bioRxiv preprint doi: https://doi.org/10.1101/2023.11.18.567670; this version posted November 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

A New Generalized Gamma-Weibull Distribution with Applications to Time-to-event Data

Kazeem Adesina Dauda[a,d,*], Rasheed Kehinde Lamidi[b], Adeshola Adediran Dauda[b], Waheed Babatunde Yahya[c]

[a] *Department of Mathematics, University of Bergen, 5007 Bergen, Norway*

[b] *Department of Mathematics and Statistics, Kwara State University, Malete, Nigeria*

[c] *Department of Statistics, University of Ilorin, Nigeria.*

[d] *alt-Address: Department of Mathematics and Statistics, Kwara State University, Malete, Nigeria*

Abstract

In this research, a new class of probability distributions referred to as Generalized Gamma Weibull (GGW) distributions was introduced within the context of parametric survival analysis. This distribution represents a modification of the gamma Weibull distribution and offers valuable insights, particularly when dealing with highly skewed lifetime data. The study extensively examined the mathematical characteristics of these distributions, encompassing hazard functions, moments, quantile functions, and order statistics. Furthermore, the research delved into parameter estimation methods for these newly proposed distributions, employing the maximum likelihood technique, Fisher Information (FI), and deriving asymptotic confidence intervals for both censored and uncensored scenarios. To illustrate the practical utility of these proposed distributions, the study applied them to analyze two sets of real-life survival data and two sets of real-life data, resulting in a total of four distinct datasets. To gauge the effectiveness of the GGW distributions in comparison to existing methods such as Generalized Weibull and Generalized gamma (G-Weibull and G-Gamma) distributions, the research employed statistical indices including the Akaike Information Criterion (AIC), Corrected Akaike Information Criterion (CAIC), and Bayesian Information Criterion (BIC). The outcomes of this comparative analysis demonstrated the superior performance of the newly introduced GGW distributions (AIC=338.6313, BIC=346.2794, and CAIC=339.5202) when contrasted with the existing methods (G-Weibull: AIC=376.1946, BIC=381.9307, and CAIC=376.5424) across all three criteria, thereby highlighting the enhanced suitability of GGW distributions for modeling and analyzing skewed lifetime data.

Keywords: generalized Weibull, Gamma-generated distributions, moments, order statistic

*Corresponding author

Email addresses: kazeem.dauda@uib.no, qdauda70@gmail.com (Kazeem Adesina Dauda), rasheedlamidi44@gmail.com (Rasheed Kehinde Lamidi), adeshola.dauda@kwasu.edu.ng (Adeshola Adediran Dauda), wbyahya@unilorin.edu.ng (Waheed Babatunde Yahya)

bioRxiv preprint doi: https://doi.org/10.1101/2023.11.18.567670; this version posted November 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

1. Introduction

Zografos and Balakrishnan [1], as well as Ristić and Balakrishnan [2], have introduced two distinct distribution categories, each generated using gamma random variables, which incorporate an additional positive shape parameter. The mathematical characteristics of this distribution were examined by [3], with a specific focus on its applicability to highly skewed time-to-event data, while disregarding the censoring information in the dataset. In separate investigations, [4] and [5] have devised a novel distribution by combining elements of the Zografos-Balakrishnan log-logistic distribution and the gamma exponentiated-Weibull distribution. When compared to the original distribution proposed by ([1]), this new distribution exhibited remarkable performance improvements. Furthermore, [6] has put forth an innovative variation within the Zografos-Balakrishnan family, introducing the Zografos-Balakrishnan-Gompertz distribution and the Zografos-Balakrishnan Lindley distribution. In their work, they meticulously established the mathematical properties, including moments, hazard rate functions, and reliability functions, for these newly introduced distributions. These distributions prove to be highly effective for modeling datasets that exhibit a pronounced skewness. Taking a broader perspective, [6] delved into a generalized form of the log-transformed inverse Weibull distribution, conducting an extensive investigation of its theoretical properties. In summary, several studies in the literature, such as those cited in references ([7, 8]), have explored and developed flexible distribution models capable of handling highly skewed time-to-event data. However, there still remains a need for the development of a distribution that can adeptly accommodate highly skewed time-to-event data while simultaneously considering censoring information.

1.1. Existing and Parent Models

The probability density function (pdf) $f(y)$ and cumulative distribution function (cdf) $F(y)$ for the distribution family introduced by [1] is expressed as follows:

$$G(y) = \frac{1}{\Gamma(\delta)} \int_0^{-log\bar{F}(y)} t^{\delta-1} e^{-t} dt, y, \delta > 0, \tag{1}$$

where $F(y)$(i.e the cdf) is the parent or based distribution.

However, the pdf of the generator is defined as;

$$g(y) = \frac{1}{\Gamma(\delta)} (-log\bar{F}(y))^{\delta-1} f(y), \delta > 0, \tag{2}$$

where $\Gamma(.)$ is the gamma function. The derivation of the new distributions relies solely on these equations. The corresponding Hazard rate function (hrf) is;

$$h(y) = \frac{\{-log\left[1 - F(y)\right]\}^{\delta-1} f(y)}{\Gamma\left(\delta, -log\left[1 - F(y)\right]\right)} \tag{3}$$

bioRxiv preprint doi: https://doi.org/10.1101/2023.11.18.567670; this version posted November 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

where $\Gamma(\delta, z) = \int_z^\infty t^{\alpha-1}e^{-t}dt$ denotes the complementary incomplete gamma function. The generator of the hazard rate function has an additional shape parameter $\delta > 0$.

This study considers introducing a new generalized gamma-weibull distribution, inspired by the findings of Zagrafos and Balakrishnan's distribution family described in equation 1. A new modified generalized gamma-Weibull distribution has been developed as a member of the Zagrofos and Balakrishnan distribution family [1]. A similar version of the generalized gamma-Weibull distribution has been introduced in the master thesis of one of the authors of this study in 2022[9]. In this new version, we derive the statistical properties of this distribution, estimate the maximum likelihood, and provide an application example. Moreover, it is a prevalent practice in the literature to focus on uncensored observations while disregarding censoring scenarios. However, our study addresses these limitations by incorporating more adaptable parameters that accommodate both censored and uncensored observations. Further elaboration on the proposed distributions can be found in the subsequent section.

The structure of the paper unfolds as follows: In Section 2, we express $g(y)$ in equation 5 as an infinite linear combination of exponentiated gamma-Weibull density functions. Section 2.3 is dedicated to providing the Quantile function for the Generalized Gamma Weibull distribution. Section 2.4 focuses on obtaining moments, while Section 2.5 delves into the distribution of order statistics. Moving forward to Section 2.6, we employ maximum likelihood estimation to determine the model parameters. The Asymptotic confidence interval for the GGW distribution is derived in Section 3. In Section 4, we use censored samples to estimate the model parameters using maximum likelihood in the Generalized Gamma Weibull distribution. The methods of model evaluation is outlined in Section 5. Finally, Section 6 illustrates the practical utility of the approach through its application to a real-life dataset. Concluding remarks are presented in Section 7.

