A NEW GENERALIZATION OF AREA-BIASED XGAMMA DISTRIBUTION WITH PROPERTIES AND ITS APPLICATION TO CANCER DATA

Athira D V[1,2], P. Pandiyan[3], and D.S. Shibu[4]

•

Research Scholar, Department of Statistics, Annamalai University
[2]Assistant Professor, Department of Statistics, Christ Nagar College Maranalloor
[3]Professor, Department of Statistics, Annamalai University
[4]Professor , Department of Statistics, University College
[1,2]athiramanuphd@gmail.com, [3]pandianau@gmail.com , [4]statshibu@gmail.com

Abstract

In this paper, we proposed a new model, "Area Biased Xgamma Distribution", a generalization of the Xgamma distribution. The various distributional properties survival function, hazard function, mean residual life ,statistical properties, moments, moment generating function and characteristic function, Bonferroni and Lorenz curve, entropy of the new model were studied. The maximum likelihood estimation method is used to estimate the parameter of the area-biased Xgamma distribution. Finally, using three cancer data sets for the illustration and application of the new model.

Keywords: area biased Xgamma distribution, moments, maximum likelihood estimation, order statistics, renyi entropy, reliability analysis

I. Introduction

Xgamma distribution (XGD) [16] is a newly developed lifetime distribution introduced by Sen. S., Maiti S. S. & Chandra N. in 2016.The Xgamma distribution is a special combination of exponential (θ) and gamma (θ, α) distributions. The exponential and gamma distributions are well known and have some special structural properties [3]. Sen. S. and others in 2016 investigated the weighted Xgamma distribution (WXGD), a relatively novel statistical model for positive-valued data and survival analysis. The length-biased Xgamma distribution (LBXG) was produced as a special case of the density function of the Xgamma distribution [15]. Fisher developed the concept of weighted distribution [9] in 1934 to describe ascertainment biases in the data, and later on, Rao [12] reinterpreted it as a unified theory for problems when the observations fail in non-experimental, non-replicated, and non-random situations. In 2015 Dey. S. Ali, S., & Park, C. are introduced Weighted exponential distribution properties and different methods of estimation [5]. The weighted distribution concept defined as $f_w(x) = \frac{w(x)f(x)}{E(w(x))}$,where w(x) is the non-negative weight function and f(x) is probability function. We have many choices of weight functions [4]. If the weight function w(x)=x is a length-biased distribution, then w(x)=x^2 is an area-biased distribution. In [9], Length and area biased exponentiated Weibull distribution. In [17], established on a length and area-biased

Maxwell distributions. In [11], a proposal was made on area-biased weighted Weibull distribution. Recently [1], discussed the characterization and estimation of area-biased quasi-Akash distribution. In [13], proposed probability and survival analysis of cancer patients

In this paper, we introduced a new generalization of area-biased Xgamma distribution, which is the extension of the Xgamma distribution. The article is divided into several sections and is structured as follows: The alternative form of the Area-biased Xgamma distribution is introduced in Section 2. The survival properties are studied in Section 3. The statistical properties such as moments, moment generating-functions, and characteristic functions are described in Section 4. In Section 5, order statistics of the distribution are derived. The likelihood ratio test is discussed in Section 6. Bonferroni and Lorenz curves and entropy are discussed in Sections 7 and 8, respectively. Methods of estimating parameters are discussed in section 9. In section 10, three actual cancer data sets are analyzed to show the application of the weighted area-biased Xgamma distribution; finally, Section 11 concludes.

II. Area Biased Xgamma distribution

Let X be a random variable follow the Xgamma distribution with scale parameter θ; then the probability density function (pdf) and cumulative distribution function (cdf) of the Xgamma distribution, respectively,

$$f(x) = \frac{\theta^2}{(1+\theta)}\left(1+\frac{\theta}{2}x^2\right)e^{-\theta x} \qquad x>0, \theta>0 \tag{1}$$

And

$$F(X) = 1-\frac{1+\theta+\theta x+\frac{\theta^2 x^2}{2}}{(1+\theta)}\, e^{-\theta x} \qquad x>0, \theta>0 \tag{2}$$

