

Research Article

Inference on the Beta Type I Generalized Half Logistic Distribution under Right-Censored Observation with Application to COVID-19

Phillip Oluwatobi Awodutire ¹, Ethelbert Chinaka Nduka,² Maxwell Azubike Ijomah,² Oluwatosin Ruth Ilori,³ and Oluwafemi Samson Balogun ⁴

¹Department of Mathematics and Computer Science, University of Africa, Toru Orua, Bayelsa State, Nigeria

²Department of Mathematics and Statistics, University of Port Harcourt, Port Harcourt, Rivers State, Nigeria

³Department of Community Medicine, Ladoko Akintola University of Technology Teaching Hospital, Ogbomosh, Nigeria

⁴School of Computing, University of Eastern Finland, Kuopio, Finland

Correspondence should be addressed to Phillip Oluwatobi Awodutire; phillip.awodutire@gmail.com

Received 23 November 2021; Revised 13 January 2022; Accepted 29 March 2022; Published 25 May 2022

Academic Editor: Niansheng Tang

Copyright © 2022 Phillip Oluwatobi Awodutire et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In real-life situations, censoring issues do arise due to the incompleteness of data. This article examined the inferences on right-censored beta type I generalized half logistic distribution. In this work, some statistical properties of the beta type I generalized half logistic distribution were derived. Furthermore, the beta type I generalized half logistic distribution was studied under a censoring situation in the presence and absence of covariates. Estimation of model parameters was conducted using the maximum likelihood estimation method. A simulation study was carried out to assess the performance of the parameters of the model in terms of efficiency and consistency. In a real-life application, the model was applied to COVID-19 data and the necessary inferences were drawn.

1. Introduction

In real-life experiments, the issue of censoring is experienced. This occurs when the subject under study is lost to follow up, survived beyond the given time of the study, or dropped out. In medical research studies, for example, to know the survival times, the first date of contact and last date of contacts are known and recorded. Censoring can be right censoring, left censoring, or interval censoring. Given T as the time of occurrence for some event and as a given value, it is right-censoring if variable T is greater than some value “ c ”. Left censoring is most likely to occur when you begin observing a sample at a time when some of the individuals may have already experienced the event. It is an interval if time T falls between two values of c . Left censoring is most likely to occur when you begin observing a sample at a time when

some of the individuals may have already experienced the event. It is interval if time T falls between two values of c .

A lot of works have been done, assessing probability distributions in the presence of right censoring conditions. Some of the recent works include [1–6], etc. For this research, the beta type I generalized half logistic distribution is studied under a right-censoring mechanism.

One of the probability distributions, which is a member of the logistic distribution, is the half logistic distribution. The half logistic distribution which is also known as folded logistic is derived from logistic distribution by truncating at point $x = 0$. The half logistic distribution has the probability distribution as follows:

$$f(x) = \frac{2e^x}{(1 + e^x)^2} 0 \leq x < \infty. \quad (1)$$

and its cumulative distribution function (cdf) is as follows:

$$F(x) = \frac{e^x - 1}{(1 + e^x)} \quad 0 \leq x < \infty. \quad (2)$$

The author of [7] in his research stated some theorems that characterized the half logistic distribution. The author of [8] derived a new probability density function of type I generalized half logistic distribution by addition of a shape parameter to half logistic distribution. He obtained the moments, median, cumulative distribution function, mode, 100p-percentage point, and order statistics of the distribution. He further estimated parameters of the distribution using the maximum likelihood method. Therefore,

$$f(x) = \frac{b2^b e^x}{(1 + e^x)^{b+1}}. \quad (3)$$

Introducing the scale parameters ψ , equation (3) becomes

$$f(x) = \frac{b2^b e^{x/\psi}}{\psi(1 + e^{x/\psi})^{b+1}}. \quad (4)$$

The type I generalized half logistic distribution by [8] is gradually gaining attention in research studies. It has been further generalized by adding parameters. These generalized distributions have been extensively studied and applied to solve real life situations both under complete and censored observations. The author of [9] obtained the four-parameter type I generalized half logistic distribution. He also obtained the cumulative distribution function (CDF), the survival function, and the hazard function, moments, the 100p-percentage point, and the mode of the distribution. The author of [10] introduced the distribution of [8] to survival analysis. The properties of the survival model such as survival function, hazard function were studied and applied to data on breast cancer patients. The author of [11] derived a parametric survival model for breast cancer patient survival data using the survival model obtained by [10]. The author of [12] further generalized the four-parameter type I generalized half logistic distribution obtained by [9] by the addition of a shifting parameter to have the five parameters generalized half logistic distribution. A further study on some of the properties and estimation of the parameter of distribution of under complete observation was carried out by [13]. The author of [14] extended the distribution of [8] by addition of a parameter to have the Lehmann type II generalized half logistic distribution. Due to the nondecreasing properties of the hazard function of the type I generalized half logistic distribution, The author of [15] derived the BTIGHLD that is not limited to only nondecreasing hazards but can take any other form of hazard function. Its properties were studied and estimation of parameters was done under complete observations. The distribution was applied

to two real life data sets. The BTIGHLD has the probability density function as follows:

$$f(x) = \frac{1}{B(a, b)} \left(1 - \left(\frac{2}{1 + e^{x/\psi}} \right)^q \right)^{(a-1)}, \quad (5)$$

$$\left(\frac{2}{1 + e^{x/\psi}} \right)^{q(b-1)} \frac{q2^q e^{x/\psi}}{(1 + e^{x/\psi})^{q+1}},$$

with CDF as follows:

$$F(x) = \frac{1}{B(a, b)} \int_0^{1 - (2/(1 + e^{x/\psi}))^q} x^{a-1} (1-x)^{b-1} dx. \quad (6)$$

Equation (6) therefore results to the following:

$$F(x) = I_{G(x)}(a, b) = I_k(a, b) = \frac{B(k; a, b)}{B(a, b)}, \quad (7)$$

where the $G(x) = k = 1 - (2/(1 + e^{x/\psi}))^q$ and $B(k; a, b)$ is the incomplete beta function.

The survival function and hazard function is as follows:

$$s(x) = \frac{1}{B(a, b)} \int_0^h x^{a-1} (1-x)^{b-1} dx = \frac{B(h; a, b)}{B(a, b)}, \quad (8)$$

and

$$h(x) = \frac{\left(1 - (2/(1 + e^{x/\psi}))^q \right)^{(a-1)} q2^{qb} e^{x/\psi}}{B(h; a, b) (1 + e^{x/\psi})^{qb+1}}, \quad (9)$$

respectively, where

$$h = \left(\frac{2}{1 + e^{x/\psi}} \right)^q, \quad (10)$$

for $a > 0$, $b > 0$, $\psi > 0$ and $q > 0$.

In literature, several studies have been conducted to examine other generalizations of the half logistic distribution under incomplete data situations. Some of these include [4–16]. The author of [16] obtained an extended generalized half logistic distribution and studied different methods of estimation of its parameters based on censored and complete data. In this paper, we examined some statistical properties of the beta type I generalized half logistic and study its inferences under right-censored observations. The work further applied the model to the survival time of COVID-19 patients in a bid to assess some factors that contribute to their survival. A preprint of this work has been previously published in [17].

