

A Weibull-Gompertz Makeham Distribution with Properties and Application to Cancer

Data

Abstract

The article presents an extension of the Gompertz Makeham distribution using the Weibull-G family of continuous probability distributions proposed by Tahir *et al.* (2016a). This new extension generates a more flexible model called Weibull-Gompertz Makeham distribution. Some statistical properties of the distribution which include the moments, survival function, hazard function and distribution of order statistics were derived and discussed. The parameters were estimated by the method of maximum likelihood and the distribution was applied to a bladder cancer data. Weibull-Gompertz Makeham distribution performed best (AIC = -6.8677, CAIC = -6.3759, BIC = 7.3924) when compared with other existing distributions of the same family to model bladder cancer data.

Keywords: Gompertz-Makeham Distribution, Weibull-Gompertz Makeham Distribution, hazard function, survival function, cancer.

1 Introduction

The Gompertz-Makeham distribution (GMD) was introduced by Makeham in 1860 (Makeham (1860)). It is an extended model of the Gompertz probability distribution that was introduced by Gompertz in 1825 (Gompertz (1825)). The GMD is a continuous probability distribution that has been widely used in survival analysis, modelling human mortality, constructing actuarial tables and growth models. It has been recently used in many fields of sciences including actuaries, biology, demography, gerontology, and computer science.

A comprehensive review of the history and theory of the GMD can be found in Marshall and Olkin (2007). Golubev (2004) emphasizes the practical importance of this probability distribution. Detailed information about the GM distribution, its mathematical and statistical properties, and its applications can be found in Johnson et al. (1995) and Dey et al. (2018).

The probability density function (pdf) and cumulative distribution function (cdf) of the Gompertz-Makeham distribution are defined as:

$$G(x) = 1 - e^{-\theta x - \frac{\alpha}{\beta}(e^{\beta x} - 1)} \tag{1}$$

and

$$g(x) = \left[\theta + \alpha e^{\beta x} \right] e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \quad (2)$$

37 respectively.

38 For $x > 0, \alpha, \beta, \theta > 0$, where θ is the scale parameter and α, β are the shape parameters of
 39 Gompertz-Makeham distribution.

40 There are families of distributions proposed by different researchers that are used in extending
 41 other distributions to produce compound distributions with better performance. These families
 42 among others include the beta generalized family (Beta-G) by Eugene et al. (2002), the
 43 Kumaraswamy-G by Cordeiro and de Castro (2011), Transmuted family of distributions by
 44 Shaw and Buckley (2007), Exponentiated T-X by Alzaghal et al. (2013), Exponentiated-G (EG)
 45 by Cordeiro et al. (2013), Logistic-G by Torabi and Montazari (2014), Logistic-X by Tahir et al.
 46 (2016b), Weibull-X by Alzaatreh et al. (2013), Weibull-G by Bourguignon et al. (2014), a new
 47 Weibull-G family by Tahir et al. (2016a), a Lomax-G family by Cordeiro et al. (2014), a new
 48 generalized Weibull-G family by Cordeiro et al. (2015) and Beta Marshall-Olkin family of
 49 distributions by Alizadeh et al. (2015) etc.

50 Recently, many authors have extended the Gompertz-Makeham distribution. Chukwu and
 51 Ogunde (2016) introduced and studied the Kumaraswamy Gompertz Makeham distribution. El-
 52 Bar (2017) used the quadratic rank transmutation map by Shaw and Buckley (2007) to defined
 53 and study the transmuted Gompertz Makeham distribution with useful discussions as well as
 54 applications.

55 Hence, this article aims to introduce another extension of the Gompertz Makeham model, a new
 56 continuous distribution called Weibull-Gompertz Makeham distribution (WGMD) from the
 57 proposed family by Tahir et al. (2016a). The rest of this article is arranged as follows: the
 58 definition of the new distribution will be presented with its plots and some properties. These are
 59 followed by the reliability functions, the order statistics for the distribution and the maximum
 60 likelihood estimates (MLEs) of the unknown parameters. The last part involves the application of
 61 the proposed model with other models to a lifetime dataset and the conclusion.

62

63

64 **2 Materials and Methods**

65 **2.1 Construction of Weibull-Gompertz Makeham Distribution (WGMD)**

66 This section defines the cdf and pdf of the Weibull-Gompertz Makeham distribution (WGMD)
 67 using the family of distributions proposed by Tahir et al. (2016a), which has been used by other
 68 authors including Ieren et al. (2018). According to Tahir et al. (2016a), the function for defining
 69 the cdf and pdf of any Weibull-based continuous distribution is given as:

70
$$F(x) = \int_0^{-\log[G(x)]} abt^{b-1} e^{-at} dt = e^{-a\{-\log[G(x)]\}^b} \quad (3)$$

71 and

72
$$f(x) = ab \frac{g(x)}{G(x)} \{-\log[G(x)]\}^{b-1} e^{-a\{-\log[G(x)]\}^b} \quad (4)$$

73 respectively, where $g(x)$ and $G(x)$ are the pdf and cdf of any continuous distribution to be
 74 generalized respectively. The parameters, α and β are the two additional new parameters
 75 responsible for the scale and shape of the distribution respectively.

76 Using equation (1) and (2) in (3) and (4) and simplifying, the cdf and pdf of the WGMD of a
 77 random variable X can be obtained as:

78
$$F(x) = e^{-a\left\{-\log\left[1 - e^{-\theta x - \frac{\alpha}{\beta}(e^{\beta x} - 1)}\right]\right\}^b} \quad (5)$$

