American Journal of Applied Mathematics and Statistics, 2021, Vol. 9, No. 3, 90-101
Available online at http://pubs.sciepub.com/ajams/9/3/3
Published by Science and Education Publishing
DOI:10.12691/ajams-9-3-3

Normal-Power Function Distribution with Logistic Quantile Function: Properties and Application

Matthew Iwada Ekum[1], Obalowu Job[2], Jimoh Taylor[1], Asimi Amalare[1], Mundher Abdullah Khaleel[3], Adeyinka Solomon Ogunsanya[2,*]

[1]Department of Mathematics & Statistics, Lagos State Polytechnic Lagos, Nigeria
[2]Department of Statistics, University of Ilorin, Ilorin, Kwara State, Nigeria
[3]College of Computer Sciences and Mathematics, Tikrit University, Tikrit, Iraq
*Corresponding author: ogunsanyaadeyinka@yahoo.com

Received August 04, 2021; Revised September 08, 2021; Accepted September 17, 2021

Abstract Developing compound probability distributions is very important in the field of probability and statistics because there are different datasets from different fields with different features. These features range from high skewness, peakedness (kurtosis), bimodality, highly dispersed, and so on. Existing distributions might not easily fit well to these emerging data of interest. So, there is a need to develop more robust and flexible distributions that are positively skewed, negatively skewed, and bathup shape, to handle some of these features in the emerging data of interest. This paper, therefore, proposed a new four-parameter distribution called the Normal-Power{logistic} distribution. The proposed distribution was characterized by its density, distribution, survival, hazard, cumulative hazard, reversed hazard, and quantile functions. Properties such as the *r*-th moment, heavy tail property, stochastic ordering, mean inactive time were obtained. A useful transformation of the proposed distribution to normal distribution was shown to help generate its quantiles. The method of Maximum Likelihood Estimation (MLE) was used to estimate the model parameters. A simulation study was carried out to test the consistency of the maximum likelihood parameter estimates. The result of the simulation shows that the biases reduce as the sample size increases for different parameter values. The importance of the new distribution was proved empirically using a real-life dataset of gauge lengths of 10mm. The proposed distribution was compared with five other competing distributions, and the results show that the proposed Normal-Power{logistic} distribution (NPLD) performed favourably than the other five distributions using the AIC, CAIC, BIC, HQIC criteria.

Keywords: *power function distribution, normal distribution, logistic distribution, Normal-Power{logistic} distribution, T-R{Y} framework*

Cite This Article: Matthew Iwada Ekum, Obalowu Job, Jimoh Taylor, Asimi Amalare, Mundher Abdullah Khaleel, and Adeyinka Solomon Ogunsanya, "Normal-Power Function Distribution with Logistic Quantile Function: Properties and Application." *American Journal of Applied Mathematics and Statistics*, vol. 9, no. 3 (2021): 90-101. doi: 10.12691/ajams-9-3-3.

1. Introduction

In the field of statistics and probability, the power function distribution is unimodal, skewed, and peaked depending on the shape parameter value [1,2]. In literature, many known probability distributions such as exponential, Weibull, gamma, and lognormal are used for modeling lifetime data and widely applied in a variety of studies because of their advantages in different situations [3], especially using them as baseline distribution in the T-R{Y} family [4,5,6]. The power function despite its flexibility as a lifetime distribution that arises in several scientific fields has not been considered in this regard [7].

The power function distribution is a special model that can be formed or related to the uniform, Weibull, Kumaraswamy distributions. The power function distribution is considered one of the simplest and handy lifetime distributions. It is a special case of the beta distribution and one may sight the importance of the distribution in statistical tests such as the likelihood ratio test. The simplicity and usefulness of the power function distribution compelled the researchers to explore its further extensions, generalizations, and applications in different areas of science [8]. It is the inverse of Pareto distribution [9]. Estimation of the power parameters has been done by various authors, for instance; [10,11] proposed the two-parameter power function distribution as a simple alternative to the exponential distribution when it comes to modeling failure data related to mortality rate and component failures. The Weibull power function distribution was suggested by [12]. [13] studied the properties of the Odd Lomax-Exponential distribution. [1] proposed the modified power function distribution.

The probability density function (pdf) and cumulative distribution function (cdf) of two-parameter power function distribution with scale parameter λ and shape parameter k are, respectively, given as

$$f(x)=\frac{kx^{k-1}}{\lambda^k};\lambda,k>0,0\le x\le\lambda$$

and

$$F(x)=\frac{x^k}{\lambda^k}$$

The normal distribution is another lifetime distribution with its various advantages. Many authors have worked on the normal distribution and it has huge applications in various fields. Also, many authors have developed and studied convoluted distributions with normal distribution parameters (μ and σ). These authors include but are not limited to Exponentiated-Normal proposed and studied by [14], Beta-Normal distribution by [15], Gamma-Normal (GN) distribution by [16], Kumaraswamy-Normal distribution by [17], Weibull-Normal by [18], Exponentiated-Generalized-Normal (EGN) distribution by [19], Weibull-Normal{log-logistics} distribution by [20], Lomax-Cauchy {Uniform} by [21], Rayleigh Cauchy distribution by [22] and Weibull-Inverse Rayleigh distribution: Classical/ Bayesian approach by [23].