In the next section, we present some properties of BTIGHLD. In Section 3, we estimated the parameters of the distribution under censored observations without covariates, and Section 4 deals with estimation with covariates. Section 5 has a simulation study of the model and assesses the performance of parameters of the model. In Section 6, the model is applied to COVID-19 data.

2. Statistical Properties of BTIGHLD

In this section, further study is carried out on some properties of type I generalized half logistic distribution. The properties under study include moments, moment generating function, incomplete moments, and order statistics. For simplicity in formulae which can be easily handled for computer software because of their ability to deal with analytic expressions of formidable size and complexity, established explicit expressions to calculate statistical measures can be more efficient than computing them directly by numerical integration. Therefore, the need to also compute the mixture representation of the PDF of the BTIGHLD by having the Exp-G form which will enable us in studying the statistical properties of the.

2.1. Mixture Representation. By using the power series,

$$(1 - z)^{b-1} = \sum_{l=0}^{\infty} (-1)^l \binom{b-1}{l} z^l,$$

$$\left[\left(\frac{2}{1 + e^{x/\psi}} \right)^q \right]^{(b-1)} = \sum_{l=0}^{\infty} (-1)^l \binom{b-1}{l} \left(1 - \left(\frac{2}{1 + e^{x/\psi}} \right)^q \right)^l. \tag{11}$$

Therefore, the p.d.f of BTIGHLD becomes as follows:

$$f(x) = \sum_{l=0}^{\infty} \frac{(-1)^l \binom{b-1}{l}}{B(a, b)} \frac{q 2^q e^{x/\psi}}{(1 + e^{x/\psi})^{q+1}} \left(1 - \left(\frac{2}{1 + e^{x/\psi}} \right)^q \right)^{l+a-1}. \tag{12}$$

Therefore, the expression in (12) can be expressed as a mixture of the exponentiated-type I generalized half logistic distribution (Lehmann type II generalized half logistic distribution) as follows:

$$f(x) = \sum_{l=0}^{\infty} j_k \Theta_{l+a}(x), \tag{13}$$

where

$$j_k = \frac{(-1)^l \binom{b-1}{l}}{(l+a)B(a, b)} = \frac{(-1)^l \Gamma(a+b)}{(l+a)\Gamma(a)\Gamma(l+1)\Gamma(b-l)}, \tag{14}$$

and

$$\Theta_{l+a}(x) = (l+a) \frac{q 2^q e^{x/\psi}}{(1 + e^{x/\psi})^{q+1}} \left(1 - \left(\frac{2}{1 + e^{x/\psi}} \right)^q \right)^{l+a-1}, \tag{15}$$

$\Theta_{l+a}(x)$ is the Exp-G PDF form of the distribution. Integrating expression in 11 gives the mixture of the Exp-G CDF as follows:

$$f(x) = \sum_{l=0}^{\infty} j_k \Phi_{l+a}(x), \tag{16}$$

where

$$\Phi_{l+a}(x) = \left(1 - \left(\frac{2}{1 + e^{x/\psi}} \right)^q \right)^{l+a}. \tag{17}$$

Further simplifying expression in 10, using the binomial expression,

$$\left(1 - \left(\frac{2}{1 + e^{x/\psi}} \right)^q \right)^{l+a-1} = \sum_{j=0}^{\infty} (-1)^j \binom{l+a-1}{j} \left(\frac{2}{1 + e^{x/\psi}} \right)^{qj}. \tag{18}$$

Therefore, expression 10 becomes as follows:

$$f(x) = \sum_{l=0}^{\infty} \sum_{j=0}^{\infty} \frac{(-1)^{l+j} \binom{b-1}{l} \binom{l+a-1}{j}}{B(a, b)} \frac{q 2^q e^{x/\psi}}{(1 + e^{x/\psi})^{q+1}} \left(\frac{2}{1 + e^{x/\psi}} \right)^{qj}, \tag{19}$$

which finally gives

$$f(x) = \sum_{l=0}^{\infty} \sum_{j=0}^{\infty} \frac{(-1)^{l+j} \binom{b-1}{l} \binom{l+a-1}{j}}{B(a, b)(j+1)} \frac{q(j+1) 2^{q(j+1)} e^{x/\psi}}{(1 + e^{x/\psi})^{q(j+1)+1}}, \tag{20}$$

(20) can be written as follows:

$$f(x) = \sum_{l=0}^{\infty} \sum_{j=0}^{\infty} r_{l+j} \beta_{q(j+1)}, \tag{21}$$

where

$$r_{l+j} = \sum_{l=0}^{\infty} \sum_{j=0}^{\infty} \frac{(-1)^{l+j} \binom{b-1}{l} \binom{l+a-1}{j}}{B(a, b)(j+1)}, \tag{22}$$

and

$$\beta_{q(j+1)} = \frac{q(j+1) 2^{q(j+1)} e^{x/\psi}}{(1 + e^{x/\psi})^{q(j+1)+1}}, \tag{23}$$

$\beta_{q(j+1)}$ is the p.d.f of the type I generalized half logistic distribution with power parameter $(q(j+1))$.

2.2. Moments and Moment Generating Function. Given that X is a random variable with BTIGHLD distribution, its s^{th} moment (ζ'_s) is given as follows:

$$\zeta'_s = E[X^s] = \sum_{l=0}^{\infty} \sum_{j=0}^{\infty} r_{l+j} E[\beta_{q(j+1)}^s], \tag{24}$$

where $E[\beta_{q(j+1)}^s]$ is the s^{th} moment of the type I generalized half logistic distribution with power parameter $(q(j+1))$. From the expression of the s^{th} moments, we obtain the second moment, skewness, and kurtosis.

For the moment generating function, say, M_s ,

$$M_s = \sum_{l=0}^{\infty} \sum_{j=0}^{\infty} r_{l+j} E[M_{s_{q(j+1)}}], \tag{25}$$

where $M_{s_{q(j+1)}}$ is the moment generating function of the type I generalized half logistic distribution with power parameter $q(j+1)$.

Tables 1 and 2 reveal the values of the moments of the BTIGHLD for different parameter values. The tables show positive values for the mean, the second moment, and the variance, indicating its ability to handle real life situations which are most times positive.

2.3. Order Statistics. Given that $X_{1:n} \leq X_{2:n} \leq X_{3:n} \leq \dots \leq X_{n:n}$ be the order statistics from a distribution with probability density function $f(x)$ and cumulative density function $F(x)$, then the PDF of $X_{v:n}$, which is the v^{th} order statistics, is given as follows:

$$f_{v:n}(x) = \frac{\sum_{b=0}^{n-v} (-1)^b \binom{n-v}{b} f(x) F(x)^{v+b-1}}{B(v, n-v+1)}, \tag{26}$$

where $B(\dots)$ is a beta function.

Introducing 11 and 12 as $f(x)$ and $F(x)$, respectively, using an equation given in page 17 of Gradshteyn and

$$C_{k,v+b-1} = (v+b-1)^{-1} (g_0)^{-1} \sum_{m=1}^{v+b-1} [m(k+1) - v - b + 1] (g_m) (C_{k,v+k-1-m}). \tag{28}$$

Therefore, we can deduce that the p.d.f of BTIGHLD order statistics is a mixture of Exp-G p.d.f.s. Hence, the properties of $X_{i:n}$ follows from properties of $a+k$.

