International Journal of Mathematical Analysis and Modelling

(Formerly Journal of the Nigerian Society for Mathematical Biology)

Volume 7 Issue 2, 2024

ISSN (Print): 2682 - 5694
ISSN (Online): 2682 – 5708
www.tnsmb.org

Half Logistic Generalized Inverse Rayleigh distribution for modeling Breast Cancer data

O.I. Oseghale*† and E. S. Micah†

Abstract

This article introduced a three-parameter extension of the Generalized Inverse Rayleigh distribution called half-logistic Generalized Rayleigh distribution, which has sub-models the Generalized Rayleigh and Rayleigh distribution. The proposed model is quite flexible and adaptable to model any kind of life-time data. Its probability density function may sometimes be unimodal and its corresponding hazard rate may be of monotone or non-monotone shape. Standard statistical properties such as it ordinary and incomplete moments, quantile function, moment generating function, reliability function, Lorenz and Bonferroni curves. The maximum likelihood method is used to obtain the estimates of the model parameters. An application of the newly developed model to breast cancer data is presented.

Keywords and phrases: maximum likelihood estimation**;** monotone shape; entropy

1 Introduction

Over the last century, many significant distributions have been introduced to serve as models in applied sciences. Among them, the so-called generalized beta distribution introduced by [18] is at the top of the list in terms of usefulness. The main feature of the generalized beta distribution is to be very rich; It includes a plethora of named distributions (to our knowledge, over thirty). In this paper, we focus our attention on one of the most attractive of these distributions, known as the inverse Lomax (IL) distribution. [22] introduced the inverse Rayleigh (IR) distribution as a model for analyzing reliability and survival data. The model underwent further investigation by [23], he observed that the lifetime distributions of various experimental units could be closely approximated with the inverse

* Corresponding author: innocentoseghale@gmail.com
† Federal University, Otuoke, Bayelsa State, Nigeria

Rayleigh distribution. Additionally, Voda explored its properties and provided a maximum likelihood (ML) estimator for the scale parameter. [12] carried an in-depth analysis of the inverse Rayleigh distribution. [17] developed a percentile estimator for the scale parameter θ of one-parameter inverse Rayleigh distribution and investigated its asymptotic efficiency. [14] established Bayesian prediction bounds for both Rayleigh and inverse Rayleigh lifetime models. [20] addressed both Bayesian and non-Bayesian issues related to parameter estimation in the IR distribution [3] propose a two-parameter extension of the inverse Rayleigh distribution, employing the half-logistic transformation to address limitations in modeling moderately right-skewed or near-symmetrical lifetime data. Their theoretical contributions encompass mathematical properties and empirical evidence, demonstrating the model's effectiveness in handling diverse right-skewed datasets.

The work in [6] present the inverse Rayleigh probability distribution as a robust model for estimating extreme wind speeds, crucial in wind power production and mechanical safety assessment. [7] introduce the compound inverse Rayleigh distribution as a model tailored for extreme wind speeds, essential in wind power generation and turbine safety evaluation. They provide a useful practical framework for real-world data analysis, accompanied by a novel Bayesian estimation approach, supported by extensive numerical simulations and robustness assessments. In a different context, [4] introduce the beta generalized inverse Rayleigh distribution, a four-parameter lifetime model, and conduct a comprehensive investigation into its properties and applications, further expanding the domain of inverse Rayleigh-based distributions.

Several generalizations of the inverse Rayleigh distribution have been recently proposed by numerous authors, with the aim of enhancing its adaptability. For example, [13] introduced an extension of the inverse Rayleigh distribution, termed the Burr XII inverse Rayleigh model, by integrating the Burr XII family framework initially introduced by [10]. [2] explored the use of the inverse Rayleigh distribution in mixture models to analyze the complex nature of engineering systems' lifetimes. [5] introduced the exponential transformed inverse Rayleigh distribution.

The generation of new distributions by adding one or more parameters to standard distributions enhances their applicability to complex data across various fields. Motivated by this approach, several authors have proposed different methods for generating new distributions. These include the Marshall–Olkin-G distribution by [15] the Beta-G distribution introduced by [11], the Kumaraswamy-G distribution by [8] and the McDonald-G distribution by [1]. [19] developed the transmuted-G class of distributions.

2 Methodology/Model formulation

The techniques for modifying the classical distributions are usually referred to as generators in literature and are capable of improving the goodness-of-fit of the modified distributions. These features have been established by the results of many generators ([8];[11]). Recently, [1] proposed an extension of the beta-generated family of distributions developed by [11]) and called it the transformed-transformer (T-X) family of distributions. According to [1] the CDF of the T-X family is defined as

$$G(x) = \int_{0}^{-log(1-F(t))} r(t)dt = R\{-log[1-F(t)]\} \tag{1}$$

Here, we define $r(t)$ as the probability density function (PDF) of the random variable T, R(t) is the CDF of T and $F(x)$ is the CDF of the random variable X. Although, the $T-X$ method have been embraced by several researchers, [24] proposed an extension of it in order to improve on some of the drawbacks of the $T-X$ method. The new family defined by [24] is called the exponentiated $T-X$ family of distributions.

In this work, our attention is focus on one of the most attractive of these distributions, known as the Generalized Rayleigh (GR) distribution. If the random variable X has a GIR distribution when its cumulative distribution function (CDF) can be represented by

$$F(x;\eta,\varphi) = 1-\left(1-e^{-(\theta x)^{-2}}\right)^{\varphi};x>0,\varphi,\theta>0 \tag{2}$$

The corresponding probability density function to equation (2) is given as

$$f(x;\eta,\varphi) = \frac{2\varphi}{\theta^{2}x^{3}}e^{-(\theta x)^{-2}}\left(1-e^{-(\theta x)^{-2}}\right)^{\varphi-1};x>0,\varphi,\theta>0 \tag{3}$$

where φ is a shape parameter and θ is a scale parameter. If $\varphi=1$, GIR reverses back to the well-known Inverse Rayleigh distribution. The main motivations behind the half logistic generated family are as follows. First, it is proved in [24] that the combined effect of the considered ratio and power transformations can significantly enrich the parental distribution, improving the flexibility of the mode, median, skewness and tail, with a positive impact on the analyses of practical data sets. However, the half logistic generated family remains underexploited and, based on the previous promising studies, needs more thorough attention.

