A NEW GENERALIZATION OF AREA BIASED DISTRIBUTION WITH PROPERTIES AND ITS APPLICATION TO REAL LIFE DATA

P. Pandiyan and R. Jothika

•

(1). Professor, Department of Statistics, Annamalai University
(2). Research Scholar, Department of Statistics, Annamalai University
Email: pandianau@gmail.com, rjothika2023@gmail.com

Abstract

This paper proposed a new generalization of the Samade distribution. The term "area biased samade distribution" refers to the recently created distribution model. After studying the various structural features, entropies, order statistics, moments, generating functions for moments, survival functions, and hazard functions were calculated. The parameters of the suggested model are estimated using the maximum likelihood estimation technique. Ultimately, a fitting of an application to a real-life blood cancer data set reveals a good fit.

Keywords: area biased samade distribution, reliability analysis, maximum likelihood estimation, order statistics, entropies.

I. Introduction

Samade distribution is a recently introduced two parametric lifetime distribution. Introduced by [1] has been discussed by samade probability distribution. Its properties and application to real lifetime data and determine the its various mathematical and statistical properties. Additionally, modelling and analysing real life data are essential in many practical fields including engineering, health, finance, and insurance. Area biased samade distribution a new parametric life time distribution is more flexible in the manner it processes lifetime data than the samade distribution. the area biased version of the samade distribution in this study exhibits greater adaptability in handling life time than the samade distribution. after fitting a real-life data set, it was discovered that the area biased samade distribution gives better performance.

In [5] firstly introduced the concept of weighted distributions to model ascertainment biases in the data and later on [14] reformulated it in a unifying theory for problems where the observations fail in non-experimental, non-replicated and non-random manner. It has been observed that when an investigator records observations in the nature according to certain stochastic model, the distribution of the recorded observations will be different from the original distribution unless every observation given an equal chance of being recorded. Suppose the original observation x_0 comes from a distribution having probability density function (pdf)

$f_0(x, \theta_1)$ where θ_1 may be a vector of parameters and the observation x is recorded according to a probability re weighted by weight function $w(x, \theta_2) > 0$, $\theta_2$being a new parameter vector, then x comes from a distribution having pdf

$$f(x; \theta_1, \theta_2) = AW(x; \theta_2) f_0(x, \theta_1)$$

Where A is a normalizing constant if should be noted that such type of distribution is known as weighted distributions. The weighted distribution with weight function $w(x, \theta_2) = x$ are called length biased distribution or size biased distribution. In [12] proposed the weighted pareto type-II distribution as a new model for handling medical science data and studied its statistical properties and applications. In [7] established a new length biased distribution. and a new length biased distribution was introduced by [15]. In [9] presented Area biased quasi-Transmuted uniform distribution. In [4] presented the size biased Zeghdoudi distribution and discuss its various statistical properties and application. In [8] provided a length and area biased exponentiated Weibull distribution. and recently [2] discussed the characterization and estimation of Area biased quasi-Akash distribution. In this study our motives are to prove that area biased version of samade distribution is more flexible and fits better in real life time data.

II. Area-Biased Samade Distribution

Let random variable be x with scale parameter alpha and shape parameter theta then the probability density function (pdf) and cumulative distribution function (cdf) of the area biased Samade distribution.

The probability density function of Samade distribution is given by

$$f(x; \alpha, \theta) = \frac{\theta^4}{\theta^4 + 6\alpha}(\theta + \alpha x^3)e^{-\theta x} \quad ; \qquad x, \theta, \alpha > 0 \tag{1}$$

The cumulative distribution function (cdf) of Samade distribution is given by

$$F(x; \alpha, \theta) = 1 - \left(\frac{\theta^4 + \alpha\left(6 + x\theta\left(6 + x\theta(3 + x\theta)\right)\right)}{6\alpha + \theta^4} \right) e^{-\theta x};\ x, \theta, \alpha > 0 \tag{2}$$

considered a random variable x with a probability density function $f(x)$. Let $w(x)$ be a non-negative weight function. Denote a new probability density function.

$$f_w(x) = \frac{w(x) f(x)}{E[w(x)]} \quad ; \quad x > 0$$

Where its non-negative weight function be $w(x)$ and $E(w(x)) = \int w(x) f(x) dx < \infty$.

Depending upon the various choices of weight function especially when $w(x) = x^c$, result is called weighted distribution. In this paper, we have to obtain the Area biased version of samade distribution, so we will consider as $w(x) = x^{\,2}$ to obtain the Area biased samade distribution. Then, the probability density function of Area biased version is given as

$$f_a(x) = \frac{x^2 f(x)}{E(X^2)} \tag{3}$$

Where

$$E(X^2) = \int_0^\infty x^2 f(x) dx$$

$$E(X^2) = \frac{2\theta^4 + 120\alpha}{\theta^4(\theta^4 + 6\alpha)} \tag{4}$$

By substituting equations (1) and (4) in equation (3), we will get the probability density function of Area biased samade distribution

$$f_a(x) = \frac{\theta^6}{2\theta^4 + 120\alpha} x^2(\theta + \alpha x^3)e^{-\theta x} \tag{5}$$