3. Estimation of BTIGHLD under Right-Censored Observation without Covariates

Let X_1, X_2, \dots, X_n be random variables of size n follows a particular probability distribution, the likelihood function under censored observation is as follows:

$$L = \prod_{i=1}^r f(x_i) \prod_{i=r+1}^n s(x_i), \tag{29}$$

where $\prod_{i=1}^r f(x_i)$ represents the joint probability of observing the uncensored survival times and $\prod_{i=r+1}^n s(x_i)$ represents the joint probability of those censored observations.

TABLE 1: Moments for BTIGHLD (1, 0.5, 0.5, a, 2.5).

a	Mean	E [X ²]	Var[X]
0.5	3.9217	18.8285	3.465919
1.0	2.976694	11.02313	2.173289
1.5	2.546119	7.364395	0.886102
2.0	2.286235	5.640117	0.415324
2.5	2.070578	4.696419	0.4111824
3.0	1.924737	3.961982	0.2586615
3.5	1.809295	3.448451	0.1757825

TABLE 2: Moments for BTIGHLD (a, 3, 0.5, 0.5, b).

b	Mean	E [X ²]	Var[X]
0.5	2.585223	9.611061	2.942398
1.0	1.885569	4.377903	0.8266652
1.5	1.527243	2.659732	0.3289052
2.0	1.435395	2.319144	0.2600847
2.5	1.3717	2.09347	0.2129731
3.0	1.275957	1.740993	0.1134955
3.5	1.274244	1.714889	0.09164902

Ryzhik for a power series raised to a positive integer n we have the order statistics of the BTIGHLD as follows:

$$f_{v:n}(x) = \frac{\sum_{b=0}^{n-v} (-1)^b \binom{n-v}{b}}{B(v, n-v+1)} \sum_{l=0}^{\infty} C_{k,v+b-1} \Theta_{l+a}(x), \tag{27}$$

where

Therefore, given a sample $X_1, X_2, X_3, \dots, X_n$ of size n from the BTIGHLD, the likelihood function under censored observation without considering the covariates is given as follows:

$$L = \left(\frac{1}{B(a, b)} \right)^r \left(1 - \left(\frac{2}{1 + e^{x_i/\psi}} \right)^q \right)^{r(a-1)}, \tag{30}$$

$$\frac{q^r 2^{rqb} \prod_{i=1}^n e^{x_i/\psi}}{\prod_{i=1}^n (1 + e^{x_i/\psi})^{qb+1}} \left(\frac{\prod_{i=r+1}^n B(k; a, b)}{B(a, b)^{n-r}} \right),$$

where

$$k = \left(\frac{2}{1 + e^{x_i/\psi}} \right)^q. \tag{31}$$

From (30), taking the logarithm of both sides, we have

$$\begin{aligned} \ln L = l = n \ln \left(\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \right) + r \ln q + \sum_{i=1}^r \frac{x_i}{\psi} + rqb \ln 2 + r(a-1) \ln \sum_{i=1}^r \left(1 - \left(\frac{2}{1+e^{x_i/\psi}} \right)^q \right) \\ - (qb+1) \sum_{i=1}^r \ln(1+e^{x_i/\psi}) + \sum_{i=r+1}^n \ln B(k; a, b), \end{aligned} \tag{32}$$

which gives

$$\begin{aligned} l = n(\ln \Gamma(a+b) - \ln \Gamma(a) - \ln \Gamma(b)) + \sum_{i=1}^r \frac{x_i}{\psi} + r(a-1) \ln \sum_{i=1}^r \left(1 - \left(\frac{2}{1+e^{x_i/\psi}} \right)^q \right) + r \ln q \\ + rqb \ln 2 - (qb+1) \sum_{i=1}^r \ln(1+e^{x_i/\psi}) + \ln \sum_{i=r+1}^n B(k; a, b). \end{aligned} \tag{33}$$

Differentiating l with respect to the parameters, we have

$$\begin{aligned} \frac{\partial l}{\partial a} &= \frac{n\Gamma'(a+b)}{\Gamma(a+b)} - \frac{n\Gamma'(a)}{\Gamma(a)} \\ &+ r \ln \sum_{i=1}^r \left(1 - \left(\frac{2}{1+e^{x_i/\psi}} \right)^q \right) + \sum_{i=r+1}^n \frac{B'(k; a, b)}{B(k; a, b)} \frac{\partial l}{\partial b} = n \left(\frac{\Gamma'(a+b)}{\Gamma(a+b)} - \frac{\Gamma'(b)}{\Gamma(b)} \right) - q \sum_{i=1}^r (1+e^{x_i/\psi}) + rqb \ln 2 + \sum_{i=r+1}^n \frac{B'(k; a, b)}{B(k; a, b)}, \\ \frac{\partial l}{\partial \psi} &= qr2^q(a-1) \sum_{i=1}^r \left(\frac{(x_i/\psi^2 e^{x_i/\psi})^q}{1-(2/1+e^{x_i/\psi})^q} \right) - \sum_{i=1}^r \frac{x_i}{\psi^2} + (qb+1) \sum_{i=1}^r e^{x_i/\psi} + \sum_{i=r+1}^n \frac{B'(k; a, b)}{B(k; a, b)}, \\ \frac{\partial l}{\partial q} &= r(a-1) \sum_{i=1}^r \left(\frac{(2/1+e^{x_i/\psi})^q \ln(2/1+e^{x_i/\psi})}{1-(2/1+e^{x_i/\psi})^q} \right) + r b \ln 2 - b \sum_{i=1}^r \ln(1+e^{x_i/\psi}) + \sum_{i=r+1}^n \frac{B'(k; a, b)}{B(k; a, b)}. \end{aligned} \tag{34}$$

Expressions for the derivatives of the incomplete beta function with respect to the model parameter can be found in Appendix. In obtaining the interval estimation, and to further carry out the test of hypothesis on model parameters (ψ, q, a, b) , we obtain a 4×4 unit information matrix, with the elements obtained by the second derivatives of the loglikelihood function.

If the assumptions that are set for fulfilment in model BTIGHLD, the asymptotic distribution in $\sqrt{n}(\hat{\Omega} - \Omega)$ is $N_4(0, K(\hat{\Omega})^{-1})$ distribution of Ω . This found use in way of constructing confidence intervals with its respective region of confidence with each of the distribution values. In obtaining or assessing the goodness of fit of BTIGHLD, we can also make use of asymptotic normality which will also be useful in comparison procedure with other distribution of sub-models using the Wald Statistics or Likelihood statistics. If Ω_i is the model parameter of BTIGHLD, we obtain the C.I for λ level of significance given by the following:

$$\begin{aligned} ACI(\Omega_i, 100(1-\lambda)) &= \hat{\Omega} - z_{\lambda/2} \sqrt{M^{\hat{\Omega}, \hat{\Omega}}}, \\ &\hat{\Omega} + z_{\lambda/2} \sqrt{M^{\hat{\Omega}, \hat{\Omega}}}. \end{aligned} \tag{35}$$

where $M^{\hat{\Omega}, \hat{\Omega}}$ is the i^{th} diagonal element of $K_n(\hat{\Omega})^{-1}$ for $i = 1, 2, 3, 4$ and $z_{\lambda/2}$ is the quantile of the standard normal distribution.