2. Methodology

The generation of new distributions relies on the use of the Generalized Weibull (GW) distribution, which was originally proposed by Mudholkar and Kumar [10]. The Generalized Weibull distribution is characterized by its probability density function (PDF) denoted as $f(y)$ and cumulative distribution function (CDF) represented as $F(y)$. The PDF and CDF expressions can be found in equations 4 and 5.

$$F(y) = \left(1 - e^{-(\alpha y)^\theta}\right)^\lambda, y > 0, \lambda > 0, \theta > 0, \alpha > 0 \tag{4}$$

$$f(y) = \lambda\theta\alpha^\theta y^{\theta-1}\left(1 - e^{-(\alpha y)^\theta}\right)^{\lambda-1} e^{-(\alpha y)^\theta}, y > 0, \lambda > 0, \theta > 0, \alpha > 0. \tag{5}$$

bioRxiv preprint doi: https://doi.org/10.1101/2023.11.18.567670; this version posted November 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

We initiate the creation of this novel distribution by substituting equation 4 into equation 1 and equation 5 into equation 2, respectively. This process results in the cumulative distribution function (CDF) of the newly modified distribution, referred to as the Generalized Gamma Weibull (GGW), which is derived and presented in equation 6 and equation 7.

$$G(y) = \frac{\gamma(\delta, -log(\bar{F}(y)))}{\Gamma(\delta)} = \frac{1}{\Gamma(\delta)} \int_0^{-log(\bar{F}(y))} t^{\delta-1} e^{-t} dt, y, \delta > 0$$

$$G(y) = \frac{1}{\Gamma(\delta)} \int_0^{-log\left[1-\left(1-e^{-(\alpha y)^\theta}\right)^\lambda\right]} t^{\delta-1} e^{-t} dt, y > 0, \lambda > 0, \theta > 0, \alpha > 0, \delta > 0 \tag{6}$$

$$\frac{\gamma\left(\delta, -log\left[1-\left(1-e^{-(\alpha y)^\theta}\right)^\lambda\right]\right)}{\Gamma(\delta)}, y > 0, \lambda > 0, \theta > 0, \alpha > 0, \delta > 0, \tag{7}$$

where $\gamma(.)$ is called power series see [11] and define as

$$\gamma(\delta, y) = \sum_{u=0}^{\infty} \frac{(-1)^u y^{u+\delta}}{(u+\delta)u!}$$

In a similar manner, the probability density function (PDF) of the Generalized Gamma Weibull (GGW) is acquired by substituting the quantity from equation 5 into equation 2. The resulting PDF is then expressed as follows:

$$g(y) = \frac{1}{\Gamma(\delta)} (-log\bar{F}(y))^{\delta-1} f(y), y \in R \in, \delta > 0$$

$$g(y) = \frac{1}{\Gamma(\delta)} \left(-log\left[1-(1-e^{-(\alpha y)^\theta}\right)^\lambda]\right)^{\delta-1} \lambda\theta\alpha^\theta y^{\theta-1} \left(1-e^{-(\alpha y)^\theta}\right)^{\lambda-1} e^{-(\alpha y)^\theta}$$

$$= \frac{1}{\Gamma(\delta)} \lambda^{\delta-1} \left((\alpha y)^\theta\right)^{\delta-1} \lambda\theta\alpha^\theta y^{\theta-1} \left(1-e^{-(\alpha y)^\theta}\right)^{\lambda-1} e^{-(\alpha y)^\theta}$$

$$= \frac{1}{\Gamma(\delta)} (\alpha y)^{(\delta-1)^\theta} \lambda\theta\alpha^\theta y^{\theta-1} \left(1-e^{-(\alpha y)^\theta}\right)^{\lambda-1} e^{-(\alpha y)^\theta}$$

$$= \frac{1}{\Gamma(\delta)} \lambda^\delta \alpha^{(\delta-1)^\theta} y^{(\delta-1)^\theta} \theta\alpha^\theta y^{\theta-1} \left(1-e^{-(\alpha y)^\theta}\right)^{\lambda-1} e^{-(\alpha y)^\theta}$$

$$= \frac{1}{\Gamma(\delta)} \lambda^\delta \alpha^{\theta\delta} \alpha^{-\theta} y^{\delta\theta} y^{-\theta} \theta\alpha^\theta y^\theta y^{-1} \left(1-e^{-(\alpha y)^\theta}\right)^{\lambda-1} e^{-(\alpha y)^\theta}$$

$$= \frac{1}{\Gamma(\delta)} \lambda^\delta \alpha^{\theta\delta} y^{\delta\theta} \theta y^{-1} \left(1-e^{-(\alpha y)^\theta}\right)^{\lambda-1} e^{-(\alpha y)^\theta}$$

$$= \frac{1}{\Gamma(\delta)} \lambda^\delta \alpha^{\theta\delta} y^{\theta\delta-1} \theta \left(1-e^{-(\alpha y)^\theta}\right)^{\lambda-1} e^{-(\alpha y)^\theta}. \tag{8}$$

The Generalized Gamma Weibull (GGW) distribution stands out for its exceptional flexibility, as it encompasses several other sub-distributions as specific instances. This remarkable adaptability allows for an

bioRxiv preprint doi: https://doi.org/10.1101/2023.11.18.567670; this version posted November 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

in-depth exploration of the diverse special cases inherent to the GGW distribution. Within the following subsections, we delve into some of the widely recognized special sub-models that fall under the Generalized Gamma Weibull distribution.

Sub-model 1:

If $\lambda = 1$ then the generalized gamma weibull distribution reduces to a generalized gamma distribution with parameter δ, α, θ, and its pdf is defined by [12]:

$$f(y) = \frac{\alpha\theta}{\Gamma(\delta)}(\alpha y)^{\theta\delta-1}e^{-(\alpha y)^{\theta}} \tag{9}$$

$$f(y) = \frac{1}{\Gamma(\delta)}\lambda^{\delta}\alpha^{\theta\delta}y^{\theta\delta-1}\theta\left(1-e^{-(\alpha y)^{\theta}}\right)^{\lambda-1}e^{-(\alpha y)^{\theta}}. \tag{10}$$

Substituting $\lambda = 1$ in equation 10 reduces to equation 9;

Sub-model 2:

If $\lambda = 1$ and $\theta = 1$ then the generalized gamma weibull distribution reduces to gamma distribution and its pdf can be written as:

$$f(y) = \frac{\alpha^{\delta}}{\Gamma(\delta)}y^{\delta-1}e^{-(\alpha y)} \tag{11}$$

$$f(y) = \frac{1}{\Gamma(\delta)}\lambda^{\delta}\alpha^{\theta\delta}y^{\theta\delta-1}\theta\left(1-e^{-(\alpha y)^{\theta}}\right)^{\lambda-1}e^{-(\alpha y)^{\theta}}. \tag{12}$$

Replacing $\lambda = \theta = 1$ in equation 12 reduces to equation 11.