2.1 Half Logistic Generalized Rayleigh distribution

Based on the work of [9], the CDF of half-logistic generalized family is defined by:

$$G(x;\lambda,\varsigma)=\frac{1-[1-F(x;\varsigma)]^{\lambda}}{1+[1-F(x;\varsigma)]^{\lambda}},\quad x\in R \tag{4}$$

where λ is a positive shape parameter and $F(x;\varsigma)$ represents a CDF of a continuous distribution, otherwise known as the baseline distribution. Here, ς represents a vector of parameter(s) related to the corresponding baseline distribution. The chief motivations behind the half-logistic family of distributions are as follows. [9] showed that the effect of the power transformation can enormously enrich the baseline distribution, improving the flexibility of the median, mode, skewness and tail, which provides a positive impact on the analyses of real-life data sets. The normal, Frechet, Weibull, and inverse Lomax distributions have been considered as baseline in the previous work using the half-logistic generated family by [12]. In recent studies, [29] developed and studied the type 1 half logistic Gompertz distribution, the Half-Logistic Generalized Weibull Distribution, type I half-logistic Topp-Leone distribution was proposed and studied by [29]. [30] proposed and studied the half logistic generalized Rayleigh distribution. However, the half logistic-g family has not been completely explored and, based on the studies carried out in the past, more work needs to be done to fully explore the richness of this generated family of distribution. In this paper, we consider the baseline distribution as Inverse Generalised Rayleigh (IGR) distribution.

Thus, we introduce a new flexible lifetime distribution with three parameters called the half-logistic Generalized Rayleigh ($HLGIR$) distribution. The CDF of the $HLGIR$ distribution with parameter vector $\zeta=(\theta,\varphi,\lambda)$ is obtained by inserting (3.1) into (3.3) as

$$G(x;\lambda,\varphi,\theta)=\frac{1-\left[1-\left(1-e^{-(\theta x)^{-2}}\right)^{\varphi}\right]^{\lambda}}{1+\left[\left[1-\left(1-e^{-(\theta x)^{-2}}\right)^{\varphi}\right]\right]^{\lambda}},\quad x\in R \tag{5}$$

The associated PDF to (5) is given by

$$g(x;\lambda,\varphi,\theta)=\frac{2\varphi e^{-(\theta x)^{-2}}\left(1-e^{-(\theta x)^{-2}}\right)^{\varphi-1}\left[1-\left(1-e^{-(\theta x)^{-2}}\right)^{\varphi}\right]^{\lambda-1}}{\theta^{2}x^{3}\left(1+\left[1-\left(1-e^{-(\theta x)^{-2}}\right)^{\varphi}\right]^{\lambda}\right)^{2}},\quad x\in R \tag{6}$$

where φ and λ are positive shape parameters and θ is a positive scale parameter.

The survival $(S(x))$, hazard $(h(x))$,and the cumulative hazard $(H(x))$ are, respectively, given as

$$S(x;\zeta) = 1 - G(x;\zeta) = \frac{2\left[1-\left(1-e^{-(\theta x)^{-2}}\right)^{\varphi}\right]^{\lambda}}{1+\left[1-\left(1-e^{-(\theta x)^{-2}}\right)^{\varphi}\right]^{\lambda}}, \qquad x \in R \tag{7}$$

and $S(x;\zeta) = 1$ for $x \leq 0$,

$$h(x;\zeta) = \frac{g(x;\zeta)}{S(x;\zeta)} = \frac{2\varphi e^{-(\theta x)^{-2}}\left(1-e^{-(\theta x)^{-2}}\right)^{\varphi-1}}{\theta^2 x^3\left[1-\left(1-e^{-(\theta x)^{-2}}\right)^{\varphi}\right]\left(1+\left[\left[1-\left(1-e^{-(\theta x)^{-2}}\right)^{\varphi}\right]\right]^{\lambda}\right)}, \tag{8}$$

and $h(x;\zeta) = 0$ for $x \leq 0$, and

$$\begin{aligned} H(x) &= -log[S(x;\zeta)] \\ &= \log(2) + \lambda log\left[1-\left(1-e^{-(\theta x)^{-2}}\right)^{\varphi}\right] - log\left[1+\left[\left[1-\left(1-e^{-(\theta x)^{-2}}\right)^{\varphi}\right]\right]^{\lambda}\right]. \end{aligned} \tag{8}$$

and $H(x;\zeta) = 0$ for $x \leq 0$. The graph of distribution, density, survival function(sf) and hazard function (hf) are respectively, given in Figures 3.1, 3.2, 3.3 and 3.4 for various hypothetical values of the parameters of the distributions.