4. Estimation of BTIGHLD under Censored Observation with Covariates

In this section, we examined the estimation of the BTIGHLD under right censored with covariates. The likelihood function is given as follows:

$$L = \prod_{i=1}^r f(x_i|d, \beta) \prod_{i=r+1}^n s(x_i^+|d, \beta), \tag{36}$$

which is equivalent to

$$L = \prod_{i=1}^r g(d|\beta) f_o(Z(x_i)) \cdot \prod_{i=r+1}^n s_o(Z(x_i^+)), \quad (37)$$

where $g(d|\beta) = \exp(-d\beta)$ and $Z(x) = x \exp(-d\beta)$. Also $f_o(Z(x))$ and $s_o(Z(x^+))$ are the baseline probability density

function and baseline survival function, respectively (Appendix).

This now gives

$$L = \prod_{i=1}^r g(d|\beta) \left(\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \right) \frac{q2^{qb} e^{Z(x_i)/\psi}}{(1 + e^{Z(x_i)/\psi})^{qb+1}} \left(1 - \left(\frac{2}{1 + e^{Z(x_i)/\psi}} \right)^q \right)^{(a-1)} \prod_{i=r+1}^n \left(\frac{B(f; a, b)}{B(a, b)} \right), \quad (38)$$

where

$$f = \left(\frac{2}{1 + e^{Z(x_i^+)/\psi}} \right)^q. \quad (39)$$

From (38), taking the logarithm of both sides, we have

$$\begin{aligned} \ln L = & r(-d'\beta) + n \ln \left(\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} \right) + r \ln q + r(a-1) \ln \sum_{i=1}^r \left(1 - \left(\frac{2}{1 + e^{Z(x_i)/\psi}} \right)^q \right) + rqb \ln 2 + \sum_{i=1}^r \frac{Z(x_i)}{\psi} \\ & - (qb+1) \sum_{i=1}^r \left(1 + e^{Z(x_i)/\psi} \right) + \ln \sum_{i=r+1}^n B(f; a, b). \end{aligned} \quad (40)$$

Differentiating l with respect to the parameters, we have

$$\begin{aligned} \frac{\partial l}{\partial a} = & n \left(\frac{\Gamma'(a+b)}{\Gamma(a+b)} - \frac{\Gamma'(a)}{\Gamma(a)} \right) + r \ln \sum_{i=1}^r \left(1 - \left(\frac{2}{1 + e^{Z(x_i)/\psi}} \right)^q \right) + \sum_{i=r+1}^n \frac{B'(f; a, b)}{B(f; a, b)} \frac{\partial l}{\partial b} = n \left(\frac{\Gamma'(a+b)}{\Gamma(a+b)} - \frac{\Gamma'(b)}{\Gamma(b)} \right) + qr \ln 2 \\ & - q \sum_{i=1}^r \left(1 + e^{Z(x_i)/\psi} \right) + \sum_{i=r+1}^n \frac{B'(f; a, b)}{B(f; a, b)}, \end{aligned}$$

$$\frac{\partial l}{\partial \psi} = qr2^q (a-1) \sum_{i=1}^r \left(\frac{\left(\frac{Z(x_i)/\psi^2 e^{Z(x_i)/\psi}}{1 - \left(\frac{2}{1 + e^{Z(x_i)/\psi}} \right)^q} \right)^q}{1 - \left(\frac{2}{1 + e^{Z(x_i)/\psi}} \right)^q} \right) - \sum_{i=1}^r \frac{Z(x_i)}{\psi^2} + (qb+1) \sum_{i=1}^r \frac{Z(x_i)}{\psi^2} + \sum_{i=r+1}^n \frac{B'(f; a, b)}{B(f; a, b)}, \quad (41)$$

$$\frac{\partial l}{\partial q} = r(a-1) \sum_{i=1}^r \left(\frac{\left(\frac{2/1 + e^{Z(x_i)/\psi}}{1 - \left(\frac{2}{1 + e^{Z(x_i)/\psi}} \right)^q} \right)^q \ln \left(\frac{2/1 + e^{Z(x_i)/\psi}}{1 - \left(\frac{2}{1 + e^{Z(x_i)/\psi}} \right)^q} \right)}{1 - \left(\frac{2}{1 + e^{Z(x_i)/\psi}} \right)^q} \right) + \frac{r}{q} + r b \ln 2 - b \sum_{i=1}^r \left(1 + e^{Z(x_i)/\psi} \right) + \sum_{i=r+1}^n \frac{B'(f; a, b)}{B(f; a, b)}.$$

For interval estimation and test of hypothesis on the parameters $(\psi, q, a, b, \beta_1, \beta_2, \dots, \beta_j)$, we obtain a $(4+j) \times (4+j)$ unit information matrix where the corresponding elements are obtained by taking the second derivatives with respect to the parameters.

Under conditions that are fulfilled for parameters, the asymptotic distribution of $\sqrt{n}(\hat{\Omega} - \Omega)$ is $N_{(4+j)}(0, K(\hat{\Omega})^{-1})$ distribution of Ω can be used to construct approximate confidence intervals and confidence regions for the parameters and for the hazard and survival

functions. This found use in way of constructing confidence intervals with its respective region of confidence with each of the distribution values. In obtaining or assessing the goodness of fit of BTIGHLD, we can also make use of asymptotic normality which will also be useful in comparison procedure with other distribution of sub-models using the Wald Statistics or Likelihood statistics. If Ω_i is the model parameter of BTIGHLD, we obtain the C.I for λ level of significance given by the following:

$$ACI(\Omega_i, 100(1 - \lambda)) = \widehat{\Omega} - z_{\lambda/2} \sqrt{M^{\widehat{\Omega}, \widehat{\Omega}}}, \widehat{\Omega} + z_{\lambda/2} \sqrt{M^{\widehat{\Omega}, \widehat{\Omega}}}, \tag{42}$$

where $M^{\widehat{\Omega}, \widehat{\Omega}}$ is the i^{th} diagonal element of $K_n(\widehat{\Omega})^{-1}$ for $i = 4 + j$ and $z_{\lambda/2}$ is the quantile of the standard normal distribution.

5. Simulation Study

In this section, a simulation will be carried out in order to assess the performance of the estimates using the maximum likelihood estimation method of the BTIGHLD. The value of scale parameter is fixed at 1 for all simulations. For, given values of a, b, q for Group I(2.5, 0.5, 3.0), Group II (0.35, 5.0, 1.8), we simulated data that are distributed to BTIGHLD. In assessing the performance of the estimates, the average estimate (AE) and MSE of $\widehat{\rho}_l$ over the r samples are given, respectively, by the following:

$$\overline{\widehat{\rho}}_l = \frac{1}{r} \sum_{i=1}^r \widehat{\rho}_{kl}, MSE(\overline{\widehat{\rho}}_l) = \frac{1}{r} \sum_{i=1}^r (\widehat{\rho}_{kl} - \rho_l)^2. \tag{43}$$

This simulation was conducted for $n = 50, 100, 200,$ and 500 and censor rate at 20% and 80%

The simulation procedure is as follows;

- (1) We generate a random sample of size $n = 50, 100, 200, 500$ from the beta distribution B_i .
- (2) From (1), the observations x_1, x_2, \dots, x_n following the BTIGHLD are as follows:

$$L = \psi \ln \left(\frac{2}{\sqrt[4]{1 - B_i}} - 1 \right) \text{ (see Appendix).} \tag{44}$$

- (3) Generate the simple sample of the censoring times c_1, c_2, \dots, c_n from a BTIGHLD and adjust the parameters of the BTIGHLD to obtain the desired censoring rates.
- (4) To get the right censored data, this is obtained from the minimum value of censoring time and survival time, that is,

$$\begin{aligned} y_i &= \min(x_i, c_i), \\ \delta_i &= I(x_i \leq c_i), \\ i &= 1, 2, \dots, n. \end{aligned} \tag{45}$$

- (5) The observed data set is $D = (y_i, \delta_i), i = 1, 2, \dots, n$.
- (6) Replicate the values from (5) r -times for the different sample sizes to be considered. For this simulation, we take $r = 1,000$.
- (7) Based on the dataset from (6), we can get the maximum likelihood estimates of the parameters.
- (8) The likelihood function of the model is maximized with respect to parameters ψ, q, a, b to obtain $\widehat{\psi}, \widehat{q}, \widehat{a}, \widehat{b}$.

