

Paper Type: Original Article

Generalized Weighted Chris-Jerry Distribution: Properties and Applications

Precious O. Ibeakuzie ¹ **, Ifeanyi C. Anabike** ² **, Peace M. Nwankwo** ³ **, EbereChukwu Q. Chinedu** ⁴ **, Divine-Favour N. Ekemezie** ⁵ **and Chinyere P. Igbokwe** ⁶*,*[∗]

Department of Statistics, Faculty of Physical Sciences, Nnamdi Azikiwe University, Awka, Anambra State, Nigeria; op.ibeakuzie@unizik.edu.ng; Department of Statistics, Faculty of Physical Sciences, Nnamdi Azikiwe University, Awka, Anambra State, Nigeria; ci.anabike@unizik.edu.ng; Department of Statistics, Faculty of Physical Sciences, Nnamdi Azikiwe University, Awka, Anambra State, Nigeria; mp.nwankwno@stu.unizik.edu.ng;

Department of Statistics, Faculty of Physical Sciences, Nnamdi Azikiwe University, Awka, Anambra State, Nigeria;eq.chinedu@stu.unizik.edu.ng; Department of Statistics, Faculty of Physical Sciences, Nnamdi Azikiwe University, Awka, Anambra State, Nigeria; de.ngozi@stu.unizik.edu.ng;6 Department of Statistics, School of Science and Industrial Technology, Abia State Polytechnic, Aba, Abia State, Nigeria; chinyere.okechukwu@abiastatepolytechnic.edu.ng;

Citation:

Abstract

This study is a generalization of the length-biased and weighted Chris-Jerry distribution. It introduces an additional scale parameter to make the distribution more flexible. The functional form of the distribution which includes the density function, the distribution function, the survival function and the hazard function together with their plots were presented. The study also encapsulates the characteristics of the model with the estimation of the parameters using the maximum likelihood method. The applicability was demonstrated using data on remission times of a sample of 128 bladder cancer patients, mortality rate of children in Japan and Ireland under the age of five. The results validate that the model is apt in describing real life events.

Keywords: Generalized Weighted Chris-Jerry distribution, Length-Biased Chris-Jerry distribution, Cancer, Mortality Rate, Goodness of Fit.

GO Licensee System Analytics. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0).

1|Introduction

No model possesses an absolute fit feature. What researchers accept to be the best fit is a relative attribute. Therefore, there is a continuous search for improved models, especially as industries and the complexities of human engagements produce big data with inherent characteristics. Another related reason for the improvement in models is the flexibility of use. No single model can suit all situations, and only a handful of models have wider applicability. It is on these premises that researchers develop generalizations of probability models.

Some of these generalizations introduce new parameters, such as scale parameters, which in turn improve model performance. Other generalizations focus on tractability, providing easy ways of deriving properties, a concept known as parsimony in the statistical literature.

The literature is rich in probability distribution models. Related works include: Kharazmi et al. [\[3\]](#page-12-0) introduced the generalized weighted exponential distribution, which is an extension of the standard exponential distribution. This generalization includes additional parameters that allow for greater flexibility in modeling data with varying hazard rates. The inclusion of weight functions helps in adjusting the shape and scale of the distribution, making it more adaptable to different types of data. Ramos and Louzada [\[4\]](#page-12-1) extends the Lindley distribution by incorporating weight functions. This generalization improves the model's ability to fit data with diverse characteristics, particularly in scenarios involving skewed data or varying hazard rates. The added parameters allow for better control over the distribution's shape and scale, enhancing its applicability in real-world situations. Kharazmi [\[9\]](#page-14-0) introduced the generalized weighted Weibull distribution, which is an extension of the Weibull distribution. This model includes additional parameters to adjust the weight function, allowing for more flexible modeling of life data and reliability analysis. The generalized version can accommodate different shapes of hazard functions, including increasing, decreasing, and bathtub-shaped hazard rates. Abbas et al. [\[8\]](#page-14-1) developed a new generalized weighted Weibull distribution, which further extends the flexibility of the Weibull distribution. This model incorporates new parameters to adjust both the scale and shape of the distribution, providing a better fit for complex data patterns. The generalized weighted Weibull distribution is particularly useful in survival analysis and reliability engineering. Beaulieu [\[10\]](#page-14-2) introduced the generalized multinomial distribution, which extends the standard multinomial distribution by incorporating additional parameters. This generalization allows for more flexible modeling of categorical data with complex dependencies. The model is particularly useful in situations where the probabilities of different categories are not fixed but vary according to some underlying factors. Domma [\[11\]](#page-14-3) proposed a new generalized weighted Weibull distribution that can model different hazard rate shapes, including decreasing, increasing, upside-down bathtub, N-shape, and M-shape hazard rates. This model provides greater flexibility in fitting data with varying hazard rate patterns, making it suitable for diverse applications in reliability and survival analysis. The extension of the generalized half-normal distribution by Acitas [\[12\]](#page-14-4) introduces weighting mechanisms that improve its applicability to real-world data. This new weighted distribution can better capture the characteristics of skewed data, making it useful in environmental studies, quality control, and other fields where data may not follow symmetric patterns. Generalized weighted exponential-Gompertz distribution by Teamah et al. [\[13\]](#page-14-5). The generalized weighted exponential-Gompertz (GWE-G) distribution combines the properties of the exponential and Gompertz distributions with additional weighting parameters. This model is particularly effective in describing data with monotonically increasing or decreasing hazard rates, commonly found in demographic studies and reliability engineering. Generalized probability weighted moments: application to the generalized Pareto distribution by Rasmussen [\[14\]](#page-14-6). Rasmussen's work provides a framework for estimating parameters of generalized distributions. GPWM is particularly useful for modeling extreme values and tail behavior of distributions, which is crucial in fields such as finance, hydrology, and environmental studies.

These generalizations highlight the ongoing efforts to create more flexible and adaptable models that can handle diverse data characteristics. By introducing new parameters and refining existing ones, researchers can develop models that provide better fits and more accurate predictions across various fields. Other useful models include [\[15,](#page-14-7) [16,](#page-14-8) [17,](#page-14-9) [18,](#page-14-10) [19,](#page--1-0) [20,](#page--1-1) [21,](#page--1-2) [22,](#page--1-3) [23,](#page--1-4) [24,](#page--1-5) [25,](#page--1-6) [26,](#page--1-7) [27,](#page--1-8) [28,](#page--1-9) [29,](#page--1-10) [31,](#page--1-11) [32,](#page--1-12) [33,](#page--1-13) [34,](#page--1-14) [35,](#page--1-15) [36,](#page--1-16) [37,](#page--1-17) [38,](#page--1-18) [39,](#page--1-19) [40,](#page--1-20) [41\]](#page--1-21).

