Journal of Biostatistics and Epidemiology

J Biostat Epidemiol. 2023;9(4): 451-460

Original Article

A New Lifetime Distribution and its Application to Cancer Data

Samuel Adewale Aderoju*, Nihimat Iyebuhola Aleshinloye, Bako Lukmon Taiwo, Bello Ishola Sanni

Department of Mathematics and Statistics, Kwara State University, Malete, Ilorin, Nigeria.

ARTICLE INFO

Received 19.04.2023
Revised 13.05.2023
Accepted 15.06.2023
Published 15.12.2023

Key words:
Weibull-generalized gamma distribution;
Hazard rate;
Maximum likelihood estimation;
Lifetime;
Cancer data.

ABSTRACT

Introduction: Recently, researchers have introduced new generated families of univariate lifetime distributions. These new generators are obtained by adding one or more extra shape parameters to the underlying distribution or compounding two distributions to get more flexibility in fitting data in different areas such as medical sciences, environmental sciences, and engineering. The addition of parameter(s) has been proven useful in exploring tail properties and for improving the goodness-of-fit of the family of the proposed distributions.
Methods: A new Three-Parameter Weibull-Generalized Gamma distribution which provides more flexibility in modeling lifetime data is developed using a two-component mixture of Weibull distribution (with parameters θ and λ) and Generalised Gamma distribution (with parameters α=4,θ and λ). Some of its mathematical properties such as the density function, cumulative distribution function, survival function, hazard rate function, moment generating function, Renyi entropy and order statistics are obtained. The maximum likelihood estimation method was used in estimating the parameters of the proposed distribution and a simulation study is performed to examine the performance of the maximum likelihood estimators of the parameters.
Results: Real life applications of the proposed distribution to two cancer datasets are presented and its fit was compared with the fit attained by some existing lifetime distributions to show how the Three-Parameter Weibull-Generalized Gamma distribution works in practice.
Conclusion: The results suggest that the proposed model performed better than its competitors and it's a useful alternative to the existing models.

Introduction

Many lifetime distributions have been developed by researchers for modeling lifetime data in the form of parametric methodology. The most popular models are Gamma, generalized gamma, Weibull, exponential and Lindley distributions. For instance, Weibull distribution[1] has attracted the attention of researchers in various fields due to its characteristics in modelling survival data. However, this distribution does not provide

*.Corresponding Author: samueladeroju77@gmail.com

a good fit to data sets with bathtub shaped or upside-down bathtub shaped (unimodal) failure rates, often encountered in reliability, engineering and biological studies.[2] In the other hand, Stacy introduced the generalized gamma distribution, which has the exponential, Weibull, gamma, and Rayleigh distributions as its special cases.[3] This distribution is known for its flexibility in modelling skewed data. The inferential procedures for the generalized gamma distribution are difficult.[2] Several researchers have studied the characteristics, properties, and properties of the distribution.[2,4] Several distributions have been proposed in the literature to extend the Weibull, generalized gamma and Lindley distributions.[5-8]

Meshkat et al.[9] introduced a new distribution named generalized gamma-Weibull (GGW) distribution. They provided mathematical properties of moments and estimate the model's parameters through the maximum likelihood estimation method. The goodness-of-fit of the GGW distribution was demonstrated through an application of a real data set. Hamed in[2] introduced another Weibull-Generalized Gamma (W-GG) distribution, which is obtained by compounding Weibull and generalized gamma distributions. The model contains some existing lifetime distributions as special cases such as the Weibull, Lindley, and exponential distributions, among others. Some statistical properties of the W-GG distribution were discussed and the maximum likelihood estimation method to obtain its parameters. The results showed that the model is capable of modelling various shapes of ageing and failure criteria.

Following a similar procedure and informed by the strengths and limitations of the Weibull and generalized gamma distributions, we introduced a new mixed distribution from the Mixture Weibull and Generalized Gamma distribution. The new Three-Parameter Weibull-Generalized Gamma Distribution (TWGG) distribution was developed using a two-component mixture of Weibull distribution (with parameters θ and λ) and Generalised Gamma distribution (with parameters $\alpha=4,\theta$ and λ). Hoping that the proposed distribution will be more flexible and serve as an alternative to other mixed models available in the literature for modelling real lifetime data in many areas. The study derived some mathematical and statistical properties of the new distribution. Finally, the distribution was evaluated by fitting real-life univariate cancer patient datasets.

