A stochastic model of area-biased Kpenadidum distribution with the characteristics and applications to real-lifetime data

R. POOVITHA AND P. PANDIYAN

Abstract

In this present research, we explore the statistical characteristics of the Area-Biased Kpenadidum Distribution (ABKD), a novel probability model. The maximum likelihood method has been used to estimate the parameters, and the asymptotic findings have been explained. The new distribution was compared to the Shanker, Lindley, and Kpenadidum distributions. When the distribution was fitted to cancer data, a good fit was observed.

Mathematics Subject Classification 2000: 62F10, 62E15, 62N05
Key Words and Phrases: Kpenadidum distribution, Area-Biased distribution, Reliability analysis, Ordered statistics, Maximum likelihood estimation, and Breast cancer.

1. INTRODUCTION

The idea of a weighted distribution was introduced by Fisher (1934), that when sample observations have an unequal probability of selection, we apply weights to the distribution to model bias. Rao (1965) subsequently and consistently modified while modeling the statistical data when the standard distributions were not appropriate to record the observations with equal probabilities. Weighted distributions can address many issues in various research areas, specifically in biological research. And also, it is essential for modeling and interpreting lifetime data in many applied fields, including engineering, medicine, behavioural science, finance, insurance, and others. The distribution that results when the weight function considers the length of the units of interest is length-biased. Thus, length-biased distribution is a special case of the more general form of weighted distribution. Cox (1969) and Zelen (1974) introduced the length-biased sampling concept.

10.2478/jamsi-2025-0003

More generally, when the sampling mechanism selects units with probability proportional to some measure of the unit size, the resulting distribution is called size-biased. A comprehensive review of weighted distributions can be found in several reliable sources. Many different weighted probability models have been evaluated and examined, and their applicability in various disciplines is illustrated by numerous authors.

Shakila & Mujahid Rasul (2016) developed the Poisson Area–Biased Lindley distribution and applied it to biological data to demonstrate that it provides a better fit than the Poisson distribution. Kilany (2016) obtained the weighted version of the Lomax distribution. Fazal and Bashir (2018) presented the discrete Poisson exponential distribution (PED), which may be area-biased to produce an area-biased variation of the single parameter PED. Khan et al. (2018) discussed the weighted modified Weibull distribution. Rajagopalan et al., (2019) obtained a new length-biased Aradhana distribution with engineering and medical science applications showing more flexibility than classical distributions. Rather et al. (2018) produced a novel size-biased Ailamujia distribution that exhibits greater flexibility than classical distributions and has applications in engineering and medical science. Rather and Subramanian (2018) discussed the characterization and estimation of length-biased weighted generalized uniform distribution. Rather and Subramanian (2019) also discussed length-biased Sushila distribution with properties and Applications. Shenbagaraja et al (2019) have discussed length-biased Garima distribution with real-life data. Anthony & Elangovan, (2020) studied a new generalization of quasi-Aradhana distribution using the length-biased technique and also the length-biased Om distribution. Elangovan & Mohanasundari, (2019) have discussed the area-biased Aradhana distribution and fitted it for lifetime data.

By transforming the Kpenadidum distribution - recently proposed by Nwikpe and Cleopas (2022) as a parametric lifetime model for various engineering and medical science applications - we have developed the Area-biased Kpenadidum distribution in this study. The distribution consists of three components: an exponential distribution with a parameter, a gamma distribution with shape parameter three and scale parameter, and a gamma distribution with shape parameter four and scale parameter.

This work proposes an Area-biased Kpenadidum distribution, which has been used to determine whether the survival dataset is appropriate. The proposed distribution was more significant and practical for lifetime data than Lindley and various other distributions. The lifetime dataset considered in this paper was cancer data.
The probability density function (pdf) of the Kpenadidum distribution is given by

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

and the cumulative density function (CDF) of Kpenadidum distribution is given by

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

2. AREA-BIASED KPENADIDUM DISTRIBUTION (ABKD)

A non-negative random variable X is with probability density function $f(x)$ and $w(x)$ be the non-negative weight function, the pdf of the weighted random variable x_w is given by

$$f_w(x) = \frac{w(x).f(x)}{E(w(x))}\,; \qquad x>0 \tag{2}$$

Where $w(x)$ be a non-negative weight function and

$$E\big(w(x)\big) = \int w(x)f(x)dx < \infty$$

In this research paper, we have considered the Area-Biased version of the Kpenadidum distribution known as Area-Biased Kpenadidum distribution (ABKD). It should be noted that there can be different choices of weight function. Therefore at $w(x) = x^2$, the resulting distribution is called Area-Biased Kpenadidum distribution.
The probability density function of Area-Biased Kpenadidum distribution (pdf) is given by

$$f_A(x) = \frac{x^2.f(x)}{E(x)}\,; \qquad x>0 \tag{3}$$

Where

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

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

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

$$E(x^2) = \frac{2(\theta^4+6\theta+60)}{\theta^2(\theta^4+\theta+6)} \tag{4}$$

On substituting equations (1) and (4) in equation (3), we get the required probability density function of Area-Biased Kpenadidum distribution.