The remainder of the work is organized as follows; section is on model specification. In section the properties are studied. In section , the parameters are estimated using maximum likelihood method. In numerical analysis are presented in section . The article is concluded in section .

2|Model Specification

Chris-Jerry distribution due to Onyekwere and Obulezi [\[37\]](#page--1-17) is a one-parameter lifetime model with probability density function (PDF) and cumulative distribution function (CDF) given respectively as

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

and

$$
Z(x,\theta) = 1 - \left[1 + \frac{\theta x (\theta x + 2)}{\theta + 2}\right] e^{-\theta x}, \quad \theta > 0, \quad x > 0
$$
 (2)

proposed a two-parameter Chris-Jerry distribution with p.d.f and CDF given as

$$
g(x,\theta,\lambda) = \frac{\theta^2}{\lambda\theta + 2} \left(\lambda + \theta x^2\right) e^{-\theta x}, \quad \theta > 0, \quad \lambda > 0, \quad x > 0
$$
 (3)

and

$$
G(x, \theta, \lambda) = 1 - \frac{1}{\lambda \theta + 2} \left(\theta^2 x^2 + 2\theta x + \theta \lambda + 2 \right) e^{-\theta x}
$$
 (4)

The PDF of the Generalized Weighted Chris-Jerry (GWCJ) distribution with parameters θ , λ , and β is given by

$$
f(x, \theta, \lambda, \beta) = \frac{\theta^{\beta+1}}{\theta\lambda + \beta(\beta+1)} \frac{x^{\beta-1}}{\Gamma(\beta)} \left(\lambda + \theta x^2\right) e^{-\theta x}, \quad \theta, \lambda, \beta > 0, \quad x > 0
$$
 (5)

where θ is the scale parameter, λ and β are the shape parameters

Proof: Tesfalem and Shanker [\[1\]](#page-12-2) provided a basis for this proof. Having

$$
f(x, \theta, \lambda, \beta) = kw(x, \theta, \lambda, \beta)g(x, \theta, \lambda)
$$
\n(6)

where *k* is the normalizing constant, $w(x, \beta) = x^{\beta-1}$ is the weight function and $g(x, \theta, \lambda)$ is the PDF of the two-parameter Chris-Jerry distribution. We require to prove that equation [\(6\)](#page-2-0) is a proper PDF by deriving the value of *k*. Hence, it is easy to see that for

$$
\int_0^\infty f(x,\theta,\lambda,\beta) dx = \int_0^\infty \frac{kx^{\beta-1}\theta^2}{\lambda\theta+2} (\lambda+\theta x^2) e^{-\theta x} dx = 1
$$
\n
$$
k = \frac{(\theta\lambda+2)\theta^{\beta}}{(\theta\lambda+\beta(\beta+1))\theta\Gamma(\beta)}
$$
\n(7)

Therefore, making appropriate substitutions in equation (6) completes the proof. \Box

Notice that if *X* ∼ *GWCJ* (*θ*, λ , β), the PDF which is in equation [\(5\)](#page-2-1) is a two-component mixture of *Gamma* (*θ*, β) and $Gamma(\theta, \beta + 2)$ with mixing proportion $\frac{\theta \lambda}{\theta \lambda + \beta(\beta + 1)}$ in the form

$$
pGamma(\theta, \beta) + (1 - p)Gamma(\theta, \beta + 2)
$$

The CDF of the GWCJ distribution with parameters θ , λ and β is given by

$$
F(x, \theta, \lambda, \beta) = 1 - \left[\frac{\Gamma(\beta, \theta x)}{\Gamma(\beta)} + \frac{1}{\theta \lambda + \beta (\beta + 1)} (1 + \beta + \theta x) \theta^{\beta} x^{\beta} e^{-\theta x} \right], \quad x, \theta, \lambda, \beta > 0
$$
 (8)

where $\Gamma(\beta, \theta x) = \int_{\theta x}^{\infty} y^{\beta - 1} e^{-y} dy$ is the upper incomplete gamma function.

Proof: We begin the proof with integrating the PDF of the proposed $GWCJ(x, \theta, \lambda, \beta)$ given in equation [\(5\)](#page-2-1).

$$
F(x,\theta,\lambda,\beta) = \int_0^x f(t,\theta,\lambda,\beta)dt = \frac{\theta^{\beta+1}}{[\theta\lambda + \beta(\beta+1)]} \left\{ \lambda \int_0^x x^{\beta-1} e^{-\theta x} dx + \theta \int_0^x x^{\beta+1} e^{-\theta x} \right\}
$$
(9)

define $y = \theta x$ then $dx = \frac{dy}{\theta}$, hence

$$
F(x,\theta,\lambda,\beta) = \frac{1}{\Gamma(\beta)(\theta\lambda + \beta(\beta+1))} \left\{ \theta\lambda \int_0^{\theta x} y^{\beta-1} e^{-y} dy + \int_0^{\theta x} y^{\beta+1} e^{-y} dy \right\} = \frac{\theta\lambda\gamma(\beta,\theta x) + \gamma(\beta+2,\theta x)}{\Gamma(\beta)(\theta\lambda + \beta(\beta+1))} \tag{10}
$$

Note: The following special derivations from the gamma function were employed;

$$
\Gamma(\beta+1,x) = \beta \Gamma(\beta,x) + x^{\beta} e^{-x}; \quad \Gamma(\beta) = \gamma(\beta,x) + \Gamma(\beta,x); \quad \Gamma(\beta+2,\theta x) = \beta(\beta+1)\Gamma(\beta,\theta x) + (\beta+1)(\theta x)^{\beta} e^{-\theta x} + (\theta x)^{\beta+1} e^{-\theta x}
$$

Making appropriate substitutions from the special derivations into equation (10), we complete the proof. \square