Methods

This study employed the use of mixing proportions to obtain the new distribution. The general form of the probability density function (pdf) of a k-component additive mixture distribution for a random variable x is defined as

$$f\left(x,\theta\right)=\sum_{j=1}^{k} f_j\left(x,\theta_j\right)\omega_{j}, \qquad (1)$$

where θ_j is the vector of parameters for the mixture models, ω_j is the mixture proportion and

$$\sum_{j=1}^{k}\omega_j = 1.$$

Theorem 1: Suppose a random variable X follows a Three-Parameter Weibull-Generalized Gamma distribution, that is,

$X \mid \alpha,\theta,\lambda \sim \text{TWGG}\left(\alpha,\theta,\lambda\right)$, then the pdf of random variable X is given as:

$$f_1(x|\alpha,\theta,\lambda)=\frac{\lambda\theta^4}{\alpha\theta^3+6}\left(\alpha+x^{3\lambda}\right)x^{\lambda-1}e^{-\theta x^{\lambda}},\ for\ x>0;\alpha,\theta,\lambda>0$$

Proof 1: Recall that the pdf of a k-component additive mixture distribution as given in (1). The pdf in (2) is a two-component mixture of Weibull (θ,λ) and Generalised Gamma (4,θ,λ) with the mixture proportion, $p=\frac{\alpha\theta^3}{\alpha\theta^3+6}$.

Therefore, the pdf of TWGG distribution can be derived as follows:

$$f_1(x|\theta,\alpha,\lambda)=pg_1(\theta,\lambda)+(1-p)g_2(4,\theta,\lambda),$$

where

$$g_1(\theta,\lambda)=\lambda\theta x^{\lambda-1}e^{-\theta x^{\lambda}}\ and$$

$$g_2(4,\theta,\lambda)=\frac{\lambda\theta^4x^{3\lambda}x^{\lambda-1}e^{-\theta x^{\lambda}}}{6}$$

$$\therefore f_1(x|\theta,\alpha,\lambda)=\frac{\alpha\theta^3}{\alpha\theta^3+6}\left(\lambda\theta x^{\lambda-1}e^{-\theta x^{\lambda}}\right)+\frac{6}{\alpha\theta^3+6}\left(\frac{\lambda\theta^4x^{3\lambda}x^{\lambda-1}e^{-\theta x^{\lambda}}}{6}\right)$$

$$=\left(\frac{6\alpha\theta^4\lambda+6\theta^4\lambda x^{3\lambda}}{6\left(\alpha\theta^3+6\right)}\right)x^{\lambda-1}e^{-\theta x^{\lambda}}$$

$$=\left(\frac{\alpha\theta^4\lambda+\theta^4\lambda x^{3\lambda}}{\alpha\theta^3+6}\right)x^{\lambda-1}e^{-\theta x^{\lambda}}$$

Therefore, the probability density function of the TWGG distribution is

$$f_1(x|\theta,\alpha,\lambda)=\frac{\theta^4\lambda}{\alpha\theta^3+6}\left(\alpha+x^{3\lambda}\right)x^{\lambda-1}e^{-\theta x^{\lambda}}$$

Theorem 2: The function in (2) is non-negative, continuous and has unit integral on its domain: $\int_0^\infty f_1(x)\,dx=1$

Proof 2: The pdf of TWGG distribution is a proper pdf since it satisfies the condition that:

$$\int_0^\infty f_1(x|\alpha,\theta,\lambda)\,dx=\frac{\lambda\theta^4}{\alpha\theta^3+6}\left[\int_0^\infty\left(\alpha+x^{3\lambda}\right)x^{\lambda-1}e^{-\theta x^{\lambda}}\right]dx$$

$$=\frac{\lambda\theta^4}{\alpha\theta^3+6}\left(\frac{\alpha\theta^3+6}{\lambda\theta^4}\right)=1$$

Note that

when λ=1,θ>0 and α>0 then the TWGG distribution reduces to the Two-Parameter Rama distribution proposed by Umeh et al.[10] Similarly,

when λ=1,θ>0 and α=1 then the TWGG distribution reduces to the One-Parameter Rama distribution proposed by Shanker et al.[11]

The corresponding cumulative density function (cdf) of the TWGG distribution was derived as follows:

$$F_1(x|\alpha,\theta,\lambda)=\int_{t=0}^{x}f_1(t|\alpha,\theta,\lambda)\,\mathrm{dt}$$

$$=\int_{t=0}^{x}\frac{\theta^4\lambda}{\alpha\theta^3+6}\left(\alpha+t^{3\lambda}\right)t^{\lambda-1}e^{-\theta t^{\lambda}}\,dt$$

$$=\frac{\theta^4\lambda}{\alpha\theta^3+6}\int_{t=0}^{x}\left(\alpha+t^{3\lambda}\right)t^{\lambda-1}e^{-\theta t^{\lambda}}\,dt$$