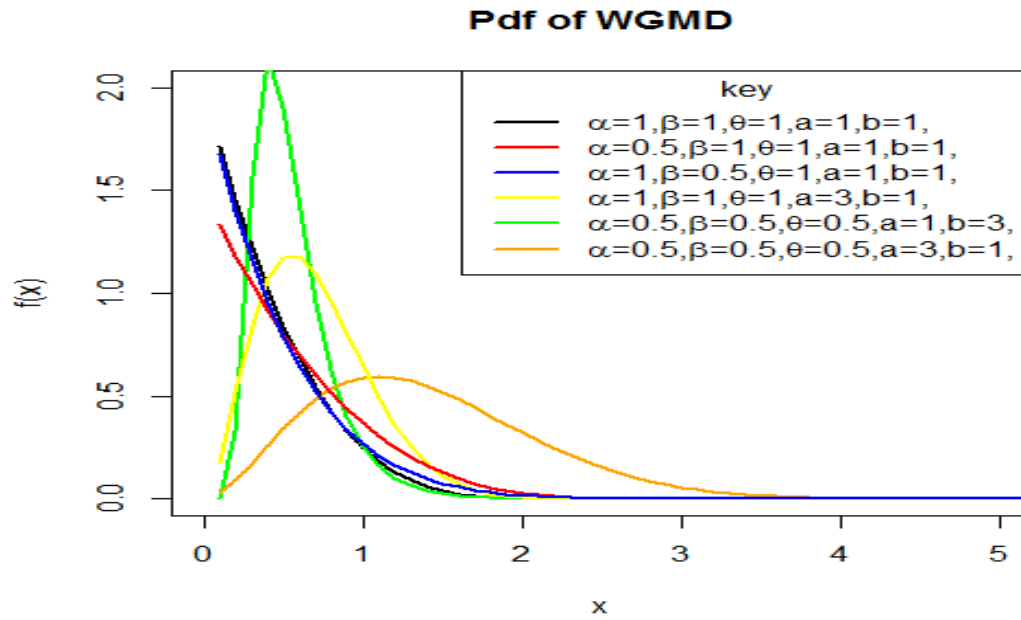
79 and

80
$$f(x) = \frac{ab\left(\theta + \alpha e^{\beta x}\right) e^{-\theta x - \frac{\alpha}{\beta}(e^{\beta x} - 1)} \left(-\log\left[1 - e^{-\theta x - \frac{\alpha}{\beta}(e^{\beta x} - 1)}\right]\right)^{b-1}}{\left[1 - e^{-\theta x - \frac{\alpha}{\beta}(e^{\beta x} - 1)}\right] e^{a\left(-\log\left[1 - e^{-\theta x - \frac{\alpha}{\beta}(e^{\beta x} - 1)}\right]\right)^b}} \quad (6)$$

81 respectively.

82 For $x > 0; a, b, \theta, \alpha, \beta > 0$; where a, b, θ, α and β are the parameters of the WGMD.

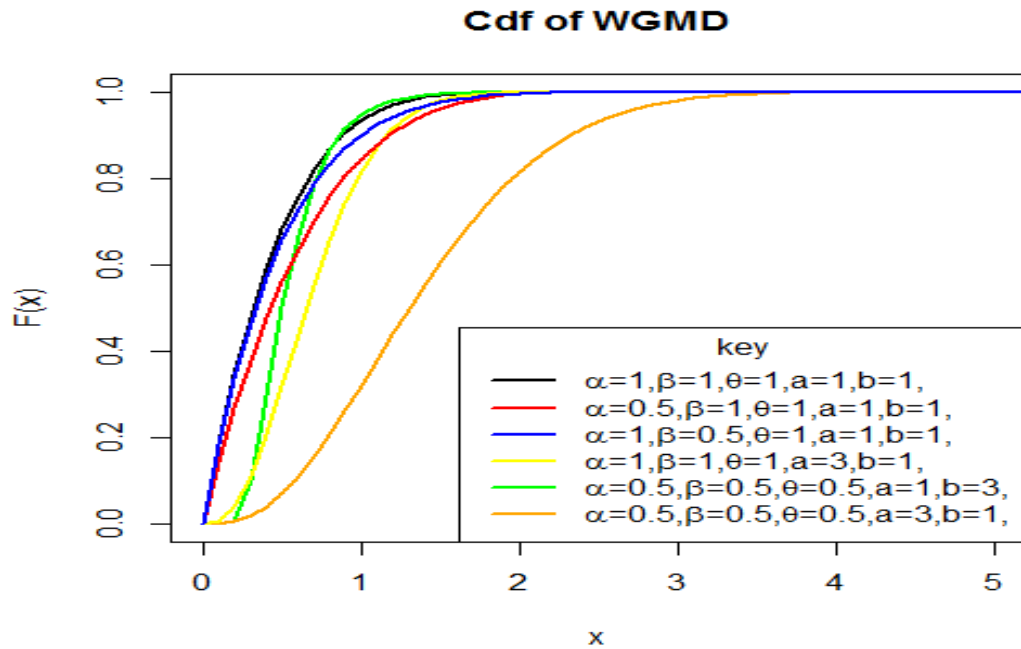
83 The following is a graphical representation of the pdf and cdf of the WGMD using arbitrary
 84 values of the parameters a, b, θ, α and β .



85

86 **Figure 1:** A plot of PDF of the WGMD for varying parameter values

87 It is observed in Figure 1 that the WGMD is a positively skewed distribution and can take
 88 various forms. This means that distribution can be very useful for datasets that are skewed.



89

90 **Figure 2:** A plot of CDF of the WGMD for varying parameter values

91 From the above cdf plot, the cdf increases when X increases, and approaches 1 when X becomes
 92 large or tends to infinity as expected.

93

94 **3 Properties**

95 In this section, we defined and discuss some properties of the WGMD distribution.

96 **3.1 Moments**

97 Let X denote a continuous random variable, the nth moment of X is given by;

$$98 \quad \mu'_n = E[X^n] = \int_0^{\infty} x^n f(x) dx \quad (7)$$

99 Considering f(x) to be the pdf of the WGMD as given in equation (6)

$$100 \quad \mu'_n = E[X^n] = \int_0^{\infty} x^n f(x) dx$$

101 Recall that from equation (6),

$$102 \quad f(x) = \frac{ab \left(\theta + \alpha e^{\beta x} \right) e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \left(-\log \left[1 - e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \right] \right)^{b-1}}{\left[1 - e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \right] e^{a \left(-\log \left[1 - e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \right] \right)^b}} \quad (8)$$

103 To simplify the pdf in (8) above, we carry out the following operations:

104 Let

$$105 \quad A = e^{-a \left(-\log \left[1 - e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \right] \right)^b}$$

Then, using a power series expansion for A, we can write A as:

$$107 \quad A = e^{-a \left(-\log \left[1 - e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \right] \right)^b} = \sum_{i=0}^{\infty} \frac{(-1)^i a^i}{i!} \left(-\log \left[1 - e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \right] \right)^{bi}$$