It is easy to define from equation [\(8\)](#page-2-3) the survival function of GWCJ distribution Let *X* ∼ *GW CJ*(*θ, λ, β*), the survival function is given as

$$
S(x, \theta, \lambda, \beta) = \frac{\Gamma(\beta, \theta x)}{\Gamma(\beta)} + \frac{1}{\theta \lambda + \beta (\beta + 1)} (1 + \beta + \theta x) \theta^{\beta} x^{\beta} e^{-\theta x}
$$
(11)

We use the survival function as a measure of system reliability. The hazard rate function which measures the likelihood that an item will survive to a certain point in time based on its survival to an earlier time *t*. Let $X \sim GWCJ(\theta, \lambda, \beta)$, the hazard rate function is given as

$$
hrf(x, \theta, \lambda, \beta) = \frac{\theta^{\beta+1} x^{\beta-1} \left(\lambda + \theta x^2\right) e^{-\theta x}}{\Gamma(\beta, \theta x) \left(\theta \lambda + \beta(\beta + 1)\right) + \Gamma(\beta) \left(1 + \beta + \theta x\right) \theta^{\beta} x^{\beta} e^{-\theta x}}
$$
(12)

Another very important measure is the reversed hazard rate function. It is the conditional probability of failures of an item in the next *dt* units of time given that it did not fail before *t*. It is essential in analyzing censored data and is commonly used in Forensic sciences (see Kayid [\[2\]](#page-12-3)). Let $X \sim GWCJ(\theta, \lambda, \beta)$, the reversed hazard rate function is given as

$$
r h r f(x, \theta, \lambda, \beta) = \frac{\frac{\theta^{\beta+1}}{\theta \lambda + \beta(\beta+1)} \frac{x^{\beta-1}}{\Gamma(\beta)} \left(\lambda + \theta x^2\right) e^{-\theta x}}{1 - \left[\frac{\Gamma(\beta, \theta x)}{\Gamma(\beta)} - \frac{1}{\theta \lambda + \beta(\beta+1)} \left(1 + \beta + \theta x\right) \theta^{\beta} x^{\beta} e^{-\theta x}\right]}
$$
(13)

The cumulative hazard rate function provides the total accumulated risk of experiencing the event of interest that has been gained by progressing to time t. While the instantaneous hazard rate $hrf(t)$ can increase or decrease with time, the cumulative hazard rate can only increase or remain the same. Let $X \sim GWCJ(\theta, \lambda, \beta)$, the cumulative hazard rate function is given as

$$
chr f(x, \theta, \lambda, \beta) = -\ln S(x, \theta, \lambda, \beta) = -\ln \left(\frac{\Gamma(\beta, \theta x)}{\Gamma(\beta)} + \frac{1}{\theta \lambda + \beta (\beta + 1)} \left(1 + \beta + \theta x \right) \theta^{\beta} x^{\beta} e^{-\theta x} \right) \tag{14}
$$

3|Mathematical Properties

In this section, we derive some important properties of the GWCJ distribution

Moment. Let $X \sim GWCJ(\theta, \lambda, \beta)$, the r^{th} non-central moment is given as

$$
\mu'_{r} = \frac{\Gamma(\beta + r)}{\Gamma(\beta)} \frac{\{\lambda \theta^{1-r} + \theta^{-r} (\beta + r + 1) (\beta + r)\}}{(\theta \lambda + \beta(\beta + 1))}; \quad r = 1, 2, ... \tag{15}
$$

Proof: The r^{th} non-central moment of a continuous distribution is given by

$$
\mu_r' = EX^r = \int_R x^r f(x) dx \tag{16}
$$

where $f(x)$ is the density function and in this case given in equation [\(5\)](#page-2-1). Therefore,

$$
\mu'_{r} = \int_{0}^{\infty} x^{r} \frac{\theta^{\beta+1}}{\theta\lambda + \beta(\beta+1)} \frac{x^{\beta-1}}{\Gamma(\beta)} (\lambda + \theta x^{2}) e^{-\theta x} dx \n= \frac{\theta^{\beta+1}}{\{\theta\lambda + \beta(\beta+1)\} \Gamma(\beta)} \left\{ \lambda \int_{0}^{\infty} x^{\beta+r-1} e^{-\theta x} dx + \theta \int_{0}^{\infty} x^{\beta+r+1} e^{-\theta x} dx \right\} \n= \frac{\theta^{\beta+1}}{\{\theta\lambda + \beta(\beta+1)\} \Gamma(\beta)} \left\{ \frac{\lambda \Gamma(\beta+r)}{\theta^{\beta+r}} + \frac{\theta \Gamma(\beta+r+2)}{\theta^{\beta+r+2}} \right\} \n= \frac{1}{\{\theta\lambda + \beta(\beta+1)\} \Gamma(\beta)} \left\{ \frac{\lambda \Gamma(\beta+r)}{\theta^{r-1}} + \frac{\Gamma(\beta+r+2)}{\theta^{r}} \right\}
$$
\n(17)

Figure 1. PDF and CDF of the GWCJ distribution

□

Let $X \sim GWCJ(\theta, \lambda, \beta)$, then the mean, 2^{nd} , 3^{rd} and 4^{th} non-central moments are respectively

$$
\mu = \frac{\theta\lambda\beta + \beta(\beta + 1)(\beta + 2)}{\theta^2\lambda + \theta\beta(\beta + 1)}
$$
\n(18)

$$
\mu_2' = \frac{\theta \lambda \beta (\beta + 1) + \beta (\beta + 1) (\beta + 2) (\beta + 3)}{\theta^3 \lambda + \theta^2 \beta (\beta + 1)}
$$
(19)

$$
\mu'_{3} = \frac{\theta\lambda\beta\left(\beta+1\right)\left(\beta+2\right)+\beta\left(\beta+1\right)\left(\beta+2\right)\left(\beta+3\right)\left(\beta+4\right)}{\theta^{4}\lambda+\theta^{3}\beta\left(\beta+1\right)}\tag{20}
$$

Figure 2. survival and hazard rate function of the GWCJ distribution

and

$$
\mu_4^{'} = \frac{\theta \lambda \beta \left(\beta + 1\right) \left(\beta + 2\right) \left(\beta + 3\right) + \beta \left(\beta + 1\right) \left(\beta + 2\right) \left(\beta + 3\right) \left(\beta + 4\right) \left(\beta + 5\right)}{\theta^5 \lambda + \theta^4 \beta \left(\beta + 1\right)}\tag{21}
$$

This follows easily from substituting 1*,* 2*,* 3 and 4 respectively for *r* in equation [\(15\)](#page-3-0).