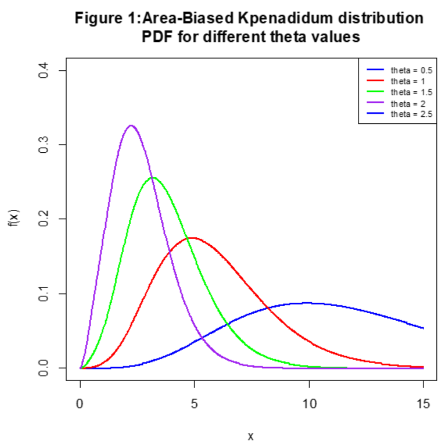
$$f_A(x;\theta) = \left[x^2.\frac{\theta^4}{2(\theta^4+\theta+6)} (2x^3+x^2+2\theta)e^{-\theta x}\right] * \left[\frac{\theta^2(\theta^4+\theta+6)}{2(\theta^4+6\theta+60)}\right]$$

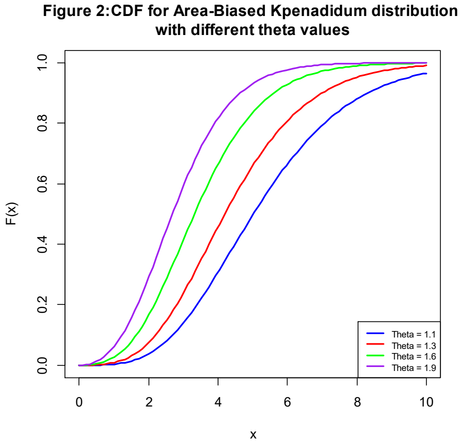
$$\therefore\ f_A(x;\theta) = \left[x^2.\frac{\theta^6}{4(\theta^4+6\theta+60)} (2x^3+x^2+2\theta)e^{-\theta x}\right] \tag{5}$$

The cumulative density function (cdf) of the Area-Biased Kpenadidium (ABKD) distribution is obtained as

$$F_A(x;\theta) = \int_0^x f_A(x;\theta)\, dx$$

$$F_A(x;\theta) = \int_0^x x^2.\frac{\theta^6}{4(\theta^4+6\theta+60)} (2x^3+x^2+2\theta)e^{-\theta x}\, dx$$





after simplification, we get the cumulative distribution function of Area-Biased Kpenadidum (ABKD) as

$$F_A(x;\theta) = \frac{\gamma(6,\theta x) + \theta\gamma(5,\theta x) + 2\theta^4\gamma(3,\theta x)}{4(\theta^4 + 6\theta + 60)} \quad (6)$$

3. RELIABILITY ANALYSIS

In this section, we discuss the survival function, failure rate, reverse hazard rate, and Mill's Ratio of the Area-Biased Kpenadidum distribution (ABKD).

3.1 Survival Function

The probability that a case survives beyond a specific time is known as the reliability function or survival function and is given by

$$S(x) = 1 - F_A(x;\theta)$$

$$S(x) = 1 - \frac{\gamma(6,\theta x) + \theta\gamma(5,\theta x) + 2\theta^4\gamma(3,\theta x)}{4(\theta^4 + 6\theta + 60)}$$

3.2 Hazard Function

The hazard function is also known as the hazard rate, instantaneous failure rate or force of mortality and is given by

$$h(x) = \frac{f_A(x;\theta)}{s(x)}$$

$$h(x) = \frac{\theta^6 x^2(2x^3 + x^2 + 2\theta)e^{-\theta x}}{4(\theta^4 + 6\theta + 60) - \ \gamma(6,\theta x) + \theta\gamma(5,\theta x) + 2\theta^4\gamma(3,\theta x)}$$

3.3 Reverse Hazard Function and Mills Ratio

The reverse hazard function of Area-Biased Kpenadidum distribution is

$$h_r(x) = \frac{f_A(x;\theta)}{F_A(x;\theta)}$$

$$h_r(x) = \frac{\theta^6 x^2(2x^3 + x^2 + 2\theta)e^{-\theta x}}{\gamma(6,\theta x) + \theta\gamma(5,\theta x) + 2\theta^4\gamma(3,\theta x)}$$

and the Mills ratio of the Area-Biased Kpenadidum distribution is

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

$$\textbf{Mills Ratio} = \frac{\boldsymbol{\gamma(6,\theta x) + \theta\gamma(5,\theta x) + 2\theta^4\gamma(3,\theta x)}}{\boldsymbol{\theta^6 x^2(2x^3 + x^2 + 2\theta)e^{-\theta x}}}$$

Figure 3: Survival Function for Area-Biased Kpenadidum distribution

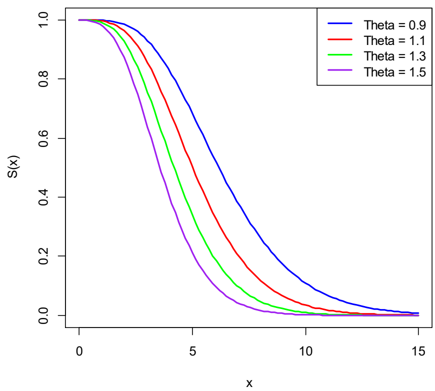
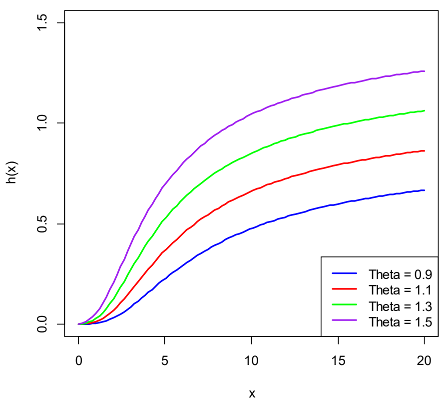