Sub-model 3:

If $\lambda = 1$,$\theta = 1$ and $\delta = 1$, then the generalized gamma weibull distribution reduces to exponential distribution and its pdf can be written as:

$$f(y) = \alpha e^{-\alpha y} \tag{13}$$

$$f(y) = \frac{1}{\Gamma(\delta)}\lambda^{\delta}\alpha^{\theta\delta}y^{\theta\delta-1}\theta\left(1-e^{-(\alpha y)^{\theta}}\right)^{\lambda-1}e^{-(\alpha y)^{\theta}}. \tag{14}$$

Replacing $\lambda = \theta = \delta = 1$ in equation 14 reduces to 13.

Sub-model 4:

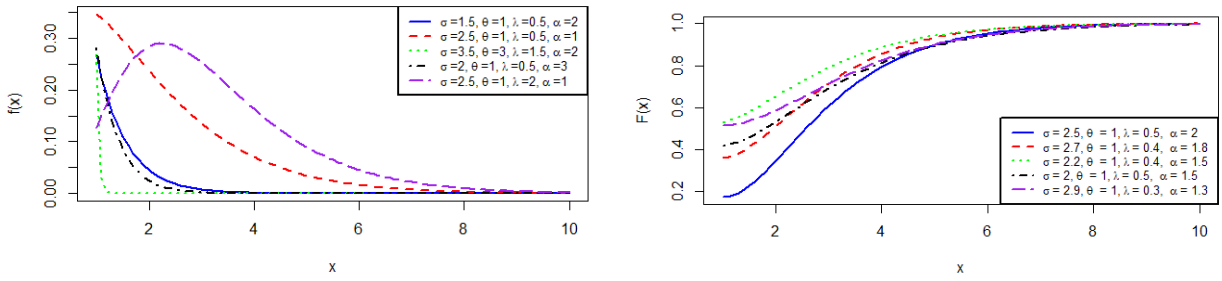
If $\delta = 1$ and $\theta = 1$, then the generalized gamma weibull distribution reduces to Exponentiated Exponential distribution which is presented by [13] as follows:

$$f(y;\lambda,\theta,\alpha) = \lambda\alpha e^{-\alpha y}\left(1-e^{-(\alpha y)}\right)^{\lambda-1} \tag{15}$$

$$f(y) = \frac{1}{\Gamma(\delta)}\lambda^{\delta}\alpha^{\theta\delta}y^{\theta\delta-1}\left(1-e^{-(\alpha y)^{\theta}}\right)^{\lambda-1}e^{-(\alpha y)^{\theta}}. \tag{16}$$

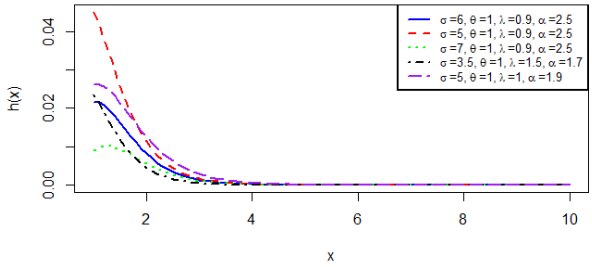
bioRxiv preprint doi: https://doi.org/10.1101/2023.11.18.567670; this version posted November 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Substituting $\delta = 1$ and $\theta = 1$ in equation 16 reduces to equation 15.

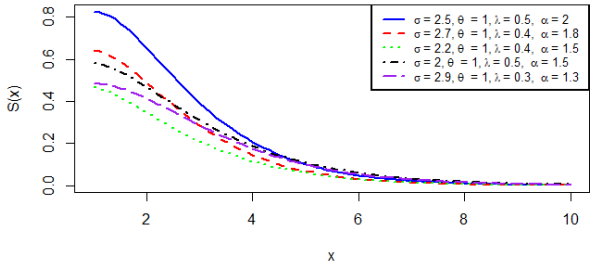


(a) probability density function of GGW

(b) cumulative distribution function of GGW



(c) Hazard function of GGW



(d) Survival function of GGW

Figure 1: Plots of the four different functions of GGW

The figures presented in Figure 1d display the (a) probability density function (pdf), (b) cumulative distribution function (Cdf), (c) hazard function, and (d) survival function of the new distribution. These plots exhibit varying shapes as the parameter values $(\delta, \theta, \lambda, \alpha)$ change. This variability is expected, as it aligns with the behavior of certain real-life datasets. Consequently, the modified distribution (GGW) demonstrates its capability to accommodate a wide range of shapes observed in real-world data.

2.1. Hazard Rate Function (HRF)

The concepts of the Hazard Rate Function (HRF), sometimes referred to as the hazard function or instantaneous failure rate, are frequently employed in the disciplines of reliability engineering, survival analysis, and statistics[14].

Thus, the hazard rate function (HRF) for the Generalized Gamma Weibull distribution, as provided in equation 17, has been meticulously developed by precisely adhering to the quantity outlined in equation 3 and incorporating all the previously mentioned functions.

bioRxiv preprint doi: https://doi.org/10.1101/2023.11.18.567670; this version posted November 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

$$h(y) = \frac{\left\{-log\left[1-\left(1-e^{-(\alpha y)^{\theta}}\right)^{\lambda}\right]\right\}^{\delta-1} \lambda\theta\alpha^{\theta}y^{\theta-1}\left(1-e^{-(\alpha y)^{\theta}}\right)^{\lambda-1} e^{-(\alpha y)^{\theta}}}{\Gamma\left(\delta, -log\left[1-\left(1-e^{-(\alpha y)^{\theta}}\right)^{\lambda}\right]\right)}. \tag{17}$$

2.2. Expansion of the GGW Density Function

For any real parameter $\delta > 0$,the following formula holds,

$$\{-log[1-F(y)]\}^{\delta-1} = (\delta-1)\sum_{w=0}^{\infty}\binom{w+1-\delta}{w}\sum_{j=0}^{w}\frac{(-1)^{j+k}P_{j,w}}{(\delta-1-j)}\binom{w}{j}\left[\left(1-e^{-}(\alpha y)^{\theta}\right)^{\lambda}\right]^{\delta+w-1},$$

where δ is a real parameter and the constants $P_{j,w}$ can be determine recursively by:

$p_{j,w} = w^{-1}\sum_{u=1}^{w}[w-u(j+1)]\,c_u P_{j,w-u}$

for $w = 1, 2, ...$ and $p_{j,0} = 1$. Furthermore, for any real parameter, we define:

$$b_w = \frac{\binom{w+1-\delta}{w}}{(\delta+w)\Gamma(\delta-1)}\sum_{j=0}^{w}\frac{(-1)^{j+w}p_{j,w}}{\delta-1-j}\binom{w}{j} \tag{18}$$

The corresponding equation 1 can be expressed as;

$$G(y) = \sum_{w=0}^{\infty} b_w H_{\delta+w}(y), \tag{19}$$

where $H_{\delta+w}(y)$ denote the cdf of the exp(G)$(\delta+w)$ distribution. So, several properties of the gamma (G) distribution can be obtained by knowing those of the exp (G) distribution, see, for example, [15], [16] and [17], among others.