Using Mathematica software to solve the integral part, we have

$$=\left(\frac{\cancel{\theta^4\lambda}}{\alpha\theta^3+6}\right)\left[\frac{6+\alpha\theta^3-e^{-x^{\lambda}\theta}\left(6+6x^{\lambda}\theta+3x^{2\lambda}\theta^2+x^{3\lambda}\theta^3+\alpha\theta^3\right)}{\cancel{\theta^4\lambda}}\right]$$

$$=1-\left(\frac{\left(6+6x^{\lambda}\theta+3x^{2\lambda}\theta^2+x^{3\lambda}\theta^3+\alpha\theta^3\right)e^{-\theta x^{\lambda}}}{\alpha\theta^3+6}\right)$$

$$=1-\left(1+\frac{\left(6\theta x^{\lambda}+3\theta^2x^{2\lambda}+\theta^3x^{3\lambda}\right)}{\alpha\theta^3+6}\right)e^{-\theta x^{\lambda}}$$

$$=1-\left[1+\frac{\theta^3x^{3\lambda}+3\theta^2x^{2\lambda}+6\theta x^{\lambda}}{\alpha\theta^3+6}\right]e^{-\theta x^{\lambda}} \qquad (3)$$

Moreover, the graphs of the pdf and cdf of TWGG are given in Figure 1 and Figure 2 respectively.

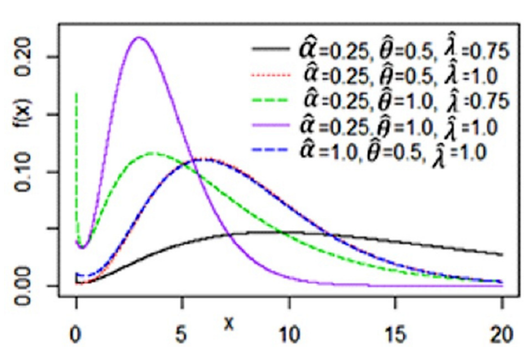

Figure 1. Graph of the pdf of TWGG Distribution of varying values of the parameters

The graphs presented in Figure 1 show the pattern of the pdf of the proposed TWGG distribution at various values of α, θ and λ. The pdf shows different shapes which makes it flexible to capture various levels of parameters.

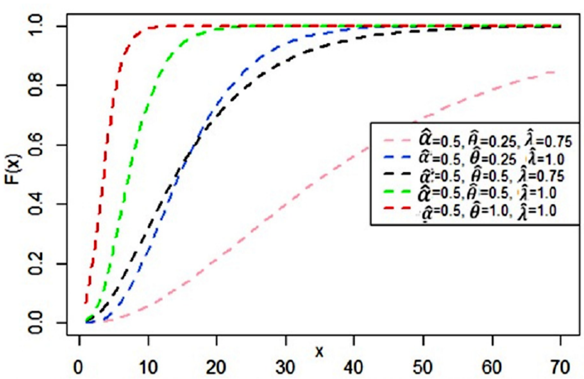

Figure 2. Graph of the cdf of TWGG distribution of varying values of papramerers

Figure 2 presented the shapes of the cdf of the proposed TWGG distribution at selected values of α, θ and λ. Just like its pdf, the cdf also show different shapes which makes it flexible to capture various levels of parameters.

Some Mathematical Properties of TWGG distribution

This section includes some properties of the TWGG distribution like, rth moment, the mean, standard deviation, skewness, kurtosis, and coefficients of variation among others.

Moments of TWGG Distribution

The rth moment of a random variable X is given as

$$\mu_r^{'} = E\left(x^r\right) = \int_0^\infty x^r f\left(x\right) dx \, , \qquad (4)$$

where r is a positive integer.

Hence, putting the (2) into (4) to give

$$\mu_r^{'} = \int_0^\infty x^r \left[\frac{\lambda\theta^4}{\alpha\theta^3+6} \left(\alpha + x^{3\lambda}\right) x^{\lambda-1} e^{-\theta x^\lambda} \right] dx$$

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

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

$$= \frac{\lambda\theta^4}{\alpha\theta^3+6} \left[\left(\alpha \frac{\tilde{A}(r+\lambda)}{\lambda\theta^{r+\lambda}} \right) + \left(\frac{\tilde{A}(r+4\lambda)}{\lambda\theta^{r+4\lambda}} \right) \right] SD$$

$$= \frac{\lambda\theta^4}{\alpha\theta^3+6} \left[\frac{\theta^{-(r+4\lambda)} \left(\alpha\theta^{3\lambda}\tilde{A}(r+\lambda) + \tilde{A}(r+4\lambda)\right)}{\lambda} \right]$$

Therefore,

$$\mu_r^{'} = \frac{\theta^{4-(r+4\lambda)} \left(\alpha\theta^{3\lambda}\tilde{A}(r+\lambda) + \tilde{A}(r+4\lambda)\right)}{\left(6+\alpha\theta^3\right)} \qquad (5)$$