108 Substituting for the expansion above in equation (8), we have;

$$109 \quad f(x) = \frac{ab \left(\theta + \alpha e^{\beta x} \right) e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \sum_{i=0}^{\infty} \frac{(-1)^i a^i}{i!} \left(-\log \left[1 - e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \right] \right)^{bi}}{\left[1 - e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \right] \left(-\log \left[1 - e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \right] \right)^{-(b-1)}}$$

$$f(x) = \frac{ab \left(\theta + \alpha e^{\beta x} \right) e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \sum_{i=0}^{\infty} \frac{(-1)^i a^i}{i!} \left(-\log \left[1 - e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \right] \right)^{bi+(b-1)}}{\left[1 - e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \right]} \quad (9)$$

$$B = \left(-\log \left[1 - e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \right] \right)^{bi+(b-1)}$$

Also, let

Now, considering the following formula from Tahir et al. (2016b) and Ieren et al. (2018) which

holds for B for $i \geq 1$, then B can be written as follows:

$$\left(-\log \left[1 - e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \right] \right)^{bi+(b-1)} = \sum_{k,l=0}^{\infty} \sum_{j=0}^k \frac{(-1)^{j+k+l} (b(i+1))}{(b(i+1)-1-j)} \binom{k-(b(i+1)-1)}{k} \binom{k}{j} \binom{(b(i-1)+1)+k}{l} P_{j,k} \left[1 - e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \right]^l \quad (10)$$

Where for (for $j \geq 0$) $P_{j,0} = 1$ and (for $k = 1, 2, \dots$)

$$P_{j,k} = k^{-1} \sum_{m=1}^k (-1)^m \frac{[m(j+1)-k]}{(m+1)} P_{j,k-m} \quad (11)$$

Combining equation (10) and (11) and inserting the above power series in equation (9) and simplifying, it gives:

$$f(x) = ab \sum_{i=0}^{\infty} \frac{(-1)^i a^i}{i!} \sum_{k,l=0}^{\infty} \sum_{j=0}^k \frac{(-1)^{j+k+l} (b(i+1))}{(b(i+1)-1-j)} \binom{k-(b(i+1)-1)}{k} \binom{k}{j} \binom{(b(i+1)-1)+k}{l} P_{j,k} \left(\theta + \alpha e^{\beta x} \right) e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \left[1 - e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \right]^{l-1}$$

$$f(x) = b \sum_{i=0}^{\infty} \sum_{k,l=0}^{\infty} \sum_{j=0}^k \frac{(-1)^{i+j+k+l} a^{i+1} (b(i+1))}{i! (b(i+1)-1-j)} \binom{k-(b(i+1)-1)}{k} \binom{k}{j} \binom{(b(i+1)-1)+k}{l} P_{j,k} \left(\theta + \alpha e^{\beta x} \right) e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \left[1 - e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \right]^{l-1} \quad (12)$$

Now, if l is a positive non-integer, we can expand the last term in (12) as:

$$\left[1 - e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \right]^{l-1} = \sum_{m=0}^{\infty} (-1)^m \binom{l-1}{m} \left[e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \right]^m \quad (13)$$

Therefore, $f(x)$ becomes:

$$f(x) = b \sum_{m=0}^{\infty} \sum_{i=0}^{\infty} \sum_{k,l=0}^{\infty} \sum_{j=0}^k \frac{(-1)^{i+j+k+l+m} a^{i+1} (b(i+1))}{i! (b(i+1)-1-j)} \binom{l-1}{m} \binom{k-(b(i+1)-1)}{k} \binom{k}{j} \binom{(b(i+1)-1)+k}{l} P_{j,k} \left(\theta + \alpha e^{\beta x} \right) e^{-(m+1)\left(\theta x + \frac{\alpha}{\beta} (e^{\beta x} - 1)\right)}$$

$$f(x) = b \sum_{m=0}^{\infty} \sum_{i=0}^{\infty} \sum_{k,l=0}^{\infty} \sum_{j=0}^k \frac{(-1)^{i+j+k+l+m} a^{i+1} (b(i+1)) (l-1) \binom{l-1}{m} \binom{k-(b(i+1)-1)}{k} \binom{k}{j} \binom{(b(i+1)-1)+k}{l}}{e^{-\frac{\alpha}{\beta}(m+1)} i! (b(i+1)-1-j)} P_{j,k} \left(\theta + \alpha e^{\beta x} \right) e^{-\theta(m+1)x} e^{-\frac{\alpha}{\beta}(m+1)e^{\beta x}} \quad (14)$$

Using power series expansion on the last term in equation (14), we have

$$e^{-\frac{\alpha}{\beta}(m+1)e^{\beta x}} = \sum_{r=0}^{\infty} \frac{(-1)^r \alpha^r (m+1)^r}{r! \beta^r} e^{r\beta x} \quad (15)$$

Now, substituting equation (15), the power series expansion in equation (14) above, one gets:

$$f(x) = b \sum_{r=0}^{\infty} \sum_{m=0}^{\infty} \sum_{i=0}^{\infty} \sum_{k,l=0}^{\infty} \sum_{j=0}^k \frac{(-1)^{i+j+k+l+m+r} \alpha^r a^{i+1} (m+1)^r (b(i+1)) (l-1) \binom{l-1}{m} \binom{k-(b(i+1)-1)}{k} \binom{k}{j} \binom{(b(i+1)-1)+k}{l}}{e^{-\frac{\alpha}{\beta}(m+1)} r! \beta^r i! (b(i+1)-1-j)} P_{j,k} \left(\theta + \alpha e^{\beta x} \right) e^{-[\theta(m+1)-r\beta]x}$$

Now, let

$$\eta_{i,j,k,l,m,r} = b \sum_{r=0}^{\infty} \sum_{m=0}^{\infty} \sum_{i=0}^{\infty} \sum_{k,l=0}^{\infty} \sum_{j=0}^k \frac{(-1)^{i+j+k+l+m+r} \alpha^r a^{i+1} (m+1)^r (b(i+1)) (l-1) \binom{l-1}{m} \binom{k-(b(i+1)-1)}{k} \binom{k}{j} \binom{(b(i+1)-1)+k}{l}}{e^{-\frac{\alpha}{\beta}(m+1)} r! \beta^r i! (b(i+1)-1-j)} P_{j,k}$$

This implies that:

$$f(x) = \eta_{i,j,k,l,m,r} \left\{ \theta e^{-[\theta(m+1)-r\beta]x} + \alpha e^{-[\theta(m+1)-\beta(1+r)]x} \right\} \quad (16)$$