Then equation 5 can be expressed as equation 20 in other to find the linear combination of GGW distribution:

$$g(y) = \sum_{w=0}^{\infty} b_w h_{\delta+w}(y), \tag{20}$$

where $h_{\delta+w}(y)$ denotes the pdf of the exp-G$(\delta+w)$ distribution. Then

$$h_{\delta+w}(y) = (\delta+w)\lambda\theta\alpha^{\theta}y^{\theta-1}\left(1-e^{-(\alpha y)^{\theta}}\right)^{\lambda-1} e^{-(\alpha y)^{\theta}}\left[\left(1-e^{-(\alpha y)^{\theta}}\right)^{\lambda}\right]^{\delta+w-1}$$

denotes the exponentiated generalized weibull (EGW) $(\tau, \delta+w)$ density function. So, the GGW density function is a linear mixture of EGW density

2.3. Quantile Function of Generalized Gamma Weibull Distribution

In other to obtain the quantile function of the GGW, consider the use of the equation developed by gradshteyn[18].

$$\left(\sum_{i=0}^{\infty}\delta_i q^i\right)^n = \sum_{i=0}^{\infty} c_{n,i} q^i, \tag{21}$$

bioRxiv preprint doi: https://doi.org/10.1101/2023.11.18.567670; this version posted November 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

where the coefficients $c_{n,i}(for$i=1,2...) are easily determined from the recurrence equation.

$$c_{n,i} = (ia_0)^{-1} \sum_{u=1}^{i} [u(n+1) - i] a_u c_{n,i-u} \tag{22}$$

where $c_{n,0} = a_0^n$. The coefficient $c_{n,i}$ can be recursively calculated base on $c_{n,0}, ..., c_{n,i-1}$. In mathematical terms, $c_{n,i}$ can be expressed explicitly as a function of $a_0, ..., a_i$, through this explicit expression might not be necessary for numerical programming using algebraic or numerical software.

Moreover, when considering a random variable y with a cumulative distribution function (CDF) denoted as F(y), we define the quantile of Y, denoted as (y_q), as the inverse of the CDF, $F^{-1}(q)$, where $0 \leq q \leq 1$. This definition holds when substituting F(y) with q in equation 4. Consequently, the quantile function of the GGW distribution can be represented as follows:

$$Q_q = \frac{1}{\alpha} \left[-log \left(1 - q^{\frac{1}{\lambda}} \right)^{\frac{1}{\theta}} \right] \tag{23}$$

Let Y be a random variable following the GGW distribution. To generate random numbers from this distribution, we can utilize the expression in equation 23. Furthermore, the median of the GGW distribution, denoted as Med(Y), can be calculated by substituting q = 0.5 in equation 24 as follows.

$$y_{0.5} = \frac{1}{\alpha} \left[-log \left(1 - (0.5)^{\frac{1}{\lambda}} \right)^{\frac{1}{\theta}} \right] \tag{24}$$

2.4. Moments of the GGW Distribution

Choudhury [19] developed a formula for calculating the jth raw moment of the EW distribution without imposing any constraints. This same approach was used to compute the expected value $E(Y_j)$. Additionally, the result $d^r(e^{ay})/dy^r = a^r e^{ay}$ was implicitly utilized for $0 \leq r \leq 1$, as demonstrated in prior research (example., [20])).

Consider transforming the generator in equation 2 using Laplace transform (LT) presented in equation 25 by letting x=F(z). Hence, the LT is then,

$$L = \frac{1}{\Gamma(\delta)} \int_0^{\infty} e^{-zs} \left[-logF(z) \right]^{\delta - 1} f(z) dz \tag{25}$$

By defining x as a function of z, we get x = F(z).

$$L = \frac{1}{\Gamma(\delta)} \int_0^{1} e^{-sQ(x)} \left[-logx \right]^{\delta - 1} dx, \tag{26}$$

where $Q(.) = Q_{GW}(.)$ indicates the quantile function of GW distribution. By expanding $e^{-sQ(m)}$,

$$L = 1 + \sum_{r=1}^{\infty} \frac{(-s)^r}{r!\Gamma(\delta)} \int_0^{1} \left[Q(x) \right]^r \left[-logx \right]^{\delta - 1} dx \tag{27}$$

bioRxiv preprint doi: https://doi.org/10.1101/2023.11.18.567670; this version posted November 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

From F(y)=$\left[1-e^{-(\alpha y)}\right]^{\lambda}$, this integral reduces to

$$\lambda^{\delta}\alpha^{-r}\frac{\partial^{\frac{r}{\theta}+\delta-1}}{\partial a^{\frac{r}{\theta}}\partial b^{\delta-1}}B(1-a,\lambda-b)\mid_{a=b=1} \tag{28}$$

where B(u,n)=$\int_0^1(1-z)^{u-1}t^{n-1}dz$ is the theta function. Hence,

$$L=1+\sum_{r=1}^{\infty}\frac{(-s)^r\lambda^{\delta}\alpha^{-r}}{r!\Gamma(\delta)}\frac{\partial^{\frac{r}{\theta}+\delta-1}}{\partial a^{\frac{r}{\theta}}\partial b^{\delta-1}}B(1-a,\lambda-b)\mid_{a=b=1} \tag{29}$$

the j^{th} ordinary moment μ_j' follows as

$$\mu_j'=(-1)^j\frac{dL}{ds^j}\mid_{s=0}=\frac{\lambda^{\delta}\alpha^{-j}}{j!\Gamma(\delta)}\frac{\partial^{\frac{j}{\theta}+\delta-1}}{\partial a^{\frac{j}{\theta}}\partial b^{\delta-1}}B(1-a,\lambda-b)\mid_{a=b=1} \tag{30}$$

2.5. Order Statistics for the GGW Distribution

Order statistics play a vital role in various aspects of statistical theory and application. Consider a random sample $Y_1,Y_2,...,Y_n$ drawn from a gamma distribution. Denote the i^{th} order statistic as $Y_{i:n}$. These relationships can be further explored using Equation 19 and Equation 20.

$$g_{i:n}=Wg(y)G^{i-1}(y)\left[1-G(y)^{n-i}\right] \tag{31}$$

$$g_{i:n}(y)=W\sum_{j=0}^{n-i}(-1)^j\binom{n-i}{j}g(y)G(y)^{j+i-1}$$

$$=W\sum_{j=0}^{n-i}(-1)^j\left[\sum_{r=0}^{\infty}b_r(\delta+r)F(y)^{\delta+r-1}f(y)\right]\left[\sum_{w=0}^{\infty}b_wF(y)^{\delta+w}\right]^{j+i-1}$$

where $W=\frac{n!}{[(i-1)!(n-i)!]}$. Using equation 21 and equation 22, we can write

$$\left[\sum_{w=0}^{\infty}b_wF(y)^{\delta+w}\right]^{j+i-1}=\sum_{w=0}^{\infty}g_{j+i-1,w}F(y)^{\delta(j+i-1)+w}$$

where $g_{j+i-1,0}=b_0^{j+i-1}$ and,

$$g_{j+i-1,w}=(wb_0)^{-1}\sum_{u=1}^{w}[u(j+i)-k]b_ug_{j+i-1,k-u}.$$

Hence,

$$g_{i:n}(y)=\sum_{j=0}^{n-i}\sum_{w=0}^{\infty}u_{j,r,w}h_{\delta(j+i)+r+w}(y)$$

$$u_{j,r,w}=\frac{(-1)^jn!}{(i-1)!(n-i-j)!j!}\frac{(\delta+r)b_rg_{j+i-1,w}}{[\delta(j+i)+r+w]} \tag{32}$$

bioRxiv preprint doi: https://doi.org/10.1101/2023.11.18.567670; this version posted November 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