From (5), we can express the first four moments as

$$\mu_1^{'} = \frac{\theta^{3-4\lambda} \left(\alpha\theta^{3\lambda}\tilde{A}(1+\lambda) + \tilde{A}(1+4\lambda)\right)}{6+\alpha\theta^3}$$

$$\mu_2^{'} = \frac{\theta^{2(1-2\lambda)} \left(\alpha\theta^{3\lambda}\tilde{A}(2+\lambda) + \tilde{A}(2+4\lambda)\right)}{6+\alpha\theta^3}$$

$$\mu_3^{'} = \frac{\theta^{1-4\lambda} \left(\alpha\theta^{3\lambda}\tilde{A}(3+\lambda) + \tilde{A}(3+4\lambda)\right)}{6+\alpha\theta^3}$$

$$\mu_4^{'} = \frac{\theta^{-4\lambda} \left(\alpha\theta^{3\lambda}\tilde{A}(4+\lambda) + \tilde{A}(4+4\lambda)\right)}{6+\alpha\theta^3}$$

Remark: the mean of TWGG distribution is $E(X)=\mu=\mu_1^{'}$ also, the variance, skewness, kurtosis, and coefficient of variation can be obtained by:

$$Variance = \sigma^2 = \mu_2^{'} - \mu^2$$

$$\sigma^2 = \frac{\theta^{2-8\lambda}(-\theta^4(\alpha\theta^{3\lambda}\Gamma(1+\lambda] + \Gamma(1+4\lambda))^2 + \theta^{4\lambda}(6+\alpha\theta^3)(\alpha\theta^{3\lambda}\Gamma(2+\lambda) + \Gamma(2+4\lambda)))}{(6+\alpha\theta^3)^2}$$

$$skewness = S_k = \frac{\mu_3'}{\sigma^3} = \frac{\mu_3'}{(\mu_2' - \mu^2)^{(3/2)}}$$

$$S_k = \frac{\theta^{1-4\lambda}\left(\alpha\theta^{3\lambda}\Gamma(3+\lambda) + \Gamma(3+4\lambda)\right)}{(6+\alpha\theta^3)\left(\frac{\theta^{2-8\lambda}\left(-\theta^4\left(\alpha\theta^{3\lambda}\Gamma(1+\lambda) + \Gamma(1+4\lambda)\right)^2 + \theta^{4\lambda}(6+\alpha\theta^3)\left(\alpha\theta^{3\lambda}\Gamma(2+\lambda) + \Gamma(2+4\lambda)\right)\right)}{(6+\alpha\theta^3)^2}\right)^{3/2}}$$

$$kurtosis = K_s = \frac{\mu_4'}{\sigma^4} = \frac{\mu_4'}{(\mu_2' - \mu^2)^{(2)}}$$

$$K_s = \frac{\theta^{-4+12\lambda}(6+\alpha\theta^3)^3(\alpha\theta^{3\lambda}\Gamma(4+\lambda) + \Gamma(4+4\lambda))}{\left(\theta^4\left(\alpha\theta^{3\lambda}\Gamma(1+\lambda) + \Gamma(1+4\lambda)\right)^2 - \theta^{4\lambda}(6+\alpha\theta^3)\left(\alpha\theta^{3\lambda}\Gamma(2+\lambda) + \Gamma(2+4\lambda)\right)\right)^2}$$

$$CV = \frac{\sigma}{\mu}$$

$$CV = \frac{\theta^{-3+4\lambda}(6+\alpha\theta^3)\sqrt{\frac{\theta^{2-8\lambda}\left(-\theta^4\left(\alpha\theta^{3\lambda}\Gamma(1+\lambda) + \Gamma(1+4\lambda)\right)^2 + \theta^{4\lambda}(6+\alpha\theta^3)\left(\alpha\theta^{3\lambda}\Gamma(2+\lambda) + \Gamma(2+4\lambda)\right)\right)}{(6+\alpha\theta^3)^2}}}{\alpha\theta^{3\lambda}\Gamma(1+\lambda) + \Gamma(1+4\lambda)}$$