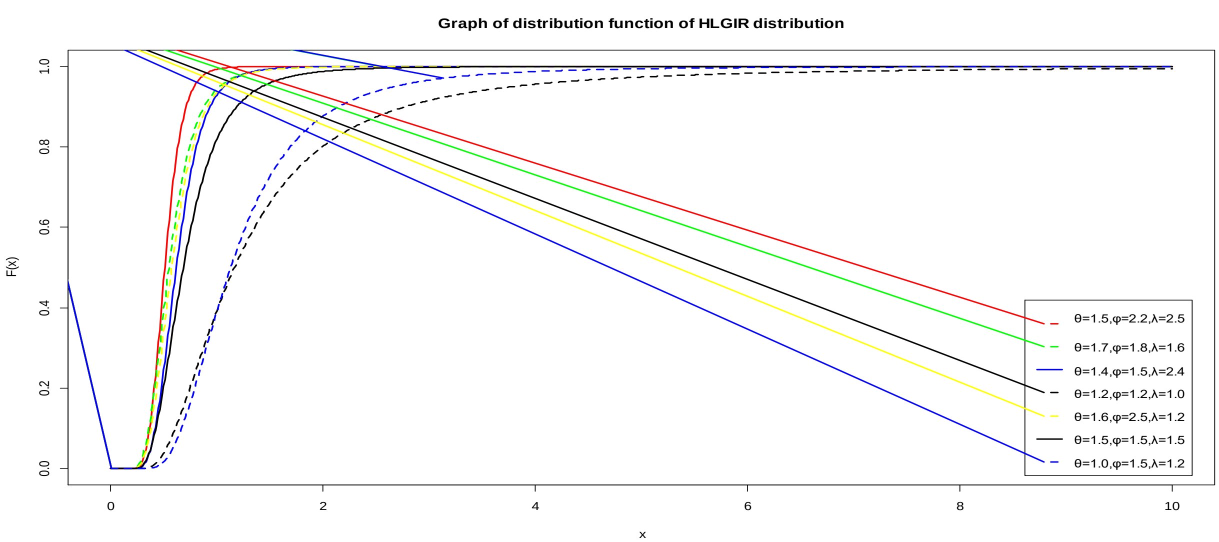


Figure 3.1. Graph of the CDF of $HLGIR$ distribution

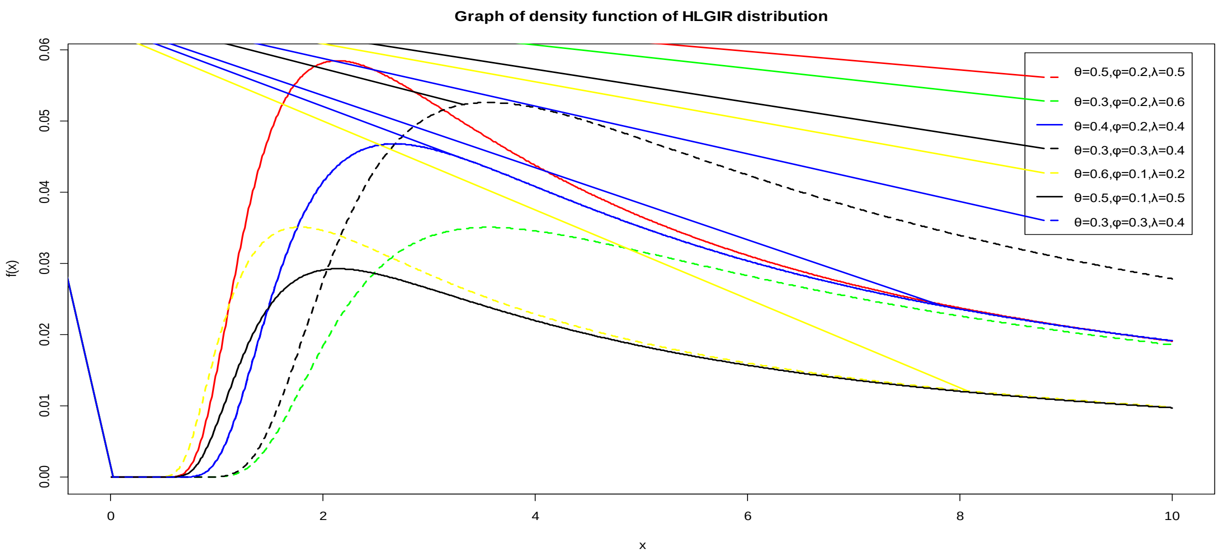


Figure 3.2. Graph of the CDF of $HLGIR$ distribution

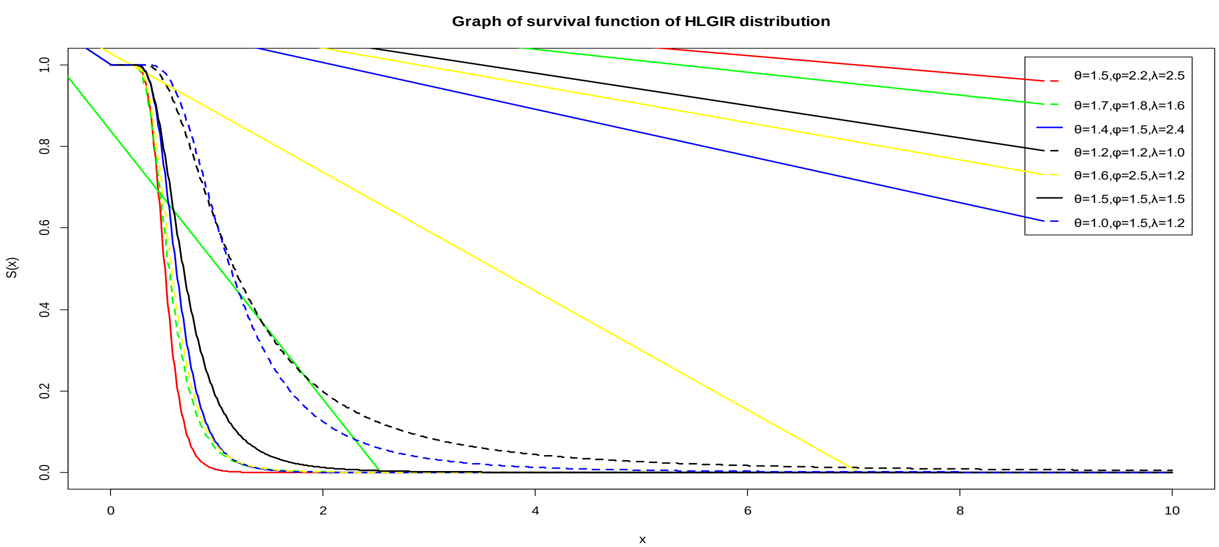
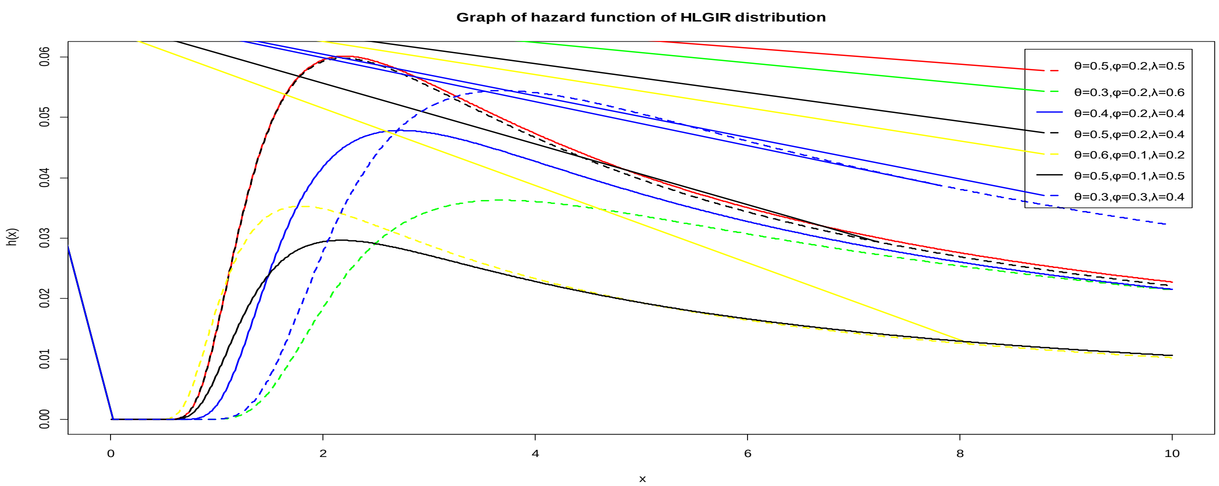


Figure 3.3. Graph of the survival function and the hazard function of $HLGIR$ distribution

- From figure 3.3 it can be observed that the failure rate function of the $HLGIR$ model exhibits the form of bathtub-shaped.



- ✓ The graph of the hazard function in figure 3.4 for various parameter values for $HLGIR$ distribution indicates that the new model exhibits various shapes such as monotonically increasing, decreasing, and upside-down bathtub shapes. This property clearly explains the flexibility of $HLGIR$ distribution and its applicability in modeling non-monotonic and monotonic hazard behavior which are often encountered in real-life data.

3.1 Important representation

In this subsection, an important tool for the expansion of the PDF for $HLGIR$ is provided. From the generalized binomial series given by

$$(1+v)^{-b} = \sum_{i=0}^{\infty} (-1)^i \binom{b+i-1}{i} v^i \tag{9}$$

For $|v| < 1$ and b is a positive real non-integer. Then, by applying the binomial theorem (9) in (6), the density function of $HLGIR$ distribution becomes

$$g(x) = \frac{2\varphi}{\theta^2} \sum_{i,j,k=0}^{\infty} (-1)^{i+j+k} \binom{i+1}{i} \binom{\lambda(i+1)-1}{j} \binom{\varphi(j+1)}{k} x^{-3} e^{-(k+1)(\varphi x)^{-2}} \tag{10}$$

3.1 Statistical characteristics of $HLGIR$ distribution

3.1.1 The quantiles, median and the upper quartile