Hence,

$$\mu_n = E(X^n) = \int_0^{\infty} x^n f(x) dx = \int_0^{\infty} x^n \left(\eta_{i,j,k,l,m,r} \left\{ \theta e^{-[\theta(m+1)-r\beta]x} + \alpha e^{-[\theta(m+1)-\beta(1+r)]x} \right\} \right) dx$$

$$\mu_n = \int_0^{\infty} x^n f(x) dx = \eta_{i,j,k,l,m,r} \left[\theta \int_0^{\infty} x^n e^{-[\theta(m+1)-r\beta]x} dx + \alpha \int_0^{\infty} x^n e^{-[\theta(m+1)-\beta(1+r)]x} dx \right] \quad (17)$$

Also, using integration by substitution method in equation (17) gives the following:

$$\text{Let } u_1 = [\theta(m+1) - r\beta]x \Rightarrow x = \frac{u_1}{\theta(m+1) - r\beta} \quad \frac{du_1}{dx} = \theta(m+1) - r\beta; \text{ and } dx = \frac{du_1}{\theta(m+1) - r\beta}$$

141 Let $u_2 = [\theta(m+1) - \beta(1+r)]x \Rightarrow x = \frac{u_2}{\theta(m+1) - \beta(1+r)} \frac{du_2}{dx} = \theta(m+1) - \beta(1+r)$; and

142
$$dx = \frac{du_2}{\theta(m+1) - \beta(1+r)}$$

143 Substituting for u , x , and dx in equation (17) and simplifying gives:

144
$$\mu'_n = \eta_{i,j,k,l,m,r} \left[\frac{\theta}{(\theta(m+1) - r\beta)^{n+1}} \int_0^\infty u_1^{n+1-1} e^{-u_1} du_1 + \frac{\alpha}{(\theta(m+1) - \beta(1+r))^{n+1}} \int_0^\infty u_2^{n+1-1} e^{-u_2} du_2 \right] \quad (18)$$

145 Again recall that $\int_0^\infty t^{n-1} e^{-t} dt = \Gamma(n)$ and that $\int_0^\infty t^n e^{-t} dt = \int_0^\infty t^{n+1-1} e^{-t} dt = \Gamma(n+1)$

146 Thus, the n^{th} ordinary moment of X for the WGMD is given as follows:

147
$$\mu'_n = \eta_{i,j,k,l,m,r} \left[\frac{\theta \Gamma(n+1)}{(\theta(m+1) - r\beta)^{n+1}} + \frac{\alpha \Gamma(n+1)}{(\theta(m+1) - \beta(1+r))^{n+1}} \right] \quad (19)$$

149 **3.2 The Mean**

150 The mean of the WGMD can be obtained from the n^{th} moment of the distribution when $n = 1$ as follows:

152
$$\mu_1 = \eta_{i,j,k,l,m,r} \left[\frac{\theta}{(\theta(m+1) - r\beta)^2} + \frac{\alpha}{(\theta(m+1) - \beta(1+r))^2} \right] \quad (20)$$

154 **3.3 The Variance**

155 The n^{th} central moment or moment about the mean of X , say μ_n , can be obtained as

156
$$\mu_n = E[X - \mu_1]^n = \sum_{i=0}^n (-1)^i \binom{n}{i} \mu_1^i \mu'_{n-i} \quad (21)$$

157 The variance of X for WGMD is obtained from the n^{th} central moment when $n = 2$, that is, the variance of X is the n^{th} central moment of order two ($n = 2$) and is given as follows:

159
$$Var(X) = E[X^2] - \{E[X]\}^2 \quad (22)$$

160
$$Var(X) = \mu_2' - \{\mu_1'\}^2$$

161
$$Var(X) = \eta_{i,j,k,l,m,r} \left[\frac{2\theta}{(\theta(m+1)-r\beta)^3} + \frac{2\alpha}{(\theta(m+1)-\beta(1+r))^3} \right] - \left\{ \eta_{i,j,k,l,m,r} \left[\frac{\theta}{(\theta(m+1)-r\beta)^2} + \frac{\alpha}{(\theta(m+1)-\beta(1+r))^2} \right] \right\}^2 \quad (23)$$

162 The coefficients variation, skewness and kurtosis measures can also be calculated from the non-
163 central moments using some well-known relationships.

164

165 3.4 Moment Generating Function

166 The mgf of a random variable X can be obtained by

167
$$M_x(t) = E[e^{tx}] = \int_0^{\infty} e^{tx} f(x) dx \quad (24)$$

168 Using power series expansion in equation (24) and simplifying the integral gives;

169
$$M_x(t) = \sum_{n=0}^{\infty} \frac{t^n}{n!} \mu_n' = \sum_{n=0}^{\infty} \frac{t^n}{n!} \left\{ \eta_{i,j,k,l,m,r} \left[\frac{\theta \Gamma(n+1)}{(\theta(m+1)-r\beta)^{n+1}} + \frac{\alpha \Gamma(n+1)}{(\theta(m+1)-\beta(1+r))^{n+1}} \right] \right\} \quad (25)$$

170 where n and t are constants, t is a real number and μ_n' denotes the nth ordinary moment of X .

171

172 3.5 Characteristics Function

173 The characteristics function of a random variable X is given by;

174
$$\phi_x(t) = E[e^{itx}] = E[\cos(tx) + i \sin(tx)] = E[\cos(tx)] + E[i \sin(tx)] \quad (26)$$

175

176 Simple algebra and power series expansion proves that

177
$$\phi_x(t) = \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n}}{(2n)!} \mu_{2n}' + i \sum_{n=0}^{\infty} \frac{(-1)^n t^{2n+1}}{(2n+1)!} \mu_{2n+1}' \quad (27)$$

178 Where μ_{2n}' and μ_{2n+1}' are the moments of X for n=2n and n=2n+1 respectively and can be
179 obtained from μ_n' in equation (20).

180

181 4 Some Reliability Functions

182 In this section, the survival and hazard functions from the WGMD are presented with adequate
183 plots and their discussions.