Renyi's Entropy

The Renyi's Entropy of a distribution can be obtained from the expression below:

$$R_H(x) = \frac{1}{1-p} log \int_0^\infty \left[f(x)\right]^p dx$$

Hence, the Renyi's Entropy of the TWGG distribution is obtained as follows:

$$R_H(x;\alpha,\theta,\lambda) = \frac{1}{1-p} log \int_0^\infty \left[\frac{\lambda\theta^4}{\alpha\theta^3+6}\left(\alpha + x^{3\lambda}\right)x^{\lambda-1}e^{-\theta x^\lambda}\right]^p dx$$

$$= \frac{1}{1-p} log \int_0^\infty \left[\frac{\lambda^p\theta^{4p}}{\left(\alpha\theta^3+6\right)^p}\left(\alpha + x^{3\lambda}\right)^p x^{\lambda p-p}e^{-\theta p x^\lambda}dx\right]$$

$$= \frac{1}{1-p} log \frac{\lambda^p\theta^{4p}}{\left(\alpha\theta^3+6\right)^p}\sum_{j=1}^\infty \binom{p}{j}\left(\frac{1}{\alpha}\right)^j \int_0^\infty x^{3\lambda j+1-1}e^{-\theta p x^\lambda}dx$$

$$= \frac{1}{1-p} log \frac{\lambda^p\theta^{4p}}{\left(\alpha\theta^3+6\right)^p}\sum_{j=1}^\infty \binom{p}{j}\left(\frac{1}{\alpha}\right)^j \frac{Ã(3\lambda j+1)}{(\theta p)^{3\lambda j+1}}$$

(6)

Order Statistics

Let $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$ be the order statistic of a random sample $X_1, X_2, \ldots, X_n$ drawn from continuous population with pdf $f_X(x)$, and cdf $F_X(x)$, then the pdf of the i^{th} order statistic $X_{(i)}$ is given by

$$f_{X_{(i)}}(x) = \frac{n!}{(i-1)!(n-i)!}f_X(x)\left[F_X(x)\right]^{i-1}\left[1-F_X(x)\right]^{n-i} \tag{7}$$

Substituting (2) and (3) into equation (7), the pdf of $X_{(i)}$ according to TWGG distribution is given as the following:

$$f_{X_{(i)}}(x)$$
$$= \frac{n!\,\lambda\theta^4}{(i-1)!\,(n-i)!\,\alpha\theta^3+6)}\left[\left(\alpha+\ x^{3\lambda}\right)x^{\lambda-1}e^{-\theta x^{\lambda}}\right]\Bigg[1$$
$$-\left(\frac{\theta^3x^{3\lambda}+\ 3\theta^2x^{2\lambda}+6\theta x^{\lambda}}{\alpha\theta^3+6}\right)e^{-\theta x^{\lambda}}\Bigg]^{i-1}\left[\left(\frac{\theta^3x^{3\lambda}+\ 3\theta^2x^{2\lambda}+6\theta x^{\lambda}}{\alpha\theta^3+6}\right)e^{-\theta x^{\lambda}}\right]^{n-i} \quad (8)$$

Now, the PDF of $X_{(1)}$ and $X_{(n)}$ respectively are given by :

$$f_{X_{(1)}}(x) = \left(\frac{n!\,\lambda\theta^4}{(i-1)!\,(n-i)!\,(\alpha\theta^3+6)}\right)e^{-\theta x^{\lambda}}\left[\left(\frac{\theta^3x^{3\lambda}+\ 3\theta^2x^{2\lambda}+6\theta x^{\lambda}}{\alpha\theta^3+6}\right)e^{-\theta x^{\lambda}}\right]^{n-1}$$

and

$$f_{X_{(n)}}(x) = \left(\frac{n!\,\lambda\theta^4}{(i-1)!\,(n-i)!\,(\alpha\theta^3+6)}\right)e^{-\theta x^{\lambda}}\left[1-\left(\frac{\theta^3x^{3\lambda}+\ 3\theta^2x^{2\lambda}+6\theta x^{\lambda}}{\alpha\theta^3+6}\right)e^{-\theta x^{\lambda}}\right]^{n-1}$$

Reliability Analysis

Aldeni et al[12] opined that the most used measures for describing the underlying distribution of a lifetime variable is the survival function or the hazard function. The Survival function, *S(x)*, and Hazard Function, *h(x)*, of the TWGG distribution is obtained respectively by

$$S(x) = 1 - F_1(x|\alpha,\theta,\lambda)$$
$$= 1-\left(1-\left[1+\frac{\theta^3x^{3\lambda}+3\theta^2x^{2\lambda}+6\theta x^{\lambda}}{\alpha\theta^3+6}\right]e^{-\theta x^{\lambda}}\right)$$
$$= \frac{e^{-x^{\lambda}\theta}\left(6+6x^{\lambda}\theta+3x^{2\lambda}\theta^2+x^{3\lambda}\theta^3+\alpha\theta^3\right)}{6+\alpha\theta^3} \quad (9)$$

$$h(x) = \frac{f_1(x|\alpha,\theta,\lambda)}{1-F_1(x|\alpha,\theta,\lambda)} = \frac{f_1(x|\alpha,\theta,\lambda)}{S(x;\theta,\alpha,\lambda)}$$
$$= \frac{\dfrac{\theta^4\lambda}{\alpha\theta^3+6}\left(\alpha+x^{3\lambda}\right)x^{\lambda-1}e^{-\theta x^{\lambda}}}{\dfrac{e^{-x^{\lambda}\theta}\left(6+6x^{\lambda}\theta+3x^{2\lambda}\theta^2+x^{3\lambda}\theta^3+\alpha\theta^3\right)}{6+\alpha\theta^3}}$$
$$= \frac{x^{\lambda-1}\left(x^{3\lambda}+\alpha\right)\theta^4\lambda}{6+6x^{\lambda}\theta+3x^{2\lambda}\theta^2+x^{3\lambda}\theta^3+\alpha\theta^3} \quad (10)$$