2.6. Maximum Likelihood Estimation for the GGW Distribution

Given a sample of size n, $y_1, ..., y_n$, drawn from the GGW $(\lambda, \theta, \alpha, \delta)$ distribution, the log-likelihood function for the parameter vector $\Phi = (\lambda, \theta, \alpha, \delta)^T$ is represented as follows:

$$L(\Phi) = \prod_{i=1}^{n} f(y_i; \lambda, \theta, \alpha, \delta) \tag{33}$$

2.6.1. Log-likelihood function for the GGW Distribution

$$l_i(\Phi) = \prod_{i=1}^{n} \left[\frac{1}{\Gamma(\delta)} \left(-log \left[1 - (1 - e^{-(\alpha y)^\theta})^\lambda \right] \right)^{\delta - 1} \lambda \theta \alpha^\theta y^{\theta - 1} \left(1 - e^{-(\alpha y)^\theta} \right)^{\lambda - 1} e^{-(\alpha y)^\theta} \right] \tag{34}$$

The log-likelihood function for Φ is

$$\begin{aligned} l_i(\Phi) = n(log(\lambda) + log(\theta) + \theta log(\alpha)) - nlog\Gamma(\delta) + (\theta - 1) \sum_{i=0}^{n} log(y_i) + (\lambda - 1) \sum_{i=0}^{n} log(1 - e^{(-\alpha y_i)^\theta}) - \\ \sum_{i=0}^{n} (\alpha y_i)^\theta + (\delta - 1) \sum_{i=0}^{n} log(-log[1 - (1 - e^{-(\alpha y_i)^\theta})^\lambda]) \end{aligned} \tag{35}$$

The partial derivatives of $l(\Phi)$ with respect to the parameters are

$$\frac{\partial l_i(\Phi)}{\partial \lambda} = \frac{n}{\lambda} + \sum_{i=0}^{n} log \left(1 - e^{-(\alpha y_i)^\theta} \right) + (-1 + \delta) \sum_{i=0}^{n} - \frac{\left(1 - e^{-(\alpha y_i)^\theta} \right)^\lambda log \left(1 - e^{-(\alpha y_i)^\theta} \right)}{\left(1 - \left(1 - e^{-(\alpha y_i)^\theta} \right)^\lambda \right) log \left(1 - \left(1 - e^{-(\alpha y_i)^\theta} \right)^\lambda \right)} \tag{36}$$

$$\begin{aligned} \frac{\partial l_i(\Phi)}{\partial \theta} = n \left(\frac{1}{\theta} + log(\alpha) \right) + \sum_{i=0}^{n} log(y_i) - \sum_{i=0}^{n} log(\alpha y_i)(\alpha y_i)^\theta + (-1 + \lambda) \sum_{i=0}^{n} \frac{e^{-(\alpha y_i)^\theta} log(\alpha y_i)(\alpha y_i)^\theta}{1 - e^{-(\alpha y_i)^\theta}} + \\ (-1 + \delta) \sum_{i=0}^{n} - \frac{e^{-(\alpha y_i)^\theta} \left(1 - e^{-(\alpha y_i)^\theta} \right)^{-1 + \lambda} \lambda log(\alpha y_i)(\alpha y_i)^\theta}{\left(1 - \left(1 - e^{-(\alpha y_i)^\theta} \right)^\lambda \right) log \left(1 - \left(1 - e^{-(\alpha y_i)^\theta} \right)^\lambda \right)} \end{aligned} \tag{37}$$

$$\begin{aligned} \frac{\partial l_i(\Phi)}{\partial \alpha} = \frac{n\theta}{\alpha} - \sum_{i=0}^{n} \theta y_i (\alpha y_i)^{-1 + \theta} + (-1 + \lambda) \sum_{i=0}^{n} \frac{e^{-(\alpha y_i)^\theta} \theta y_i (\alpha y_i)^{-1 + \theta}}{1 - e^{-(\alpha y_i)^\theta}} + \\ (-1 + \delta) \sum_{i=0}^{n} - \frac{e^{-(\alpha y_i)^\theta} (1 - e^{-(\alpha y_i)^\theta})^{-1 + \lambda} \lambda \theta y_i (\alpha y_i)^{-1 + \theta}}{\left(1 - \left(1 - e^{-(\alpha y_i)^\theta} \right)^\lambda \right) log \left(1 - \left(1 - e^{-(\alpha y_i)^\theta} \right)^\lambda \right)} \end{aligned} \tag{38}$$

$$\frac{\partial l_i(\Phi)}{\partial \delta} = \sum_{i=0}^{n} log \left(-log \left(1 - \left(1 - e^{-(\alpha y_i)^\theta} \right)^\lambda \right) \right) - \frac{n\Gamma'(\delta)}{\Gamma(\delta)} \tag{39}$$

The maximum likelihood estimates (MLEs) for the parameters λ, θ, α and δ, denoted as $\widehat{\lambda}, \widehat{\theta}, \widehat{\alpha}$ and $\widehat{\delta}$ respectively, are determined by solving the following equations $\frac{\partial l}{\partial \lambda} = \frac{\partial l}{\partial \theta} = \frac{\partial l}{\partial \alpha} = \frac{\partial l}{\partial \delta} = 0$. Since there is no closed-form solution for these equations, an optimization technique must be applied to find the optimal parameter values.

bioRxiv preprint doi: https://doi.org/10.1101/2023.11.18.567670; this version posted November 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

3. Asymptotic Confidence Intervals for the GGW Distribution

Numerical computation facilitates the derivation of expectations within the Fisher Information Matrix (FIM). Consider $\widehat{\Phi}$, denoted as $(\widehat{\lambda}, \widehat{\theta}, \widehat{\alpha}, \widehat{\delta})$, representing the maximum likelihood estimate of the parameter set Φ, which consists of $(\lambda, \theta, \alpha, \delta)$. Provided the standard regularity conditions hold and the parameters reside within the interior of the parameter space (without touching its boundary), we can establish that $\sqrt{n(\widehat{\Phi} - \Phi)}$ converges asymptotically to a normal distribution $N_4(0, I(\Phi)^{-1})$, where $I\Phi$ represents the expected information matrix. To enhance clarity, we'll replace $I\Phi$ with $J\widehat{\Phi}$, evaluated at $\widehat{\Phi}$, thus yielding a multivariate normal distribution $N_4(0, I(\Phi)^{-1})$. In this distribution, the mean vector is $\underline{0} = (0, 0, 0, 0,)^T$. This statistical framework enables the construction of confidence intervals and regions for various model parameters, including λ, θ, α and δ. These intervals are expressed as $100\,(1-\gamma)\,\%$ confidence intervals. Furthermore, the subsequent section formulates and presents the maximum likelihood estimation with censored samples.