And the cumulative distribution function of area biased samade distribution can be obtained as

$$F_a(x) = \int_0^x f_a(x)\, dx$$

$$F_a(x) = \int_0^x \frac{\theta^6}{2\theta^4 + 120\alpha} x^2(\theta + \alpha x^3)e^{-\theta x}\ dx$$

$$F_a(x) = \frac{\theta^6}{2\theta^4 + 120\alpha} \int_0^x x^2(\theta + \alpha x^3)e^{-\theta x}\ dx$$

$$F_a(x) = \frac{\theta^6}{2\theta^4 + 120\alpha} \int_0^x (x^2\theta + \alpha x^5)e^{-\theta x}\ dx$$

$$F_a(x) = \frac{\theta^6}{2\theta^4 + 120\alpha} \left[\int_0^x x^2\theta\, e^{-\theta x}\, dx + \int_0^x \alpha x^5\, e^{-\theta x}\, dx \right]$$

$$F_a(x) = \frac{\theta^6}{2\theta^4 + 120\alpha} \left[\theta \int_0^x x^2\, e^{-\theta x}\, dx + \alpha \int_0^x x^5\, e^{-\theta x}\, dx \right]$$

Put $\theta x = t, \quad x = \frac{t}{\theta}, \qquad dx = \frac{dt}{\theta}$

When $x \to 0, t \to 0, and\ x \to x,\ \ t \to \theta x$

$$F_a(x) = \frac{\theta^6}{2\theta^4 + 120\alpha} \left[\theta \int_0^{\theta x} \left(\frac{t^2}{\theta^3}\right) e^{-t}\, dt + \alpha \int_0^{\theta x} \left(\frac{t^5}{\theta^6}\right)\, e^{-t}\, dt \right] \tag{6}$$

After simplification of equation (6), we obtain the cumulative distribution function of area biased samade distribution.

$$F_a(x) = \frac{1}{2\theta^4 + 120\alpha} [\theta^4 \gamma(3, \theta x) + \alpha \gamma(6, \theta x)] \tag{7}$$

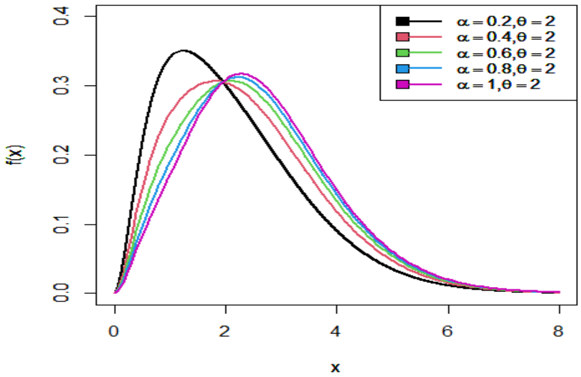

Figure 1: *Pdf plot of area biased samade distribution*

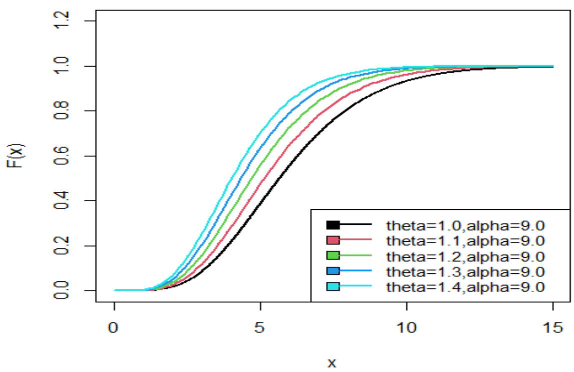

Figure 2: *Cdf plot of area biased Samade distribution*

III. Reliability Analysis

In this section, we will discuss the survival function, hazard function, reverse hazard function, cumulative hazard function, Odds rate, Mills ratio and, Mean Residual function for the proposed area biased samade distribution.

I. Survival Function

The survival function or the reliability function of the area biased Samade distribution is given by

$$S(x) = 1 - F_a(x)$$

$$S(x) = 1 - \frac{1}{2\theta^4 + 120\alpha}\left[\theta^4\gamma(3, \theta x) + \alpha\gamma(6, \theta x)\right]$$

II. Hazard Function

The corresponding hazard function or failure rate of the area biased Samade distribution is given by

$$h(x) = \frac{f_a(x)}{S(x)}$$

$$h(x) = \frac{\frac{\theta^6}{2\theta^4 + 120\alpha} x^2(\theta + \alpha x^3)e^{-\theta x}}{1 - \frac{1}{2\theta^4 + 120\alpha}[\theta^4\gamma(3, \theta x) + \alpha\gamma(6, \theta x)]}$$

$$h(x) = \frac{x^2\theta^6(\theta + \alpha x^3)e^{-\theta x}}{(2\theta^4 + 120\alpha) - [\theta^4\gamma(3, \theta x) + \alpha\gamma(6, \theta x)]}$$