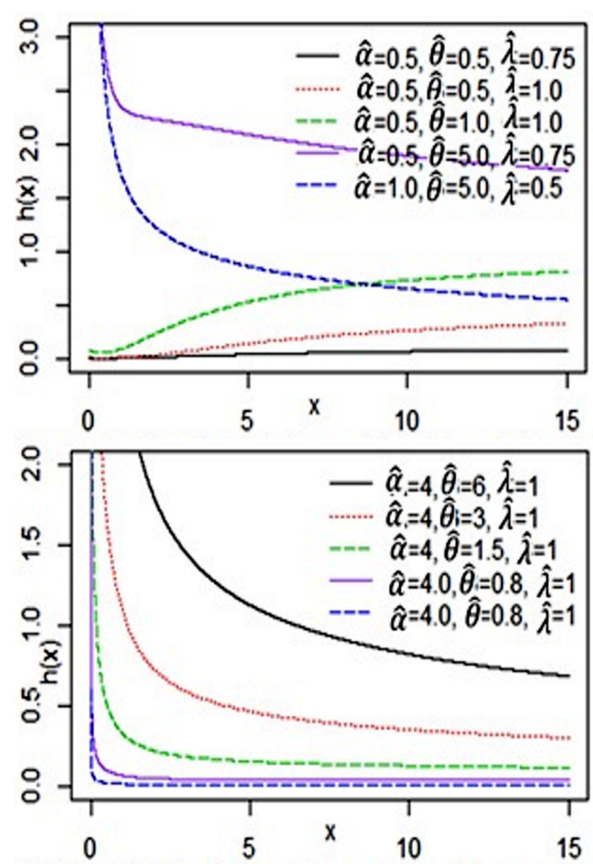

Figure 3. Graph of the hazard function of the TWGG Distribution for varying values of the parameters

From Figure 3, it is clearly observed that the hazard rate function of the TWGG distribution at different values of α,θ and λ can be constant, increasing, decreasing and bathtub-shaped, which offer more flexibility.

Maximum Likelihood Estimator

Given that $X_1, X_2, \ldots, X_n$ is a random sample of size n from TWGG probability distribution, the Maximum Likelihood Estimation (MLE) function of the distribution can be written as

$$L=\prod_{i=1}^{n} f_1\left(x_i|\alpha,\theta,\lambda\right) \tag{11}$$

Substituting (2) into (11) and taking the logarithm function for both sides we found that the Loglikelihood function is given by

$$logL = \text{nlog}\lambda + 4\text{nlog}\theta - n\log\left(\alpha\theta^3+6\right) + \sum_{i=1}^{n}\log\left(\alpha+x_i^{3\lambda}\right)+\sum_{i=1}^{n}\log\left(x_i^{\lambda-1}\right)-\sum_{i=1}^{n}\theta x_i^{\lambda} \tag{12}$$

By taking the partial derivative for logL in (12) with respect to α,θ and λ respectively, we have

$$\frac{\partial logL}{\partial\alpha} = -\frac{n\theta^3}{6+\alpha\theta^3}+\sum_{i=1}^{n}\frac{1}{\alpha+x_i^{3\lambda}}$$

$$\frac{\partial logL}{\partial\theta} = \frac{4n}{\theta}-\frac{3n\alpha\theta^2}{6+\alpha\theta^3}-\sum_{i=1}^{n}x_i^{\lambda}$$

$$\frac{\partial logL}{\partial\lambda} = \frac{n}{\lambda}+\sum_{i=1}^{n}\log\left(x_i\right)-\sum_{i=1}^{n}\theta\log\left(x_i\right)x_i^{\lambda} + \sum_{i=1}^{n}\frac{3\log\left(x_i\right)x_i^{3\lambda}}{\alpha+x_i^{3\lambda}}$$

The Maximum Likelihood Estimator of α,θ and λ can be obtained by solving the above three nonlinear equations when

$$\frac{\partial logL}{\partial\alpha}=0, \frac{\partial logL}{\partial\theta}=0 \text{ and } \frac{\partial logL}{\partial\lambda}=0$$

numerically using R software.[13]

Results

Simulation Study

In this subsection, we evaluate $\hat{\theta}_{MLE}$, $\hat{\lambda}_{MLE}$ and $\hat{\alpha}_{MLE}$ through a brief simulation study. The simulation study of the TWGG distribution is carried out by choosing random samples, say n = 50, 100, . . ., 1000. These samples are obtained using the inverse cdf.