184

185 4.1 The Survival Function

186 The survival function as the name implies describes the probability that a component or an
187 individual will not fail after a given time. It is mathematically given as:

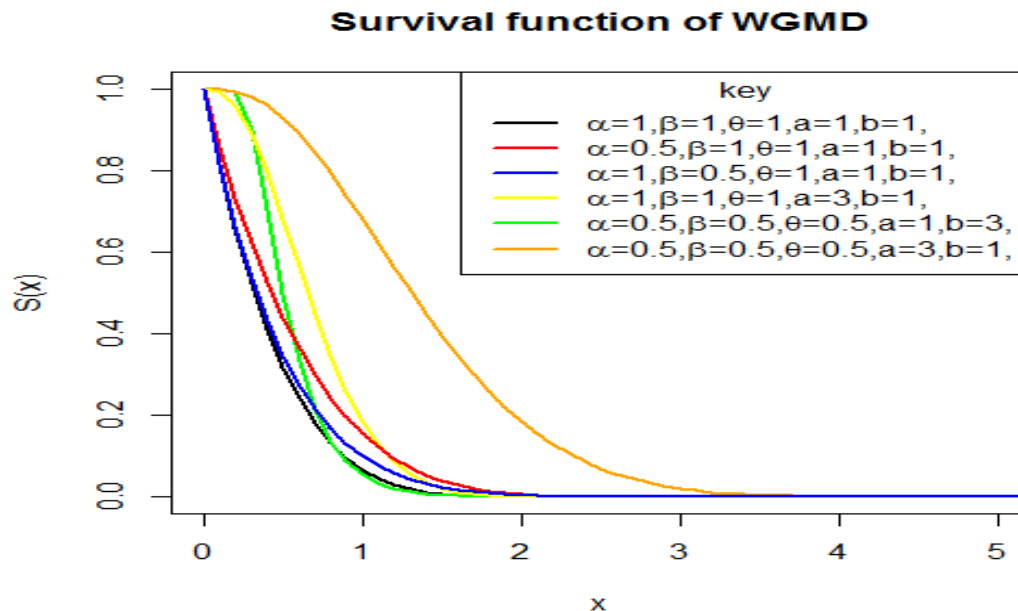
188
$$S(x) = 1 - F(x) \tag{28}$$

189 Taking $F(x)$ to be the cdf of the WGMD, substituting and simplifying (28) above, we get the
190 survival function for the WGMD as:

191
$$S(x) = 1 - e^{-a \left(-\log \left[1 - e^{-\frac{\theta x - \alpha}{\beta} (e^{\beta x} - 1)} \right] \right)^b} \tag{29}$$

192

The following is a plot of the survival function for arbitrary parameter values.



193

194 **Figure 3:** A plot of the survival function of WGMD.

195 The figure above reveals that the probability of survival for any random variable following a
196 WGMD which decreases as the time increases, that is, as time goes on, probability of life

200 decreases as it is expected. This shows that the WGMD would be useful for modelling the most
 201 real-life situations.

202

203 **4.2 The Hazard Function**

204 The hazard function is also called failure or risk function. It describes the probability failure for a
 205 component given an interval of time. The hazard function is defined mathematically as;

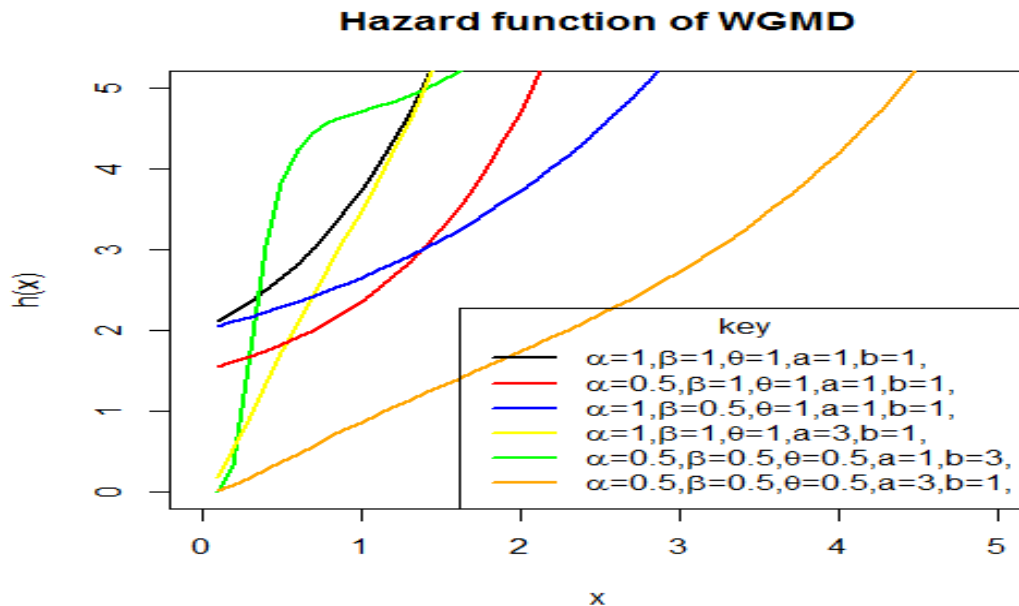
$$206 \quad h(x) = \frac{f(x)}{1-F(x)} = \frac{f(x)}{S(x)} \quad (30)$$

207 Considering $f(x)$ and $F(x)$ to be the pdf and cdf of the proposed WGMD gave previously, we
 208 obtain the hazard function as:

$$209 \quad h(x) = \frac{ab \left(\theta + \alpha e^{\beta x} \right) e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \left(-\log \left[1 - e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \right] \right)^{b-1} e^{-a \left(-\log \left[1 - e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \right] \right)^b}}{\left[1 - e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \right] \left(1 - e^{-a \left(-\log \left[1 - e^{-\theta x - \frac{\alpha}{\beta} (e^{\beta x} - 1)} \right] \right)^b} \right)} \quad (31)$$

209

The following is a plot of the hazard function at chosen parameter values.



210

211 **Figure 4:** A plot of the hazard function of the WGMD.

210 Figure 4 above shows the behaviour of the hazard function of the WGMD. It means that the
 211 probability of failure for any WGM random variable increases as the time or age of the subject
 212 increases, that is, as time goes on, the probability of failure or death increases.