[Variance of the GWCJ distribution] The variance of a random variance is a measure of spread which is generally denoted by $\sigma^2 = \mu_2' - (\mu_1)^2$. From a random variable *X* that is GWCJ distributed, the variance is given as

$$
\sigma^{2} = \frac{\theta\lambda\beta\left(\beta+1\right)+\beta\left(\beta+1\right)\left(\beta+2\right)\left(\beta+3\right)}{\theta^{3}\lambda+\theta^{2}\beta\left(\beta+1\right)} - \left(\frac{\theta\lambda\beta+\beta\left(\beta+1\right)\left(\beta+2\right)}{\theta^{2}\lambda+\theta\beta\left(\beta+1\right)}\right)^{2} \tag{22}
$$

[Kurtosis] Kurtosis is another measure that explains the peakness of a distribution. Suppose *X* ∼ GWCJ (θ, λ, β) , the kurtosis can be measured as follows;

$$
\rho = \frac{\mu_4^{'} - 4\mu\mu_3^{'} + 6(\mu)^2\mu_2^{'} - 4(\mu)^4}{\sigma^4} = \tag{23}
$$

$$
\frac{\theta\lambda\beta(\beta+1)(\beta+2)(\beta+3)+\beta(\beta+1)(\beta+2)(\beta+3)(\beta+4)(\beta+5)}{\theta^{5}\lambda+\theta^{4}\beta(\beta+1)} - 4\left(\frac{\theta\lambda\beta+\beta(\beta+1)(\beta+2)}{\theta^{2}\lambda+\theta\beta(\beta+1)}\right)\left(\frac{\theta\lambda\beta(\beta+1)(\beta+2)+\beta(\beta+1)(\beta+2)(\beta+3)(\beta+4)}{\theta^{4}\lambda+\theta^{3}\beta(\beta+1)}\right)}{(\theta^{4}\lambda+\theta^{3}\beta(\beta+1))} + 6\left(\frac{\theta\lambda\beta+\beta(\beta+1)(\beta+2)}{\theta^{2}\lambda+\theta\beta(\beta+1)}\right)^{2}\left(\frac{\theta\lambda\beta(\beta+1)+\beta(\beta+1)(\beta+2)(\beta+3)}{\theta^{3}\lambda+\theta^{2}\beta(\beta+1)}\right) - 4\left(\frac{\theta\lambda\beta+\beta(\beta+1)(\beta+2)}{\theta^{2}\lambda+\theta\beta(\beta+1)}\right)^{4}\right)
$$
\n(24)

[Skewness] Skewness is an important measure that indicates where the weight of the distribution is. Given *X* ∼ $\text{GWCJ}(\theta, \lambda, \beta)$, the measure of skewness is expressed as;

$$
\kappa = \frac{\mu_{3}^{'} - 3\mu_{2}^{'} + 2(\mu)^{2}}{\sigma^{\frac{3}{2}}}
$$
\n
$$
= \frac{\frac{\theta\lambda\beta(\beta+1)(\beta+2)+\beta(\beta+1)(\beta+2)(\beta+3)(\beta+4)}{\theta^{4}\lambda+\theta^{3}\beta(\beta+1)}}{2\pi\sigma^{\frac{3}{2}}}
$$
\n
$$
= \frac{2\theta\lambda\beta(\beta+1)(\beta+2)(\beta+3)(\beta+4)}{\theta^{2}\lambda+\theta\beta(\beta+1)} - 3\frac{\theta\lambda\beta+\beta(\beta+1)(\beta+2)}{\theta^{2}\lambda+\theta\beta(\beta+1)}\frac{\theta\lambda\beta(\beta+1)+\beta(\beta+1)(\beta+2)(\beta+3)}{\theta^{3}\lambda+\theta^{2}\beta(\beta+1)} + 2\left(\frac{\theta\lambda\beta+\beta(\beta+1)(\beta+2)}{\theta^{2}\lambda+\theta\beta(\beta+1)}\right)^{2}}{2\pi\sigma^{\frac{3}{2}}} \tag{25}
$$

3.1|Odd Function

An odd function is a reliability tool for modeling a data set that shows a non-monotone hazard rate. It is defined to be the ratio of the CDF to the Survival function

$$
O_{GWCJ}(x; \theta, \lambda, \beta) = \frac{F_{GWCJ}(x; \theta, \lambda, \beta)}{S_{GWCJ}(x; \theta, \lambda, \beta)}
$$
(26)

$$
O_{GWCJ}(x; \theta, \lambda, \beta) = \frac{1 - \left[\frac{\Gamma(\beta, \theta x)}{\Gamma(\beta)} - \frac{1}{\theta \lambda + \beta(\beta + 1)} \left(1 + \beta + \theta x\right) \theta^{\beta} x^{\beta} e^{-\theta x}\right]}{\frac{\Gamma(\beta, \theta x)}{\Gamma(\beta)} + \frac{1}{\theta \lambda + \beta(\beta + 1)} \left(1 + \beta + \theta x\right) \theta^{\beta} x^{\beta} e^{-\theta x}} = \frac{\Gamma(\beta)}{\Gamma(\beta, \theta x)} + \frac{\theta \lambda + \beta(\beta + 1)}{1 + \beta + \theta x} \theta^{-\beta} x^{-\beta} e^{\theta x} - 1
$$
\n(27)

3.2|Index of Dispersion

Index of dispersion is a normalized measure of the dispersion of a probability distribution which is used to quantify whether a set of observed occurrences are clustered or dispersed compared to a standard statistical model.

$$
\xi = \frac{\sigma^2}{\mu} = \left[\left(\frac{\theta \lambda \beta \left(\beta + 1\right) + \beta \left(\beta + 1\right) \left(\beta + 2\right) \left(\beta + 3\right)}{\theta^3 \lambda + \theta^2 \beta \left(\beta + 1\right)} - \left(\frac{\theta \lambda \beta + \beta \left(\beta + 1\right) \left(\beta + 2\right)}{\theta^2 \lambda + \theta \beta \left(\beta + 1\right)} \right)^2 \right) \div \frac{\theta \lambda \beta + \beta \left(\beta + 1\right) \left(\beta + 2\right)}{\theta^2 \lambda + \theta \beta \left(\beta + 1\right)} \right] \tag{28}
$$