Figure 4: Hazard Function for Area-Biased Kpenadidum distribution



4. STATISTICAL PROPERTIES

4.1 Moments

Let X denote the random variable of Area-Biased Kpenadidum distribution with parameter θ, then the r^{th} order moment $E(x^r)$ of Area-Biased Kpenadidum distribution (ABKD) can be obtained as

$$E(x^r) = \mu_r' = \int_0^\infty x^r . f_A(x;\theta)\, dx$$

$$E(x^r) = \mu_r' = \int_0^\infty x^r . \frac{\theta^6 . x^2}{4(\theta^4 + 6\theta + 60)} (2x^3 + x^2 + 2\theta) e^{-\theta x}\, dx$$

$$E(x^r) = \frac{\theta^6}{4(\theta^4 + 6\theta + 60)} \int_0^\infty x^{r+2} . (2x^3 + x^2 + 2\theta) e^{-\theta x}\, dx$$

$$E(x^r) = \frac{\theta^6}{4(\theta^4 + 6\theta + 60)} \left[2 \int_0^\infty x^{r+5}\, e^{-\theta x} dx + \int_0^\infty x^{r+4}\, e^{-\theta x} dx + 2\theta \int_0^\infty x^{r+2}\, e^{-\theta x} dx\, e^{-\theta x} \right]$$

$$E(x^r) = \frac{\theta^6}{4(\theta^4 + 6\theta + 60)} \left[\frac{2\Gamma(r+6) + \theta\, \Gamma(r+5) + 2\theta^4 \Gamma(r+3)}{\theta^{r+6}} \right]$$

$$E(x^r) = \mu_r' = \left[\frac{2\Gamma(r+6) + \theta\, \Gamma(r+5) + 2\theta^4 \Gamma(r+3)}{\theta^r 4\, (\theta^4 + 6\theta + 60)} \right] \quad (7)$$

Where $\Gamma(.)$ is the gamma function. Followed by substituting r = 1, 2, 3, and 4 in equation (7), we get the first four moments of ABKD.

$$E(x^1) = \mu_1' = \left[\frac{2\Gamma(7) + \theta\, \Gamma(6) + 2\theta^4 \Gamma(4)}{\theta^1 4\, (\theta^4 + 6\theta + 60)} \right]$$

$$\mu_1' = \left[\frac{3(\theta^4 + 10\theta + 120)}{\theta(\theta^4 + 6\theta + 60)} \right]$$

$$E(x^2) = \mu_2' = \left[\frac{2\Gamma(8) + \theta\, \Gamma(7) + 2\theta^4 \Gamma(5)}{\theta^2 4\, (\theta^4 + 6\theta + 60)} \right]$$

$$\mu_2' = \left[\frac{12(\theta^4 + 15\theta + 210)}{\theta^2\, (\theta^4 + 6\theta + 60)} \right]$$

Therefore,

$$\text{variance } = \left[\frac{12(\theta^4+15\theta+210)}{\theta^2\,(\theta^4+6\theta+60)}\right]-\left[\frac{3(\theta^4+10\theta+120)}{\theta(\theta^4+6\theta+60)}\right]^2$$

$$\mu_2=\frac{3(\theta^8+24\theta^5+360\theta^4+60\theta^2+1440\theta+7200)}{\theta^2\,(\theta^4+6\theta+60)^2}$$

The standard deviation (σ), Coefficient of variation (C.V.) and index of dispersion (r) can be obtained as

$$\text{Standard Deviation } (\sigma)=\frac{\sqrt{3(\theta^8+24\theta^5+360\theta^4+60\theta^2+1440\theta+7200)}}{\theta\,(\theta^4+6\theta+60)}$$

$$\text{Index of dispersion } (r)=\frac{\sigma^2}{\mu_1}$$

$$\text{Index of dispersion } (r)=\frac{3(\theta^8+24\theta^5+360\theta^4+60\theta^2+1440\theta+7200)}{3\theta\,\,(\theta^4+6\theta+60)(\theta^4+10\theta+120)}$$

The Harmonic mean of the proposed distribution is obtained as

$$\mathrm{H.M}=E\left[\frac{1}{x}\right]=\int_0^\infty\frac{1}{x}.f_A(x;\theta)dx$$

$$\mathrm{H.M}=\int_0^\infty\frac{1}{x}.\frac{\theta^6.x^2}{4(\theta^4+6\theta+60)}(2x^3+x^2+2\theta)\,e^{-\theta x}dx$$

$$=\frac{\theta^6}{4(\theta^4+6\theta+60)}\int_0^\infty\frac{1}{x}.x^2(2x^3+x^2+2\theta)\,e^{-\theta x}dx$$

$$\mathrm{H.M}=\frac{\theta\,(\theta^4+3\theta+12)}{2(\theta^4+6\theta+60)}$$

4.2 Moment generating function and characteristic function

Let X have an ABK distribution, and then the MGF of X is obtained as

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

Using Taylor's Series,

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

$$= \sum_{j=0}^{\infty} \frac{t^j}{j!} \int_0^{\infty} x^j \, f_A(x;\theta)\, dx$$

$$= \sum_{j=0}^{\infty} \frac{t^j}{j!} \, \mu_j$$

$$= \sum_{j=0}^{\infty} \frac{t^j}{j!} \left[\frac{2\Gamma(j+6) + \theta\, \Gamma(j+5) + 2\theta^4 \Gamma(j+3)}{\theta^j 4\,(\theta^4 + 6\theta + 60)} \right]$$

$$M_X(t) = \frac{1}{2\,(\theta^4 + 3\theta + 24)} \sum_{j=0}^{\infty} \frac{t^j}{j!\ \theta^j} (2\Gamma(j+5) + \theta\, \Gamma(j+4) + 2\theta^4 \Gamma(j+2))$$