213

214 5 Order Statistics

215 Suppose X_1, X_2, \dots, X_n is a random sample from a distribution with pdf, $f(x)$, and
 216 $X_{1:n} < X_{2:n} < \dots < X_{i:n}$ denote the corresponding order statistic obtained from this sample. Then
 217 the pdf, $f_{i:n}(x)$ of the i^{th} order statistic can be defined as;

$$218 \quad f_{i:n}(x) = \frac{n!}{(i-1)!(n-i)!} \sum_{k=0}^{n-i} (-1)^k \binom{n-i}{k} f(x) F(x)^{k+i-1} \quad (32)$$

219 where $f(x)$ and $F(x)$ are the pdf and cdf of the Weibull Gompertz Makeham distribution
 220 respectively.

221 Using (5) and (6), the pdf of the i^{th} order statistics $X_{i:n}$ can be expressed from (32) as;

$$222 \quad f_{i:n}(x) = \frac{n!}{(i-1)!(n-i)!} \sum_{k=0}^{n-i} (-1)^k \binom{n-i}{k} \frac{ab(\theta + \alpha e^{\beta x}) \left(-\log \left[1 - e^{-\frac{\theta x - \alpha}{\beta} (e^{\beta x} - 1)} \right] \right)^{b-1}}{\left[1 - e^{-\frac{\theta x - \alpha}{\beta} (e^{\beta x} - 1)} \right] e^{a \left(-\log \left[1 - e^{-\frac{\theta x - \alpha}{\beta} (e^{\beta x} - 1)} \right] \right)^b} e^{\frac{\theta x + \alpha}{\beta} (e^{\beta x} - 1)}} \left[e^{-a \left(-\log \left[1 - e^{-\frac{\theta x - \alpha}{\beta} (e^{\beta x} - 1)} \right] \right)^b} \right]^{i+k-1} \quad (33)$$

223 Hence, the pdf of the minimum order statistic $X_{(1)}$ and maximum order statistic $X_{(n)}$ of the
 224 WGMD are given by;

$$225 \quad f_{1:n}(x) = n \sum_{k=0}^{n-1} (-1)^k \binom{n-1}{k} \frac{ab(\theta + \alpha e^{\beta x}) \left(-\log \left[1 - e^{-\frac{\theta x - \alpha}{\beta} (e^{\beta x} - 1)} \right] \right)^{b-1}}{\left[1 - e^{-\frac{\theta x - \alpha}{\beta} (e^{\beta x} - 1)} \right] e^{a \left(-\log \left[1 - e^{-\frac{\theta x - \alpha}{\beta} (e^{\beta x} - 1)} \right] \right)^b} e^{\frac{\theta x + \alpha}{\beta} (e^{\beta x} - 1)}} \left[e^{-a \left(-\log \left[1 - e^{-\frac{\theta x - \alpha}{\beta} (e^{\beta x} - 1)} \right] \right)^b} \right]^k \quad (34)$$

226 and

$$f_{n:n}(x) = n \frac{ab(\theta + \alpha e^{\beta x}) \left(-\log \left[1 - e^{-\theta x - \frac{\alpha}{\beta}(e^{\beta x} - 1)} \right] \right)^{b-1}}{\left[1 - e^{-\theta x - \frac{\alpha}{\beta}(e^{\beta x} - 1)} \right] e^{a \left(-\log \left[1 - e^{-\theta x - \frac{\alpha}{\beta}(e^{\beta x} - 1)} \right] \right)^b} e^{\theta x + \frac{\alpha}{\beta}(e^{\beta x} - 1)}} \left[e^{-a \left(-\log \left[1 - e^{-\theta x - \frac{\alpha}{\beta}(e^{\beta x} - 1)} \right] \right)^b} \right]^{n-1} \quad (35)$$

228 respectively.

229

230 6 Estimation of unknown Model Parameters using the Maximum Likelihood Method

231 Let X_1, X_2, \dots, X_n be a sample of size 'n' independently and identically distributed random
 232 variables from the WGMD with unknown parameters a, b, α, β and θ defined previously. The pdf
 233 of the WGMD is given from (6) as:

$$f(x) = \frac{ab(\theta + \alpha e^{\beta x}) e^{-\theta x - \frac{\alpha}{\beta}(e^{\beta x} - 1)} \left(-\log \left[1 - e^{-\theta x - \frac{\alpha}{\beta}(e^{\beta x} - 1)} \right] \right)^{b-1} e^{-a \left(-\log \left[1 - e^{-\theta x - \frac{\alpha}{\beta}(e^{\beta x} - 1)} \right] \right)^b}}{\left[1 - e^{-\theta x - \frac{\alpha}{\beta}(e^{\beta x} - 1)} \right]}$$

235 The likelihood function is given by;

$$L(\underline{X} | a, b, \alpha, \beta, \theta) = \frac{(ab)^n \prod_{i=1}^n \left((\theta + \alpha e^{\beta x_i}) e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right) \prod_{i=1}^n \left(-\log \left[1 - e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right] \right)^{b-1} e^{-a \sum_{i=1}^n \left(-\log \left[1 - e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right] \right)^b}}{\prod_{i=1}^n \left[1 - e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right]} \quad (36)$$

237 Let the log-likelihood function, $l = \log L(\underline{X} | a, b, \alpha, \beta, \theta)$ therefore

$$l = n \log a + n \log b + \sum_{i=1}^n \log \left(\theta + \alpha e^{\beta x_i} \right) - \theta \sum_{i=1}^n x_i - \frac{\alpha}{\beta} \sum_{i=1}^n (e^{\beta x_i} - 1) + (b-1) \sum_{i=1}^n \log \left(-\log \left[1 - e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right] \right) - a \sum_{i=1}^n \left(-\log \left[1 - e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right] \right)^b - \sum_{i=1}^n \left(1 - e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right) \quad (37)$$

240 Differentiating l partially a, b , concerning α, β and θ respectively gives;

$$\frac{\partial l}{\partial a} = \frac{n}{a} - \sum_{i=1}^n \left(-\log \left[1 - e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right] \right)^b \quad (38)$$

242

$$\frac{\partial l}{\partial b} = \frac{n}{b} + \sum_{i=1}^n \log \left(-\log \left[1 - e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right] \right) - a \sum_{i=1}^n \left(-\log \left[1 - e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right] \right)^b \log \left(-\log \left[1 - e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right] \right) \quad (39)$$