III. Reverse Hazard Function

Reverse hazard function of area biased samade distribution is given by

$$h_r(x) = \frac{f_a(x)}{F_a(x)}$$

$$h_r(x) = \frac{x^2\theta^6(\theta + \alpha x^3)e^{-\theta x}}{\theta^4\gamma(3, \theta x) + \alpha\gamma(6, \theta x)}$$

IV. Odds Rate Function

Odds Rate function of area biased samade distribution is given by

$$O(x) = \frac{F_a(x)}{1 - F_a(x)}$$

$$O(x) = \frac{\theta^4\gamma(3, \theta x) + \alpha\gamma(6, \theta x)}{2\theta^4 + 120\alpha - \theta^4\gamma(3, \theta x) + \alpha\gamma(6, \theta x)}$$

V. Cumulative Hazard Function

Cumulative hazard function of area biased samade distribution is given by

$$H(x) = -\ln(1 - F_a(x))$$

$$H(x) = -\ln\left(1 - \frac{\theta^4\gamma(3, \theta x) + \alpha\gamma(6, \theta x)}{2\theta^4 + 120\alpha}\right)$$

VI. Mills Ratio

$$\text{Mills ratio} = \frac{1}{\mathrm{h_r(x)}}$$

$$\text{Mills ratio} = \frac{\theta^4\gamma(3, \theta x) + \alpha\gamma(6, \theta x)}{x^2\theta^6(\theta + \alpha x^3)e^{-\theta x}}$$

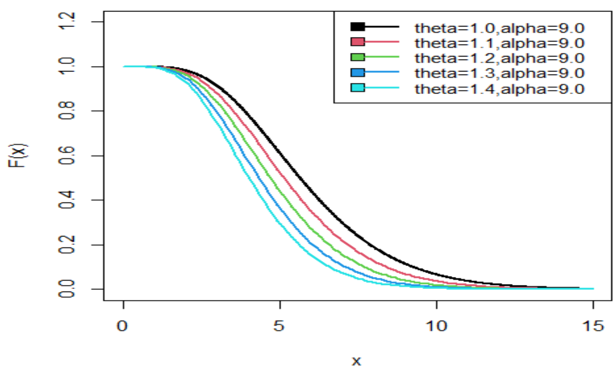

Figure 3: *Survival plot of area biased Samade distribution*

IV. Statistical Properties

In this section, discuss about the different statistical properties of area biased samade distribution especially its moments, harmonic mean, moment generating function and characteristic function.

I. Moments

let the random variable X represents area biased samade distribution with parameters θ and α, then the r^{th} order moment $E(X^r)$ of X about origin can be obtained as

$$E(X^r) = \mu_r{}' = \int_0^\infty x^r f_a(x)dx$$

$$E(X^r) = \int_0^\infty x^r \frac{\theta^6}{2\theta^4 + 120\alpha} x^2(\theta + \alpha x^3)e^{-\theta x}dx$$

$$E(X^r) = \frac{\theta^6}{2\theta^4 + 120\alpha}\int_0^\infty x^{r+2}(\theta + \alpha x^3)e^{-\theta x}dx$$

$$E(X^r)$$

$$= \frac{\theta^6}{2\theta^4 + 120\alpha}\left(\theta\int_0^\infty x^{(r+3)-1}e^{-\theta x}dx + \alpha\int_0^\infty x^{(r+6)-1}e^{-\theta x}dx\right) \quad (8)$$

After simplification of equation (14) get

$$E(X^r) = \mu_r{}' = \frac{\theta^4\Gamma r + 3 + \alpha\Gamma r + 6}{\theta^r(2\theta^4 + 120\alpha)} \quad (9)$$

Putting r = 1, 2, 3 and 4 in equation (9), obtain the first four moments of Area biased samade distribution.

$$E(X^1) = \mu_1{}' = \frac{6\theta^4 + 720\alpha}{\theta(2\theta^4 + 120\alpha)}$$

$$E(X^2) = \mu_2{}' = \frac{24\theta^4 + 5040\alpha}{\theta^2(2\theta^4 + 120\alpha)}$$

$$E(X^3) = \mu_3{}' = \frac{120\theta^4 + 40320\alpha}{\theta^3(2\theta^4 + 120\alpha)}$$

$$E(X^4) = \mu_4{}' = \frac{720\theta^4 + 362880\alpha}{\theta^4(2\theta^4 + 120\alpha)}$$

$$\text{variance} = \mu_2{}' - (\mu_1{}')^2$$

$$\text{variance} = \frac{(24\theta^4 + 5040\alpha)(2\theta^4 + 120\alpha) - (6\theta^4 + 720\alpha)^2}{\theta^2(2\theta^4 + 120\alpha)^2}$$

Standard Deviation

$$S.\mathrm{D}(\sigma) = \frac{\sqrt{(24\theta^4 + 5040\alpha)(2\theta^4 + 120\alpha) - (6\theta^4 + 720\alpha)^2}}{\theta(2\theta^4 + 120\alpha)}$$