Similarly, the characteristic function of ABK distribution can be obtained as

$$\varphi_X(t) = M_X(it)$$

$$\varphi_X(t) = \frac{1}{4(\theta^4 + 6\theta + 60)} \sum_{j=0}^{\infty} \frac{(it)^j}{j!\ \theta^j} [2\Gamma(j+6) + \theta\, \Gamma(j+5) + 2\theta^4 \Gamma(j+3)]$$

5. ORDERED STATISTICS

Let $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ be the order statistics of a random sample $X_1, X_2, \dots, X_n$ from the continuous population with probability density function $f_X(x)$ and cumulative density function with $F_X(x)$ then the pdf of r^th order statistics $X_{(r)}$ can be written as

$$f_{r:n}(x) = \frac{n!}{(r-1)!\,(n-r)!} [F(x;\theta)]^{r-1} [1 - F(x;\theta)^{n-r} f(x;\theta), \;\; x > 0 \qquad (8)$$

Substituting the value of equations (5) and (6) in equation (8), we get the pdf of r^th order statistics $X_{(r)}$ for Area-Biased Kpenadidum distribution and is given by

$$f_{r:n}(x) = \frac{n!}{(r-1)!\,(n-r)!} \left[\frac{\gamma(6,\theta x) + \theta\gamma(5,\theta x) + 2\theta^4 \gamma(3,\theta x)}{4(\theta^4 + 6\theta + 60)} \right]^{r-1}$$

$$\times \left[1 - \frac{\gamma(6,\theta x) + \theta\gamma(5,\theta x) + 2\theta^4 \gamma(3,\theta x)}{4(\theta^4 + 6\theta + 60)} \right]^{n-r}$$

$$\times \left[x^2 . \frac{\theta^6}{4(\theta^4 + 6\theta + 60)} \; (2x^3 + x^2 + 2\theta) e^{-\theta x} \right]$$

$$f_{r:n}(x) = \frac{n!}{(r-1)!\,(n-r)!}\sum_{j=0}^{\infty}(-1)^j\binom{n-r}{j}\left[\frac{\gamma(6,\theta x)+\theta\gamma(5,\theta x)+2\theta^4\gamma(3,\theta x)}{4(\theta^4+6\theta+60)}\right]^{r+j-1}$$
$$\times\left[x^2.\frac{\theta^6}{4(\theta^4+6\theta+60)}\,(2x^3+x^2+2\theta)e^{-\theta x}\right] \tag{9}$$

Therefore, the probability density function of 1st order statistics $x_{(1)}$ can be obtained as

$$f_{1:n}(x) = \frac{n!}{(n-1)!}\sum_{j=0}^{\infty}(-1)^j\binom{n-1}{j}\left[\frac{\gamma(6,\theta x)+\theta\gamma(5,\theta x)+2\theta^4\gamma(3,\theta x)}{4(\theta^4+6\theta+60)}\right]^{j}$$
$$\times\left[x^2.\frac{\theta^6}{4(\theta^4+6\theta+60)}\,(2x^3+x^2+2\theta)e^{-\theta x}\right]$$

And the probability density function of nth order statistics is

$$f_{n:n}(x) = \frac{n!}{(n-1)!}\sum_{j=0}^{\infty}(-1)^j\binom{n-1}{j}\left[\frac{\gamma(6,\theta x)+\theta\gamma(5,\theta x)+2\theta^4\gamma(3,\theta x)}{4(\theta^4+6\theta+60)}\right]^{n+j-1}$$
$$\times\left[x^2.\frac{\theta^6}{4(\theta^4+6\theta+60)}\,(2x^3+x^2+2\theta)e^{-\theta x}\right]$$

6. LIKELIHOOD RATIO TEST

Let $X_1, X_2, \ldots, X_n$ be a random sample from the Area-biased Kpenadidum distribution. To test the hypothesis

$H_0: f(x) = f(x;\theta)\ against\ H_1: f(x) = f_A(x;\theta)$

The following test statistics are used to test whether the random sample of size n comes from the Kpenadidum distribution or an Area-Biased Kpenadidum distribution.

$$\Delta = \frac{L_1}{L_0} = \prod_{i=1}^{n}\frac{f_A(x;\theta)}{f(x;\theta)}$$

$$\Delta = \frac{L_1}{L_0} = \prod_{i=1}^{n}\frac{\left[x^2.\frac{\theta^6}{4(\theta^4+6\theta+60)}\,(2x^3+x^2+2\theta)e^{-\theta x}\right]}{\left[\frac{\theta^4}{2(\theta^4+\theta+6)}\,(2x^3+x^2+2\theta)e^{-\theta x}\right]}$$

$$\Delta = \frac{L_1}{L_0} = \prod_{i=1}^{n}{x_i}^2\left[\frac{\theta^2(\theta^4+\theta+6)}{2(\theta^4+6\theta+60)}\right]^n$$

Thus, we reject the null hypothesis if

$$\left[\frac{\theta^2(\theta^4+\theta+6)}{2(\theta^4+6\theta+60)}\right]^n \prod_{i=1}^{n} {x_i}^2 > K$$

Equivalently we reject the null hypothesis when

$$\Delta^* = \prod_{i=1}^{n} {x_i}^2 > \; K^*$$

Where

$$K^* = K\left[\frac{\theta^2(\theta^4+\theta+6)}{2(\theta^4+6\theta+60)}c\right]^n > 0$$

Thus, for a large sample size of n, 2 $log\Delta$ is distributed as a chi-square distribution with one degree of freedom and the p value is obtained from the chi-square distribution.