Many classical distributions have been extensively used for modeling real data in many areas. However, in many situations; there is a clear need for extended forms of these distributions to improve the flexibility and goodness of fit of these distributions. For that reason, generated families of continuous distributions are developed by introducing one or more additional shape parameter(s) to the baseline distribution or by combining two or more distributions to produced new ones. [24] described such new distributions as convoluted distributions.

In this paper, the aim is to develop a novel univariate continuous probability distribution called the normal-power function {logistic} distribution {NPLD}. It is derived from the T-Power function {Y} family proposed and studied by [7]. The distribution has four parameters, two from the normal distribution and the other two from the power function distribution. The scope covers different characterizations of the NPLD are to be established such as its density, distribution, survival, hazard, cumulative hazard, reversed hazard, and quantile function. To achieve this aim, different properties of the proposed distribution such as the r-th moment, heavy tail property, stochastic ordering, mean inactive time are to be obtained. The relationship between the proposed distribution and normal distribution is to be established. Thus, the proposed distribution can be positively skewed, negatively skewed, symmetric, and has a bathtub shape. The method of Maximum Likelihood Estimation (MLE) is used to estimate the model parameters. The importance of the new distribution will be proved empirically using a real-life dataset of gauge lengths of 10mm and compared with other convoluted distributions with normal distribution parameters (μ and σ). The proposed distribution will be very useful in engineering, medicine, and all filed of life. It is expected to perform well when normal distribution fails to fit the data of interest.

The rest of the paper contains the following sections. The formation of the pdf and cdf of the NPLD is performed in Section 2. Some statistical properties of NPLD distribution are provided in Section 3. Estimation of parameters, simulation study, and application to real data are presented in Section 4. The article ends with concluding remarks in section 5.

2. Derivation of Normal - Power{logistic} Distribution

Ekum et al. [7] defined the cdf and pdf of T-Power {logistic} family as

$$F_X(x)=F_T\left[\log\left(\frac{x^k}{\lambda^k-x^k}\right)\right] \quad (1)$$

and

$$f_X(x)=\left[\frac{k^k}{x\left(\lambda^k-x^k\right)}\right]f_T\left[\log\left(\frac{x^k}{\lambda^k-x^k}\right)\right] \quad (2)$$

respectively, where T can follow any probability distribution that has the same support with the logistic distribution. The quantile function of the logistic distribution is used, which is the log of the odd function of the power function distribution.

In this research, let T follow normal distribution with parameters μ and σ, i.e. $T\sim N(\mu,\sigma)$. This proposed distribution is called the Normal-Power Function Distribution.

Let $T\in(-\infty,\infty)$ be any random variable. Here, T follows the normal distribution with cdf and pdf given by

$$F_T(x)=\Phi\left(\frac{x-\mu}{\sigma}\right)=\frac{1}{2}\left[1+erf\left(\frac{x-\mu}{\sigma\sqrt{2}}\right)\right] \quad (3)$$

and

$$f_T(x)=\frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{1}{2\sigma^2}(x-\mu)^2\right]; \quad (4)$$
$$\sigma>0,\mu\in\Re,-\infty\le x\le\infty$$

respectively.

By substituting equation (3) into (1) we have

$$F_X(x)=\Phi\left[\frac{\log\left(\frac{x^k}{\lambda^k-x^k}\right)-\mu}{\sigma}\right]$$
$$=\frac{1}{2}\left\{1+\text{erf}\left[\frac{\log\left(\frac{x^k}{\lambda^k-x^k}\right)-\mu}{\sigma\sqrt{2}}\right]\right\}. \quad (5)$$

where

$$\operatorname{erf}(x)=\frac{2}{\sqrt{\pi}}\int_{0}^{x}e^{-t^{2}}dt$$

is called the error function.