Coefficient Of Variation

$$C.V\left(\frac{\sigma}{\mu}\right) = \frac{\sqrt{(24\theta^4 + 5040\alpha)(2\theta^4 + 120\alpha) - (6\theta^4 + 720\alpha)^2}}{6\theta^4 + 720\alpha}$$

Dispersion

$$\text{Dispersion} = \frac{\sigma^2}{\mu}$$

$$\text{Dispersion} = \left[\frac{(24\theta^4 + 5040\alpha)(2\theta^4 + 120\alpha) - (6\theta^4 + 720\alpha)^2}{\theta(2\theta^4 + 120\alpha)(6\theta^4 + 720\alpha)}\right]$$

II. Harmonic Mean

The harmonic mean for the proposed model can be obtained as

$$H.M = E\left(\frac{1}{x}\right) = \int_0^\infty \frac{1}{x} f_a(x)dx$$

$$H.M = E\left(\frac{1}{x}\right) = \int_0^\infty \frac{1}{x}\frac{\theta^6}{2\theta^4 + 120\alpha} x^2(\theta + \alpha x^3)e^{-\theta x}dx$$

After simplification, obtain

$$H.M = \frac{\theta^6}{2\theta^4 + 120\alpha}\left[\theta\int_0^\infty x\, e^{-\theta x}dx + \int_0^\infty x^4 e^{-\theta x}dx\right]$$

$$H.M = \frac{\theta^6}{2\theta^4 + 120\alpha}\left[\theta\left(\frac{1!}{\theta^{1+1}}\right) + \alpha\left(\frac{4!}{\theta^{4+1}}\right)\right]$$

After simplification get

$$H.M = \frac{\theta(\theta^4 + 6\alpha)}{(2\theta^4 + 120\alpha)}$$

III. Moment Generating Function and Characteristic Function

Suppose the random variable X follows Area biased samade distribution with parameters θ and α, then the MGF of X can be obtained as:

$$M_X(t) = E(e^{tx}) = \int_0^\infty e^{tx} f_a(x)dx$$

Using Taylor's series, can be obtain

$$M_X(t) = E(e^{tx}) = \int_0^\infty \left(1 + tx + \frac{(tx)^2}{2!} + \cdots\right) f_a(x)dx$$

$$M_X(t) = \int_0^\infty \sum_{j=0}^\infty \frac{t^j}{j!} x^j f_a(x)dx$$

$$M_X(t) = \sum_{j=0}^\infty \frac{t^j}{j!}\, \mu_j'$$

$$M_X(t) = \sum_{j=0}^\infty \frac{t^j}{j!}\left(\frac{\theta^4\Gamma j + 3 + \alpha\Gamma j + 6}{\theta^r(2\theta^4 + 120\alpha)}\right)$$

$$M_X(t) = \frac{1}{2\theta^4 + 120\alpha} \sum_{j=0}^{\infty} \frac{t^j}{j!\ \theta^j}\ (\theta^4 \Gamma j + 3 + \alpha \Gamma j + 6)$$

Similarly, the characteristic function area biased samade distribution can be obtained as
$\phi_X(t) = M_X(it)$

$$M_X(it) = \frac{1}{2\theta^4 + 120\alpha} \sum_{j=0}^{\infty} \frac{it^j}{j!\ \theta^j}\ (\theta^4 \Gamma j + 3 + \alpha \Gamma j + 6)$$

V. Order statistics

In this section, derived the distributions of order statistics from the area biased Samade distribution.
Let $X_{(1)}, X_{(2)}, X_{(3)}, \ldots.. X_{(n)}$ be the order statistics of the random sample$X_1, X_2, X_3, \ldots.. X_n$ selected from area biased samade distribution. Then the probability density function of the r^{th}order statistics X(r) is defined as.

$$f_{X(r)}(x) = \frac{n!}{(r-1)!\,(n-r)!} f_X(x)[F_X(x)]^{r-1}[1-F_X(x)]^{n-r} \tag{10}$$

Using equations (5) and (7) in equation, get the probability density function of r^{th} order statistics of area biased samade distribution.

$$f_{X(r)}(x) = \frac{n!}{(r-1)!\,(n-r)!} \left(\frac{\theta^6}{2\theta^4 + 120\alpha} x^2(\theta + \alpha x^3) e^{-\theta x} \right)$$
$$\times \left(\frac{1}{2\theta^4 + 120\alpha} [\theta^4 \gamma(3, \theta x) + \alpha \gamma(6, \theta x)] \right)^{r-1}$$
$$\times \left(1 - \frac{1}{2\theta^4 + 120\alpha} [\theta^4 \gamma(3, \theta x) + \alpha \gamma(6, \theta x)] \right)^{n-r}$$

$$f_{X(1)}(x)$$
$$= \frac{n\theta^6}{2\theta^4 + 120\alpha} x^2(\theta + \alpha x^3) e^{-\theta x} \times \left(1 - \frac{1}{2\theta^4 + 120\alpha} [\theta^4 \gamma(3, \theta x) + \alpha \gamma(6, \theta x)] \right)^{n-1}$$

$$f_{X(n)}(x) = \frac{n\theta^6}{2\theta^4 + 120\alpha} x^2(\theta + \alpha x^3) e^{-\theta x} \times \left(\frac{1}{2\theta^4 + 120\alpha} [\theta^4 \gamma(3, \theta x) + \alpha \gamma(6, \theta x)] \right)^{n-1}$$