Thus, we reject the null hypothesis, when the probability value is given by, $P(\Delta^* > \alpha^*)$ where $\alpha^* = \prod_{i=1}^{n} {x_i}^2$ is the observed value of statistics Δ^*

7. BONFERRONI AND LORENZ CURVE

The Bonferroni and Lorenz curve is used in economics to study income and poverty and in other fields like reliability, medicine, demography, and actuarial statistics. The Bonferroni and Lorenz curves are given by

$$B(p) = \frac{1}{p\mu_1'}\int_0^q x.f_A(x;\theta)dx \text{ and}$$

$$L(p) = \frac{1}{\mu_1'}\int_0^q x.f_A(x;\theta)dx$$

where,

$$\mu_1' = \left[\frac{3(\theta^4+10\theta+120)}{\theta\,(\theta^4+6\theta+60)}\right] \; and \; q = F^{-1}(p)$$

$$B(p) = \frac{\theta\,(\theta^4+6\theta+60)}{p3\,(\theta^4+10\theta+120)}\int_0^q x^3.\frac{\theta^6}{4(\theta^4+6\theta+60)}\,(2x^3+x^2 + 2\theta)e^{-\theta x}dx$$

$$B(p) = \frac{\theta^7}{p.12\,(\theta^4+10\theta+120)}\int_0^q x^3.(2x^3+x^2+2\theta)e^{-\theta x}dx$$

$$B(p) = \frac{2\gamma(7,\theta q) + \theta\gamma(6,\theta q) + 2\theta^4\gamma(4,\theta q)}{12p(\theta^4 + 10\theta + 120)}$$

And Lorenz curve is

$$L(p) = \frac{2\gamma(7,\theta q) + \theta\gamma(6,\theta q) + 2\theta^4\gamma(4,\theta q)}{12(\theta^4 + 10\theta + 120)}$$

8. STOCHASTIC ORDERING

Stochastic ordering is an important tool in finance and reliability theory to assess the comparative performance of the models. Let *X* and *Y* be two random variables with pdf, cdf, and reliability function

$$f(x;\theta), f(y;\theta), F(x;\theta), F(y;\theta), S(x) = 1 - F(x;\theta) \; and \; S(y) = 1 - F(x;\theta)$$

Then,

- Mean residual life order denoted by

$$X \leq_{MRLO} Y, If \; m_x(x) \leq m_y(y), \forall x$$

- Hazard rate order denoted as

$$X \leq_{HRO} Y, If \; \frac{S_X(x)}{S_Y(y)} \; is \; decreasing \; if \; x \geq 0$$

- Stochastic order denoted as

$$X \leq_{SO} Y, If \; S_X(x) \; \leq_{SO} S_Y(x), \;\; \forall x$$

- Likelihood ratio order denoted as

$$X \leq_{LRO} Y, If \; \frac{f_X(x;\theta)}{f_Y(x;\theta)}$$

Assume that *X* and *Y* are two independent random variables with pdf $f_X(x;\theta)$ and $f_Y(x;\lambda)$ If $\theta < \lambda$

$$\Lambda = \frac{f_x(x;\theta)}{f_y(x;\lambda)}$$

$$\Lambda = \frac{\left(x^2 . \frac{\theta^6}{4(\theta^4 + 6\theta + 60)} \; (2x^3 + x^2 + 2\theta)e^{-\theta x}\right)}{\left(x^2 . \frac{\lambda^6}{4(\lambda^4 + 6\lambda + 60)} \; (2x^3 + x^2 + 2\lambda)e^{-\lambda x}\right)}$$

$$\Lambda = \left(\frac{\theta^6(\lambda^4 + 6\lambda + 60)}{\lambda^6(\theta^4 + 6\theta + 60)}\right) \times \left(\frac{(2x^3 + x^2 + 2\theta)}{(2x^3 + x^2 + 2\lambda)}\right) \times e^{(\lambda - \theta)x}$$

Thus,

$$\log(\Lambda) = log(\theta^6(\lambda^4 + 6\lambda + 60)) - log(\lambda^6(\theta^4 + 6\theta + 60)) + log(2x^3 + x^2 + 2\theta) - log(2x^3 + x^2 + 2\lambda) + (\lambda - \theta)x$$

Differentiating to x we get,

$$\frac{\partial \log(\Lambda)}{\partial x} = \left(\frac{6x^2 + 2x}{(2x^3 + x^2 + 2\theta)}\right) - \left(\frac{6x^2 + 2x}{(2x^3 + x^2 + 2\lambda)}\right) + (\lambda - \theta)$$

$$\frac{\partial \log(\Lambda)}{\partial x} < 0 \; if \; \lambda < \theta$$

Thus,

$$X \leq_{LRO} Y, X \leq_{HRO} Y, X \leq_{MRLO} Y, and\ X \leq_{SO} Y$$

9. ENTROPIES

An entropy can be considered as a measure of variability for absolutely continuous random variables or a measure of variation or diversity of the possible values of a discrete random variable, it can be used for risk assessment in various domains.