3.3|Distribution of the Order Statistics

Given an ordered sample $X_{(1)}, X_{(2)}, \cdots, X_{(n)}$ from a GWCJ distribution. The distribution of the statistics is

$$
f_{r:n}(x; \theta, \lambda, \beta) = \frac{n!}{(r-1)!(n-r)!} f_{GWCJ}(x; \theta, \lambda, \beta) \left[F_{GWCJ}(x; \theta, \lambda, \beta) \right]^{r-1} \left[1 - F_{GWCJ}(x; \theta, \lambda, \beta) \right]^{n-r}
$$
(29)

Figure 3. Mean of GWCJ distribution

Figure 4. Variance of GWCJ distribution

Figure 5. Skewness of GWCJ distribution

Figure 6. Kurtosis of GWCJ distribution

where $f_{GWCJ}(x; \theta, \lambda, \beta)$ and $F_{GWCJ}(x; \theta, \lambda, \beta)$ are the pdf and cdf of GWCJ distribution respectively. Hence, we have

 $f_{r:n}(x; \theta, \lambda, \beta) =$

$$
\frac{n!}{(r-1)!(n-r)!} \frac{\theta^{\beta+1}}{\theta\lambda+\beta(\beta+1)} \frac{x^{\beta-1}}{\Gamma(\beta)} (\lambda+\theta x^2) e^{-\theta x}
$$

$$
\left\{ 1 - \left[\frac{\Gamma(\beta,\theta x)}{\Gamma(\beta)} - \frac{1}{\theta\lambda+\beta(\beta+1)} (1+\beta+\theta x) \theta^{\beta} x^{\beta} e^{-\theta x} \right] \right\}^{r-1}
$$

$$
\left\{ \frac{\Gamma(\beta,\theta x)}{\Gamma(\beta)} - \frac{1}{\theta\lambda+\beta(\beta+1)} (1+\beta+\theta x) \theta^{\beta} x^{\beta} e^{-\theta x} \right\}^{n-r}
$$
(30)

The PDF of the largest order statistics is gotten by setting $r = n$

$$
\frac{n\theta^{\beta+1}}{\theta\lambda+\beta(\beta+1)}\frac{x^{\beta-1}}{\Gamma(\beta)}\left(\lambda+\theta x^2\right)e^{-\theta x}\left\{1-\left[\frac{\Gamma(\beta,\theta x)}{\Gamma(\beta)}-\frac{1}{\theta\lambda+\beta(\beta+1)}\left(1+\beta+\theta x\right)\theta^{\beta}x^{\beta}e^{-\theta x}\right]\right\}^{n-1}\tag{31}
$$

The PDF of the smallest order statistics is gotten by setting $r = 1$

$$
\frac{n\theta^{\beta+1}}{\theta\lambda+\beta(\beta+1)}\frac{x^{\beta-1}}{\Gamma(\beta)}\left(\lambda+\theta x^2\right)e^{-\theta x}\left\{\frac{\Gamma(\beta,\theta x)}{\Gamma(\beta)}-\frac{1}{\theta\lambda+\beta(\beta+1)}\left(1+\beta+\theta x\right)\theta^{\beta}x^{\beta}e^{-\theta x}\right\}^{n-1}\tag{32}
$$

3.4|Generating Functions

An endless series of numbers can be encoded using a generating function in mathematics by treating them as the coefficients of a formal power series.

3.4.1|Moment Generating Function

The moment-generating function of a real-valued random variable is an alternative specification of its probability distribution in probability theory and statistics. Instead of using probability density function or cumulative distribution functions directly, it provides the foundation for an alternative path to analytical solutions.

The moment generating function of a *X* ~ GWCJ (θ, λ, β) is given by

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

\n
$$
= \frac{\theta^{\beta+1}}{\theta\lambda + \beta(\beta+1)\Gamma(\beta)} \int_0^\infty e^{tx} x^{\beta-1} (\lambda + \theta x^2) e^{-\theta x} dx
$$

\n
$$
= \frac{\theta^{\beta+1}}{\theta\lambda + \beta(\beta+1)\Gamma(\beta)} \int_0^\infty x^{\beta-1} (\lambda + \theta x^2) e^{-(\theta-t)x} dx
$$

\n
$$
= \frac{\theta^{\beta+1}}{\theta\lambda + \beta(\beta+1)\Gamma(\beta)} \left[\lambda \int_0^\infty x^{\beta-1} e^{-(\theta-t)x} dx + \theta \int_0^\infty x^{\beta+1} e^{-(\theta-t)x} dx \right]
$$

\n
$$
= \frac{\theta^{\beta+1}}{\theta\lambda + \beta(\beta+1)\Gamma(\beta)} \left[\frac{\lambda \Gamma(\beta)}{(\theta-t)^{\beta}} + \frac{\theta \Gamma(\beta+2)}{(\theta-t)^{\beta+2}} \right]
$$

\n(33)

3.4.2|Characteristic Generating Function

In probability theory and statistics, the characteristic function of any real-valued random variable completely defines its probability distribution. Unlike the moment-generating function, the characteristic function always exists when treated as a function of a real-valued argument.

The characteristic function of a *X* ~ GWCJ(θ , λ , β) is given by

$$
\phi_X(it) = \frac{\theta^{\beta+1}}{\theta\lambda + \beta(\beta+1)\Gamma(\beta)} \left[\frac{\lambda\Gamma(\beta)}{(\theta-it)^{\beta}} + \frac{\theta\Gamma(\beta+2)}{(\theta-it)^{\beta+2}} \right]
$$
(34)

4|Maximum Likelihood Estimation of the Parameters

Let $(x_1, x_2, ..., x_n)$ be a random sample of size *n* drawn from GWCJ distribution, then the likelihood function is given as

$$
L(f_{GWCJ}(x; \theta, \lambda, \beta)) = \prod_{i=1}^{n} \frac{\theta^{\beta+1}}{\theta\lambda + \beta(\beta+1)} \frac{x_i^{\beta-1}}{\Gamma(\beta)} (\lambda + \theta x_i^2) e^{-\theta x_i} = \frac{\theta^{n(\beta+1)} e^{-\theta \sum x_i}}{(\theta\lambda + \beta(\beta+1))^n \Gamma(\beta)^n} \prod_{i=1}^{n} x_i^{\beta-1} (\lambda + \theta x_i^2)
$$
\n(35)