So, equation (5) is the cdf of the proposed normal-power function {logistic} distribution (NPLD) and the corresponding pdf is derived by substituting (4) into (2) and we have

$$f_X(x)=\frac{k^{k}}{x\left(^{k}-x^{k}\right)\sqrt{2\pi\sigma^{2}}}\times\exp\left\{-\frac{1}{2\sigma^{2}}\left[\log\left(\frac{x^{k}}{^{k}-x^{k}}\right)-\mu\right]^{2}\right\} \qquad (6)$$

$$\lambda,k,\sigma>0,\left(0\le\mu,x\le\lambda\right)$$

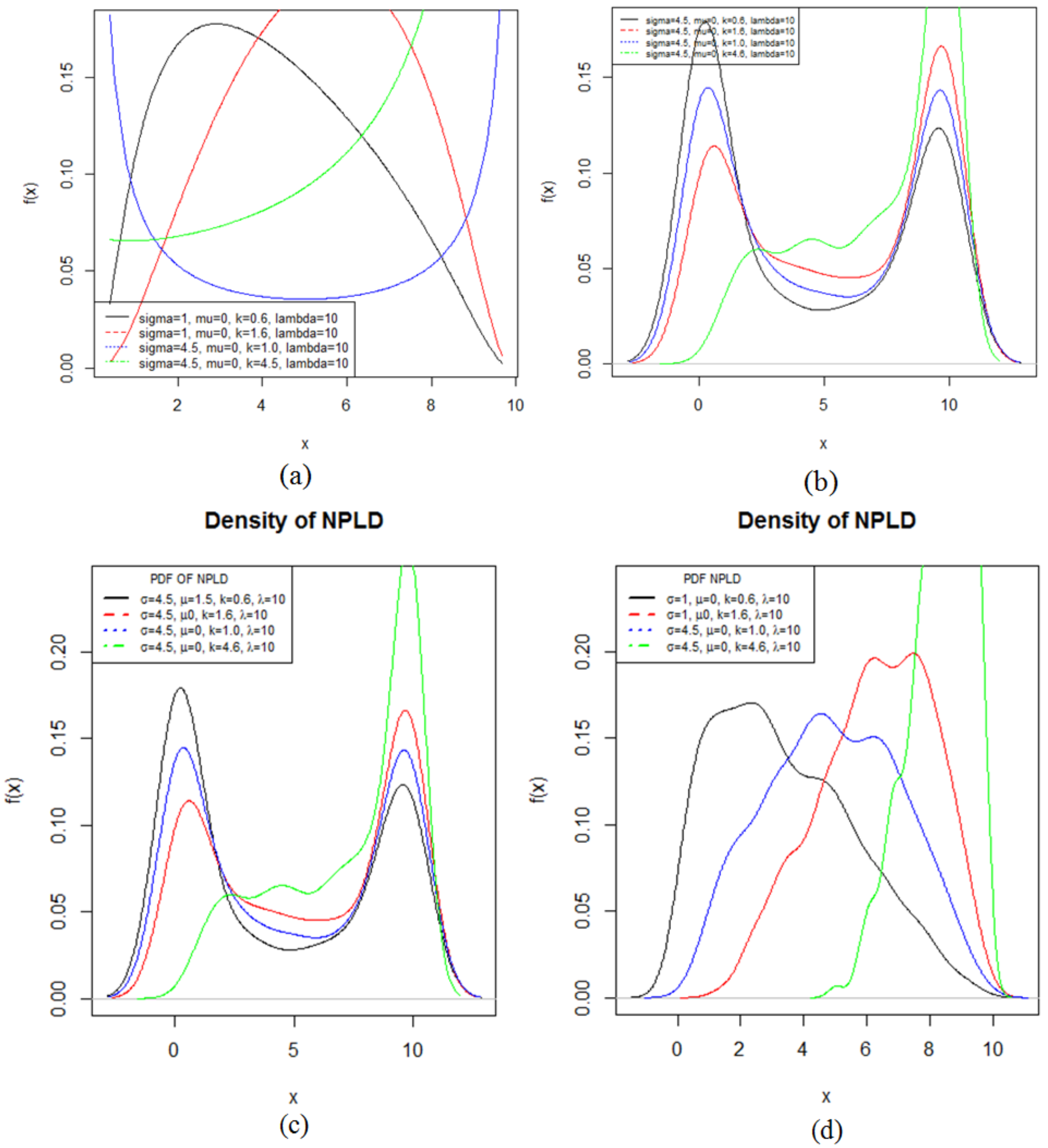

Figure 1. PDFs of NPLD for various values of μ, σ, k and λ

Equation (6) can be written as

$$f_X(x)=k\lambda^{k}x^{-1}\left(\left(\lambda^{k}-x^{k}\right)\sqrt{2\pi\sigma^{2}}\right)^{-1}\exp\left\{-\frac{1}{2\sigma^{2}}\left[\log\left(\frac{x^{k}}{\lambda^{k}-x^{k}}\right)-\mu\right]^{2}\right\}.$$