An expression for the quantile and the median of $HLGR$ are obtained in this subsection.

The quantile x_q of the $HLGR$ is represented as follows

$$x_q = \frac{1}{\theta}\left\{-log\left(1-\left[1-\left(\frac{1-q}{1+q}\right)^{1/\lambda}\right]^{1/\varphi}\right)\right\}^{-1/2} \tag{11}$$

The median and the upper quartile of HLGIR are found by putting $q = 0.5$ and 0.75 in (11), respectively, as follows:

$$x_{0.5} = \frac{1}{\theta}\left\{-log\left(1-\left[1-\left(\frac{0.5}{1.5}\right)^{1/\lambda}\right]^{1/\varphi}\right)\right\}^{-1/2} \tag{12}$$

and

$$x_{0.75} = \frac{1}{\theta}\left\{-log\left(1-\left[1-\left(\frac{0.25}{1.25}\right)^{1/\lambda}\right]^{1/\varphi}\right)\right\}^{-1/2} \tag{13}$$

3.2 The ordinary and incomplete moments of $HLGIR$ distribution

Moments are the anticipated outcome of a particular function of a random variable. Due to the mathematical tractability of the various kinds of moments of $HLGIR$ distribution, the moments can be calculated directly. In a similar vein, Bonferroni and Lorenz curves, the mean waiting time, and the mean residual life can be computed using the first incomplete moment. Thus, the r^{th} ordinary moment of a distribution is given by

$$\mu_r' = E(x)^r$$

Thus the r^{th} moment of $HLGIR$ distribution is given by

$$\mu_r' = \int_{-\infty}^{\infty} x^r f(x;,\theta\lambda)dx \tag{14}$$

Putting equation (10) in (14), we have

$$\mu_r' = \frac{2\varphi}{\theta^2}\sum_{i,j,k=0}^{\infty}(-1)^{i+j+k}\binom{i+1}{i}\binom{\lambda(i+1)-1}{j}\binom{\varphi(j+1)}{k}\int_{-\infty}^{\infty} x^{r-3}e^{-(k+1)(\varphi x)^{-2}}dx \tag{15}$$