9.1 Renyi Entropy

The entropy of a random variable x measures the variation of the uncertainty. The Renyi entropy is defined by

$$R_\gamma = \frac{1}{(1-\gamma)} \log \int_0^\infty \left(f_A(x;\theta)\right)^\gamma dx\,, \qquad \gamma > 0, \gamma \neq 1$$

$$R_\gamma = \frac{1}{(1-\gamma)} \log \int_0^\infty \left(x^2 . \frac{\theta^6}{4(\theta^4 + 6\theta + 60)} (2x^3 + x^2 + 2\theta) e^{-\theta x}\right)^\gamma dx$$

$$R_\gamma = \frac{1}{(1-\gamma)} \log \left[\frac{\theta^6}{4(\theta^4 + 6\theta + 60)}\right]^\gamma \int_0^\infty (2x^3 + x^2 + 2\theta)^\gamma x^{2\gamma} . e^{-\theta\gamma x}\, dx$$

Using a Binomial series of expansion,

$$R_\gamma = \frac{1}{(1-\gamma)} \log \left[\frac{\theta^6}{4(\theta^4 + 6\theta + 60)}\right]^\gamma \times \sum_{k=0}^{\gamma} \binom{\gamma}{k} (2\theta)^{\gamma-k} \sum_{n=0}^{\infty} \binom{k}{n} 2^n \left[\frac{\Gamma(2(\gamma + k) + n + 1}{(\theta\gamma)^{2(\gamma+k)+n+1}}\right]$$

9.2 Tsallis Entropy

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

$$T_\gamma = \frac{1}{\gamma - 1}\left[1 - \int_0^\infty \left(f_A(x;\theta)\right)^\gamma dx\right]; \gamma > 0, \gamma \neq 1$$

$$T_\gamma = \frac{1}{\gamma - 1}\left(1 - \int_0^\infty \left(x^2 . \frac{\theta^6}{4(\theta^4 + 6\theta + 60)} \left(2x^3 + x^2 + 2\theta\right)e^{-\theta x}\right)^\gamma dx\right)$$

$$T_\gamma = \frac{1}{\gamma - 1}\left(1 - \frac{\theta^{\gamma 6}}{\left(4(\theta^4 + 6\theta + 60)\right)^\gamma}\int_0^\infty x^{\gamma 2} . e^{-\theta\gamma x}\left(2x^3 + x^2 + 2\theta\right)^\gamma dx\right)$$

Again, using a Binomial series of expansion,

$$T_\gamma = \frac{1}{\gamma - 1}\left(1 - \frac{\theta^{\gamma 6}}{\left(4(\theta^4 + 6\theta + 60)\right)^\gamma} \times \sum_{k=0}^{\gamma}\binom{\gamma}{k}(2\theta)^{\gamma - k}\sum_{n=0}^{\infty}\binom{k}{n}2^n\left(\frac{\Gamma(2(\gamma + k) + n + 1)}{(\theta\gamma)^{2(\gamma + k) + n + 1}}\right)\right)$$

10. MAXIMUM LIKELIHOOD ESTIMATION AND FISHER'S INFORMATION MATRIX

In this section, we discuss the maximum likelihood estimator of the parameter of Area-Biased Kpenadidum distribution. The maximum likelihood method is one of the most useful methods for estimating the different parameters of the distribution.

Let $X_1, X_2, \ldots, X_n$ be the random sample of size n from the Area-biased Kpenadidum distribution, then the likelihood function can be written as

$$L(x;\theta) = \prod_{i=1}^{n} f_A(x;\theta)$$

$$L(x;\theta) = \prod_{i=1}^{n}\left[x_i{}^2 . \frac{\theta^6}{4(\theta^4 + 6\theta + 60)} \left(2x_i^3 + x_i^2 + 2\theta\right)e^{-\theta x_i}\right]$$

The log-likelihood function is

$$\log L(x;\theta) = 6n\, log\theta - n\log(4(\theta^4 + 6\theta + 60)) + 2\sum_{i=1}^{n}\log x_i + \sum_{i=1}^{n}\log(2{x_i}^3 + {x_i}^2 + 2\theta) - \theta\sum_{i=1}^{n}x_i \qquad (10)$$

The maximum likelihood estimate of θ can be obtained by differentiating equation (10) to θ and must satisfy the normal equation.

$$\frac{\partial \log L}{\partial\theta} = \frac{6n}{\theta} - n\left[\frac{16\theta^3 + 24}{4\theta^4 + 24\theta + 240}\right] + \sum_{i=1}^{n}\frac{2}{(2{x_i}^3 + {x_i}^2 + 2\theta)} - \sum_{i=1}^{n}x_i = 0$$

Because of the complicated form of the likelihood equation, it is difficult to solve the system of non-linear equations. Whence, we have used R and Wolfram Mathematica to estimate the required parameter of the proposed distribution. To get the confidence interval, we used the asymptotic normality test. We have that if $\hat{\lambda} = (\widehat{\theta})$ denotes the maximum likelihood estimates of $\lambda = (\theta)$, we state the result as follows

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

Where I(λ) is Fisher's Information matrix i.e.,

$$I(\lambda) = -\frac{1}{n}E\left(\frac{\partial^2 \log L}{\partial\theta^2}\right)$$

Where

$$E\left(\frac{\partial^2 \log L}{\partial\theta^2}\right) = \frac{-6n}{\theta^2} + \frac{4(4\theta^6 - 6\theta^3 - 180\theta^2 + 9}{(\theta^4 + 6\theta + 60)^2} - \sum_{i=1}^{n}\frac{4}{(2{x_i}^3 + {x_i}^2 + 2\theta)^2}$$

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

11. APPLICATION

In this section, we fit the real lifetime data set (Secondary data) to discuss the goodness of fit of Area-Biased Kpenadidum distribution with an example related to cancer data.

Dataset 1: The data represents the age (in years) of the 100 patients with lung cancer as reported by Rashad Mammadov. (2024).