243

244

$$\begin{aligned}
\frac{\partial l}{\partial \alpha} &= \sum_{i=1}^n \left\{ \frac{e^{\beta x_i}}{\theta + \alpha e^{\beta x_i}} \right\} - \frac{1}{\beta} \sum_{i=1}^n (e^{\beta x_i} - 1) + \frac{(b-1)}{\beta} \sum_{i=1}^n \left\{ \frac{(e^{\beta x_i} - 1) e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)}}{\left(-\log \left[1 - e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right] \right) \left(1 - e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right)} \right\} \\
&- \frac{1}{\beta} \sum_{i=1}^n \left\{ \frac{(e^{\beta x_i} - 1) e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)}}{\left(1 - e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right)} \right\} + ab \sum_{i=1}^n \left\{ \frac{\left(-\log \left[1 - e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right] \right)^{b-1} (e^{\beta x_i} - 1) e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)}}{\left(1 - e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right)} \right\} \quad (40)
\end{aligned}$$

248

$$\begin{aligned}
\frac{\partial l}{\partial \beta} &= \sum_{i=1}^n \left\{ \frac{\alpha e^{\beta x_i}}{\left(\theta + \alpha e^{\beta x_i} \right)} \right\} - \frac{\alpha}{\beta} \sum_{i=1}^n \left\{ x_i e^{\beta x_i} - \beta^{-1} (e^{\beta x_i} - 1) \right\} + \frac{\alpha(b-1)}{\beta} \sum_{i=1}^n \left\{ \frac{\left(x_i e^{\beta x_i} - \beta^{-1} (e^{\beta x_i} - 1) \right) e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)}}{\left(-\log \left[1 - e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right] \right) \left(1 - e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right)} \right\} \\
&- \frac{\alpha}{\beta} \sum_{i=1}^n \left\{ \frac{\left(x_i e^{\beta x_i} - \beta^{-1} (e^{\beta x_i} - 1) \right) e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)}}{\left(1 - e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right)} \right\} + \frac{ab\alpha}{\beta} \sum_{i=1}^n \left\{ \frac{\left(-\log \left[1 - e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right] \right)^{b-1} \left(x_i e^{\beta x_i} - \beta^{-1} (e^{\beta x_i} - 1) \right) e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)}}{\left(1 - e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right)} \right\} \quad (41)
\end{aligned}$$

251

$$\begin{aligned}
\frac{\partial l}{\partial \theta} &= \sum_{i=1}^n \left\{ \frac{1}{\left(\theta + \alpha e^{\beta x_i} \right)} \right\} - \sum_{i=1}^n x_i - (b-1) \sum_{i=1}^n \left\{ \frac{x_i e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)}}{\left(-\log \left[1 - e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right] \right) \left(1 - e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right)} \right\} \\
&- \sum_{i=1}^n \left\{ \frac{x_i e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)}}{\left(1 - e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right)} \right\} + ab \sum_{i=1}^n \left\{ \frac{\left(-\log \left[1 - e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right] \right)^{b-1} x_i e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)}}{\left[1 - e^{-\theta x_i - \frac{\alpha}{\beta}(e^{\beta x_i} - 1)} \right]} \right\} \quad (42)
\end{aligned}$$

253

254

255

256

Equating equation (38), (39), (40), (41) and (42) to zero and solving for the solution of the non-linear system of equations gives the maximum likelihood estimates of the parameters a, b, α, β and θ respectively. However, the solution cannot be obtained analytically except numerically with the aid of suitable statistical software like Python, R, SAS, etc., when data sets are given.

257

258

259

260 **7 Application to a real-life dataset**

261 This section presents a dataset on the remission times (in months) of a random sample of 128
262 bladder cancer patients with its descriptive statistics and application to some selected extensions
263 of the Gompertz-Makeham distribution together with the classical Gompertz distribution. The
264 performance of the Weibull Gompertz-Makeham distribution (WGMD) is compared to some
265 families of Makeham distribution such as Kumaraswamy Gompertz Makeham distribution
266 (KGMD), Transmuted Gompertz-Makeham distribution (TGMD), Gompertz-Makeham
267 distribution (GMD) and the Gompertz distribution (GD).

268 The performance of the above-listed models is ranked using some criteria such as the AIC
269 (Akaike Information Criterion), CAIC (Consistent Akaike Information Criterion) and BIC
270 (Bayesian Information Criterion). It is considered that the model with the smallest values of
271 these statistics will be the best model to fit the data.

272 **Data set:** This data set represents the remission times (in months) of a random sample of 128
273 bladder cancer patients. It has previously been used by Ieren and Chukwu (2018), Ieren and
274 Kuhe (2018) and Abdullahi et al. (2018). It is summarized as follows:

275 **Table 1: Summary Statistics for the dataset**

Parameter	n	Min	Q_1	Median	Q_3	Mean	Max	Var	Skew	Kurt
Values	128	0.0800	3.348	6.395	11.840	9.366	79.05	110.425	3.3257	19.1537

n-sample size, Min-Minimum, Q_1 -First Quartile, Q_3 -Third Quartile, Max-Maximum, Var-Variance, Skew-Skewness, Kurt-Kurtosis

276

277 From the descriptive statistics in table 1, it is observed that the data set is positively skewed with
278 a very high coefficient of kurtosis and therefore suitable for flexible and skewed distributions.

279

280

281

282

283

284

285

286

287

288 **Table 2:** The strength of the selected models using the AIC, CAIC and BIC values of the models
 289 evaluated from the maximum likelihood estimations based on the bladder cancer data.

Distributions	Parameter estimates	AIC	CAIC	BIC	Ranks of models
WGMD	$\alpha = 0.006233$ $\beta = 0.005620$ $\theta = 0.006119$ $a = 0.006070$ $b = 0.004811$	-6.8677	-6.3759	7.3924	1
KGMD	$\alpha = 1.4996$ $\beta = 0.0008286$ $\theta = 3.2157$ $a = 0.1934$ $b = 0.09415$	17.3139	17.8057	31.5741	3
TGMD	$\alpha = 9.7913$ $\beta = 9.6327$ $\theta = 9.4246$ $\lambda = 0.8682$	51.5307	51.8559	62.9388	5
GMD	$\alpha = 2.04152$ $\beta = 7.8924$ $\theta = 7.9969$	29.2259	29.4194	37.7820	4
GD	$\alpha = 4.5814$ $\beta = 4.5809$	12.8378	12.9338	18.5418	2

290 From Table 2, comparing the values of the AIC, CAIC and BIC for each model, the WGMD has
 291 the best performance compared to the KGMD, TGMD GMD and GD. This is due to the decision
 292 rule which says that the distribution or model with the smallest values of the test statistics (AIC,
 293 CAIC and BIC) is taken as the most adequate or efficient model. These values also agree with
 294 the fact that generalizing any continuous distribution provides a compound distribution with a
 295 better fit than the baseline distribution (Koleoso et al. (2019)).