Letting $m=(k+1)(\varphi x)^{-2}$, $x=\frac{1}{\varphi}(k+1)^{1/2}m^{-1/2}$ $dx=-\frac{1}{2\varphi}(k+1)^{1/2}m^{-3/2}dm$ and substitute it in (15), we have

$$\mu_r' = \frac{1}{\theta^2}\sum_{i,j,k=0}^{\infty}(-1)^{i+j+k}\binom{i+1}{i}\binom{\lambda(i+1)-1}{j}\binom{\varphi(j+1)}{k}\frac{(k+1)^{1-r/2}}{\varphi^{r-3}}\int_{0}^{\infty} m^{\frac{-r}{2}}e^{-m}dx$$

Finally, we have,

$$\mu_r' = \frac{1}{\theta^2}\sum_{i,j,k=0}^{\infty}(-1)^{i+j+k}\binom{i+1}{i}\binom{\lambda(i+1)-1}{j}\binom{\varphi(j+1)}{k}\frac{(k+1)^{1-r/2}}{\varphi^{r-3}}\Gamma\left(1-r/2\right) \tag{16}$$

Where $\Gamma(.)$ is a gamma function. It should be noted that the mean (μ_1') and the variance $(\sigma^2=\mu_2'-\{\mu_1'\}^2)$ of $HLGIL$ can respectively, be obtained by taking $r=1,2$.
The mean is given by

$$\mu_1' = \frac{1}{\theta^2}\sum_{i,j,k=0}^{\infty}(-1)^{i+j+k}\binom{i+1}{i}\binom{\lambda(i+1)-1}{j}\binom{\varphi(j+1)}{k}\frac{(k+1)}{\varphi^{-2}}\Gamma\left(1/2\right) \tag{17}$$

The variance is given as

$$\sigma^2 = \frac{1}{\theta^2}\sum_{i,j,k=0}^{\infty}(-1)^{i+j+k}\binom{i+1}{i}\binom{\lambda(i+1)-1}{j}\binom{\varphi(j+1)}{k}\frac{1}{\varphi^{-1}} - \left[\frac{1}{\theta^2}\sum_{i,j,k=0}^{\infty}(-1)^{i+j+k}\binom{i+1}{i}\binom{\lambda(i+1)-1}{j}\binom{\varphi(j+1)}{k}\frac{(k+1)}{\varphi^{-2}}\Gamma\left(1/2\right)\right]^2 \tag{18}$$

3.3 Incomplete Moments 0f $HLGIL$ distribution

Thus the r^{th} moment of $HLGIR$ distribution is given by

$$v_r' = \int_0^t x^r f(x;\varphi,\theta,\lambda)dx \tag{19}$$

Putting equation (10) in (19), we have

$$v_r' = \frac{2\varphi}{\theta^2}\sum_{i,j,k=0}^{\infty}(-1)^{i+j+k}\binom{i+1}{i}\binom{\lambda(i+1)-1}{j}\binom{\varphi(j+1)}{k}\int_0^t x^{r-3}e^{-(k+1)(\varphi x)^{-2}}dx \tag{20}$$

Letting $m=(k+1)(\varphi x)^{-2}$, $x=\frac{1}{\varphi}(k+1)^{1/2}m^{-1/2}$ $dx=-\frac{1}{2\varphi}(k+1)^{1/2}m^{-3/2}dm$ and substitute it in (20), we have

$$v_r' = \frac{1}{\theta^2}\sum_{i,j,k=0}^{\infty}(-1)^{i+j+k}\binom{i+1}{i}\binom{\lambda(i+1)-1}{j}\binom{\varphi(j+1)}{k}\frac{(k+1)^{1-r/2}}{\varphi^{r-3}}\int_0^t m^{\frac{-r}{2}}e^{-m}dx \tag{21}$$

Finally, we have,

$$v_r' = \frac{1}{\theta^2}\sum_{i,j,k=0}^{\infty}(-1)^{i+j+k}\binom{i+1}{i}\binom{\lambda(i+1)-1}{j}\binom{\varphi(j+1)}{k}\frac{(k+1)^{1-r/2}}{\varphi^{r-3}}$$
$$\times\Gamma\left(1-r/2,(k+1)(\varphi t)^{-2}\right) \tag{22}$$

The first incomplete moment is obtained by taking $r=1$, and we have

$$v_1' = \frac{1}{\theta^2}\sum_{i,j,k=0}^{\infty}(-1)^{i+j+k}\binom{i+1}{i}\binom{\lambda(i+1)-1}{j}\binom{\varphi(j+1)}{k}\frac{(k+1)}{\varphi^{-2}}\Gamma((k+1)(\varphi t)^{-2}) \tag{23}$$

3.4 Moments generating function

The moment generating function (MGF) of a random variable X provides an alternative method that can be used in describing the characteristics of a distribution. Mathematically, the MGF is defined as

$$\mathcal{M}_X(t) = E(e^{tX}) = \sum_{r=0}^{\infty} \frac{t^r}{r!} E(X^r). \tag{24}$$

Putting (23) in (24) for $E(X^r)$ for $HLGIL$ model, we obtain

$$\mathcal{M}_X(t) = \frac{1}{\theta^2} \sum_{i,j,k,r=0}^{\infty} \frac{t^r}{r!} (-1)^{i+j+k} \binom{i+1}{i} \binom{\lambda(i+1)-1}{j} \binom{\varphi(j+1)}{k} \frac{(k+1)^{1-r/2}}{\varphi^{r-3}} \Gamma\left(1 - {r}/{2}\right) \tag{25}$$