The probability density function of the weighted random variable X_w is given by

$$f_w(x) = \frac{w(x)f(x)}{E(w(x))} \qquad x>0$$

We have considered the weight function $w(x)=x^2$; then the probability function of area-biased distribution is given

$$f_\alpha(x) = \frac{x^2 f(x)}{E(x^2)} \tag{3}$$

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

$$E(x^2) = \frac{2(\theta+6)}{\theta^2(1+\theta)} \tag{4}$$

Substituting (1) and (4) in (3) we get the probability density function of area-biased Xgamma distribution

Definition: A continuous random variable X is said to follow the area-biased Xgamma (ABXG) distribution with parameter θ if its probability density function(pdf)

$$f_a(x) = \frac{x^2\theta^4}{2(\theta+6)}\left(1+\frac{\theta}{2}\,x^2\right)e^{-\theta x} \qquad x>0\,, \theta>0 \tag{5}$$

It is denoted by X~ABXGD(θ)
The cumulative density function (cdf) X is given by

$$F_a(x) = \frac{2\theta\gamma(3,\theta x) + \gamma(5,\theta x)}{4(\theta+6)} \quad ; x > 0, \theta > 0 \tag{6}$$

Where, $\gamma(3,\theta x)$ and $\gamma(5,\theta x)$ are lower incomplete gamma functions.
The plots of the corresponding pdf and cdf of the area-biased Xgamma distribution are shown in **Figure 1** and **Figure 2**. It is possible to see from **Figure 1** that the area-biased Xgamma distribution is a positively skewed distribution.

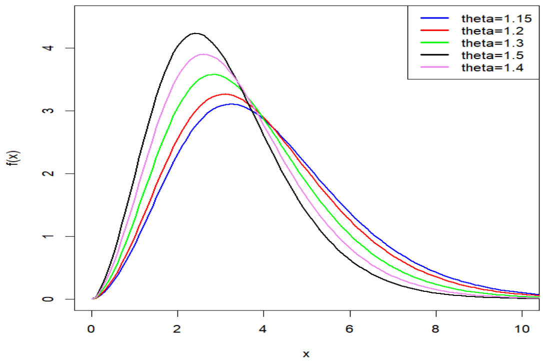

Figure 1: *probability density function of Area-biased Xgamma distribution*

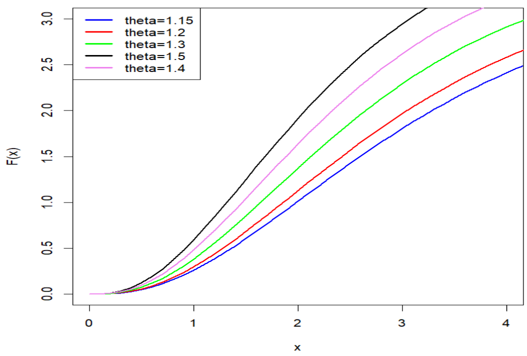

Figure 2: *cumulative distribution function area-biased Xgamma distribution*

III. Reliability analysis

In this section, we will cover the survival function, hazard function, reverse hazard function, cumulative hazard function, odds rate, Mills ratio, and, mean residual life for the area-biased Xgamma distribution**.**

I. Survival Function

Let X be a random variable with pdf f(x) and cdf F(x). The general formula for the survival function is $S(x) = 1 - F_a(x)$. Then the survival function, or the reliability function of the area-biased Xgamma distribution, is given by

$$S(x) = 1 - \frac{2\theta\gamma(3,\theta x) + \gamma(5,\theta x)}{4(\theta+6)} \quad x > 0 \tag{7}$$

II. Hazard function

The corresponding hazard function / failure rate of the area-biased Xgamma distribution is provide as

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

$$h(x) = \frac{\frac{x^2\theta^4}{2(\theta+6)}\left(1+\frac{\theta}{2}x^2\right)e^{-\theta x}}{1-\frac{2\theta\gamma(3,\theta x) + \gamma(5,\theta x)}{4(\theta+6)}}$$

after simplification, we get

$$h(x) = \frac{2x^2\theta^4\left(1+\frac{\theta}{2}x^2\right)e^{-\theta x}}{4(\theta+6) - 2\theta\,\gamma(3,\theta x) - \gamma(5,\theta x\,)} \tag{8}$$