68, 58, 44, 72, 37, 50, 68, 48, 52, 40, 40, 53, 65, 69, 53, 32, 51, 31, 53, 73, 59, 67, 31, 50, 62, 41, 51, 73, 54, 78, 56, 71, 57, 45, 44, 76, 73, 32, 66, 36, 50, 38, 68, 47, 33, 54, 43, 79, 38, 55, 31, 49, 57, 76, 36, 73, 37, 76, 64, 43, 46, 65, 79, 69, 33, 31, 35, 71, 33, 58, 47, 55, 73, 63, 39, 65, 43, 60, 77, 44, 37, 43, 52, 69, 50, 45, 74, 47, 76, 53, 55, 54, 74, 70, 58, 44, 74, 30, 54, 36, 38

Dataset 2: The data representing the size of the Tumour (in mm) of 100 patients with node-positive breast cancer, as reported by Royston and Altman (2013).

18, 20, 40, 25, 30, 52, 21, 20, 20, 30, 12, 30, 21, 25, 57, 25, 21, 27, 12, 45, 18, 18, 30, 30, 25, 14, 21, 30, 23, 30, 55, 11, 40, 25, 15, 29, 15, 45, 40, 30, 60, 21, 20, 58, 30, 28, 25, 18, 40, 35, 30, 21, 45, 40, 20, 19, 52, 80, 40, 20, 80, 30, 33,16, 25, 39, 52, 50, 39, 42, 65, 28, 25, 23, 32, 15, 40, 35, 25, 30, 22, 45, 20, 27, 29, 22, 24, 25, 23, 30, 50, 45, 7, 8, 18, 16, 28, 21, 15, 20, 12

Dataset 3: The data representing the recurrence-free time from the treatment (in months) of 206 patients with node-positive breast cancer, as reported by Royston and Altman (2013).

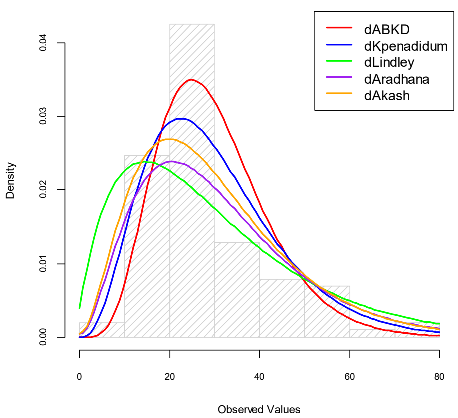
60, 13, 53, 6, 61, 28, 10, 1, 19, 36, 14, 8, 24, 17, 15, 12, 57, 43, 31, 12, 32, 18, 13, 23, 8, 67, 45, 17, 38, 31, 22, 60, 33, 40, 14, 11, 16, 61, 58, 52, 15, 11, 2, 49, 27, 10, 56, 66, 67, 24, 9, 6, 48, 9, 26, 70, 39, 48, 36, 48, 18, 47, 6, 17, 6, 15, 12, 38, 28, 15, 25, 18, 6, 10, 61, 65, 18, 39, 6, 12, 26, 12, 26, 11, 19, 15, 17, 17, 18, 24, 57, 9, 67, 2, 1, 14, 67, 29, 44, 39, 24, 31, 36, 12, 54, 24, 21, 65, 24, 65, 21, 21, 24, 6, 40, 18, 58, 17, 48, 8, 15, 12, 7, 20, 14, 7, 36, 56, 65, 24, 47, 32, 16, 18, 21, 58, 5, 25, 47, 14, 57,2, 46, 8, 26, 16, 54, 11, 24, 14, 25, 62, 39, 12, 12, 66, 78, 77, 46, 39, 12, 6, 45, 38,73, 20, 17, 48, 78, 13, 9, 55, 35, 8, 18, 39, 28, 17, 18, 38, 18, 67, 60, 36, 1, 74, 52, 17, 10, 22, 49, 21, 19, 10, 4, 18, 6, 55, 27, 4, 81, 28, 21, 30, 39, 9, 73, 18, 30

Table 1: The MLEs of the Area-Biased Kpenadidum distribution parameters and -2log L, AIC and BIC, values for the given data sets.

Datasets	Distributions	ML Estimates ($\hat{\theta}$)	S.E.	-2log L	AIC	BIC	AICC
Dataset 1 Age of the patients in years	**Area-Biased Kepenadidum**	**0.1110**	**0.0045**	**851.086**	**853.086**	**855.701**	**853.126**
	Kepenadidum	0.0739	0.0036	878.495	880.495	883.110	880.535
	Lindley	0.0364	0.0025	940.292	942.292	944.907	942.333
	Aradhana	0.5464	0.0031	911.927	913.927	916.542	913.968
	Akash	0.0555	0.0031	901.107	903.107	905.722	903.147
Dataset 2 Tumor size in mm	**Area-Biased Kepenadidum**	**0.1995**	**0.0081**	**800.024**	**802.024**	**804.639**	**802.065**
	Kepenadidum	0.1325	0.0064	838.814	840.814	843.477	840.852
	Lindley	0.0647	0.0045	836.751	838.751	841.366	838.792
	Aradhana	0.0968	0.0055	828.898	830.898	833.513	830.938
	Akash	0.0997	0.0057	810.048	812.048	814.664	812.089
Dataset 3 Recurrence-free Survival time of the patients in months	**Area-Biased Kepenadidum**	**0.2085**	0.0060	1763.308	1765.308	1768.581	1765.329
	Kepenadidum	0.1316	0.0045	1816.011	1818.011	1821.334	1818.03
	Lindley	0.0650	0.0031	1786.673	1788.673	1792.011	1788.693
	Aradhana	0.0968	0.0038	1836.067	1838.067	1841.4	1838.087
	Akash	0.0997	0.0039	1797.418	1799.418	1802.751	1799.438

The best distribution corresponds to lower values of -2log L, AIC, AICC, and BIC. It can be easily seen from the above Table 1 that the ABK distribution provides a better fit as compared to the Kepenadidum, Lindley, and Shanker distributions for all three datasets. And fitting the distribution curves for dataset 2, Figure 5, provides a good illustration of the optimal distribution.