- (9) .If $\widehat{\rho}_{kl}$ is a MLE of ρ_l , such that $l = 1, 2, 3, 4$ (i.e. $\rho_1 = a, \rho_2 = b, \rho_3 = q, \rho_4 = \psi$), based on sample sizes $k, k = 1, \dots, r$. All simulation results are summarized in Tables 3 and 4.

Clearly, from Tables 3 and 4, they show that the values of the average estimates decrease as the sample size increases for all parameters across the two censoring proportions. It indicates that MSE reduces as samples size increases. This indicated efficiency and consistency of the estimates of the parameters.

6. Application of the BTIGHLD to COVID-19 Data

In this section, we considered a data that were culled from information released by the Mexican Ministry of Health (Secretaría de Salud, SS) through the Epidemiological Surveillance System for Viral Respiratory Diseases on COVID-19. Some of the confirmed cases of COVID-19 registered from February 21st to 18th February, 2021, were used in the present analysis. The database included all positive, negative, and suspected cases of COVID-19 registered by 475 Viral Respiratory Disease Monitoring Units (Unidades Monitoras de Enfermedad Respiratoria viral; USMER by its Spanish acronym) and by the medical units that attended the cases. The data can be found using the link [18]. For this research, the variables considered for every subject includes: sex, age, immunosuppression, pneumonia, and asthma. The survival time(in days) is taken as time to death, which is the time difference between the date on admission and the date of last contact. The survival time was right censored at 20 days. The Kaplan–Meier method was used to plot survival curves. These graphs served to test the proportional hazard assumption. The BTIGHLD as a parametric model was used to fit the data with the covariates under consideration. The result obtained was compared with that of the cox proportional regression model using the Akaike Information Criterion(AIC) and the Corrected AIC(CAIC). The significance level for the contributions of the covariate was set at $\tau = 0.05$. The data were analyzed with the statistical package software Statistical Packages for Social Sciences SPSS version 20.0 and R 4.0.4.

After analysis, Figure 1 shows the histogram of the survival times of the patients, with the fit of the BTIGHLD. It reveals that the data are rightly skewed which from [14] had shown that the BTIGHLD can give good fit.

Table 5 gives the descriptive statistics of the survival time.

The graph of the distribution of the age of the patients as at the time of report is in Figure 2. It shows that most patients were between 55 and 60 years.

Furthermore, Figure 3 reveals the KM plot of the survival data.

For variables with dichotomous responses, we plotted KM curves to compare their survival rates. For pneumonia, Figure 4 shows the probability of survival is higher in those who do not have pneumonia than those who have it.

TABLE 3: Results of the simulation of BTIGHLD (2.5, 0.5, 3.0) at 80% censoring.

Sample size	Parameter	AE	MSE
50	\hat{a}	3.265	3.725
	\hat{b}	0.618	4.628
	\hat{q}	2.924	3.884
100	\hat{a}	3.038	3.167
	\hat{b}	0.564	4.606
	\hat{q}	2.651	3.526
200	\hat{a}	2.231	2.026
	\hat{b}	0.563	4.590
	\hat{q}	2.218	3.016
500	\hat{a}	1.878	1.803
	\hat{b}	0.546	4.434
	\hat{q}	1.971	2.674

TABLE 4: Results of simulation of BTIGHLD (0.35, 5.0, 1.8) at 20% censoring.

Sample size	Parameter	AE	MSE
50	\hat{a}	0.424	0.444
	\hat{b}	6.961	5.195
	\hat{q}	2.567	1.044
100	\hat{a}	0.281	0.295
	\hat{b}	5.531	4.339
	\hat{q}	1.668	0.947
200	\hat{a}	0.242	0.234
	\hat{b}	4.028	3.391
	\hat{q}	1.496	0.821
500	\hat{a}	0.210	0.166
	\hat{b}	2.367	2.767
	\hat{q}	1.075	0.603

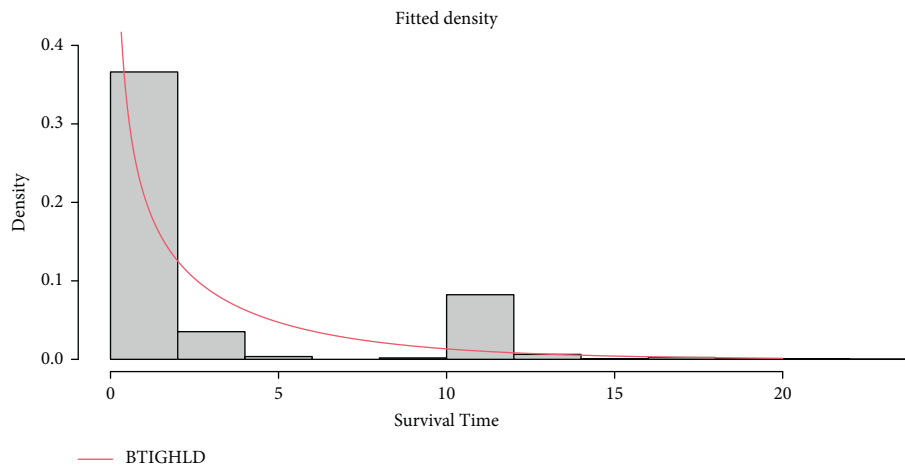


FIGURE 1: Histogram with fitted density.

TABLE 5: Descriptive statistics of COVID-19 survival data.

Min	1st qu.	Median	Mean	3rd qu.	Max.
1	2	6	20	16	168

Also, in Figure 5 shows that those that do not have asthma have a higher chance of survival than those who have asthma.

Figure 6 shows that those who have immunosuppression and those who do not have almost equal chance of survival.

Figure 7 shows that the probability of survival in male is higher than that of the female.