III. Reverse Hazard function

Reverse hazard function of area-biased Xgamma distribution is indicated as

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

$$h_r(x) = \frac{2x^2\theta^4\left(1+\frac{\theta}{2}\,x^2\right)e^{-\theta x}}{2\theta\gamma(3,\theta x) + \gamma(5,\theta x)} \tag{9}$$

IV. Odds Rate Function

Odds Rate function of Area-biased Xgamma distribution is given by

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

$$O(x) = \frac{2\theta\gamma(3,\theta x) + \gamma(5,\theta x)}{4(\theta+6) - \theta\gamma(3,\theta x) - \gamma(5,\theta x)} \tag{10}$$

V. Cumulative Hazard Function

Cumulative hazard function of Area-biased distribution is expressed as

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

$$H(x) = -\ln\left(1 - \frac{2\theta\gamma(3,\theta x) + \gamma(5,\theta x)}{4(\theta+6)}\right)$$

VI. Mills Ratio

The mills ratio of Area-biased area distribution is provided by

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

$$\text{Mills ratio} = \frac{2\theta\gamma(3,\theta x) + \gamma(5,\theta x)}{2x^2\theta^4\left(1+\frac{\theta}{2}\,x^2\right)e^{-\theta x}} \tag{11}$$

VII. Mean residual life

The mean residual life (MRL) at time t is the expected remaining life given that the random variable x survived up to time t.

$$MRL(x) = \frac{1}{S(x)} \int_x^\infty t\, f_a(t)\, dt - x$$

after simplification, we get

$$MLR(x) = \frac{2\theta\Gamma(4,\theta t) + \Gamma(6,\theta t)}{S(x)\, 4\theta\, (\theta + 6)} \tag{12}$$

Where S(x) is the survival function (7).

Figures 3 and 4 are hazard function plot and survival function plot of the area-biased Xgamma distribution respectively.

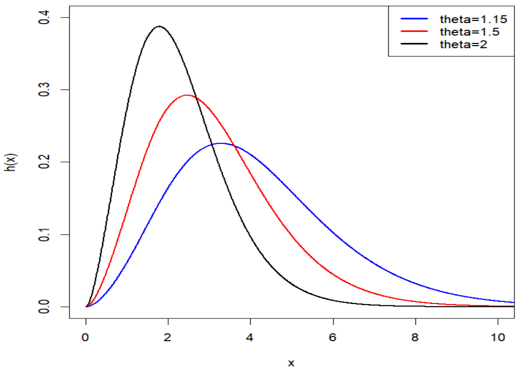

Figure 3: *Hazard function of area-biased Xgamma distribution*

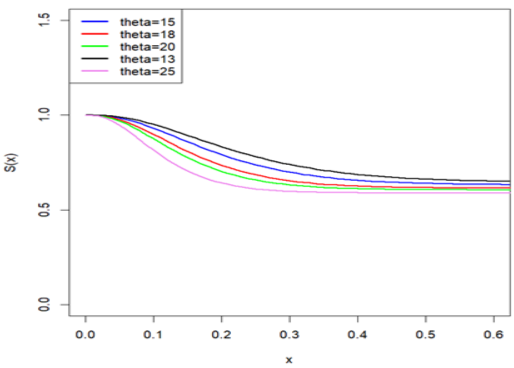

Figure 4: *survival function area-biased Xgamma distribution*

IV. Statistical Properties

In this section, discuss the different statistical properties of the area-biased Xgamma distribution that are moments, moment-generating functions, and characteristic functions.