By setting $t = it$ in (25), we derived an expression for the characteristics function of the $HLGIL$ model as

$$\zeta_X(t) = \frac{1}{\theta^2} \sum_{i,j,k,r=0}^{\infty} \frac{t^r}{r!} (-1)^{i+j+k} \binom{i+1}{i} \binom{\lambda(i+1)-1}{j} \binom{\varphi(j+1)}{k} \frac{(k+1)^{1-r/2}}{\varphi^{r-3}} \Gamma\left(1 - {r}/{2}\right) \tag{26}$$

3.5 Mean deviation about the Mean and the Median

The mean deviation, about the mean and the median, are used to determine the degree of spread (scatter) in a population. Let μ = (X) and M be the median and the mean of the $HLGIL$ distribution given by (12) and (23) respectively

The mean deviation of $HLGIL$ model about the mean can be calculated as

$$\Gamma_1(X) = E|X-\mu| = \int_0^{\infty} |X-\mu| f(x;\varphi,\theta,\lambda)dx, \tag{27}$$

$$= 2\mu F(\mu; v, w, \lambda) - 2\mu + 2\int_{\mu}^{\infty} x f(x;\varphi,\theta,\lambda)dx$$

$$= 2\mu \left\{ \frac{1 - \left[1 - \left(1 - e^{-(\theta x)^{-2}}\right)^{\varphi}\right]^{\lambda}}{1 + \left[\left[1 - \left(1 - e^{-(\theta x)^{-2}}\right)^{\varphi}\right]\right]^{\lambda}} \right\} - 2\mu + \frac{2}{\theta^2} \sum_{i,j,k=0}^{\infty} (-1)^{i+j+k} \binom{i+1}{i} \binom{\lambda(i+1)-1}{j}$$

$$\times \binom{\varphi(j+1)}{k} \frac{(k+1)^{1-r/2}}{\varphi^{r-3}} \Gamma\left(1 - r/2, (k+1)(\varphi\mu)^{-2}\right) \quad (28)$$

The mean deviation about the median can also be calculated as

$$\Gamma_2(X) = E|X - M| = \int_0^\infty |X - M| f(x) dx, \quad (29)$$

$$= -\mu + 2\int_m^\infty x f(x; \varphi, \theta, \lambda) dx$$

$$= -\boldsymbol{\mu} + \frac{2}{\theta^2} \sum_{i,j,k=0}^{\infty} (-1)^{i+j+k} \binom{i+1}{i} \binom{\lambda(i+1)-1}{j} \binom{\varphi(j+1)}{k} \frac{(k+1)^{1-r/2}}{\varphi^{r-3}}$$

$$\Gamma\left(1 - r/2, (k+1)(\varphi m)^{-2}\right) \quad (30)$$

3.6 Bonferroni and Lorenz curves

The Bonferroni and Lorenz curves have been found suitable to have applications not only in economics to study income and poverty analysis, but also in other fields like demography, insurance, and reliability. The Bonferroni and Lorenz curves are defined by

$$B(p) = \frac{1}{p\mu} \int_0^q x f(x; \lambda, \varphi, \theta) dx \quad (31)$$

and

$$L(p) = \frac{1}{\mu} \int_0^q x f(x; \lambda, \varphi, \theta) dx \quad (32)$$

Respectively, where $\mu = E(X)$ and $q = F^{-1}(p)$. In the case of $HLGIR$ distribution, we obtain

$$B(p) = \frac{1}{\mu\theta^2 p}\sum_{i,j,k=0}^{\infty}(-1)^{i+j+k}\binom{i+1}{i}\binom{\lambda(i+1)-1}{j}\binom{\varphi(j+1)}{k}\frac{(k+1)}{\varphi^{-2}}\Gamma((k+1)(\varphi t)^{-2})$$

(33)

and

$$L(p) = \frac{1}{\mu\theta^2}\sum_{i,j,k=0}^{\infty}(-1)^{i+j+k}\binom{i+1}{i}\binom{\lambda(i+1)-1}{j}\binom{\varphi(j+1)}{k}\frac{(k+1)}{\varphi^{-2}}\Gamma((k+1)(\varphi t)^{-2})$$

(34)

4.0 Application of $HLGIR$ to Breast cancer data

Here, we demonstrate the flexibility of Half logistic Generalized Inverse Power Lomax distribution using two lifetime data sets. The first data represent 121 breast cancer patients' survival times during a specific period from 1929 to 1938. The data source [26] studied these datasets. The observations are listed as follows: 0.3, 0.3, 4.0, 5.0, 5.6, 6.2, 6.3, 6.6, 6.8, 7.4, 7.5, 8.4, 8.4,10.3, 11.0, 11.8, 12.2, 12.3, 13.5, 14.4, 14.4, 14.8, 15.5, 15.7,16.2, 16.3, 16.5, 16.8, 17.2, 17.3, 17.5, 17.9, 19.8, 20.4, 20.9,21.0, 21.0, 21.1, 23.0, 23.4, 23.6, 24.0, 24.0, 27.9, 28.2, 29.1,30.0, 31.0, 31.0, 32.0, 35.0, 35.0, 37.0, 37.0, 37.0, 38.0, 38.0,38.0, 39.0, 39.0, 40.0, 40.0, 40.0, 41.0, 41.0, 41.0, 42.0, 43.0, 43.0, 43.0, 44.0, 45.0, 45.0, 46.0, 46.0, 47.0, 48.0, 49.0, 51.0, 51.0, 51.0, 52.0, 54.0, 55.0, 56.0, 57.0, 58.0, 59.0, 60.0, 60.0, 60.0, 61.0, 62.0, 65.0, 65.0, 67.0, 67.0, 68.0, 69.0, 78.0, 80.0, 83.0, 88.0, 89.0, 90.0, 93.0, 96.0, 103.0, 105.0, 109.0, 109.0, 111.0, 115.0, 117.0, 125.0, 126.0, 127.0, 129.0, 129.0, 139.0,154.0. Then the fit of the Half logistic Generalized Inverse Rayleigh model is compared with those of some competitive model such as Half logistic Inverse Rayleigh distribution (HLIR) by [27] and the inverse Rayleigh model. The maximum likelihood estimation procedure was used to determine the estimates of the parameters of the models considered. We compute statistical measures which represent the measures of goodness of fits and this includes Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), Consistent Akaike Information Criterion (CAIC), and Hana Quin Information Criterion (HQIC). The model with the smallest AIC, BIC, CAIC, and HQIC is assumed to be the best fit model in the class of models considered. Table 4.1 shows that the breast cancer data is over-dispersed, positively skewed. Figure 4.1 shows that the Breast cancer data contains no outlier observation and it has a wider spread.