VI. Likelihood Ratio Test

In this section, derive the likelihood ratio test from the area biased samade distribution. Let $X_1, X_2, \ldots, X_n$ be a random sample from the area biased samade distribution. To test the hypothesis.

$$H_0 : f(x) = f(x{:}\,\theta) \text{ against}$$

$$H_1 : f(x) = f_a(x{:}\,\theta)$$

In test whether the random sample of size n comes from the samade distribution or area biased samade distribution, the following test statistics is used.

$$\Delta = \frac{L_1}{L_0} = \prod_{i=1}^{n} \frac{f_a(x{:}\,\theta)}{f(x{:}\,\theta)}$$

$$\Delta = \frac{L_1}{L_0} = \prod_{i=1}^{n} \frac{\theta^2(\theta^4 + 6\alpha)}{2\theta^4 + 120\alpha} \, {x_i}^2$$

$$\Delta = \frac{L_1}{L_0} = \left(\frac{\theta^2(\theta^4 + 6\alpha)}{2\theta^4 + 120\alpha}\right)^n \prod_{i=1}^{n} {x_i}^2$$

We should reject the null hypothesis, if

$$\Delta = \left(\frac{\theta^2(\theta^4 + 6\alpha)}{2\theta^4 + 120\alpha}\right)^n \prod_{i=1}^{n} {x_i}^2 > k$$

Or equivalently, reject the null hypothesis

$$\Delta^* = \prod_{i=1}^{n} {x_i}^2 > k \left(\frac{2\theta^4 + 120\alpha}{\theta^2(\theta^4 + 6\alpha)}\right)^n$$

$$\Delta^* = \prod_{i=1}^{n} {x_i}^2 > k^* \;\; where \; k^* = k \left(\frac{2\theta^4 + 120\alpha}{\theta^2(\theta^4 + 6\alpha)}\right)^n$$

For large sample size n, 2 log Δ is distribution as chi-square variates with one degree of freedom. Thus, we rejected the null hypothesis, when the probability value is given by $p(\Delta^* > \alpha^*)$,where $\alpha^* = \prod_{i=1}^{n} {x_i}^2$is less than level of significance and $\prod_{i=1}^{n} {x_i}^2$is the observed value of the statistics Δ^*.

VII. Bonferroni and Lorenz Curves

In this section, derived the Bonferroni and Lorenz curves and from the area biased samade distribution.

The Bonferroni and Lorenz curve is a powerful tool in the analysis of distributions and has applications in many fields, such as economies, insurance, income, reliability, and medicine. The Bonferroni and Lorenz cures for a X be the random variable of a unit and $f(x)$ be the probability density function of x. $f(x)dx$ will be represented by the probability that a unit selected at random is defined as

$$B(p) = \frac{1}{p{\mu_1}'} \int_0^q x \, f_a(x) dx$$

And

$$L(p) = \frac{1}{{\mu_1}'} \int_0^q x \, f_a(x) dx$$

Where ${\mu_1}' = E(X) = \frac{6\theta^4 + 720\alpha}{\theta(2\theta^4 + 120\alpha)}$ and $q = F^{-1}(p)$

$$B(p) = \frac{\theta(2\theta^4 + 120\alpha)}{p(6\theta^4 + 720\alpha)} \int_0^q x \, \frac{\theta^6}{2\theta^4 + 120\alpha} x^2(\theta + \alpha x^3) e^{-\theta x} dx$$

$$B(p) = \frac{\theta^7}{p(6\theta^4 + 720\alpha)} \int_0^q x^3(\theta + \alpha x^3) e^{-\theta x} dx$$

$$B(p) = \frac{\theta^7}{p(6\theta^4 + 720\alpha)} \int_0^q x^3 \theta e^{-\theta x} dx + \int_0^q \alpha x^6 e^{-\theta x} dx$$

After simplification we get

$$B(p) = \frac{\theta^4 \gamma(4, \theta q) + \alpha \gamma(7, \theta q)}{p(6\theta^4 + 720\alpha)}$$

Where $L(p) = pB(p)$

$$L(p) = \frac{\theta^4 \gamma(4, \theta q) + \alpha\gamma(7, \theta q)}{(6\theta^4 + 720\alpha)}$$

VIII. Entropies

The concept of entropy is important in various fields such as probability, statistics, physics, communication theory and economics. Entropies quantify the diversity, uncertainty, or randomness of a system. Entropy of a random variable X is a measure of variation of the uncertainty.