4. Maximum Likelihood Estimation in the Generalized Gamma Weibull Distribution with Censored Samples

Suppose a dataset with sample size 'n' consisting of independent positive random variables, denoted as $T_1, ..., T_n$. Each T_i is associated with an indicator variable Δ_i, which equals 0 if T_i represents a censoring time. Let's define Φ as a parameter vector, $\Phi = (\lambda, \theta, \alpha, \delta)^T$. In this context, the likelihood function is denoted as $L(\Phi)$, for a dataset that includes right-censored observations, represented as $(t_1, \Delta_1), ..., (t_n, \Delta_n)$, originating from a Generalized Gamma Weibull (GGW) distribution. The GGW distribution is characterized by its probability density function (pdf) and survival function, which would be employed to formulate the likelihood function.

$$L\left(\Phi; t, \Delta\right) = \prod_{i=1}^{n} g\left(\Phi; t_i\right)^{\Delta_i} S\left(\Phi; t_i\right)^{1-\Delta_i} \tag{40}$$

where $S(t_i) = 1 - G(t_i)$.

The log-likelihood function, $l(\Phi)$, based on data, from equation 7 and equation 8 is

$$L\left(\Phi; t, \Delta\right) = \prod_{i=1}^{n} \left[\frac{1}{\Gamma(\delta)} \left(-log\left[1 - \left(1 - e^{-(\alpha t)^{\theta}}\right)^{\lambda} \right] \right)^{\delta - 1} \lambda \theta \alpha^{\theta} t^{\theta - 1} \left(1 - e^{-(\alpha t)^{\theta}}\right)^{\lambda - 1} e^{-(\alpha t)^{\theta}} \right]^{\Delta_i} \left[\frac{\gamma\left(\delta, -log\left[1 - \left(1 - e^{-(\alpha t)^{\theta}}\right)^{\lambda}\right]\right)}{\Gamma(\delta)} \right]^{1-\Delta_i} \tag{41}$$

bioRxiv preprint doi: https://doi.org/10.1101/2023.11.18.567670; this version posted November 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

$$l(\Phi) = \sum_{i=0}^{n} \Delta_i (n(log(\lambda) + log(\theta) + \theta log(\alpha)) - nlog\Gamma(\delta) + (\theta - 1)log(t_i) + (\lambda - 1)log\left(1 - e^{-(\alpha t_i)^{\theta}}\right) - (\alpha t_i)^{\theta} + (\delta - 1)log\left(-log\left(1 - \left(1 - e^{-(\alpha t_i)^{\theta}}\right)^{\lambda}\right)\right)) + \sum_{i=0}^{n}(1 - \Delta_i)\left(log(\gamma(\delta, -log\left(1 - \left(1 - e^{-(\alpha t_i)^{\theta}}\right)^{\lambda}\right) - nlog\Gamma(\delta)))\right) \quad (42)$$

5. Methods of Evaluation

The aim of the real-life application is to select the best distribution when comparisons are made with respect to related existing distributions. To achieve this, the following criteria were used; Log-likelihood, Akaike Information Criteria (AIC), Bayes Information Criteria (BIC) and Corrected Akaika Information Criterion (CAIC).

AIC is defined as

$$AIC = -2log\ell(\widehat{\theta}) + 2p \quad (43)$$

While BIC is defined as

$$BIC = -2log\ell(\widehat{\theta}) + p \times logn \quad (44)$$

In the context of statistical modeling, $\widehat{\theta}$ symbolizes the optimal parameter estimates derived from maximum likelihood estimation. $\ell(\widehat{\theta})$ corresponds to the log-likelihood associated with these estimates. In both equation 43 and equation 44, the final components introduce a regularization term into the log-likelihood, which is contingent on the number of parameters involved. In linear models, these parameters are typically associated with regression coefficients. These regularization terms are employed to mitigate the risk of overfitting, and it's worth noting that, for practical sample sizes, the regularization effect is more pronounced in the Bayesian Information Criterion (BIC) compared to the Akaike Information Criterion (AIC). Furthermore, for sufficiently large datasets, there exists the Corrected Akaika Information Criterion (CAIC).

$$CAIC = -2log[L(\widehat{\theta})] + K[log(n) + 1]. \quad (45)$$

The theoretical framework of the Akaike Information Criterion (AIC) falls short when it comes to addressing situations in which the genuine model exhibits both significant effects and more gradual, diminishing influences (referenced as [21]). These particular aspects do not exhibit a direct correlation with the Kullback-Leibler discrepancy.

bioRxiv preprint doi: https://doi.org/10.1101/2023.11.18.567670; this version posted November 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

6. Application

The study delved into the application of the GGW distribution in the analysis of both censored and uncensored datasets. This exploration encompassed the examination and discussion of two real-life survival datasets and two real-life time datasets, constituting a total of four distinct datasets. The intent was to demonstrate that the GGW distribution holds promise as a robust lifetime model, particularly when juxtaposed with several well-known distributions, including but not limited to the Generalized Weibull, Generalized gamma, Gamma Weibull, Exponentiated Weibull, and Gamma distributions (referred to as G. Weibull, G. Gamma, Gamma W., E. Weibull, and Gamma, respectively). Furthermore, to gauge the performance of the GGW distribution, various statistical measures were employed, including the calculation of log-likelihood (L), Akaike information criterion (AIC), Corrected Akaike Information Criterion (CAIC), and Bayesian information criterion (BIC) values.

Table 1: GGW Estimation for Life time of 50 devices ([22]).

	Estimate				**Statistic**			
Distribution	λ	θ	α	δ	**-2 log Likelihood**	**AIC**	**BIC**	**CAIC**
GGW	0.6316	0.0777	0.0037	0.3286	330.6313	**338.6313**	**346.2794**	**339.5202**
G Weibull	1.5438	0.7646	0.0090		370.1946	376.1946	381.9307	376.5424
Gamma Weibull		2.435	0.650	0.843	387.2605	393.2605	398.9965	393.7822
E W	6.360	0.690	3.205		411.0439	417.0439	422.7800	417.5657
Gamma			0.0175	0.8002	480.3805	484.3805	488.2045	484.6358
G Gamma		0.500	1.1466	1.462	1040.847	1046.847	1052.583	1047.369

The statistical values under consideration encompass the Akaike Information Criterion (AIC), Corrected Akaike Information Criterion (CAIC), and Bayesian Information Criterion (BIC). Upon analyzing these values, it becomes evident that the GGW model demonstrates a superior fit to the provided data.