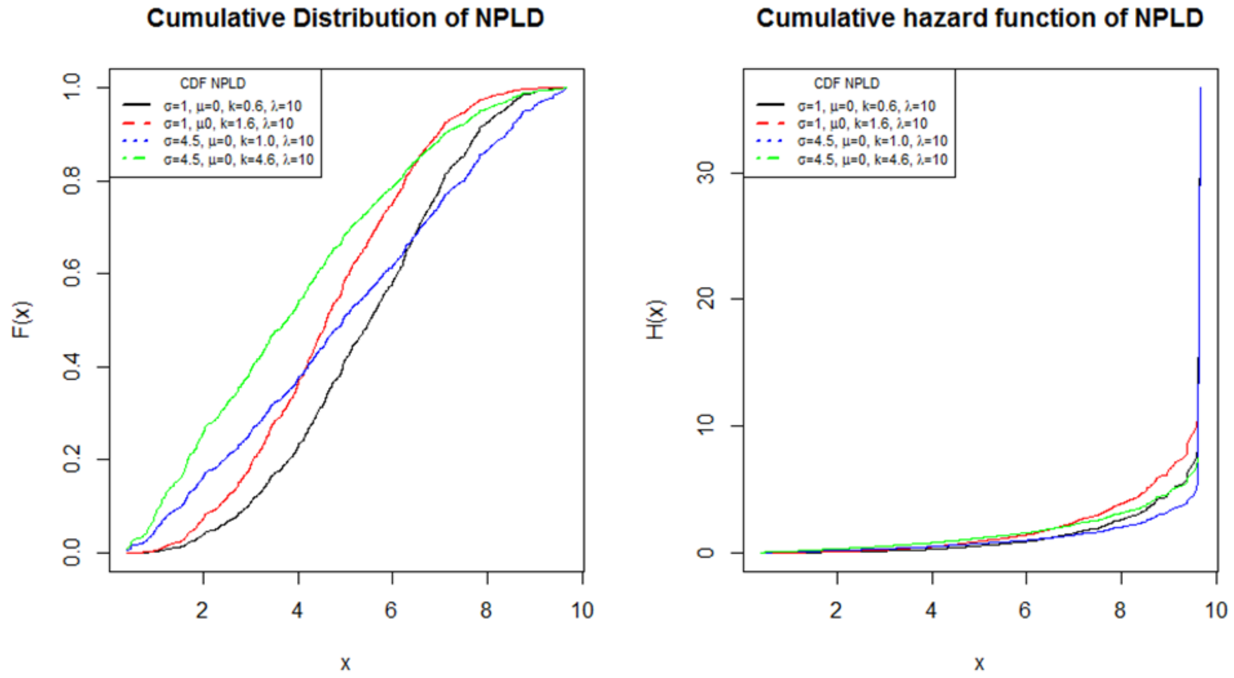

Figure 2. CDFs and Cum. Hazard Functions of NPLD for various values of μ, σ, k and λ

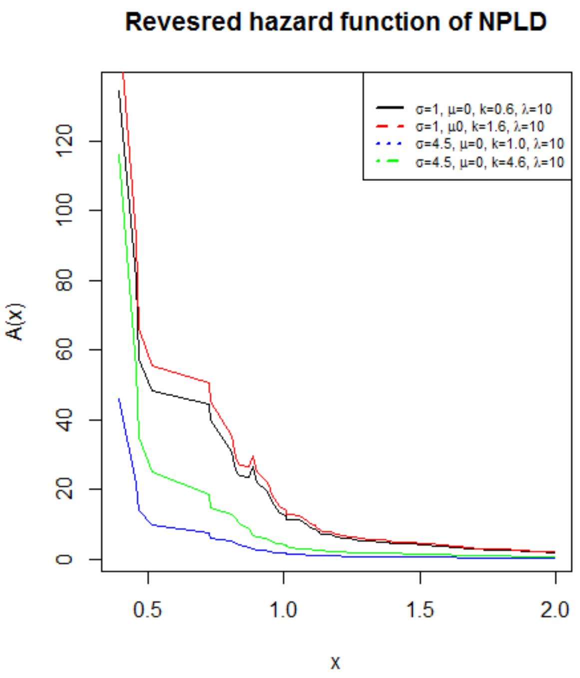

Figure 3. Reversed Hazard Functions of NPLD for various values of μ, σ, k and λ

The NPLD is a novel four-parameter distribution. Parameters λ and σ are scale parameters, k is a shape parameter, while μ is a location parameter. The pdf plots in Figure 1 show that NPLD can be positively skewed, negatively skewed, symmetric (see also the histogram in Figure 4), and can have a bath up shape, depending on the parameter values. The cdf plots in Figure 2 shows that for increasing values of x, the cdfs increase as well, increasing from zero and flattened at lambda since lambda is an upper bound. Figure 2 shows the cumulative hazard plots, which rise from zero slowly and asymptotic to the vertical axis at lambda. The reversed hazard decreases with increasing values of x and it is asymptotic to both axes. The reversed hazard plot is adjusted to display on x values less than 2 for clearer curves.