Analyzing the survival time using the BTIGHLD to assess the contribution of the prognostic factors to survival of the COVID-19 patients, we have the result in Table 6 in which it revealed that pneumonia, age, asthma, and

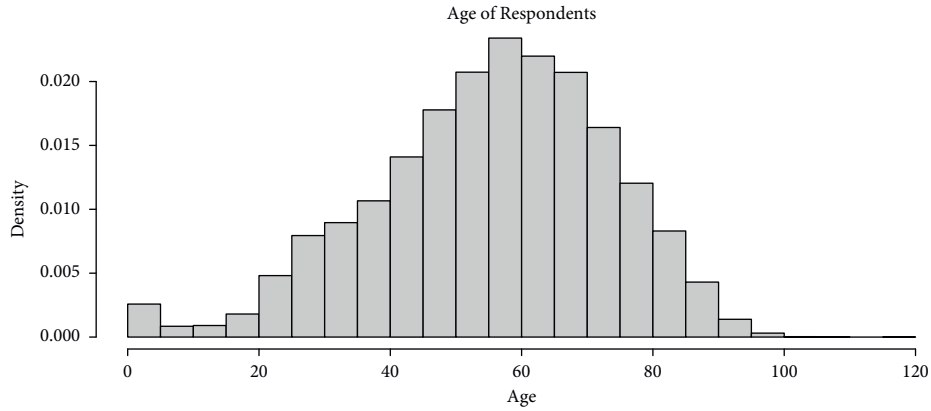


FIGURE 2: Histogram of the distribution of age.

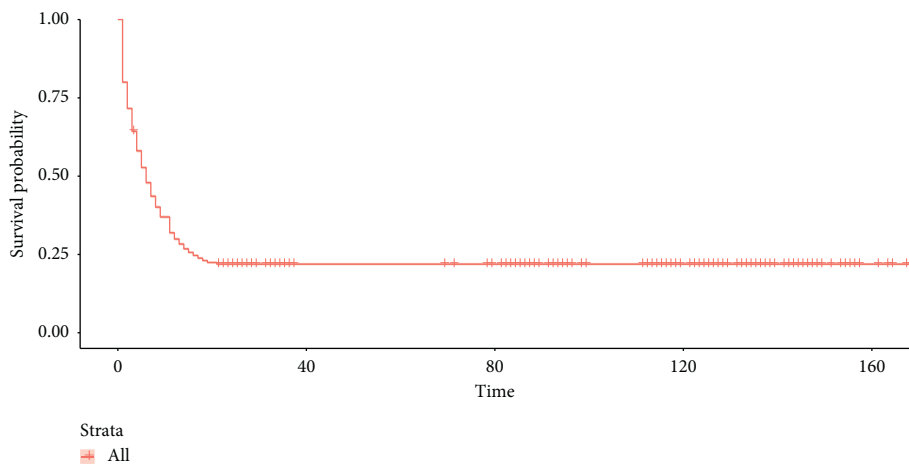


FIGURE 3: KM plot of survival time.

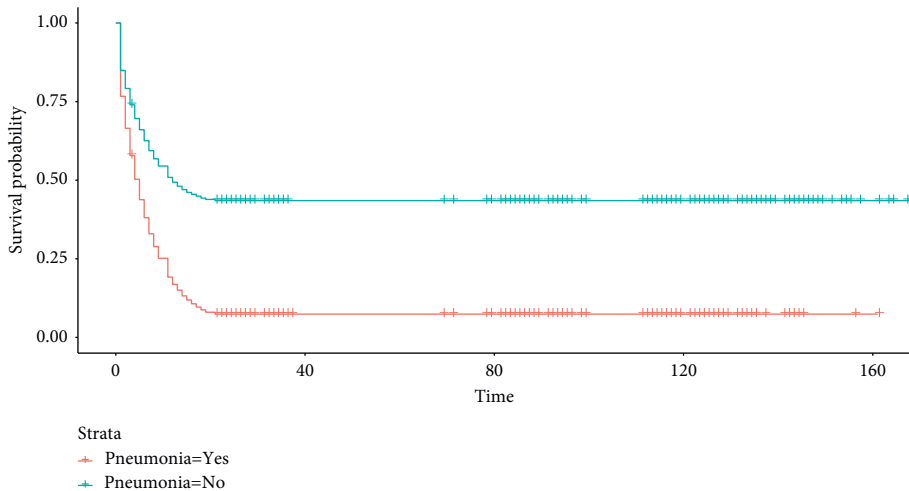


FIGURE 4: KM plot for pneumonia.

immunosuppression are significant, with a gender insignificant at level 0.05.

In this study, it was discovered that about two third of respondents had pneumonia comorbidity followed by

immunosuppression. This is not farfetched as the symptoms of COVID-19 are very similar to those of pneumonia. There could even be a transposition of symptoms, more patients presenting with pneumonia-like symptoms. This finding is

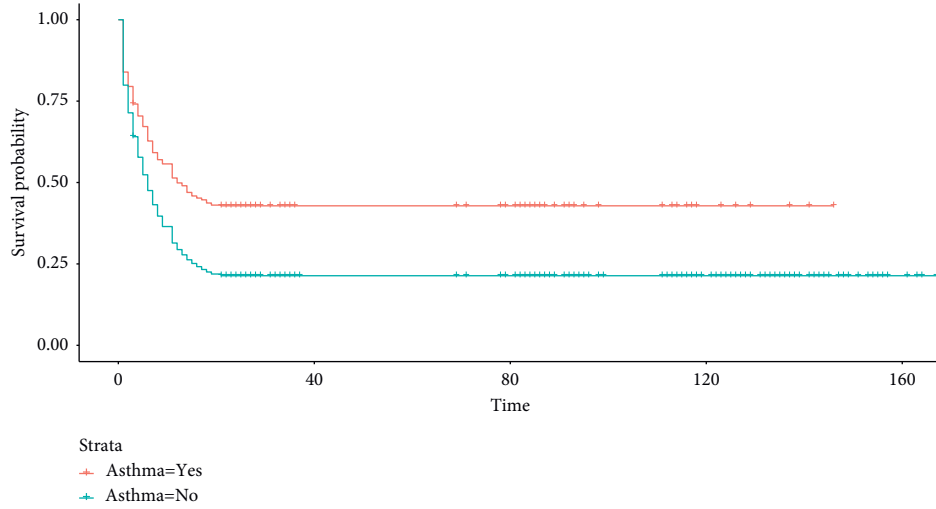


FIGURE 5: KM plot for asthma.

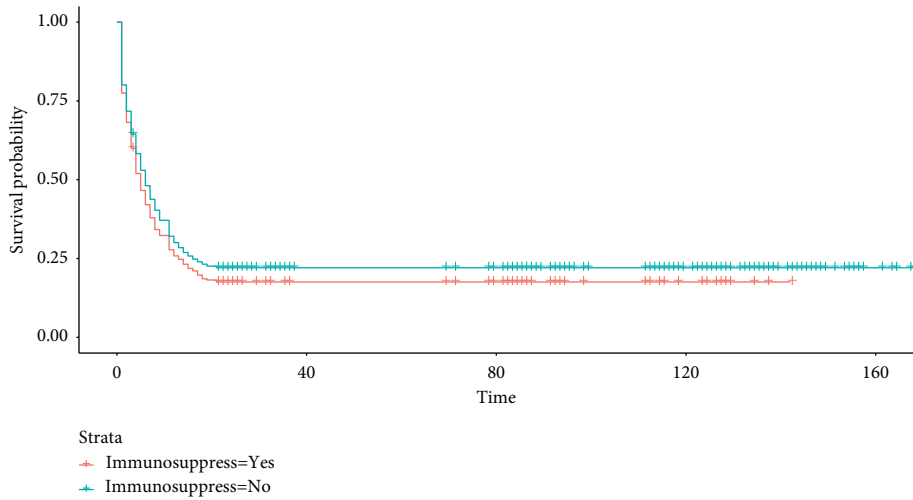


FIGURE 6: KM plot for immunosuppression.