Table 1 provides estimates of the GGW distribution parameters and its related sub-models, accompanied by standard errors in parentheses. Additionally, it provides the AIC, CAIC, and BIC values for the Aarset dataset. Furthermore, the estimated covariance matrix for the GGW model can be shown below:

$$\begin{pmatrix} 145.1327 & -234.3283 & -5037.143 & 260.6122 \\ -234.3283 & 1690.3617 & -12857.996 & -416.6699 \\ -5037.1431 & -12857.9958 & 3112518.962 & -8963.2406 \\ 260.6122 & -416.6699 & -8963.241 & 518.0060 \end{pmatrix}$$

bioRxiv preprint doi: https://doi.org/10.1101/2023.11.18.567670; this version posted November 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The 95% asymptotic confidence intervals are: $\lambda \epsilon$ 0.6316 $\pm$ 1.96(12.0471), $\theta \epsilon$ 0.0777 $\pm$ 1.96(41.1140), $\alpha \epsilon$ 0.0037 $\pm$ 1.96(1764.2333) and $\delta \epsilon$ 0.3286 $\pm$ 1.96(22.7598). The second example consists of lifetime data relating to relief times (in minutes) of patients receiving an analgesic.

Table 2: GGW Estimation for lifetime data relating to relief times (in minutes) of patients receiving an analgesic.

	Estimate				**Statistic**			
Distribution	λ	θ	α	δ	**-2 log Likelihood**	**AIC**	**BIC**	**CAIC**
GGW	0.6027	0.0789	0.0450	0.3446	33.9009	**41.9009**	**45.8839**	**44.5676**
Gamma			2.1589	4.1020	41.5398	45.5398	47.5312	46.2457
G Weibull	2.3000	2.4200	0.2967		66.2335	72.2335	75.2207	73.7335
Gamma W		2.4350	0.6500	1.8209	83.0377	89.0377	92.0249	90.5377
E Weibull	6.3600	2.7612	12.3400		282.8704	288.8704	291.8576	290.3704
G Gamma		2.2500	0.3461	1.4616	447.8106	453.8106	456.7978	455.3106

Table 2 presents parameter estimates for the GGW distribution and its associated sub-models, as well as the AIC, CAIC, and BIC values for lifetime data pertaining to the relief times (measured in minutes) of patients administered an analgesic [23]. Additionally, the estimated covariance matrix for the GGW model is given by:

$$\begin{pmatrix} 63.6880 & -95.9834 & -168.3381 & 103.6132 \\ -95.9834 & 655.5238 & -400.4383 & -156.0997 \\ -168.3381 & -400.4383 & 7127.2915 & -273.7726 \\ 103.6132 & -156.0910 & -273.7726 & 190.1124 \end{pmatrix}$$

The 95% asymptotic confidence intervals are: $\lambda \epsilon$ 0.6027 $\pm$ 1.96(7.9805), $\theta \epsilon$ 0.0789 $\pm$ 1.96(25.6032), $\alpha \epsilon$ 0.0450 $\pm$ 1.96(84.4233) and $\delta \epsilon$ 0.3446 $\pm$ 1.96(13.7881)

6.1. Long-term survival with random censoring

The incorporation of long-term survival considerations necessitates adjustments to both the probability density function (PDF) and the survival function. To address this, two distinct datasets were analyzed: one from the Veterans' Administration Lung Cancer study [24] and the other from the Statistical Analysis of Failure Time Data and North Central Cancer Treatment Group for Lung Cancer [25]. Table 3 and Table 4 present estimates of various parameters for right-censored models, alongside metrics such as AIC (Akaike Information Criterion), CAIC (Consistent Akaike Information Criterion), and BIC (Bayesian Information Criterion).

bioRxiv preprint doi: https://doi.org/10.1101/2023.11.18.567670; this version posted November 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 3: GGW Estimation for Veterans' Administration Lung Cancer study.

	Estimate					Statistic			
Distribution	λ	θ	α	δ	Φ	**-2 log Likelihood**	**AIC**	**BIC**	**CAIC**
GGW	46.9411	0.2296	26.2710	2.2000		1322.985	**1326.98**	**133.824**	**1327.07**
Weibull			1.1736	4.4.7932		1496.1820	1500.18	1506.02	1500.27
G Gamma		0.2529	0.4280	1.5678		1731.2010	1737.20	1741.04	1735.29
EKD	0.2060	0.9641	0.08147	0.7230	0.5000	3608.9610	3618.96	3633.56	3619.41
G Weibull	0.9760	0.02120	2.6200			4643.7480	4649.74	4658.50	4649.92

$$\begin{pmatrix} 0.0374 & -43.5481 & -0.0621 & 2.2295 \\ -43.548 & 91143.3342 & 104.3615 & -3125.9218 \\ -0.0621 & 104.3615 & 0.1352 & -4.0622 \\ 2.2295 & -3125.9218 & -4.0622 & 78.6294 \end{pmatrix}$$

The 95% asymptotic confidence intervals are: $\lambda \epsilon$ 46.9411 $\pm$ 1.96(0.1934), $\theta \epsilon$ 0.2296 $\pm$ 1.96(301.8910), $\alpha \epsilon$ 26.2710 $\pm$ 1.96(0.3677) and $\delta \epsilon$ 2.2000 $\pm$ 1.96(8.8673)

Table 4: GGW Estimation for North Central Cancer Treatment Group for Lung Cancer.

	Estimate					Statistic			
Distribution	λ	θ	α	δ	Φ	**-2 log Likelihood**	**AIC**	**BIC**	**CAIC**
GGW	0.9285	0.2000	19.2000	5.2000		4824.7575	**4832.757**	**4846.474**	**4832.936**
Weibull			1.0900	5.1085		4885.9930	4891.993	4891.140	4896.263
G Gamma		0.1078	0.4280	1.3599		6845.0630	6851.063	6861.351	6851.170
EKD	0.0504	0.0018	0.0450	0.0049	0.0062	6985.0832	6995.084	7012.231	7000.135
G Weibull	4.8000	0.2500	5.7600			7841.1200	7841.1800	7841.2830	7841.1181

$$\begin{pmatrix} 384737552 & -2359810.7 & -11945478.0 & -14545606.8 \\ -2359811 & 735336306.2 & 170710.9 & 205698.7 \\ -11945478 & 170710.9 & 167246127.7 & 223983494.8 \\ -14545607 & 205698.7 & 223983494 & 183655381.3 \end{pmatrix}$$

bioRxiv preprint doi: https://doi.org/10.1101/2023.11.18.567670; this version posted November 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

The 95% asymptotic confidence intervals are: $\lambda \epsilon$ 0.9285 $\pm$ 1.96(19614.728), $\theta \epsilon$ 0.2000 $\pm$ 1.96(27117.085), $\alpha \epsilon$ 19.2000 $\pm$ 1.96(12932.367) and $\delta \epsilon$ 5.2000 $\pm$ 1.96(13551.951)

The GGW distribution emerges as a formidable contender when juxtaposed with its counterparts featuring three and four parameters distributions, as indicated in Tables 1-4. In this evaluation, we rely on key values such as the AIC, BIC, and CAIC. Remarkably, the GGW distribution consistently yields the most favorable results, exhibiting the lowest values for AIC, BIC, and CAIC across both censored and uncensored datasets. When assessing model performance based on these information criteria, the clear preference leans towards selecting the model with the lowest information criteria value.

7. Conclusions

This study introduces a new modified distribution known as GGW, which represents a significant advancement in statistical modeling. Extensive efforts have been dedicated to exploring and substantiating the estimations and properties associated with this new modified distribution. Additionally, the practical applicability of GGW has been demonstrated using real-life data in the R programming environment. Notably, the outcomes of this investigation showcase the superior performance of GGW in comparison to existing distributions, as evidenced by the AIC, BIC, and CAIC criteria.