I. Shannon Entropy

Shannon entropy of the random variable X such that area biased samade distribution is defined as

$$S_\lambda = -\int_0^\infty f(x)\log(f(x))dx \qquad \lambda > 0, \lambda \neq 1$$

$$S_\lambda = -\int_0^\infty f(x)\log(f_a(x))dx$$

$$S_\lambda = -\int_0^\infty \frac{\theta^6}{2\theta^4 + 120\alpha} x^2(\theta + \alpha x^3)e^{-\theta x}dx \times \log\left(\frac{\theta^6}{2\theta^4 + 120\alpha} x^2(\theta + \alpha x^3)e^{-\theta x}\right)dx$$

II. Renyi Entropy

The Renyi entropy is important in ecology and statistics as an index of diversity. It is also important in quantum information, where it can be used as a measure of entanglement. For a given probability distribution, Renyi entropy is given by

$$R(\beta) = \frac{1}{1-\beta}\log\int_0^\infty ({f_a}^\beta(x))dx$$

Where $\beta > 0 \ and \ \beta \neq 1$

$$R(\beta) = \frac{1}{1-\beta}\log\int_0^\infty \left(\frac{\theta^6}{2\theta^4 + 120\alpha} x^2(\theta + \alpha x^3)e^{-\theta x}dx\right)^\beta dx$$

$$R(\beta) = \frac{1}{1-\beta}\log\left(\frac{\theta^6}{2\theta^4 + 120\alpha}\right)^\beta \int_0^\infty x^{2\beta}(\theta + \alpha x^3)^\beta e^{-\theta\beta x}dx$$

Using binomial expansion in above equation and can be obtain

$$R(\beta) = \frac{1}{1-\beta}\log\left(\frac{\theta^6}{2\theta^4 + 120\alpha}\right)^\beta \sum_{j=0}^{\beta}\sum_{k=0}^{\infty}\binom{\beta}{j}\frac{\log(\alpha)^k j^k}{k!}\int_0^\infty x^{2\beta+j}e^{-\theta\beta x}dx$$

$$R(\beta) = \frac{1}{1-\beta}\log\left(\frac{\theta^6}{2\theta^4+120\alpha}\right)^{\beta}\sum_{j=0}^{\beta}\sum_{k=0}^{\infty}\binom{\beta}{j}\frac{\log(\alpha)^k j^k}{k!}\frac{\Gamma 2\beta+j+1}{(\theta\beta)^{2\beta+j+1}}$$

III. Tsallis Entropy

A generalization of Boltzmann-Gibbs (B-G) statistical mechanics initiated by Tsallis has attracted a great deal of attention. This generalization of (B-G) statistics was proposed firstly by introducing the mathematical expression of Tsallis entropy (Tsallis, 1988) for a continuous random variable which is defined as follow.

$$T_\lambda = \frac{1}{\lambda-1}\left(1-\int_0^\infty f_a{}^{\lambda}(x)dx\right)$$

$$T_\lambda = \frac{1}{\lambda-1}\left(1-\int_0^\infty\left(\frac{\theta^6}{2\theta^4+120\alpha}x^2(\theta+\alpha x^3)e^{-\theta x}dx\right)^{\lambda}dx\right)$$

$$T_\lambda = \frac{1}{\lambda-1}\left(1-\left(\frac{\theta^6}{2\theta^4+120\alpha}\right)^{\lambda}\int_0^\infty x^{2\lambda}(\theta+\alpha x^3)^{\lambda}e^{-\theta x}dx\right)$$

Using by binomial expansion in above equation

$$T_\lambda = \frac{1}{\lambda-1}\left(1-\left(\frac{\theta^6}{2\theta^4+120\alpha}\right)^{\lambda}\sum_{j=0}^{\lambda}\sum_{k=0}^{\infty}\binom{\lambda}{j}\frac{\log(\alpha)^k j^k}{k!}\frac{\Gamma 2\lambda+j+1}{(\theta\lambda)^{2\lambda+j+1}}\right)$$

IX. Estimations of Parameter

In this section, the maximum likelihood estimates and Fisher's information matrix of the area biased samade distribution parameter is given.

Maximum Likelihood estimation (MLE)and Fisher's Information Matrix

Consider $(x_1, x_2, x_3, \ldots x_n)$be a random sample of size n from the area biased samade distribution. Then the likelihood function is given by

$$L(x) = \prod_{i=1}^{n} f_a(x)$$

$$L(x) = \prod_{i=1}^{n}\left(\frac{\theta^6}{2\theta^4+120\alpha}x_i{}^2(\theta+\alpha x_i{}^3)e^{-\theta x_i}\right)$$

$$L(x) = \frac{\theta^{6n}}{(2\theta^4+120\alpha)^n}\prod_{i=1}^{n}\left(x_i{}^2(\theta+\alpha x_i{}^3)e^{-\theta x_i}\right)$$

The log likelihood function is given by

$$\log L = 6n \log\theta\,(2\theta^4 + 120\alpha) + 2\sum_{i=1}^{n} \log x_i + \sum_{i=1}^{n} \log(\theta + \alpha {x_i}^3)$$
$$- \theta \sum_{i=1}^{n} x_i \qquad (11)$$