Table 4.1 Exploratory data analysis of Breast cancer data

$Min.$	q_1	$median$	$mean$	q_3	max	$Range$	$Skewness$	$Kurtosis$	$variance$
0.0004	0.0095	0.7405	0.9765	1.4122	3.8474	3.8470	1.0432	3.4021	1244.46

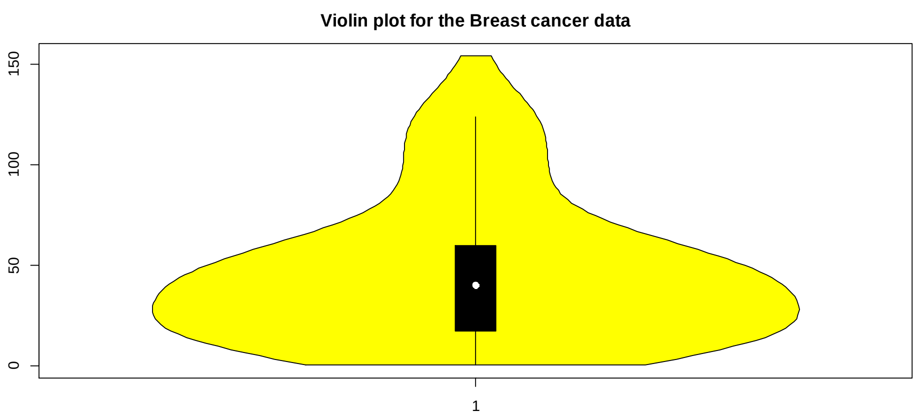


Figure 4.1. Violin plot of Breast cancer data

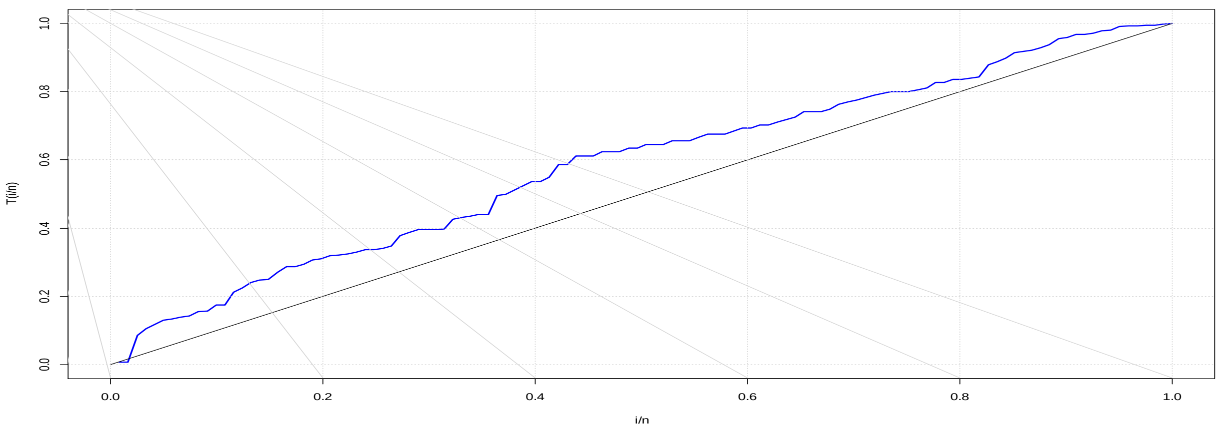


Figure 4.2. **Total time on test (TTT) Plot for Breast cancer data.**

Table 4.2 **Measures of goodness of fit of Breast cancer data**

	Parameter estimate			*Measures of goodness of fit*				
$Model$	θ	φ	λ	$-l$	AIC	$CAIC$	$HQIC$	BIC
$HLGIR$	1.213 (0.012)	3.569 (0.012)	0.003 (0.005)	276.76	559.52	559.72	562.92	567.90
$HLIR$	7.070 (5.897)	− (−)	0.0015 (0.0117)	419.60	843.20	843.29	845.47	848.79
IR	5.338 (0.485)	− (−)	− (−)	1092.55	2187.10	2187.12	2188.23	2189.89

From Table 4.2, we observe that $HLGIR$ model has a better fit than its existing sub-model models which includes $HLIR$, and IR model because it possesses the smallest $AIC, CAIC, HQIC$, and BIC. This result shows that the $HLGIR$ is more suitable for modeling breast cancer data than other models considered in the study.

5 Discussions and conclusion

It is evident in this study that the new model developed show the potentiality of a process or phenomenon under consideration. In the field of statistics many of these models were developed through generalization, combination, or through modification of existing models to develop new one. In this research, we have presented a new generalized distribution called the Half Logistic Generalized Inverse Rayleigh distribution. The new distribution was developed using the Half logistic generalized family of distributions introduced in the literature. Some mathematical properties of the new developed generalized distribution were studied which includes: Renyi entropy, Bonferroni and Lorenz curve, ordinary moments and incomplete moments, moment generating function, mean deviations, quantile function, and also the parameters of the developed model are estimated using the maximum likelihood estimation procedure. An application of the newly developed model to breast cancer data demonstrates its attractiveness in modeling lifetime data.