Taking the natural log of $L(.)$ and set $log L\phi$, we have

$$
\phi = n(\beta + 1)\log \theta - \theta \sum x_i - n\log \Gamma(\beta) - n\log(\theta \lambda + \beta(\beta + 1)) + \sum \log x_i^{\beta - 1} + \sum \log(\lambda + \theta x_i^2) \tag{36}
$$

Differentiating partially with respect to θ , λ and β yields the following;

$$
\frac{\partial \phi}{\partial \theta} = \frac{n(\beta + 1)}{\theta} - \sum_{i=1}^{n} x_i - \frac{n\lambda}{\theta \lambda + \beta(\beta + 1)} + \sum_{i=1}^{n} \frac{x_i^2}{\lambda + \theta x_i^2}
$$

$$
\frac{\partial \phi}{\partial \lambda} = -\frac{n\theta}{\theta \lambda + \beta(\beta + 1)} + \sum_{i=1}^{n} \frac{1}{\lambda + \theta x_i^2}
$$

$$
\frac{\partial \phi}{\partial \beta} = n \log \theta - \frac{n \Gamma'(\beta)}{\Gamma(\beta)} - \frac{n(2\beta + 1)}{\theta \lambda + \beta(\beta + 1)} + \sum_{i=1}^{n} \log x_i
$$

(37)

The system of non-linear equations in [\(37\)](#page-9-0) do not possess analytical solution hence **optim()** function in R will be used to obtain the solution, see [\[42\]](#page--1-22).

5|Numerical Analysis

This section is used to illustrate the applicability of the proposed distribution to life time data. Data-I is the remission times (in months) of a sample of 128 bladder cancer patients documented by Lee and Wang [\[5\]](#page-12-4), see table [1.](#page-10-0) Data-II and III consist of the mortality rate of children in Japan and Ireland under five years of age from 1969 to 2021 respectively. The data are presented in table [2](#page-10-1) and [3](#page-10-2) which were obtained https://data.worldbank.org/indicator/SP.DYN.IMRT.IN https://data.worldbank.org/indicator/SP.DYN.IMRT.IN (accessed on 21 July 2024). We provide the model comparison using the following measures; log-likelihood (LL) statistic, Akaike Information Criterion (AIC), Corrected Akaike Information Criterion (CAIC), Bayesian Information Criterion (BIC) and the Hannan Quinn Information Criterion (HQIC). Further, we provide the fitness of the model to the data sets using the Cramér von Mises (W) statistics, Anderson-Darling (A) statistics, Kolmogorov-Smirnov (KS) statistics with its p-value. From the results in table [4,](#page-10-3) the GWCJ distribution fits the data better than the rest of the distributions compared and also performed adequately in estimating the parameters. However, from the bloated value of $\hat{\lambda}_{MLE}$, there is indication of poor estimation. In table [5,](#page-10-4) the GWCJ competes favourably while it is better than the competitors in table [6.](#page-11-0) Figures [7,](#page-11-1) [8](#page-12-5) and [9](#page-14-11) represent the histogram, CDF, survival function, TTT plot, PP and QQ plots of the three data sets respectively. They visually tell how well the distribution fits the associated data sets.

6|Conclusion

Generalization of distributions has inadvertently proved useful in the statistical literature especially in shaping the decision on the choice of the number of parameters and weighting a random variable. In this article, we generalized the length-biased Chris-Jerry distribution proposed by Subramanian and Subhashree [\[6\]](#page-14-12) and the weighted Chris-Jerry distribution by Praseeja et al. [\[7\]](#page-14-13). The functional form of the distribution including the

Table 1. Data-I The remission times (in months) of a sample of 128 bladder cancer patients

Table 2. Data-II The mortality rate of children in Japan under five years of age from 1969 to 2021

	39.7 36.2 32.9 29.8 27.0 24.6 22.7 21.0 19.6 18.5 17.5 16.5 15.7 14.8 14.0 13.3 12.5							
	11.8 11.1 10.5 9.9 9.3 8.8 8.3 7.9 7.5 7.1 6.8 6.6 6.5 6.3 6.2 6.1 6.0							
	5.9 5.7 5.5 5.2 5.0 4.7 4.5 4.3 4.1 4.0 3.9 3.7 3.6 3.5 3.4 3.3 3.2							
	3.2 3.0 2.9 2.8 2.8 2.7 2.6 2.5 2.5 2.4 2.3							

Table 3. Data-III The mortality rate of children in Ireland under five years of age from 1969 to 2021

							35.3 33.8 32.5 31.3 30.0 28.6 27.1 25.5 24.1 23.0 22.2 21.7 21.4 21.1 20.6 20.0 19.1	
							18.0 16.8 15.6 14.3 13.2 12.4 11.7 11.1 10.7 10.4 10.1 9.9 9.6 9.1 8.6 8.1 7.7	
							7.4 7.3 7.3 7.3 7.3 7.3 7.1 6.9 6.5 6.0 5.6 5.2 4.9 4.6 4.4 4.3 4.2	
	4.1 4.0 3.9 3.8 3.7 3.6 3.5 3.4 3.3 3.2 3.1							

Table 4. Measures of Model Adequacy using Data-I

Dist		AIC-	CAIC	BIC	HQIC	W A	KS	p-value	$\theta_{\rm MLE}$	$\lambda_{\rm MLE}$	$\beta_{\rm MLE}$
GWCJ	-412.8			831.594 831.788 840.150 835.071 0.104 0.625 0.066				0.626	0.142	-307.490	1.211
Perks	-414.32	832.634 832.730			838.338 834.952 0.125 0.747 0.079			0.395	17.923 0.110		
		Gamma -413.36 830.727 830.823 836.431 833.044 0.119 0.716 0.073 0.498							1.172	0.125	
		Weibull -414.08 832.159 832.255 837.863 834.477 0.131 0.782 0.070						0.559	9.558	1.047	
	Gumbel 1-432.25	868.498 868.594 874.202 870.816					0.446 2.753 0.112	0.081	5.644	5.422	

Table 5. Measures of Model Adequacy using Data-II

PDF, CDF, Survival function and hazard were presented. The graphical illustration of these functions were also provided. The characteristics of the distribution were studied. The parameters of the model were estimated using the maximum likelihood estimation procedure. To illustrate the importance of the model, three lifetime data sets were used and the model proved better than some classical models.