296
 297
 298 **8 Conclusion**
 299 This article proposed a new distribution called Weibull-Gompertz Makeham distribution. The
 300 statistical properties of the distribution have been derived and studied extensively. The model
 301 parameters were estimated using the maximum likelihood method. The distribution (WGMD)
 302 has the best fit compared to the other four models considered in this study.

303 **Competing Interests**
 304 Authors have declared that no competing interests exist.

306 **References**

- 307 1. Abd El-Bar AMT. An extended Gompertz-Makeham distribution with application to lifetime
308 data. *Communications in Statistics-Simulation and Computation*. 2017;1-48.
309 DOI:10.1080/03610918.2017.1348517
310
- 311 2. Abdullahi J, Abdullahi UK, Ieren TG, Kuhe DA, Umar AA. On the properties and
312 applications of transmuted odd generalized exponential-exponential distribution. *Asian Journal*
313 *of Probability and Statistics*. 2018;1(4):1-14.
314 DOI: 10.9734/AJPAS/2018/44073
315
- 316 3. Alizadeh M, Cordeiro GM, de Brito E, Demetrio CGB. The Beta Marshall-Olkin family of
317 distributions. *Journal of Statistical Distributions and Applications*. 2015;2(4):1-18.
318
- 319 4. Alzaatreh A, Famoye F, Lee C. Weibull-Pareto distribution and its applications.
320 *Communications in Statistics-Theory and Methods* 2013;42(9):1673-1691.
321
- 322 5. Alzaghal A, Famoye F., Lee C. Exponentiated T-X family of distributions with some
323 applications. *International Journal of Statistics and Probability* 2013;2(3):31-49.
324
- 325 6. Bourguignon M, Silva RB, Cordeiro GM. The Weibull-G family of probability distributions.
326 *Journal of Data Science*. 2014;12:53-68.
327
- 328 7. Chukwu AU, Ogunde AA. On Kumaraswamy Gompertz Makeham distribution. *American*
329 *Journal of Mathematics and Statistics*. 2016;6(3):122-127.
330
- 331 8. Cordeiro GM, Ortega EMM, da Cunha DCC. The Exponentiated generalized class of
332 distribution. *Journal of Data Science*. 2013;11:1-27.
333
- 334 9. Cordeiro GM, Ortega EMM, Ramires TG. A new generalized Weibull family of distributions:
335 mathematical properties and applications. *Journal of Statistical Distributions and Applications*.
336 2015;2(13):1-25.
337 DOI 10.1186/s40488-015-0036-6
338
- 339 10. Dey S, Moala FA, Kumar D. Statistical properties and different methods of estimation of
340 Gompertz distribution with the application. *Journal of Statistics and Management Systems*.
341 2018;21(5):839-876.
342
- 343 11. Eugene N, Lee C, Famoye F. Beta-Normal distribution and its applications. *Communications*
344 *in Statistics-Theory and Methods*. 2002;31(4):497-512.
345
- 346 12. Golubev A. Does Makeham make sense? *Biogerontology*. 2004;5:159-167.
347
- 348 13. Gompertz B. On the nature of the function expressive of the law of human mortality, and on
349 a new mode of determining the value of life contingencies. *Philosophical Transactions of the*
350 *Royal Society of London*. 1825;115:513-583.
351

- 352 14. Ieren TG, Chukwu AU. Bayesian analysis of a shape parameter of the Weibull-Frechet
353 distribution. *Asian Journal of Probability and Statistics*. 2018;2(1):1-19.
354 DOI: 10.9734/AJPAS/2018/44184
355
- 356 15. Ieren TG, Kuhe DA. On the properties and applications of Lomax-Exponential distribution.
357 *Asian Journal of Probability and Statistics*. 2018;1(4):1-13.
358 DOI: 10.9734/AJPAS/2018/42546
359
- 360 16. Ieren TG, Oyamakin SO, Chukwu AU. Modeling lifetime data with Weibull-Lindley
361 distribution. *Biometrics and Biostatistics International Journal*. 2018;7(6):532–544.
362
- 363 17. Johnson NL, Kotz S, Balakrishnan N. *Continuous Univariate Distributions, volume 2*. John
364 Wiley & Sons, New York;1995.
365
- 366 18. Koleoso PO, Chukwu AU, Bamiduro TA. A Three-Parameter Gompertz-Lindley
367 Distribution: Its Properties and Applications. *Journal of Mathematical Theory and Modeling*.
368 2019;9(4):29-42.
369 DOI: 10.7176/MTM
370
- 371 19. Lawless JF. *Statistical Models and Methods for Lifetime Data, 2nd Edition*, Wiley, New
372 Jersey; 2003.
373
- 374 20. Makeham WM. On the law of mortality and the construction of annuity tables. *Journal of the*
375 *Institute of Actuaries*; 1860;8(6):301-310.
376
- 377 21. Marshall A, Olkin I. *Life Distributions. Structure of Nonparametric, Semiparametric, and*
378 *Parametric Families*. Springer, New York; 2007.
379
- 380 22. Shaw WT, Buckley IRC. The alchemy of probability distributions: beyond Gram-Charlier
381 expansions and a skew-kurtotic-normal distribution from a rank transmutation map. *Research*
382 *report*; 2007.
383
- 384 23. Tahir MH, Cordeiro GM, Alzaatreh A, Mansoor M, Zubair M. The logistic-X family of
385 distributions and its applications. *Communications in Statistics-Theory and Methods*.
386 2016;45(24):7326-7349.
387 DOI: 10.1080/03610926.2014.980516
388
- 389 24. Tahir MH, Zubair M, Mansoor M, Cordeiro GM, Alizadeh M, Hamedani GG. A new
390 Weibull-G family of distributions. *Hacettepe Journal of Mathematics and Statistics*.
391 2016;45(2):629-647.
392
- 393 25. Torabi H, Montazeri NH. The Logistic-Uniform distribution and its application.
394 *Communications in Statistics-Simulation and Computation*. 2014;43(10):2551–2569.
395 DOI: 10.1080/03610918.2012.737491