The simulation study is conducted for the combination values θ = 0.25, λ = 0.75, and α = 0.50. The judgement about the performances of the $\hat{\theta}_{MLE}$, $\hat{\lambda}_{MLE}$ and $\hat{\alpha}_{MLE}$ are made by considering two evaluation criteria. These criteria are given by.

Mean square error (MSE)

$$MSE\left(\hat{\theta}_{MLE}\right)=\frac{1}{N}\sum_{i=1}^{N}\left(\hat{\theta}_i-\theta\right)^2$$

Bias

$$Bias\left(\hat{\theta}_{MLE}\right)=\frac{1}{N}\sum_{i=1}^{N}\left(\hat{\theta}_i-\theta\right)$$

The above evaluation criteria are also computed for $\hat{\lambda}_{MLE}$ and $\hat{\alpha}_{MLE}$. The simulation study is performed using the optim()R-function with argument method = ”L-BFGSB”. The simulation results are presented numerically in Tables 1. From the results of the simulation of the TWGG distribution presented in Tables 1, we can see that.

1. As n increases (i.e., as n → ∞), the values of $\hat{\theta}_{MLE}$, $\hat{\lambda}_{MLE}$ and $\hat{\alpha}_{MLE}$ tend to become stable.
2. As n → ∞, the MSEs of $\hat{\theta}_{MLE}$, $\hat{\lambda}_{MLE}$ and $\hat{\alpha}_{MLE}$ dwindle to zero.
3. As n → ∞, the Biases of $\hat{\theta}_{MLE}$, $\hat{\lambda}_{MLE}$ and $\hat{\alpha}_{MLE}$ tend toward zero.

Applications

In this section, two real data sets are used to compare the performance of proposed TWGG distribution with three existing models: Generalized Gamma (GGD) distribution,[3] Two-Parameter Rama (TPRD) which is as special case of our proposed distribution and Generalized Lindley ("GLD5" as named by the authors) distribution.[14] To compare the performance of our model with the others models the following criterions are used: Akaike information criterion (AIC), Bayesian information criterion (BIC) and Corrected Akaike information criterion (CAIC). The distribution with the smallest values of AIC, BIC, and AICc is considered as the best model for the given data. In this section the numerical results are obtained by using R software.

Dataset 1: The data set represents the remission times (in months) of a random sample of 128 bladder cancer patients reported in.[15] See[16] for more details.

Dataset 2: The data represent the survival times of 121 patients with breast cancer obtained from a large hospital in a period from 1929 to 1938.[17]

Table 1. The numerical illustration of the SS of the TWGG distribution for θ = 0.25, λ = 0.75, and α = 0.50

n	Parameters	MLEs	MSEs	Biases
50	θ	0.2427	0.0056	-0.0072
	λ	0.7698	0.0072	0.0198
	α	0.7850	2.7356	0.2850
100	θ	0.2481	0.0027	-0.0018
	λ	0.7585	0.0031	0.0085
	α	0.6873	2.4773	0.1873
150	θ	0.2485	0.0019	-0.0015
	λ	0.7558	0.0021	0.0058
	α	0.6797	2.0457	0.1797
300	θ	0.2495	0.0010	-0.0005
	λ	0.7528	0.0011	0.0028
	α	0.6390	1.6917	0.1390
	θ	0.2505	0.0005	-0.0005
500	λ	0.7521	0.0006	0.0021
	α	0.6139	1.2811	0.1139
	θ	0.2504	0.0003	0.0004
800	λ	0.7504	0.0003	0.0004
	α	0.5189	0.2263	0.0189
	θ	0.2496	0.0003	-0.0004
1000	λ	0.7510	0.0003	0.0010
	α	0.5053	0.0975	0.0053

Table 2. Goodness-of-fit of the distributions for dataset 1 (n=128)

Distribution	MLE Estimates			-2InL	AIC	CAIC	BIC
	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\lambda}$				
TWGG	0.0246	1.3937	0.5178	821.395	827.395	827.588	831.099
TPRD	13356.9	0.1158	-	827.407	831.407	831.503	837.111
GLD5	1.5690	0.1533	-	832.314	836.314	836.410	842.018
GGD	0.1719	2.2686	0.4391	821.891	827.708	827.901	831.412

Table 3. Goodness-of-fit of the distributions for dataset 2 (n=121)