Furthermore, the Generalized Gamma Weibull (GGW) distribution, an encompassing framework that subsumes the generalized gamma Weibull, gamma, exponential, and exponentiated exponential distributions as special cases, is proposed. This comprehensive approach delves into diverse aspects of GGW, including moments, the quantile function, median, and order statistics, with a focus on deriving these properties. The parameter estimation for GGW is accomplished through the maximum likelihood method, and its efficacy is scrutinized across varying sample sizes and parameter combinations. The overarching characteristics of the new modified distribution can be revealed by representing its density function as a linear combination of gamma exponentiated Weibull density functions. Additionally, alternative expressions for moments have been derived using fractional derivatives, an aspect that often goes overlooked in statistical studies, despite its widespread utility in Engineering and Physics. As an illustrative example, we refer to Lemma 5 (b) from Zografos and Balakrishnan [1], which pertains to gamma-F distribution and emphasizes the relevance of these derivatives in statistical analysis. It states that

$$E\left\{[logF(y)]^{v}\right\} = \frac{(-1)^{\delta-1}}{\Gamma(\delta)} \frac{\partial^{v+\delta-1}}{\partial a^{v} \partial b^{\delta-1}} B(a,b) \mid_{a=b=1}$$

The condition of natural $\delta > 1$ and the presence of variable v lead to a significant outcome. This result remains consistent for any positive real value of v, achieved through the utilization of fractional derivatives. It is highly probable that these derivatives can be applied to compute raw moments for various distributions belonging to the gamma (G) and gamma($\bar{G}$) classes.

bioRxiv preprint doi: https://doi.org/10.1101/2023.11.18.567670; this version posted November 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

for natural $\delta > 1$ and v. This result holds for any positive real v by using fractional derivatives. It is very likely that these derivatives can be used for many members of the gamma (G) and gamma($\bar{G}$) classes of distributions as a way of obtaining their raw moments.

References

[1] K. Zografos, N. Balakrishnan, On families of beta-and generalized gamma-generated distributions and associated inference, Statistical methodology 6 (4) (2009) 344–362.

[2] M. M. Ristić, N. Balakrishnan, The gamma-exponentiated exponential distribution, Journal of Statistical Computation and Simulation 82 (8) (2012) 1191–1206.

[3] S. Nadarajah, G. M. Cordeiro, E. M. Ortega, The zografos–balakrishnan-g family of distributions: Mathematical properties and applications, Communications in Statistics-Theory and Methods 44 (1) (2015) 186–215.

[4] M. W. A. Ramos, P. R. D. Marinho, R. V. da Silva, G. M. Cordeiro, The exponentiated lomax poisson distribution with an application to lifetime data, Advances and Applications in Statistics 34 (2) (2013) 107.

[5] L. Pinho, G. Cordeiro, J. Nobre, The gamma-exponentiated weibull distribution, Journal of Statistical Theory and Applications 11 (4) (2012) 379–395.

[6] M. R. Irshad, V. D'cruz, R. Maya, The zografos-balakrishnan lindley distribution: properties and applications, Statistica 81 (1) (2021) 45–64.

[7] F. N. Haredasht, K. A. Dauda, C. Vens, Exploiting censored information in self-training for time-to-event prediction, IEEE Access 11 (2023) 96831–96840. doi:10.1109/ACCESS.2023.3312310.

[8] K. A. Dauda, Optimal tuning of random survival forest hyperparameter with an application to liver disease, Malays J Med Sci 29 (6) (2022) 67–76. doi:https://doi.org/10.21315/mjms2022.29.6.7.

[9] R. K. Lamidi, New generalized gamma-weibull and exponentiated-uniform distributions with applications to time-to-event data, Master's thesis, Kwara State University, Malete, Nigeria (2022).
URL https://tinyurl.com/y9jyuzbc

[10] G. S. Mudholkar, D. K. Srivastava, Exponentiated weibull family for analyzing bathtub failure-rate data, IEEE transactions on reliability 42 (2) (1993) 299–302.

[11] M. Abramowitz, I. A. Stegun, Handbook of mathematical functions with formulas, graphs, and mathematical tables. national bureau of standards applied mathematics series 55. tenth printing.

[12] E. W. Stacy, A generalization of the gamma distribution, The Annals of mathematical statistics (1962) 1187–1192.

[13] D. Kundu, R. D. Gupta, An extension of the generalized exponential distribution, Statistical Methodology 8 (6) (2011) 485–496.

bioRxiv preprint doi: https://doi.org/10.1101/2023.11.18.567670; this version posted November 18, 2023. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

[14] K. A. Dauda, B. Pradhan, B. Uma Shankar, S. Mitra, Decision tree for modeling survival data with competing risks, Biocybernetics and Biomedical Engineering 39 (3) (2019) 697–708. doi:https://doi.org/10.1016/j.bbe.2019.05.001.
URL https://www.sciencedirect.com/science/article/pii/S0208521619300245

[15] G. S. Mudholkar, D. K. Srivastava, M. Freimer, The exponentiated weibull family: A reanalysis of the bus-motor-failure data, Technometrics 37 (4) (1995) 436–445.

[16] R. D. Gupta, D. Kundu, Exponentiated exponential family: an alternative to gamma and weibull distributions, Biometrical Journal: Journal of Mathematical Methods in Biosciences 43 (1) (2001) 117–130.

[17] S. Nadarajah, S. Kotz, The exponentiated type distributions, Acta Applicandae Mathematica 92 (2006) 97–111.

[18] I. S. Gradshteyn, I. Ryzhik, Table of integrals, series, and products. seventh, Elsevier/Academic Press, Amsterdam 48 (2007) 1171.

[19] A. Choudhury, A simple derivation of moments of the exponentiated weibull distribution, Metrika 62 (1) (2005) 17–22.

[20] K. Trenčevski, Ž. Tomovski, On fractional derivatives of some functions of exponential type, Publikacije Elektrotehničkog fakulteta. Serija Matematika (2002) 77–84.

[21] H. Akaike, On the likelihood of a time series model, Journal of the Royal Statistical Society: Series D (The Statistician) 27 (3-4) (1978) 217–235.

[22] M. V. Aarset, How to identify a bathtub hazard rate, IEEE transactions on reliability 36 (1) (1987) 106–108.

[23] A. Clark, G. Mankikar, I. Gray, Diogenes syndrome: a clinical study of gross neglect in old age, The Lancet 305 (7903) (1975) 366–368.

[24] J. D. Kalbfleisch, R. Prentice, Survival analysis, Wiley, New York, 1980.

[25] C. L. Loprinzi, J. A. Laurie, H. S. Wieand, J. E. Krook, P. J. Novotny, J. W. Kugler, J. Bartel, M. Law, M. Bateman, N. E. Klatt, Prospective evaluation of prognostic variables from patient-completed questionnaires. north central cancer treatment group., Journal of Clinical Oncology 12 (3) (1994) 601–607.