I. Moments

Let the random variable X follow the area-biased Xgamma distribution with parameters θ; then the r^{th} order moment $E(X^r)$ of X about the origin can be obtained as

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

$$E(X^r) = \int_0^\infty x^r \frac{x^2\theta^4}{2(\theta+6)} \left(1 + \frac{\theta}{2} x^2\right) e^{-\theta x}\, dx$$

$$E(X^r) = \frac{\theta^4}{2(\theta+6)} \int_0^\infty x^2\, x^r \left(1 + \frac{\theta}{2} x^2\right) e^{-\theta x}\, dx \tag{13}$$

After simplification (13) get

$$E(X^r) = \mu_r{}' = \frac{2\theta\, (r+2)! + (r+4)!}{4\, (\theta+6)\, \theta^r} \tag{14}$$

Putting r=1,2,3 and 4 in equation (14), obtain the first four moments of the area-biased Xgamma distribution.

$$E(X^1) = \mu_1' = \frac{3\,\theta + 30}{\theta\,(\theta + 6)} \tag{15}$$

$$E(X^2) = \mu_2' = \frac{12\,\theta + 180}{\theta^2(\theta + 6)} \tag{16}$$

$$E(X^3) = \mu_3' = \frac{60\,\theta + 1260}{\theta^3(\theta + 6)} \tag{17}$$

$$E(X^4) = \mu_4' = \frac{360\theta + 10080}{\theta^4(\theta + 6)} \tag{18}$$

Property 1: Mean of the area-biased Xgamma distribution

$$mean = \frac{3\,\theta + 30}{\theta\,(\theta + 6)}$$

Property 2: Variance of the area-biased Xgamma distribution is given by

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

$$\text{variance} = \frac{3\theta^2 + 72\theta + 180}{\theta^2(\theta + 6)^2}$$

Property 3: Standard Deviation

$$\sigma = \frac{\sqrt{3\theta^2 + 72\theta + 180}}{\theta\,(\theta + 6)}$$

Property 4: Coefficient Of Variation

$$C.V = \left(\frac{\sigma}{\mu}\right)100$$

$$C.V = \frac{\sqrt{3\theta^2 + 72\theta + 180}}{3\theta + 30}\;100$$

II. Moment Generating Function and Characteristic Function

In this section, we discuss the moment-generating function and characteristic function of area-biased Xgamma distribution.

Theorem 1:
Let X be an area-biased Xgamma random variable with parameters θ; then the MGF of X is

$$M_X(t) = \frac{\theta^4(\theta - t)^2 - 6\theta^5}{(\theta + 6)(\theta - t)^5}$$

Proof: Let X~ABXG(θ),then the moment-generating function of X is obtained by

$$M_X(t) = E(e^{tx}) = \int_0^\infty e^{tx} f_a(x)dx \tag{19}$$

$$M_X(t) = E(e^{tx}) = \int_0^\infty e^{tx} \frac{x^2\theta^4}{2(\theta+6)} \left(1 + \frac{\theta}{2} x^2\right) e^{-\theta x} \, dx$$

$$M_X(t) = \frac{\theta^4}{2(\theta+6)} \int_0^\infty e^{-x(\theta-t)} x^2 \left(1 + \frac{\theta}{2} x^2\right) dx$$

After simplification we get the moment-generating function

$$M_X(t) = \frac{\theta^4(\theta-t)^2 - 6\theta^5}{(\theta+6)(\theta-t)^5} \tag{20}$$

Similarly, the characteristic function area-biased Xgamma distribution can be obtained as

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

$$\phi_X(t) \quad = \frac{\theta^4(\theta-it)^2 - 6\theta^5}{(\theta+6)(\theta-it)^5} \tag{21}$$

V. Order statistics

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

$$f_{X(r)}(x) = \frac{n!}{(r-1)!\,(n-r)!} f_X(x)[F_X(x)]^{r-1}[1-F_X(x)]^{n-r} \tag{22}$$

Using equations (5) and (6) in equation (22), we get the probability density function of r^{th} order statistics of the area-biased Xgamma distribution.

$$f_{x(r)}(x) = \frac{n!}{(r-1)!\,(n-r)!} \left(\frac{x^2\theta^4}{2(\theta+6)} \left(1 + \frac{\theta}{2} x^2\right) e^{-\theta x}\right) \left(\frac{2\theta\gamma(3,\theta x) + \gamma(5,\theta x)}{4(\theta+6)}\right)^{r-1} \left(1 - \frac{2\theta\gamma(3,\theta x) + \gamma(5,\theta x)}{4(\theta+6)}\right)^{n-r} \tag{23}$$