References

[1] Alexander, C.; Cordeiro, G.M.; Ortega, E.M.M.; Sarabia, J.M. (2012). Generalized beta-generated distributions. Comput. Stat. Data Anal. 2, 56, 1880–1897.

[2] Ali, S. (2015). Mixture of the inverse Rayleigh distribution: Properties and estimation in a Bayesian framework. Appl. Math. Model., 39, 515–530.

[3] Almarashi, A.M.; Badr, M.M.; Elgarhy, M.; Jamal, F.; Chesneau, C. (2020). Statistical inference of the half-logistic inverse Rayleigh distribution. Entropy, 22, 449.

[4] Bakoban, R.A.; Al-Shehri, A.M. (2021). A new generalization of the generalized inverse Rayleigh distribution with applications. Symmetry, 13, 711.

[5] Banerjee, P.; Bhunia, S. (2022). Exponential transformed inverse Rayleigh distribution: Statistical properties and different methods of estimation. Austrian J. Stat., 51, 60–75.

[6] Chiodo, E.; Noia, L.P.D. (2020). Stochastic extreme wind speed modeling and bayes estimation under the inverse Rayleigh distribution. Appl. Sci., 10, 5643.

[7] Chiodo, E.; Fantauzzi, M.; Mazzanti, G. (2022). The compound inverse Rayleigh as an extreme wind speed distribution and its bayes estimation. Energies, 15, 861.

[8] Cordeiro, G. M. and de Castro, M. (2011). A new family of generalized distributions. Journal of Statistical Computation and Simulation, 81(7):883–898

[9] Cordeiro, G. M., Alizadeh, M. and Marinho, M.R.P.D. (2015). The type I half-logistic family of distributions, Journal of Statistical Computation and Simulation, 86, 707-728.

[10] Cordeiro, G.M.; Yousof, H.M.; Ramires, T.G.; Ortega, E.M. (2018). The Burr XII system of densities: Properties, regression model and applications. J. Stat. Comput. Simul., 88, 432–456.

[11] Eugene, N., Lee, C., and Famoye, F. (2002). The beta-normal distribution and its applications. Communications in Statistics-Theory and Methods, 31(4):497–512.

[12] Gharraph, M. K. (1993). Comparison of estimators of location measures of an inverse Rayleigh distribution. Egypt. Stat. J., 37, 295–309.

[13] Goual, H.; Yousof, H.M. (2020). Validation of Burr XII inverse Rayleigh model via a modified chi-squared goodness-of-fit test. J. Appl. Stat. 2020, 47, 393–423.

[14] Howlader, H.A.; Hossain, A.M.; Makhnin, O. (2009). Bayesian prediction bounds for Rayleigh and inverse Rayleigh lifetime models. J. Appl. Stat. Sci., 17, 131.

[15] Marshall, A.W. and Olkin, I. (1997). A new method for adding a parameter to a family of distributions with application to the exponential and Weibull families. Biometrika, 84, 641–652.

[16] Mudholkar, G. S. and Srivastava, D. K. (1993). Exponentiated Weibull family for analyzing bathtub failure-rate data. IEE Transactions on Reliability, 42:299–302.

[17] Mukherjee, S.P.; Maiti, S.S. (1996). A percentile estimator of the inverse Rayleigh parameter. IAPqR Trans. 1996, 21, 63–66.

[18] McDonald, J.B.; Xu, Y.J. (1995). A generalization of the beta distribution with applications. J. Econom., 66, 133–152.

[19] Shaw, W.T.; Buckley, I.R. (2009). The alchemy of probability distributions: Beyond Gram-Charlier expansions, and a skew-kurtotic-normal distribution from a rank transmutation map. arXiv, arXiv:0901.0434

[20] Soliman, A.; Essam, A.; Amin, A.; Abd-El Aziz, A.A. (2010). Estimation and Prediction from Inverse Rayleigh Distribution Based on Lower Record Values. Appl. Math. Sci., 4, 3057–3066.

[22] Treyer, V.N. (1964). Inverse Rayleigh (IR) model. Proc. USSR Acad. Sci.

[23] Voda, V.G. (972). On the inverse Rayleigh distributed random variable. Rep. Statis. App. Res. JUSE, 19, 13–21.

[24] Alzaatreh, A. Lee, C. and Famoye, F. 2013). A new method for generating families of continuous distributions. Metron, 71, 63-79 (2013).

[25] Fayomi, A. (2019). Type I half logistic power Lomax distribution: Statistical properties and application. Adv. Appl. 54, 85–98.

[26] Tahir, M.H.; Cordeiro, G.M.; Mansoor, M.; Zubair, M. (2013). The Weibull-Lomax distribution: Properties and applications. Hacet. J. Math. Stat., 44, 461–480.

[27] Al-Marzouki, S.A. New Generalization of Power Lomax Distribution. (2018). Int. J. Math. Appl. 7, 59–68.

[28] Anwar,M.; Bibi, A. (2018). The Half-Logistic Generalized Weibull Distribution. J. Probab. Stat.,12, 8767826.

[29] Ogunde, A. A., Oseghale, O. I. and Audu, A. T. (2017). Performance Rating of the Type Half Logistic Gompertz Distribution: An Analytical Approach, American Journal of Mathematics and Statistics, Vol. 7 (3), 93-98.

[30] Ogunde, A. A., Subhankar, D. and Almetawally, E. M. (2024). Half logistic Generalized Distribution for modeling Hydrological Data. Annals of Data Science, Springer-Verlag GmbH Germany.