Dist	LL	AIC		CAIC BIC HQIC W A KS p-value $\hat{\theta}_{MLE}$ $\hat{\lambda}_{MLE}$ $\hat{\beta}_{MLE}$				
GWCJ –				85.91 -165.830 -165.416 -159.449 -163.324 0.263 1.630 0.133 0.222 20.390 1.528				- 1.757
				Perks 81.99 -159.982 -159.778 -155.728 -158.312 0.330 1.998 0.171 0.054 1.389 13.189				
				Gamma 85.39 -166.771 -166.567 -162.517 -165.100 0.276 1.702 0.142 0.167 1.672 16.825				
				Weibull 83.68 -163.368 -163.164 -159.113 -161.697 0.316 1.924 0.135 0.208			0.108 1.250	
				Gumbel 78.11 -152.213 -152.213 -147.959 -150.543 0.413 2.455 0.164 0.071			0.063 0.053	

Table 6. Measures of Model Adequacy using Data-III

Figure 7. Plots for Data-I

Figure 9. Plots for Data-III

Acknowledgments

The authors would like to express their sincere gratitude to the editors and anonymous reviewers for their invaluable comments and constructive feedback, which significantly contributed to the enhancement of this paper.

Author Contribution

Ibeakuzie, Igbokwe & Ekemezie: methodology, software, and editing. Anabike: conceptualization. Nwankwo & Chinedu: writing and editing. All authors have read and agreed to the published version of the manuscript.

Funding

The authors declare that no external funding or support was received for the research presented in this paper, including administrative, technical, or in-kind contributions. Data Availability All data supporting the reported findings in this research paper are provided within the manuscript.

Conflicts of Interest

The authors declare that there is no conflict of interest concerning the reported research findings. Funders played no role in the study's design, in the collection, analysis, or interpretation of the data, in the writing of the manuscript, or in the decision to publish the results.