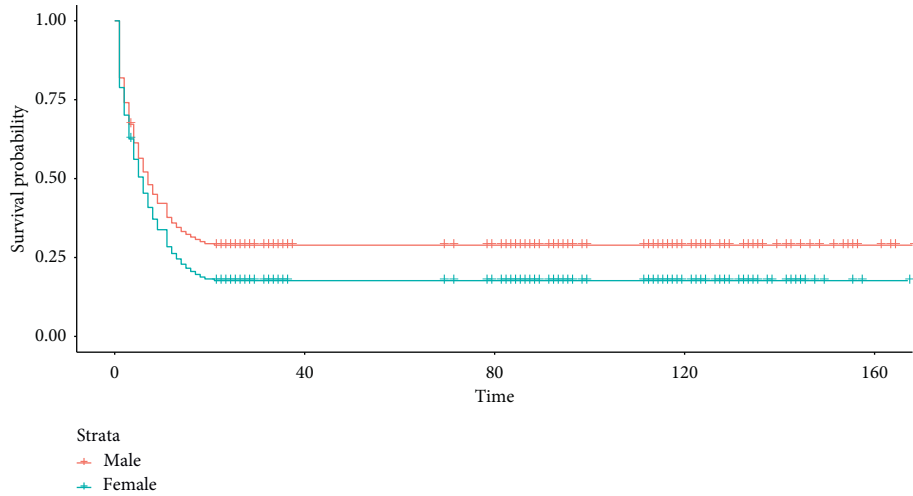


FIGURE 7: KM plot for gender.

TABLE 6: Results for analysis of COVID-19 data.

Parameter	Estimate	SError	Z	p
Sex	0.008	0.016	0.504	0.614
Pneumonia	0.072	0.017	4.133	0.000
Age	0.004	0.0005	9.498	0.000
Asthma	0.705	0.032	21.804	0.000
Immuno	0.511	0.029	17.512	0.000

similar to another study done by Hernandez–Vasquez where pneumonia is one of the presenting complaints of COVID-19 patients.

As also found in this study, asthma comorbidity with COVID-19 is minute, which is comparable to another study conducted in Saudi Arabia where only 1.7 of respondents had asthma. However, it was concluded in the study that patients with pneumonia-like symptoms could develop a life-threatening situation to COVID-19. Furthermore, pneumonia, asthma, and immunosuppression were found to statistically influence the survival times of patients, right from the day of admission till day of last contact. This is in tandem with another study done in Mexico by [19] where there was a statistically significant incremental gradient between COVID-19 and pneumonia as well as death among respondents with comorbidity like pneumonia and asthma. Also, immunosuppression was found out in the same study to increase the likelihood of dying from COVID-19. This is not farfetched as comorbidity as well as immunosuppression weakens the immune system of individuals, thereby making it difficult for immune responses to be mounted against diseases that should not have been able to overcome the patient. It has been postulated that most comorbidities with COVID-19 are associated with the ACE-2 receptor expression subsequently leading to a copious release of pro-protein convertase that ultimately increases the entry of the virus into the cells as found by the author of [20].

7. Conclusion

In this work, some properties of the BTIGHLD have been further studied which include the moments, moment generating function, and order statistics. Furthermore, the BTIGHLD was studied under right-censored observations, both in presence and absence of covariates. Using the maximum likelihood estimation method, the properties of estimates of the model were discussed using simulation studies with various parameters under different censoring rates. It was revealed from the result that the estimates performed well due to reducing values of average estimates coupling with the mean square error as the sample sizes increase. We further applied this model to COVID-19 data [8, 18, 21, 22].

Appendix

The AFT model is of the following form:

$$\ln x = \nu + d\beta + \epsilon, \tag{A.1}$$

where ϵ is said to follow a particular distribution, β' s are the estimates of the covariates x , and ν is the intercept of the model. Let $Y = \ln x$.

$$\begin{aligned} Y &= \nu + d\beta + \epsilon, \\ Y - d\beta &= \nu + \epsilon. \end{aligned} \tag{A.2}$$

From the definition of survival function with covariates,

$$\begin{aligned} s(x) &= \Pr(X > x|d), \\ s_o(x) &= \Pr(X > x|d = 0). \end{aligned} \tag{A.3}$$

The above equation is the baseline survival function

$$s_o(x) = \Pr(\exp(\nu + \epsilon) > x). \tag{A.4}$$

Now,

$$\begin{aligned} s(x) &= \Pr(X > x|d), \\ s(x) &= \Pr(Y > \ln x|d), \\ s(x) &= \Pr(Y - d'\beta > \ln x - d'\beta|d), \\ s(x) &= \Pr(\exp(Y - d'\beta) > \exp(\ln x - d'\beta)|d), \\ s(x) &= \Pr(\exp(\nu + \epsilon) > x \exp(-d'\beta)|d), \\ s(x|d, \beta) &= s_o(x \exp(-d'\beta)). \end{aligned} \tag{A.5}$$

From the relationship between probability distribution function and survival function,

$$\begin{aligned} f(x|d, \beta) &= -\frac{d}{dt} s(x|d, \beta), \\ f(x|d, \beta) &= \frac{d}{dt} s_o(x \exp(-d'\beta)), \\ f(x|x, \beta) &= \exp(-d'\beta) f_o(x \exp(-d'\beta)). \end{aligned} \tag{A.6}$$

From the relationship between hazard function, probability distribution function, and survival function,

$$\begin{aligned} h(x|d, \beta) &= \frac{f(x|d, \beta)}{s(x|d, \beta)}, \\ h(x|d, \beta) &= \frac{\exp(-d\beta f_o(x \exp(-d\beta)))}{s_o(x \exp(-d\beta))}, \\ h(x|d, \beta) &= \exp(-d\beta) h_o(x \exp(-d\beta)). \end{aligned} \tag{A.7}$$

Let $g(d|\beta) = \exp(-d\beta)$ and $Z(x) = x \exp(-d\beta)$, thus,

$$\begin{aligned} h(x|d, \beta) &= g(d|\beta) h_o(Z(x)), \\ f(x|d, \beta) &= g(d|\beta) f_o(Z(x)), \\ s(x|d, \beta) &= s_o(Z(x)). \end{aligned} \tag{A.8}$$

Let

$$X \in \text{Beta}(a, b) \tag{A.9}$$

Then,

$$\begin{aligned}
 X &= 1 - \left(\frac{2}{1 + e^{Q/\psi}} \right)^q, \\
 \left(\frac{2}{1 + e^{Q/\psi}} \right) &= 1 - X, \\
 \left(\frac{2}{1 + e^{Q/\psi}} \right) &= (1 - X)^{1/q}, \\
 1 + e^{Q/\psi} &= \frac{2}{(1 - X)^{1/q}}, \quad (\text{A.10}) \\
 e^{Q/\psi} &= \frac{2}{(1 - X)^{1/q}} - 1, \\
 \frac{Q}{\psi} &= \ln \left(\frac{2}{(1 - X)^{1/q}} - 1 \right), \\
 Q &= \psi \ln \left(\frac{2}{(1 - X)^{1/q}} - 1 \right).
 \end{aligned}$$