Put r=1 in equation (23), we get first order statistics X $_{(1)}$=mini (x_1, x_2,…… x_n) and m=n, we get nth order statistics X$_{(n)}$=max (x_1, x_2,…… x_n) respectively.

$$f_{X(1)}(x) = \frac{nx^2\theta^4}{2(\theta+6)} \left(1 + \frac{\theta}{2} x^2\right) e^{-\theta x} \left(1 - \frac{2\theta\gamma(3,\theta x) + \gamma(5,\theta x)}{4(\theta+6)}\right)^{n-1}$$

$$f_{X(n)}(x) = \frac{nx^2\theta^4}{2(\theta+6)} \left(1 + \frac{\theta}{2} x^2\right) e^{-\theta x} \left(\frac{2\theta\gamma(3,\theta x) + \gamma(5,\theta x)}{4(\theta+6)}\right)^{n-1}$$

VI. Likelihood Ratio Test

In this section, derive the likelihood ratio test from the area-biased Xgamma distribution. The likelihood-ratio test is a hypothesis test that compares two competing statistical models' goodness of fit. Let $X_1, X_2, \ldots, X_n$ be a random sample from the area biased samade distribution. To test the hypothesis

$$H_0 : f(x) = f(x{:}\theta) \qquad \text{against} \qquad H_1 : f(x) = f_a(x{:}\theta)$$

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

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

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

We should reject the null hypothesis, if

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

Or equivalently, reject the null hypothesis

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

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

For a 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 the 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, we derived the Bonferroni and Lorenz curves from the area-biased Xgamma distribution. The Bonferroni and Lorenz curves are powerful tools in the analysis of distributions and have applications not only in economics but also in many fields, such as reliability analysis, medicine, insurance, and income. The Bonferroni 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

$${\mu_1}' = E(X) = \frac{3\,\theta + 30}{\theta\,(\theta+6)}$$

and $q = F^{-1}(p)$

$$B(p) = \frac{\theta\,(\theta+6)}{p(3\,\theta+30)}\int_0^q x\,\frac{x^2\theta^4}{2(\theta+6)}\left(1+\frac{\theta}{2}\,x^2\right)e^{-\theta x}\,dx$$

$$B(p) = \frac{\theta^5\,(\theta+6)}{p(3\theta+30)2(\theta+6)}\int_0^q x^2\left(1+\frac{\theta}{2}\,x^2\right)e^{-\theta x}\,dx$$

After simplification, we get

$$B(p) = \frac{(\theta+6)[\theta^2\gamma(3,\theta q)+\frac{\theta}{2}\gamma(5,\theta q)}{2p(3\theta+30)(\theta+6)} \tag{24}$$

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

$$L(p) = \frac{(\theta+6)[\theta^2\gamma(3,\theta q)+\frac{\theta}{2}\gamma(5,\theta q)}{2(3\theta+30)(\theta+6)} \tag{25}$$

Figure 5 show the Lorence curve of area-biased Xgamma distribution

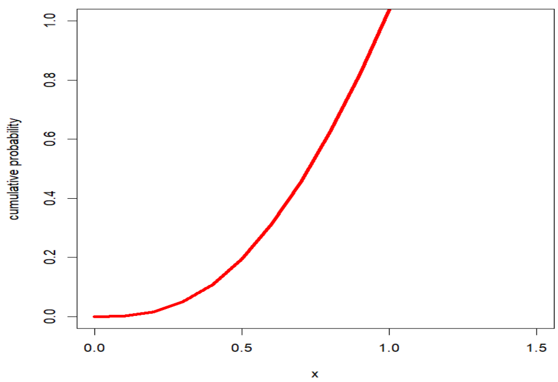

Figure 5: *Lorence curve of Area biased Xgamma distribution*

VIII. Entropies

In this section, we derived the entropies of the area-biased Xgamma distribution. The idea of entropy is essential in many domains, including probability, statistics, physics, communication theory, and economics. Entropies measure a systems diversity, uncertainty, or randomness. Entropy of a random variable X is a measure of variation in uncertainty.