References

- [1] Tesfalem, E., Shanker, R. (2019). A Two-Parameter Weighted Rama Distribution with Properties and Application. *Universal Journal of Mathematics and Applications*, 2(2), 48–58.
- [2] Kayid, M., Al-Nahawati, H., Ahmad, I. A. (2011). Testing behavior of the reversed hazard rate. *Applied Mathematical Modelling*, 35(5), 2508–2515. Elsevier.
- [3] Kharazmi, O., Mahdavi, A., Fathizadeh, M. (2015). Generalized weighted exponential distribution. *Communications in Statistics-Simulation and Computation*, 44(6), 1557–1569. Taylor Francis.
- [4] Ramos, P. L., Louzada, F. (2016). The generalized weighted Lindley distribution: Properties, estimation, and applications. *Cogent Mathematics*, 3(1), 1256022. Taylor Francis.
- [5] Lee, E. T., Wang, J. (2003). Statistical methods for survival data analysis. (Vol. 476). John Wiley Sons.
- [6] Subramanian, C., Subhashree, M. (2024). A new extension of Chris-Jerry distribution with its properties. *Journal of Applied Mathematics, Statistics and Informatics*, 20(1), 5–16.
- [7] Praseeja, C. B., Prasanth, C. B., Subramanian, C., Unnikrishnan, T. (2024). Characteristics of Srimin-H distribution and its biomedical application. *JP Journal of Biostatistics*, 24(1), 47–60.
- [8] Abbas, S., Ozal, G., Shahbaz, S. H., Shahbaz, M. Q. (2019). A new generalized weighted Weibull distribution. *Pakistan Journal of Statistics and Operation Research*, 161–178. College of Statistical and Actuarial Sciences.
- [9] Kharazmi, O. (2016). Generalized weighted Weibull distribution. *Journal of Mathematical Extension*, 10, 89–118.
- [10] Beaulieu, N. C. (1991). On the generalized multinomial distribution, optimal multinomial detectors, and generalized weighted partial decision detectors. *IEEE Transactions on Communications*, 39(2), 193–194. IEEE.
- [11] Domma, F., Condino, F., Popović, B. V. (2017). A new generalized weighted Weibull distribution with decreasing, increasing, upside-down bathtub, N-shape and M-shape hazard rate. *Journal of Applied Statistics*, 44(16), 2978–2993. Taylor Francis.
- [12] Acitas, S. (2018). A new weighted distribution as an extension of the generalized half-normal distribution with applications. *Journal of Statistical Computation and Simulation*, 88(12), 2325–2341. Taylor Francis.
- [13] Teamah, A. E. A. M., El-Damrawy, H. H., Swan, S. M. T. (2020). Generalized weighted exponential-gompertez distribution. *Applied Mathematics*, 11(02), 97. Scientific Research Publishing.
- [14] Rasmussen, P. F. (2001). Generalized probability weighted moments: application to the generalized Pareto distribution. *Water Resources Research*, 37(6), 1745–1751. Wiley Online Library.
- [15] Obulezi, O. J., Nwankwo, B. C., Nwankwo, M. P., Anabike, I. C., Igbokwe, C. P., & Igwebudu, C. N. (2023). Modeling Mortality Rate of COVID-19 Patients in the United Kingdom, Canada, and the Survival Rate of COVID-19 Patients in Spain. *Journal of Xidian University*, 17(11), 520–538.
- [16] Ekemezie, D. F. N., & Obulezi, O. J. (2024). The Fav-Jerry Distribution: Another Member in the Lindley Class with Applications. *Earthline Journal of Mathematical Sciences*, 14(4), 793–816.
- [17] Oha, F. C., Etaga, H. O., Ibeakuzie, P. O., Anabike, I. C., & Obulezi, O. J. (2024). Power XShanker Distribution: Properties, Estimation, and Applications. *Eng OA*, 2(1), 01–20.
- [18] Nwankwo, B. C., Obiora-Ilouno, H. O., Almulhim, F. A., SidAhmed Mustafa, M., & Obulezi, O. J. (2024). Group acceptance sampling plans for type-I heavy-tailed exponential distribution based on truncated life tests. *AIP Advances*, 14(3). AIP Publishing.
- [19] Etaga, H. O., Onyekwere, C. K., Oramulu, D. O., & Obulezi, O. J. (2023). A New Modification of Rani Distribution with More Flexibility in Application. *Sch J Phys Math Stat*, 7, 160–176.
- [20] Etaga, H. O., Nwankwo, M. P., Oramulu, D. O., Anabike, I. C., & Obulezi, O. J. (2023). The Double XLindley Distribution: Properties and Applications. *Sch J Phys Math Stat*, 10, 192–202.
- [21] Etaga, H. O., Onyekwere, C. K., Lydia, O. I., Nwankwo, M. P., Oramulu, D. O., & Obulezi, O. J. (2023). Estimation of the Xrama distribution parameter under complete and progressive type-II censored schemes. *Sch J Phys Math Stat*, 10, 203–219.
- [22] Chukwuma, P. O., Harrison, E. O., Ibeakuzie, P., Anabike, I. C., & Obulezi, O. J. (2024). A new reduced quantile function for generating families of distributions. *Annals of Mathematics and Physics*, 7(1), 001–015.
- [23] Nwankwo, B. C., Orjiakoh, J. N., Nwankwo, M. P., Chukwu, E. I. M. I., & Obulezi, O. J. (2024). A new distribution for modeling both blood cancer data and median effective dose (ED50) of Artemether-Lumefantrine against P. falciparum. *Earthline Journal of Mathematical Sciences*, 14(1), 41–62.
- [24] Nwankwo, M. P., Nwankwo, B. C., & Obulezi, O. J. (2023). The Exponentiated Power Akash Distribution: Properties, Regression, and Applications to Infant Mortality Rate and COVID-19 Patients' Life Cycle. *Annals of Biostatistics and Biometric Applications*, 5(4), 1–12.
- [25] Obulezi, O. J., Ujunwa, O. E., Anabike, I. C., & Igbokwe, C. P. (2023). The Exponentiated Power Chris-Jerry Distribution: Properties, Regression, Simulation and Applications to Infant Mortality Rate and Lifetime of COVID-19 Patients. *TWIST*, 18(4), 328–337.
- [26] Etaga, H. O., Nwankwo, M. P., Oramulu, D. O., & Obulezi, O. J. (2023). An improved XShanker distribution with applications to rainfall and vinyl chloride data. *Sch J Eng Tech*, 9, 212–224.
- [27] Onuoha, H. C., Osuji, G. A., Etaga, H. O., & Obulezi, O. J. (2023). The Weibull distribution with estimable shift parameter. *Earthline Journal of Mathematical Sciences*, 13(1), 183–208.
- [28] Oramulu, D. O., Etaga, H. O., Onuorah, A. J., & Obulezi, O. J. (2023). A new member in the Lindley class of distributions with flexible applications. *Sch J Phys Math Stat*, 7, 148–159.
- [29] Etaga, H. O., Celestine, E. C., Onyekwere, C. K., Omeje, I. L., Nwankwo, M. P., Oramulu, D. O., & Obulezi, O. J. (2023). A new modification of Shanker distribution with applications to increasing failure rate data. *Earthline Journal of Mathematical Sciences*, 13(2), 509–526.
- [30] Obulezi, O. J., Anabike, I. C., Okoye, G. C., Igbokwe, C. P., Etaga, H. O., & Onyekwere, C. K. (2023). The Kumaraswamy Chris-Jerry distribution and its applications. *Journal of Xidian University*, 17(6), 575–591.
- [31] Oramulu, D. O., Igbokwe, C. P., Anabike, I. C., Etaga, H. O., & Obulezi, O. J. (2023). Simulation study of the Bayesian and non-Bayesian estimation of a new lifetime distribution parameters with increasing hazard rate. *Asian Research Journal of Mathematics*, 19(9), 183–211.
- [32] Chinedu, E. Q., Chukwudum, Q. C., Alsadat, N., Obulezi, O. J., Almetwally, E. M., & Tolba, A. H. (2023). New lifetime distribution with applications to single acceptance sampling plan and scenarios of increasing hazard rates. *Symmetry*, 15(10), 1881.
- [33] Tolba, A. H., Onyekwere, C. K., El-Saeed, A. R., Alsadat, N., Alohali, H., & Obulezi, O. J. (2023). A New Distribution for Modeling Data with Increasing Hazard Rate: A Case of COVID-19 Pandemic and Vinyl Chloride Data. *Sustainability*, 15(17), 12782.
- [34] Onyekwere, C. K., Okoro, C. N., Obulezi, O. J., Udofia, E. M., & Anabike, I. C. (2022). Modification of Shanker distribution using quadratic rank transmutation map. *Journal of Xidian University*, 16(8), 179–198.
- [35] Musa, A., Onyeagu, S. I., & Obulezi, O. J. (2023). Exponentiated Power Lindley-Logarithmic distribution and its applications. *Asian Research Journal of Mathematics*, 19(8), 47–60.
- [36] Abuh, M., Onyeagu, S. I., & Obulezi, O. (2023). Comparative Study Based on Simulation of Some Methods of Classical Estimation of the Parameters of Exponentiated Lindley–Logarithmic Distribution. *Asian Journal of Probability and Statistics*, 22(4), 14–30.
- [37] Onyekwere, C. K., & Obulezi, O. J. (2022). Chris-Jerry distribution and its applications. *Asian Journal of Probability and Statistics*, 20(1), 16–30.
- [38] Obi, C. D., Chukwuma, P. O., Igbokwe, C. P., Ibeakuzie, P. O., & Anabike, I. C. (2024). A Novel Extension of Rayleigh Distribution: Characterization, Estimation, Simulations and Applications. *Journal of Xidian University*, 18(7), 177–188.
- [39] Anabike, I. C., Okoro, E. S., Igbokwe, C. P., Ibeakuzie, P. O., & Ekemezie, D. F. (2024). New Extended Rayleigh Distribution: Properties, Estimation, Simulation and Application. *Journal of Xidian University*, 18(7), 179–207.
- [40] Anabike, I. C., Igbokwe, C. P., Onyekwere, C. K., Obulezi, O. J. (2023). Inference on the parameters of Zubair-Exponential distribution with application to survival times of Guinea Pigs. *Journal of Advances in Mathematics and Computer Science*, 38(7), 12–35.
- [41] Obulezi, O. J., Anabike, I. C., Oyo, O. G., Igbokwe, C., Etaga, H. (2023). Marshall-Olkin Chris-Jerry distribution and its applications. *International Journal of Innovative Science and Research Technology*, 8(5), 522–533.
- [42] Team, R. C. (2020). R: A language and environment for statistical computing. *R Foundation for Statistical Computing*.