2.1. Properties of NPLD

This section considered some major properties of the new proposed distribution.

2.1.1. Survival, Hazard, Cumulative Hazard, and Reversed Hazard Distributions

The survival, hazard, cumulative hazard, and the reversed hazard functions are given in equations (7 – 10) respectively.

$$S_X(x) = \frac{1}{2}\left\{1 - \operatorname{erf}\left[\frac{\log\left(\frac{x^k}{\lambda^k - x^k}\right) - \mu}{\sigma\sqrt{2}}\right]\right\} \tag{7}$$

$$h_X(x) = \frac{2k\lambda^k \exp\left\{-\frac{1}{2\sigma^2}\left[\log\left(\frac{x^k}{\lambda^k - x^k}\right) - \mu\right]^2\right\}}{\sigma x\left(\lambda^k - x^k\right)\sqrt{2\pi}\left\{1 - \operatorname{erf}\left[\frac{\log\left(\frac{x^k}{\lambda^k - x^k}\right) - \mu}{\sigma\sqrt{2}}\right]\right\}} \tag{8}$$

$$H_X(x) = -\log\left\{\frac{1}{2} - \frac{1}{2}\operatorname{erf}\left[\frac{\log\left(\frac{x^k}{\lambda^k - x^k}\right) - \mu}{\sigma\sqrt{2}}\right]\right\} \tag{9}$$

$$A_X(x) = \frac{2k\lambda^k \exp\left\{-\frac{1}{2\sigma^2}\left[\log\left(\frac{x^k}{\lambda^k - x^k}\right) - \mu\right]^2\right\}}{\sigma x\left(\lambda^k - x^k\right)\sqrt{2\pi}\left\{1 + \text{erf}\left[\frac{\log\left(\frac{x^k}{\lambda^k - x^k}\right) - \mu}{\sigma\sqrt{2}}\right]\right\}} \quad (10)$$

2.2. Useful Transformation

Lemma 1. If X follows the NPLD with parameters μ, σ, k, λ then random variable $Y = \log\left(\frac{X^k}{\lambda^k - X^k}\right)$ follows the normal distribution with parameters μ and σ.

Proof.

Let

$$y = \log\left(\frac{x^k}{\lambda^k - x^k}\right) \quad (11)$$

Differentiating equation (11) with respect to x gives

$$\frac{dy}{dx} = \frac{k^k}{x\left(\lambda^k - x^k\right)}$$

$$dx = \frac{x\left(\lambda^k - x^k\right)dy}{k\lambda^k}$$

$$\int g(y)\,dy = \int \frac{k^k}{x\left(\lambda^k - x^k\right)\sqrt{2\pi\sigma^2}} \times \exp\left[-\frac{1}{2\sigma^2}(y-\mu)^2\right]\frac{x\left(\lambda^k - x^k\right)dy}{k\lambda^k}$$

$$\int g(y)\,dy = \int \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{1}{2\sigma^2}(y-\mu)^2\right]dy$$

Therefore, the pdf of random variable Y is given by

$$g(y) = \frac{1}{\sqrt{2\pi\sigma^2}}\exp\left[-\frac{1}{2\sigma^2}(y-\mu)^2\right] \quad (12)$$

Equation (12) completes the proof.

2.3. Quantile Function of NPLD

The quantile function of NPLD is given by

$$Q_X(p) = \left(\frac{e^{-Q_T(p)}}{1 + e^{-Q_T(p)}}\right)^{1/k} \quad (13)$$

where $Q_T(p)$ is the quantile function of the normal distribution with parameters μ and σ. Note that $Q_T(p)$ is used to generate random variate that is normally distributed with mean μ and variance σ^2. This can be achieved by using R codes [T = qnorm(p, mean = μ, sd = σ)], where p is vector of probabilities. It can be a sequence of numbers between 0 and 1 or uniformly distributed random values between 0 and 1. We can have in R as p = runif(n), where n is the number of observations to be generated.