From the Wolfram statistics,

$$\frac{dB(k; a, b)}{dk} = (1 - k)^{b-1} k^{a-1}, \quad (\text{A.11})$$

where $k = 1 - (2/1 + e^{x/\psi})$ Therefore,

$$\frac{dB(k; a, b)}{dq} = \frac{dB(k; a, b)}{dk} \frac{dk}{dq}. \quad (\text{A.12})$$

This gives

$$\begin{aligned}
 B'(k; a, b) &= \frac{dB(k; a, b)}{dq} = \left(\frac{2}{1 + e^{x/\psi}} \right)^{q(b-1)} \left(1 - \left(\frac{2}{1 + e^{x/\psi}} \right)^q \right)^{a-1} \left(\frac{2}{1 + e^{x/\psi}} \right)^q + \ln \left(\frac{2}{1 + e^{x/\psi}} \right), \\
 \frac{dB(k; a, b)}{da} &= B_k(a, b) \log(k) - k^a \Gamma(a) {}_3\tilde{F}_2(a, a, 1 - b; a + 1, a + 1; k), \\
 \frac{d^2B(k; a, b)}{da^2} &= 2\Gamma(a)^2 (\Gamma(a) {}_4\tilde{F}_3(a, a, a, 1 - b; a + 1, a + 1, a + 1; k) \\
 &\quad - {}_3\tilde{F}_2(a, a, 1 - b; a + 1, a + 1; k) \log(k)) k^a + B(k; a, b) \log^2(k), \quad (\text{A.13}) \\
 \frac{dB(k; a, b)}{db} &= \Gamma(b)^2 (1 - k) {}_3\tilde{F}_2(b, b, 1 - a; b + 1, b + 1; 1 - k) \\
 &\quad - {}_3\tilde{F}_2(b, b, 1 - a, b + 1, b + 1; 1 - k) - B(1 - k; b, a) \log(1 - k) + (\phi(b) \\
 &\quad - \phi(a + b)) B(a, b), \frac{dB'(k; a, b)}{db^2} = B(a, b) 0
 \end{aligned}$$

Data Availability

The data are available using this link: https://drive.google.com/file/d/115uEhRYUyvQqTYcy_-f2mB2y7GRiS_R/view?usp=sharing.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

This manuscript is supported by Digiteknologian TKI-ymparisto project A74338(ERDF, Regional Council of Pohjois-Savo)

References

- [1] R. Elmorsy and A. Hassan, "Parameter estimation for inverted exponentiated Lomax distribution with right censored data," *Gazi University Journal of Science*, vol. 32, no. 4, pp. 1370–1386, 2019.
- [2] M. Kayid, "One generalized mixture pareto distribution and estimation of the parameters by the EM algorithm for complete and right-censored data," *IEEE Access*, vol. 9, Article ID 149372, 2021.
- [3] Q. Yu, "The MLE of the uniform distribution with right-censored data," *Lifetime Data Analysis*, vol. 27, no. 4, pp. 662–678, 2021.
- [4] U. Usman, S. Suleiman, B. M. Arkilla, and Y. Aliyu, "Nadarajah-haghighi model for survival data with long term

- survivors in the presence of right censored data,” *Pakistan Journal of Statistics and Operation Research*, pp. 695–709, 2021.
- [5] M. S. Eliwa, M. El-Morshedy, and M. Ibrahim, “Inverse Gompertz distribution: properties and different estimation methods with application to complete and censored data,” *Annals of data science*, vol. 6, no. 2, pp. 321–339, 2019.
- [6] F. Jamal, C. Chesneau, and K. Aidi, “The sine extended odd Fréchet-G family of distribution with applications to complete and censored data,” *Mathematica Slovaca*, vol. 71, no. 4, pp. 961–982, 2021.
- [7] A. K. Olapade, “On characterizations of the half logistic distribution,” 2003, <http://interstat.stat.vt.edu/InterStat/ARTICLES/2003articles/F06002.pdf>.
- [8] A. K. Olapade, “The type I generalized half logistic distribution,” *JIRSS*, vol. 13, no. 1, pp. 69–82, 2014.
- [9] A. K. Olapade, “On a four-parameter type I generalized half logistic distribution,” *Proceedings of the Jangieon Mathematical Society*, vol. 14, no. 2, pp. 189–198, 2011.
- [10] P. O. Awodutire, A. K. Olapade, and O. A. Kolawole, “The type I generalized half logistic survival model,” *International Journal of Theoretical and Applied Mechanics*, vol. 2, no. 2, pp. 74–78, 2016.
- [11] P. Awodutire, A. Olapade, O. Kolawole, and O. Ilori, “Assessing survival times of breast cancer patients using type I generalized half logistic survival model,” *Journal of Advances in Medicine and Medical Research*, vol. 25, no. 2, pp. 1–7, 2018.
- [12] B. Olalekan Akanji, I. Sule, P. O. Awodutire, and A. K. Olapade, “A study of five parameter type I generalized half logistic distribution,” *Pure and Applied Mathematics Journal*, vol. 6, no. 6, pp. 177–181, 2017.
- [13] S. Ibrahim, O. A. Bello, A. P. Oluwatobi, and H. O. Lawal, “The characteristic property of five parameter type II generalized logistic distribution,” *International Journal of Statistical Distributions and Applications*, vol. 6, no. 4, pp. 71–74, 2020.
- [14] P. O. Awodutire, E. C. Nduka, and M. A. Ijomah, “Lehmann type II generalized half logistic distribution: properties and applications,” *Mathematical Theory and Modelling*, vol. 10, no. 2, pp. 103–115, 2020.
- [15] P. Awodutire, E. Nduka, and M. Ijomah, “The beta type I generalized half logistic distribution: properties and application,” *Asian Journal of Probability and Statistics*, vol. 6, no. 2, pp. 27–41, 2020b.
- [16] H. Torabi and F. Bagheri, “Estimation of parameters for an extended generalized half logistic distribution based on complete and censored data,” *JIRSS*, vol. 9, no. 2, pp. 171–195, 2010.
- [17] P. O. Awodutire, E. Nduka, M. Ijomah, O. R. Ilori, and O. S. Balogun, “Inference on the beta type I generalized half logistic distribution under right censored observation with application to covid-19,” 2021, <https://ssrn.com/abstract=3933572>.
- [18] A. Hernández-Vásquez, D. Azañedo, R. Vargas-Fernández, and G. Bendezu-Quispe, “Association of comorbidities with pneumonia and death among COVID-19 patients in Mexico: a nationwide cross-sectional study,” *J Prev Med Public Health*, vol. 53, no. 4, pp. 211–219, 2020.
- [19] H. Ejaz, A. Alsrhani, A. Zafar et al., “COVID-19 and comorbidities: deleterious impact on infected patients,” *Journal of infection and public health*, vol. 13, no. 12, pp. 1833–1839, 2020.
- [20] Gob, “Bases de datos COVID-19 [Internet],” 2021, <https://www.gob.mx/salud/documentos/datos-abiertos-152127>.
- [21] X. Shang and H. K. T. Ng, “On parameter estimation for the generalized gamma distribution based on left-truncated and right-censored data,” *Computational and Mathematical Methods*, vol. 3, no. 1, Article ID e1091, 2021.
- [22] K. Wu, L. Wang, L. Yan, and Y. Lio, “Statistical inference of left truncated and right censored data from marshall–olkin bivariate Rayleigh distribution,” *Mathematics*, vol. 9, no. 21, p. 2703, 2021.