematical properties, as well as the estimation of parameters using the maximum likelihood method. lastly, we suggested that the ASRP distribution is a better one for modeling right-skewed data sets, which also perform better when compared with other competitors models.

REFERENCES

Abubakar, U., Osi, A. A., Shuaibu, A., Abubakar, A., Salisu, I. A., and Muhammad, Y. I. (2024). Truncated exponential log-topp-leone generalized family of distributions: properties and application to real data sets. *International Journal of Research and technopreneural innovation*, 1(1):147–162.

Ahmad, Z., Mahmoudi, E., Roozegarz, R., Hamedani, G., and Butt, N. S. (2022). Contributions towards new families of distributions: An investigation, further developments, characterizations and comparative study. *Pakistan Journal of Statistics and Operation Research*, pages 99–120.

Al-Babtain, A. A., Elbatal, I., Chesneau, C., and Elgarhy, M. (2020). Sine topp-leone-g family of distributions: Theory and applications. *Open Physics*, 18(1):574–593.

Al-Kadim, K. A. and Mohammed, B. D. (2018). Rayleigh pareto distribution. *Journal of University of Babylon for Pure and Applied Sciences*, 26(1):84–95.

Aryal, G. R., Ortega, E. M., Hamedani, G., and Yousof, H. M. (2017). The topp-leone generated weibull distribution: regression model, characterizations and applications. *International Journal of Statistics and Probability*, 6(1):126–141.

Benatmane, C., Zeghdoudi, H., Shanker, R., and Lazri, N. (2021). Composite rayleigh-pareto distribution: Application to real fire insurance losses data set. *Journal of Statistics and Management Systems*, 24(3):545–557.

Emam, W. and Tashkandy, Y. (2022). The arcsine kumaraswamy-generalized family: Bayesian and classical estimates and application. *Symmetry*, 14(11):2311.

Nadarajah, S. (2005). Exponentiated pareto distributions. *Statistics*, 39(3):255–260.

Nanga, S., Nasiru, S., and Dioggban, J. (2022). Tangent topp-leone family of distributions. *Scientific African*, 17:e01363.

Nanga, S., Nasiru, S., and Dioggban, J. (2023). Cosine topp–leone family of distributions: properties and regression. *Research in Mathematics*, 10(1):2208935.

Selim, M. A. and Badr, A. M. (2016). The kumaraswamy generalized power weibull distribution. *Math. Theo. Model*, 6:110–124.

Shawky, A. and Abu-Zinadah, H. H. (2009). Exponentiated pareto distribution: different method of estimations. *International Journal of Contemporary Mathematical Sciences*, 4(14):677–693.