I. Shannon Entropy

Shannon entropy of the area-biased Xgamma distribution is defined as

$$S_\tau = -\int_0^\infty f_a(x)\ln(f_a(x))dx \qquad \tau>0,\tau\neq 1$$

$$S_\tau = -\int_0^\infty \frac{x^2\theta^4}{2(\theta+6)}\left(1+\frac{\theta}{2}\,x^2\right)e^{-\theta x}\ln\left(\frac{x^2\theta^4}{2(\theta+6)}\left(1+\frac{\theta}{2}\,x^2\right)e^{-\theta x}\right)dx$$

II. Renyi Entropy

The Renyi entropy is an essential diversity in ecology and statistics. It is also useful in quantum information, where it can be used to quantify entanglement. The Renyi entropy of the area-biased

Xgamma distribution is given by

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

Where $\gamma > 0$ *and* $\gamma \neq 1$

$$R(\gamma) = \frac{1}{1-\gamma} \ln \int_0^\infty \left(\frac{x^2\theta^4}{2(\theta+6)} \left(1 + \frac{\theta}{2}\, x^2\right) e^{-\theta x} \right)^\gamma dx$$

$$R(\gamma) = \frac{1}{1-\gamma} \ln \left(\frac{\theta^4}{2(\theta+6)} \right)^\gamma \int_0^\infty \left(x^2 \left(1 + \frac{\theta}{2}\, x^2\right) e^{-\theta x} \right)^\gamma dx$$

IX. Estimations of Parameter

The parameter estimation of the area-biased Xgamma distribution is obtained in this section. There are two estimation methods used in this section to estimate unknown parameters: the method of moment and maximum likelihood estimation.

I. Method of moment

Let $X_1, X_2 \ldots\ldots X_n$ be a random sample taken from area-biased Xgamma distribution with parameter θ. By the method of moment

$$\mu_r{}' = m_r{}'$$

Hence, we get

$$\mu_1{}' = m_1{}'$$

$$\frac{3\,\theta + 30}{\theta\,(\theta+6)} = \frac{1}{n} \sum_{i=1}^{n} x_i$$

After simplification ,we get

$$\overline{x}\,\theta^2 + \theta(6\overline{x} - 3) - 30 = 0 \tag{26}$$

Since the equation (26) is complicated, we can use the gmm package [10] from the R statistical software [14] to solve the equation (26).

II. Maximum likelihood estimation

Let $x_1, x_2 \ldots\ldots x_n$ be the random sample of size n taken from area-biased Xgamma distribution with parameter θ; then the likelihood function is

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

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

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

The log-likelihood function is given by

$$\ln L = n \ln \theta^4 - n\, ln2 - ln\,(6+\theta) + \sum_{i=1}^{n} \ln x_i^2 + \sum_{i=1}^{n} \ln(1+\frac{\theta}{2}x_i^2) - \theta \sum_{i=1}^{n} x_i \tag{27}$$

The normal equation is

$$\frac{\partial \ln L}{\partial \theta} = \frac{4n}{\theta} - \frac{n}{\theta+6} + \sum_{i=1}^{n} \frac{x_i^2}{2(1+\frac{\theta}{2}\,x_i^2)} - \sum_{i=1}^{n} x_i = 0 \tag{28}$$

The numerical method is used to solve the equations (28). We used the optimx package [10] from R statistical software [14] to obtain the maximum likelihood estimator of the area-biased Xgamma distribution.

X. Applications

In this section, we have considered three real data sets for the purpose of showing that the distribution of area-biased Xgamma distribution shows a better fit over Xgamma distribution and area-biased Aradhana distribution.

The Akaike information criterion (AIC), Bayesian information criterion (BIC), Akaike information criterion corrected (AICC), Hannan-Quinn information criterion (HQIC), and -2logl are used for model selection. It can be evaluated by using the formula as follows:

$AIC = 2k - 2logL$ $\qquad$ $AICC = AIC + \frac{2k(k+1)}{(n-k-1)}$ $\qquad$ $BIC = klog(n) - 2logL$

$HQIC = 2klog(\log(n)) - 2logL$

Where n is the sample size, k is the number of parameters, and -2logL is the maximal value of the log likelihood function. After the calculation of AIC, AICC, HQIC, and -2logL, the model with the minimum value is chosen as the best model to fit the data.