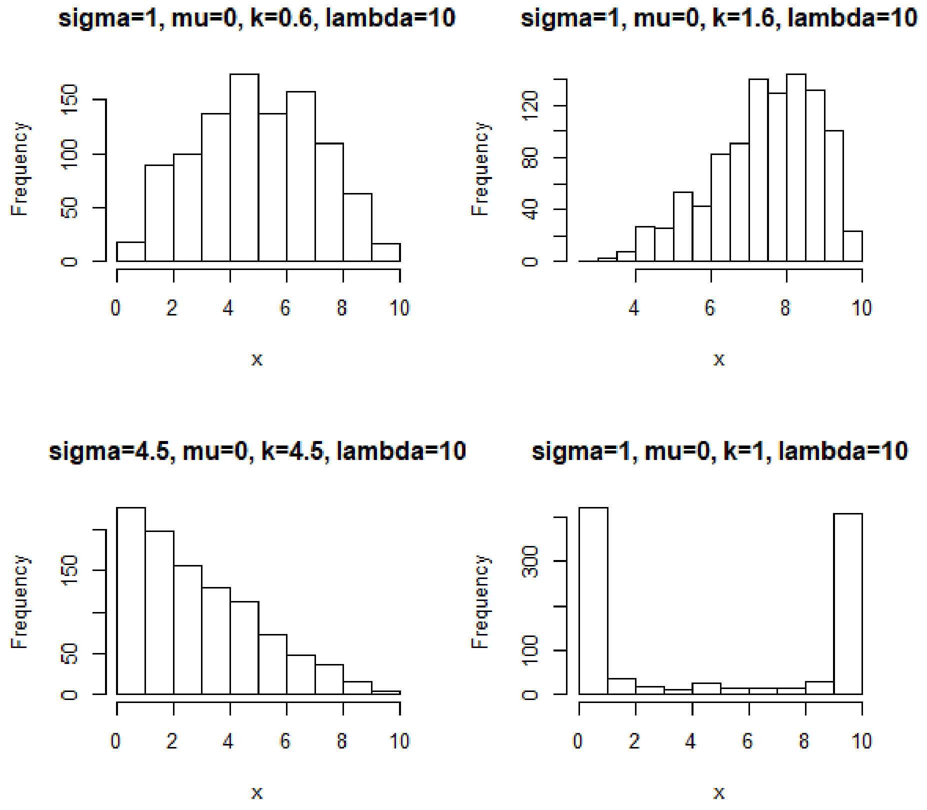

Figure 4. Histogram showing different shapes of NPLD for various values of μ, σ, k and λ

The quantile function of a normal distribution is given by

$$Q_T(p) = \mu + \sigma\sqrt{2}erf^{-1}(2p-1) \qquad (14)$$

Substitute equation (14) into (13) to have the quantile function of NPLD given by

$$Q_X(p) = \left\{\frac{\exp\left[-\mu+\sigma\sqrt{2}erf^{-1}(2p-1)\right]}{1+\exp\left[-\mu+\sigma\sqrt{2}erf^{-1}(2p-1)\right]}\right\}^{1/k} \qquad (15)$$

The quantile function in (15) can be applied by using the quantile function in (13).

2.3. The r-th Moment

Consider a random variable $X \sim NPLD(\lambda, k, \sigma, \mu)$. The *r*-th moment of the random variable X is defined as

$$E\left(X^r\right) = \int_0^\infty x^r f_X(x)dx \qquad (16)$$

where $f_X(x)$ is the pdf of the NPLD defined in equation (6). Substituting equation (6) into (16) and using integration by parts, we obtained the *r*-th moment of the NPLD as

$$E\left(X^r\right) = \frac{\sum_{j=0}^{\infty}\left[\left(\frac{\lambda}{x}\right)^r\right]^j \lambda^r r}{(2\pi)^2\sigma}\int_0^\infty x^{2r-1} \times\exp\left\{-\frac{1}{2\sigma^2}\left[\log\left(\frac{x^r}{\lambda^r - x^r}\right)-\mu\right]^2\right\}.$$

$$E\left(X^r\right) = \sum_{j=0}^{\infty}\left[\left(\frac{\lambda}{x}\right)^r\right]^j \lambda^r r\Gamma(r+1)(2r-1)(2\pi)^{-2}\sigma^{-3}. \qquad (17)$$

Thus, equation (17) is the *r*-th moment of GPLD.

2.4. The Heavy-tail Property

Supposing X is a continuous random variable with a probability density function $f_X(x)$. Then, by definition, $f_X(x)$ is said to be heavy-tailed if and only if

$$\lim_{x\to\infty}\sup\left[f_X(x)\right]e^{tx} = \infty \quad \forall t > 0. \qquad (18)$$

To examine the heavy-tail property of NPLD, substitute equation (6) into (18) to obtain

$$\lim_{x\to\infty}\sup\left[f_X(x)\right]e^{tx} = \lim_{x\to\infty}\sup\left[\begin{array}{l} k\lambda^k x^{-1}\left(\left(\lambda^k - x^k\right)\sqrt{2\pi\sigma^2}\right)^{-1} \\ \times\exp\left\{-\frac{1}{2\sigma^2}\left[\log\left(\frac{x^k}{\lambda^k - x^k}\right)-\mu\right]^2\right\}\end{array}\right]e^{tx}.$$

Using the infinity property

$$\lim_{x\to\infty}\frac{1}{x} = 0.$$

Thus,

$$\lim_{x\to\infty}\sup\left[f_X(x)\right]e^{tx} = 0.$$

Therefore, $f_X(x)$ is not a heavy-tailed function. This can be seen from the pdf plots in Figure 1.