Distribution	MLE Estimates			-2InL	AIC	CAIC	BIC
	$\hat{\alpha}$	$\hat{\theta}$	$\hat{\lambda}$				
TWGG	5.1209	0.3506	0.6455	1157.422	1162.422	1162.628	1167.014
TPRD	10692.2	0.0651	-	1166.687	1170.687	1170.788	1176.278
GLD5	1.6470	0.0348	-	1157.704	1161.704	1161.806	1167.296
GGD	61.685	1.1921	1.5216	1157.774	1163.774	1163.980	1167.366

Discussion

Note that in Tables 2 and 3, the parameter estimates for the TWGG, TPRD, GLD5 and GGD models are estimated by using the maximum likelihood method. The -2lnL, AIC, CAIC and BIC are presented for the two different data sets. For the first dataset, based on the lowest values of the AIC, CAIC, and BIC, the TWGG model turns out to be a better model than the TPRD, GLD5 and GGD models. However, for the second dataset, based on the values of the AIC, CAIC, and BIC, the TWGG model turns out to be a better model than the TPRD and GGD models.

Conclusion

In this paper, a new three-parameter Weibull-Generalized Gamma distribution is proposed and the mathematical properties such as the shape of the density, hazard rate function, moments, Renyi entropy, skewness, kurtosis, and order statistics have been discussed. The maximum likelihood estimation method for estimating its parameters has been discussed. A simulation study was carried out to evaluate the bias and mean square error of the maximum likelihood estimates of the parameters. The application of the proposed distribution to a real lifetime dataset (cancer patients' data) reveals its superiority over the TPRD and GGD models.

Conflicts of Interest

Authors declare that there is no conflict of interest regarding this article.

References

1. Weibull, W. A statistical theory of the strength of materials, IVA Handlingar (Royal Swedish Academy of Engineering Sciences, Proceedings). (1939). 151.

2. Hamed, M.S. The Mixture Weibull-Generalized Gamma Distribution. Advances and Applications in Statistics. (2020). 62(2), 139-171: http://dx.doi.org/10.17654/AS062020139

3. Stacy, E. W. A generalization of the gamma distribution, The Annals of Mathematical Statistics. (1962). 1187-1192.

4. Maswadah, M., Seham, A. M. and Ahsanullah, M. Bayesian inference on the generalized gamma distribution based on generalized order statistics, Journal of Statistical Theory and Applications. (2013). 12(4), 356-377.

5. Pham, H. and Lai, C. D. On recent generalizations of the Weibull distribution, IEEE Transactions on Reliability. (2007). 56(3), 454-458.

6. Aderoju, S. Samade Probability Distribution: Its Properties and Application to Real Lifetime Data. Asian Journal of Probability and Statistics. (2021). 14(1), 1-11. https://doi.org/10.9734/ajpas/2021/v14i130317.

7. Aderoju, S. and Adeniyi, I., On Power Generalized Akash Distribution with Properties and Applications. Journal of Statistical Modeling &Amp; Analytics (JOSMA). (2022). 4(1). https://ejournal.um.edu.my/index.php/JOSMA/article/view/36340

8. Aderoju, S.A. and Jolayemi, E.T. Power Hamza distribution and its applications to model survival time. Journal of the Nigerian Statistical Association. (2022). Vol.34; 1-9.

9. Meshkat, R.S., Torabi, H. and Hamedani, G.G. A Generalized Gamma-Weibull Distribution: Model, Properties and Applications. Pakistan Journal of Statistics and Operation Research. (2016). Vol. 12, No. 2.

10. Umeh, E.U., Umeokeke, E.T. and Ibenegbu, H.A. A two-parameter Rama distribution. Earthline3 Journal of Mathematical Sciences. (2019). 2, 365- 382.

11. Shanker, R. Rama distribution and its application. International Journal of Statistics and Application. (2017). 7(1):26-35.

12. Aldeni M., Famoye, F.& Lee C. A generalized Family of Lifetime Distributions and Survival Models. Journal of Modern Applied Statistical Methods. (2020). 18(2): 1-34.

13. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. (2022). URL https://www.R-project.org/.

14. Zakerzadeh, H. and Dolati, A. Generalized Lindley distribution. Journal of Mathematical Extension. (2010). 3(2), 13-25.

15. Lee, E.T. and Wang, J.W. (2003). Statistical Methods for Survival Data Analysis. Third Edition, Wiley, New York. https://doi.org/10.1002/0471458546.

16. Aderoju, S.A. and Babaniyi, O. Power Samade distribution: its properties and application to real lifetime data. Nigerian Journal of Science and Environment (NJSE). 2023: 21 (1), 237–248. https://delsunjse.com/index.php/njse.

17. Lee, E.T. (1992). Statistical Methods for Survival Data Analysis. John Wiley: New York.