Data set 1: Breast cancer

The data set is the age of 155 patients with breast cancer from June to November 2014, (Breast Tumours Early Detection Unit, Benha Hospital University, Egypt).[8] Data are as follows

46, 32, 50, 46, 44, 42, 69, 31, 25, 29, 40, 42, 24, 17, 35, 48, 49, 50, 60, 26, 36, 56, 65, 48, 66, 44, 45, 30, 28, 40, 40, 50, 41, 39, 36, 63, 40, 42, 45, 31, 48, 36, 18, 24, 35, 30, 40, 48, 50, 60, 52, 47, 50, 49, 38, 30, 52, 52, 12, 48, 50, 45, 50, 50, 50, 53, 55, 38, 40, 42,42, 32, 40, 50, 58, 48, 32, 45, 42, 36, 30, 28, 38, 54, 90, 80, 60, 45, 40, 50,50, 40, 50, 50, 50, 60, 39, 34, 28, 18, 60, 50, 20, 40, 50, 38, 38, 42, 50, 40, 36, 38, 38, 50, 50, 31, 59, 40, 42, 38, 40, 38, 50, 50, 50, 40, 65, 38, 40, 38, 58, 35, 60, 90, 48, 58, 45, 35, 38, 32, 35, 38, 34, 43, 40, 35, 54, 60, 33, 35, 36, 43, 40, 45, 56

Table 1: *The summary of Brest cancer data set*

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
12.00	36.00	42.00	43.65	50.00	90.00

Data set 2: **blood cancer (leukemia)**

The following real lifetime (in years) data set consists of 40 patients suffering from blood cancer (leukemia) reported from one of the ministries of health hospitals in Saudi Arabia (see Abouammah et al. [2]). Data Set 2 is given by:

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, 3.858, 2.805, 2.91, 2.912, 3.192, 3.263, 3.348, 3.348, 3.427, 3.499, 3.534, 3.767, 3.751, 3.986, 4.049, 4.244, 4.323, 4.381, 4.392, 4.397, 4.753, 4.929, 4.973, 5.074, 5.381

Table 2: *The summary of blood cancer data set*

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.315	2.187	3.348	3.102	4.146	5.381

Data set 3: Head and Neck cancer data

The following real life data set of survival time of 44 patients diagnosed by head and Neck cancer disease are available in [6]. Data set 2 given by:

12.20, 23.56, 23.74, 25.87, 31.98, 37, 41.35, 47.38, 55.46, 58.36, 63.47, 68.46, 78.26, 74.47, 81.43, 84, 92, 94, 110, 112, 119, 127, 130, 133, 140, 146, 155, 159, 173, 179, 194, 195, 209, 249, 281, 319, 339, 432, 469, 519, 633, 725, 817, 1776

Table 3: *The summary of Head and Neck cancer data set*

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
12.20	67.21	128.50	223.48	219.00	1776.00

Table 4: *MLEs AIC, BIC, AICC, HQIC and -2log L of the fitted distribution for the given data set 1*

Distribution	MLE	-2logL	AIC	BIC	AICC	HQIC
Area biased Xgamma	$\hat{\theta}$=1.4815838 (0.1101315)	139.7668	141.7668	143.4303	141.8749	142.3637
Xgamma	$\hat{\theta}$=0.73392828 (0.07638865)	155.4983	157.4983	159.1619	157.6064	158.0952
Area biased Aradhana	$\hat{\theta}$=1.4311764 (0.1038796)	141.485	143.485	145.1485	143.5931	144.0819

Table 5: *MLEs AIC, BIC, AICC, HQIC and -2log L of the fitted distribution for the given data set2*

Distribution	MLE	-2logL	AIC	BIC	AICC	HQIC
Area biased Xgamma	$\hat{\theta}$=0.100000 (0.006751215)	1554.055	1556.055	1557.84	1556.151	1556.717
Xgamma	$\hat{\theta}$=0.100000 (0.008736029)	1790.128	1792.128	1793.128	1792.223	1792.79
Area biased Aradhana	$\hat{\theta}$=0.100000 (0.006741885)	1555.373	1557.373	1559.157	1557.468	1558.035