2.5. Stochastic Ordering

The comparative study of the behavior of a continuous random variable can be investigated using stochastic ordering [25,26].

Supposing $M(x) = x - \frac{1}{f_X(x)}\int_0^x t f_T(t)dt, \quad x > 0.$ and Y be continuous random variables. X is said to be smaller than Y in the [25], if the following conditions hold.

$$\begin{array}{l} X \le_L Y \Rightarrow X \le_{hr} Y \Rightarrow X \le_m Y \\ \Downarrow \\ X \le_{st} Y \end{array}$$

i. Stochastic order $(X \le_{st} Y)$ if $F_X(X) \ge F_Y(Y)$ for all x

ii. Hazard rate order $(X \le_{hr} Y)$ if $f_X(x) \ge f_y(y)$ for all x

iii. Mean residual life order $(X \le_m Y)$ if $m_X(x) \le m_Y(y)$ for all x

iv. Likelihood ratio order $(X \le_L Y)$ if $\frac{f_X(x)}{f_Y(x)}$ decreases in x

Theorem 1: Let X and $Y \sim$ NPLD with $\lambda_1, k_1, \sigma_1, \mu_1$ and $\lambda_2, k_2, \sigma_2, \mu_2$ respectively. If

$$\lambda_1, k_1, \sigma_1, \mu_1 \ge \lambda_2, k_2, \sigma_2, \mu_2,$$

then $X \le_L Y$. Hence $X \le_{hr} Y$, $X \le_m Y$ and $X \le_{st} Y$.

Proof

The NPLD will be ordered based on the strongest likelihood ratio order.

$$\frac{f_X(x,\lambda_1,k_1,\sigma_1,\mu_1)}{f_Y(x,\lambda_2,k_2,\sigma_2,\mu_2)} = \frac{\left\{\begin{array}{l}\frac{k_1\lambda_1^{k_1}}{x\left(\lambda_1^{k_1}-x^{k_1}\right)\sqrt{2\pi\sigma_1^2}} \\ \times\exp\left\{-\frac{1}{2\sigma_1^2}\left[\log\left(\frac{x^{k_1}}{\lambda_1^{k_1}-x^{k_1}}\right)-\mu_1\right]^2\right\}\end{array}\right\}}{\left\{\begin{array}{l}\frac{k_2\lambda_2^{k_2}}{x\left(\lambda_2^{k_2}-x^{k_2}\right)\sqrt{2\pi\sigma_2^2}} \\ \times\exp\left\{-\frac{1}{2\sigma_2^2}\left[\log\left(\frac{x^{k_2}}{\lambda_2^{k_2}-x^{k_2}}\right)-\mu_2\right]^2\right\}\end{array}\right\}}. \qquad (19)$$

$$\frac{d}{dx}\ln\frac{f_X(x,\lambda_1,k_1,\sigma_1,\mu_1)}{f_Y(x,\lambda_2,k_2,\sigma_2,\mu_2)} = \frac{\sigma_2 k_1 x^{k_1-1}}{\left(\lambda_1^{k_1}-x^{k_1}\right)} - \frac{\sigma_1 k_2 x^{k_2-1}}{\left(\lambda_2^{k_2}-x^{k_2}\right)} + \mu_2^1 - \mu_1^2. \qquad (20)$$

Thus, for $\lambda_1^{k_1}, \lambda_2^{k_2} > x^{k_1}, x^{k_2}$ and $\mu_2^1 > \mu_1^1$,

$$\frac{d}{dx}\ln\frac{f_X\left(x,\lambda_1,k_1,\sigma_1,\mu_1\right)}{f_Y\left(x,\lambda_2,k_2,\sigma_2,\mu_2\right)}<0\Rightarrow X\leq_L Y,$$

for $\lambda_1^{k_1}, \lambda_2^{k_2} < x^{k_1}, x^{k_2}$ and $\mu_2^1 < \mu_1^1$,

$$\frac{d}{dx}\ln\frac{f_X\left(x,\lambda_1,k_1,\sigma_1,\mu_1\right)}{f_Y\left(x,\lambda_2,k_2,\sigma_2,\mu_2\right)}>0$$

and if $\lambda_1^{k_1}, \lambda_2^{k_2} = x^{k_1}, x^{k_2}$ and $\mu_2^1 = \mu_1^1$

$$\frac{d}{dv}\ln\frac{f_X\left(x,\lambda_1,k_1,\sigma_1,\mu_1\right)}{f_Y\left(x,\lambda_2,k_2,\sigma_2,\mu_2\right)}=0.$$

Hence, $X\leq_{hr} Y$, $X\leq_m Y$ and $X\leq_{st} Y$.