Now differentiating the log likelihood equation (11) with respect to parameters θ and α we must satisfy the normal equations as

$$\frac{\partial \log L}{\partial \theta} = \frac{6n}{\theta} - n\left(\frac{8\theta^3}{2\theta^4 + 120\alpha}\right) + n\left(\frac{1}{(\theta + \alpha {x_i}^3)}\right) - \sum_{i=1}^{n} x_i = 0 \qquad (12)$$

$$\frac{\partial \log L}{\partial \alpha} = -n\left(\frac{120}{2\theta^4 + 120\alpha}\right) + n\left(\frac{{x_i}^3}{(\theta + \alpha {x_i}^3)}\right) = 0 \qquad (13)$$

The equation (12) and (13) gives the maximum likelihood estimation of the parameters for the area biased samade distribution. However, the equation cannot be solved analytically, thus solved numerically using R programming with data set.

To obtain confidence interval use the asymptotic normality results. have that if $\hat{\lambda} = (\hat{\theta}, \hat{\alpha})$ denotes the MLE of $\lambda = (\theta, \alpha)$. state the results as follows.

$$\sqrt{n}(\hat{\lambda} - \lambda) \longrightarrow N_2(0, I^{-1}(\lambda))$$

Where $I(\lambda)$ is fisher's information matrix. i.e.

$$I(\lambda) = \begin{bmatrix} E\left[\frac{\partial^2 \log L}{\partial \theta^2}\right] & E\left[\frac{\partial^2 \log L}{\partial \theta \partial \alpha}\right] \\ E\left[\frac{\partial^2 \log L}{\partial \theta \partial \alpha}\right] & E\left[\frac{\partial^2 \log L}{\partial \alpha^2}\right] \end{bmatrix}$$

Here we see

$$E\left[\frac{\partial^2 \log L}{\partial \theta^2}\right] = -\frac{6n}{\theta^2} - n\left(\frac{2880\theta^2\alpha - 16\theta^2}{(2\theta^4 + 120\alpha)^2}\right) + n\left(\frac{(\theta + \alpha {x_i}^3) - 1}{(\theta + \alpha {x_i}^3)^2}\right)$$

$$E\left[\frac{\partial^2 \log L}{\partial \alpha^2}\right] = n\left(\frac{14400}{(2\theta^4 + 120\alpha)^2}\right) - n\left(\frac{\alpha {x_i}^3}{(\theta + \alpha {x_i}^3)^2}\right)$$

$$E\left[\frac{\partial^2 \log L}{\partial \theta \partial \alpha}\right] = -n\left(\frac{960\theta^3}{(2\theta^4 + 120\alpha)^2}\right) - n\left(\frac{\alpha {x_i}^3}{(\theta + \alpha {x_i}^3)^2}\right)$$

Since λ being unknown, estimate $I^{-1}(\lambda)$ by $I^{-1}(\hat{\lambda})$and this can be used to obtain asymptotic confidence interval for θ and α.

X. Applications

In this section, we have fitted a real lifetime data set in area biased Samade distribution to discuss its goodness of fit and the fit has been compared over Samade and Aradhana distributions. The real lifetime data set is given below as

The following real lifetime data set consists of 40 patients suffering from blood cancer (leukemia) reported from one of the ministry of health hospitals in Saudi Arabia (see Abouammah et al.). The ordered lifetimes (in year) is given below in Data set 1 as:

Data set 1: Data regarding the blood cancer (leukemia) patients reported by [3]

(0.315, 0.496, 0.616, 1.145, 1.208, 1.263, 1.414, 2.025, 2.036, 2.162, 2.211, 2.37, 2.532, 2.693, 2.805, 2.91, 2.912, 3.192, 3.263, 3.348, 3.348, 3.427, 3.499, 3.534, 3.767, 3.751, 3.858 ,3.986, 4.049, 4.244, 4.323, 4.381, 4.392, 4.397, 4.647, 4.753, 4.929, 4.973 ,5.074 ,5.381)

To compare to the goodness of fit of the fitted distribution, the following criteria: Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC), Akaike Information Criteria Corrected (AICC) and $-2 \log L$.
AIC, BIC, AICC and $-2 \log L$ can be evaluated by using the formula as follows.

$$AIC = 2k - 2\log L\,, \quad BIC = k \log n - 2\log L \; and \; AICC = AIC + \frac{2k(k+1)}{(n-k-1)}$$

Where, k = number of parameters, n sample size and $-2 \log L$ is the maximized value of loglikelihood function.