Table 6: *MLEs AIC, BIC, AICC, HQIC and -2log L of the fitted distribution for the given data set*3

Distribution	MLE	-2logL	AIC	BIC	AICC	HQIC
Area biased Xgamma	$\hat{\theta}$=0.113843128 (0.004098989)	1262.845	1264.845	1267.888	1264.871	1266.081
Xgamma	$\hat{\theta}$=0.100000 (0.004664461)	1418.24	1420.24	1423.284	1420.266	1421.476
Area biased Aradhana	$\hat{\theta}$=0.113273264 (0.004069012)	1262.981	1264.981	1268.024	1265.007	1266.217

From tables 4, 5, and 6 it can be clearly observed and seen from the results that the area-biased Xgamma distribution has the lesser AIC, BIC, AICC, HQIC, -2logL, and values as compared to the Xgamma and area-biased Aradhana, which indicates that the area-biased Xgamma distribution better fits than the Xgamma and area-biased Aradhana distributions. Hence, it can be concluded that the area-biased Xgamma distribution leads to a better fit over the other distributions.

XI. Conclusion

Area-Biased Xgamma distribution is a generalization of Xgamma distribution, was derived, and studied the different properties.Which is the lifetime model for a real-life data set. In the section 3, 4, 5, & 6 discussing the statistical property area-biased Xgamma distribution. The effectiveness and applicability of the suggested model is demonstrated by an examination of three real cancer data sets and compared with the Xgamma distribution and area-biased Aradhana distribution. The result indicates that the area-biased Xgamma distribution is more flexible and practical than the Xgamma distribution and area-biased Aradhana distribution.

Reference

[1] Ade, R. B., Teltumbade, D. P., & Tasare, P. W. (2021). Charecterizaton And Estimation of Area Biased Quasi Akash Distribution. *International Journal of Mathematics Trends and Technology-IJMTT, 67*

[2]Atikankul, Y., Thongteeraparp, A., Bodhisuwan, W., & Volodin, A. (2020). The length-biased weighted Lindley distribution with applications. Lobachevskii Journal of Mathematics, 41, 308-319

[3] Bakoban, R. A. (2010). A study on mixture of exponential and exponentiated gamma distributions. Advances and Applications in Statistical Sciences, 2(1), 101-127.

[4] Bartoszewicz, J. (2009). On a representation of weighted distributions. Statistics & probability letters, 79(15), 1690-1694

[5] Dey, S., Ali, S., & Park, C. (2015). Weighted exponential distribution: properties and different methods of estimation. Journal of Statistical Computation and Simulation, 85(18), 3641-3661.

[6] Efron, B. (1988). Logistic regression, survival analysis, and the Kaplan-Meier curve. Journal of the American statistical Association, 83(402), 414-425

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

[8] Hassan, A. S., Elbatal, I., & Hemeda, S. E. (2016). Weibull quasi-Lindley distribution and its statistical properties with applications to lifetime data. International Journal of Applied Mathematics and Statistics, 55(3), 63-80..

[9] 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

[10] P. Chauss e, Computing generalized method of moments and generalized empirical likelihood with R, J. Stat. Software 34 (11), 1 35 (2010).

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

[12] 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.

[13] Ray, M., & Shanker, R. (2024). A PROBABILITY MODEL FOR SURVIVAL ANALYSIS OF CANCER PATIENTS. Reliability: Theory & Applications, 19(3 (79)), 78-94.

[14]RCoreTeam,R:ALanguageandEnvironmentforStatisticalComputing,RFoundationforStatistical Computing (Vienna, Austria, 2018). https://www.R-project.org/.Accessed 2019

[15] Sen, S., Chandra, N., & Maiti, S. S. (2017). The weighted xgamma distribution: properties and application.

[16] Sen, S., Maiti, S. S., & Chandra, N. (2016). The xgamma distribution: statistical properties and application. Journal of Modern Applied Statistical Methods, 15(1), 38.

[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