Figure 5: Fitted Distributions to Tumor Size (mm)



11. CONCLUSION

In this study, we explored the area-based continuous distribution and discussed some of the model's structural characteristics. This distribution is a good option for modeling lifespan data, particularly when well-liked substitutes like the Kpenadidum, Lindley, and Shanker distributions can't adequately capture the observed events. The proposed generalized model exhibits greater flexibility and better fitting performance when compared to competing distributions. We are confident that this model, which provides a powerful and flexible tool for statistical analysis in various situations, will find widespread use in many fields that affect people's lives.

REFERENCES

1. Anthony, M., & Elangovan, R. (2020). Length Biased Quasi Aradhana Distribution and its Applications to Survival Data. *Journal of Xidian University, 14*(1001), 893–908.
2. Bashir S, Rasul M. Poisson Area–Biased Lindley Distribution and Its Applications on Biological Data. *BBIJ.* 2016;3(1):1–10.
3. Celik, Nuri. 2022. "Applications in Some Area-Biased Distribution". *Asian Journal of Probability and Statistics* 19 (3):42-50 https://doi.org/10.9734/ajpas/2022/v19i330472.
4. Chouia, S. Zeghdoudi, H. Raman, V. and Beghriche, A. (2021). A new size biased distribution with the application. *Journal of Applied Probability and Statistics*, 16:111-125.
5. Cox D. R. (1969). Some sampling problems in technology, In New Development in Survey Sampling, Johnson, N. L. and Smith, H., Jr . (eds.) New York Wiley-Inter science, 506-527.
6. Elangovan, R., & Anthony, M. (2020). Length-biased OM distribution with properties and Applications to Survival Times. *High Technology Letters, 26*(7), 204–217.
7. Elangovan, R., & Mohanasundari, R. (2019). A New Area-Biased Distribution with Applications in Cancer Data. *Science, Technology and Development. VIII*(XII), 1-14.
8. Fazal A. Area–biased Poisson exponential distribution with applications. *Biom Biostat Int J.* 2018;7(3):256–261. DOI: 10.15406/bbij.2018.07.00216
9. Fisher, R. A., (1934), The effects of methods of ascertainment upon the estimation of frequencies. *Annals of Eugenics*, 6, 13-25.
10. Jaimole VM, Venkatesan D and M Raghavender Sharma (2022). A new area biased two-parameter Pranav distribution and its application to medical sciences. *International Journal of Statistics and Applied Mathematics*; 7(4): 236-246
11. Shakila Bashir and Huma Mahmood (2019). Multivariate area-biased Lindley distribution: properties and applications. *Journal of Reliability and Statistical Studies*; Vol. 12, Issue 1 (2019): 205-223
12. Kilany, N. M. (2016). Weighted Lomax distribution. *Springer Plus*, *5*, 1-18.

13. Nwikpe, B. J., & Cleopas, I. E. (2022). Kpenadidum Distribution: Statistical Properties and Application. *Asian Journal of Pure and Applied Mathematics*, *4*(3), 624–630.
14. Mir KA, Ahmed M. Size–Biased Distribution and there Applications. *Pakistan Journal of Statistics.* 2009;25(3):283–294.
15. Oluwafemi OS, Olalekan DM. Length and area-biased exponentiated Weibull distribution based on forest inventories. Biometrics & Biostatistics International Journal. 2017;6(2). Available from: https://doi.org/10.15406/bbij.2017.06.00163
16. Perveen, Z. Munir, M. Ahmed, Z. and Ahmad, M. (2016). On area-biased weighted Weibull distribution. *Sci. Int. (Lahore)*, 28:3669- 3679.
17. Rajakopalan, V., Ganaie, R.A., & Rather, A.A. (2019). A New Length Biased Distribution with Applications. *Science, Technology and Development. VIII*(VIII), 161–174.
18. Royston P, Altman DG. (2013) External validation of a Cox prognostic model: principles and methods. *BMC Med Res Methodol*. Mar 6; 13:33. doi: 10.1186/1471-2288-13-33. PMID: 23496923; PMCID: PMC3667097.
19. Rashad Mammadov. (2024). Lung Cancer Prediction [Data set]. Kaggle. https://doi.org/10.34740/KAGGLE/DS/5109540
20. 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.
21. 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.
22. Shenbagaraja, R., Aafaq, A. R., & Subramanian, C. (2019). On Some Aspects of Length Biased Technique with Real Life Data, *Science, Technology and Development*, *VIII*(IX), 326–335.
23. Zelen, M. (1974). Problems in cell kinetic and the early detection of disease, in Reliability and Biometry, F. Proschan & R. J. Sering, eds, SIAM, Philadelphia, 701-706

R. Poovitha, Research Scholar,
Department of Statistics,
Annamalai University,
Annamalai Nagar, Tamil Nadu, India-608002.

P. Pandiyan, Professor,
Department of Statistics,
Annamalai University,
Annamalai Nagar, Tamil Nadu, India-608002.

Received