Table 1: *MLEs AIC, BIC, AICC, and -2log L of the fitted distribution for the given data set 1*

Distribution	ML Estimates	-2 log L	AIC	BIC	AICC
Area biased Samade distribution	$\hat{\boldsymbol{\alpha}} = \mathbf{0.5475209\ (0.4673469)}$ $\hat{\boldsymbol{\theta}} = \mathbf{1.7122824\ (0.1423247)}$	**141.1354**	**145.1354**	**148.5131**	**145.4597**
Samade distribution	$\hat{\alpha} = 4.4717985\ (4.5855767)$ $\hat{\theta} = 1.2041767\ (0.1058379)$	144.0608	148.0680	151.4385	148.3923
Aradhana	$\hat{\theta} = 0.74955550\ (0.07008456)$	153.1011	155.1011	156.79	155.4254

From table 1 it can be clearly observed and seen from the results that the area biased samade distribution have the lesser AIC, BIC, AICC, -2log L, and values as compared to the Samade and Aradhana, which indicates that the area biased samade distribution better fits than the Samade and Aradhana distributions. Hence, it can be concluded that the area biased samade distribution leads to a better fit over the other distributions.

XI. Conclusion

Research has focused a great deal of attention on selecting an appropriate model for fitting survival data. In this paper, the samade distribution is extended to provide a new distribution called the area biased samade distribution for the model's lifetime data. It has various special cases that have been presented in the paper. The statistical properties of the distribution have been studied, including survival and hazard functions, moments, mean and, median deviations, moment generating functions, Entropies, Bonferroni and Lorenz curve, and order statistics. The Inference of parameters for a area biased Samade distribution was obtained using the method of maximum likelihood estimates. When the parameters have been estimated using the maximum likelihood

method, a good performance is seen. The application of statistical distributions is critical for medical research and can significantly affect public health, especially for cancer patients.
Thus, the usefulness of this distribution is illustrated through its applications to the survival of some cancer patients, including both complete and censored cases. In using various goodness-of-fit criteria, including AIC, BIC, AICC, and -2logL, the results demonstrate the superior performance of the area biased samade distribution. Overall, it is intended that the area biased Samade distribution that is given in table 1 will offer a better fit than other existing distributions for simulating real life data in survival analysis, specifically for cancer data.

References

[1] Aderoju, S. (2021). Samade probability distribution: its Properties and Application to real lifetime data. *Asian Journal of Probability and Statistics*, 11:1-11.

[2] Ade, R. B. et al. (2021). Characterization and estimation of area biased quasi-Akash distribution. *International Journal of Mathematics Trends and Technology,* 67:53-59.

[3] Abouammoh, A. M. Ahmed, R. and Khalique, A. (2000). On new renewal better than used classes of life distribution. *Statistics and Probability Letters*, 48:189-194.

[4] Chouia, S. Zeghdoudi, H. Raman, V. and Beghriche, A. (2021). A new size biased distribution with application. *Journal of Applied Probability and Statistics*, 16:111-125.

[5] Fisher, R. A. (1934). The effects of methods of ascertainment upon the estimation of frequencies. Annals of Eugenics, 6:13-25.

[6] Gupta, R. C. and Keating, J. P. (1986). Relations for reliability measures under length biased sampling. *Scandinavian Journal of Statistics*, 13:49-56.

[7] Jayakumar. and Elangovan, R. (2019). A New Length Biased Distribution with Blood Cancer Data. *Journal of Information and Computational Science*, 9:463-471.

[8] Oluwafemi, OS. Olalekan, DM. (2017). Length and area biased exponentiated Weibull distribution based on forest inventories. *Biometrics & Biostatistics International Journal*, 6:311-320.

[9] Osowole, OI. Amaefula, CG. and Ngozi, N. (2020). On the area biased quasi transmuted uniform distribution. *International Journal of Innovative Mathematics, Statistics and Energy Policies*, 8:13-23.

[10] Patil, G. P. and Rao, C. R. (1978). Weighted distributions and Size biased sampling with applications to wildlife populations and human families. Biometrics, 34:179-189.

[11] Perveen, Z. Munir, M. Ahmed, Z. and Ahmad, M. (2016). On area biased weighted weibull distribution. *Sci. Int. (Lahore)*, 28:3669- 3679.

[12] Para, B. A. and Jan, T. R. (2018). On Three Parameter Weighted Pareto Type II Distribution: Properties and Applications in Medical Sciences. *Applied Mathematics & Information Sciences Letters*, 6:13-26.

[13] Shanker, R. and Shukla, K. K. (2017). Weighted Shanker distribution and its applications to model lifetime data‖. *Journal of Applied Quantitative Methods*, 12:1 – 17.

[14] Rao, C. R. and Patil, G. P. (1965). On discrete distributions arising out of methods of ascertainment, in Classical and Contagious Discrete Distribution. *Pergamon Press and Statistical Publishing Society, Calcutta,* 320-332.

[15] Rajagopalan, V. Rashid, A, Ganaie. and Aafaq, A, Rather. (2019). A New Length Biased Distribution with Applications. *Science Technology and Development*, 8:161-174.

[16] Shanker, R. (2022). Uma Distribution with properties and applications, *Biometrics & Biostatistics International Journal*, 11: 165-169.

[17] Sharma, VK. Dey, S. Singh, SK. and Manzoor, U. (2017). On length and area-biased Maxwell distributions. *Communications in Statistics Simulation and Computation*, 47:1506-1528.