3. Mean Inactivity Time

The mean inactivity time function of a random variable X with a known pdf and cdf is defined as

$$M(x)=x-\frac{1}{f_X(x)}\int_0^x t f_T(t)dt,\ x>0,t>0. \quad (21)$$

where t is the time, $f_T(t)$ the probability function of the NPLD. From the above definition, we obtain the mean inactive function of the NPLD as

$$M(x)=x-\frac{1}{f_X(x)}\int_0^x t\frac{k\lambda^k}{t\left(\lambda^k-t^k\right)\sqrt{2\pi\sigma^2}}\times\exp\left\{-\frac{1}{2\sigma^2}\left[\log\left(\frac{t^k}{\lambda^k-t^k}\right)-\mu\right]^2\right\}dt.$$

$$M(x)=x-\frac{k\lambda^k}{(2\pi)^2\sigma f_X(x)}\times\sum_{j=0}^{\infty}\left(\frac{x}{\lambda}\right)^{jk}\int_0^x\exp\left\{-\frac{1}{2\sigma^2}\left[\log\left(\frac{t^k}{\lambda^k-t^k}\right)-\mu\right]^2\right\}dt.$$

Thus,

$$M(x)=x-\frac{k\lambda^k}{(2\pi)^2\sigma f_X(x)}\times\sum_{j=0}^{\infty}\left(\frac{x}{\lambda}\right)^{jk}\left[\frac{\exp\left\{-\frac{1}{2}\left(\mu/\sigma\right)^2\right\}}{\frac{1}{2}\left(\mu/\sigma\right)^2}-\frac{\exp\left\{-\frac{1}{2\sigma^2}\left[\log\left(\frac{x^k}{\lambda^k-x^k}\right)-\mu\right]^2\right\}}{\frac{1}{2\sigma^2}\left[\log\left(\frac{x^k}{\lambda^k-x^k}\right)-\mu\right]^2}\right]. \quad (22)$$

4. Results and Discussion

In this section, we derived the maximum likelihood parameter estimates, carried out simulation studies to test the consistency of the parameters, and applied the distribution to two real datasets.

4.1. Maximum Likelihood Estimation

From the pdf in (6), we derived the likelihood function as

$$L(\mu,\sigma,k,\lambda/x)=k^n\lambda^{kn}\left(2\pi\sigma^2\right)^{-n/2}\times\exp\left\{-\frac{1}{2\sigma^2}\left[\ln\left(\frac{x_i^k}{\lambda^k-x_i^k}\right)-\mu\right]^2\right\}\prod_{i=1}^{n}\frac{1}{x_i\left(\lambda^k-x_i^k\right)}$$

Take the log of the likelihood function to have the log-likelihood given by

$$LogL=nlnk+knln\lambda-\frac{n}{2}\ln\left(2\pi\sigma^2\right)-\frac{1}{2\sigma^2}\left[\ln\left(\frac{x_i^k}{\lambda^k-x_i^k}\right)-\mu\right]^2-\sum_{i=1}^{n}\ln x_i-\sum_{i=1}^{n}\ln\left(\lambda^k-x_i^k\right) \quad (23)$$

Each of the parameters of the distribution can be estimated by differentiating the $LogL=l$, with respect to each parameter, and equating the result to zero.

$$\frac{\partial l}{\partial\mu}=\frac{1}{\sigma^2}\left[\ln\left(\frac{x_i^k}{\lambda^k-x_i^k}\right)-\mu\right] \quad (24)$$

$$\frac{\partial l}{\partial\sigma}=-\frac{n}{\sigma}+\frac{1}{\sigma^3}\sum_{i=1}^{n}\left\{\left[\ln\left(\frac{x_i^k}{\lambda^k-x_i^k}\right)-\mu\right]^2\right\} \quad (25)$$

$$\frac{\partial l}{\partial k}=\frac{\partial}{\partial k}\left\{\begin{array}{l}nlnk+knln\lambda-\frac{n}{2}\ln\left(2\pi\sigma^2\right)\\-\frac{1}{2\sigma^2}\left[\ln\left(\frac{x_i^k}{\lambda^k-x_i^k}\right)-\mu\right]^2\\-\sum_{i=1}^{n}\ln x_i-\sum_{i=1}^{n}\ln\left(\lambda^k-x_i^k\right)\end{array}\right\} \quad (26)$$

$$\frac{\partial l}{\partial\lambda}=\frac{\partial}{\partial\lambda}\left\{\begin{array}{l}nlnk+knln\lambda-\frac{n}{2}\ln\left(2\pi\sigma^2\right)\\-\frac{1}{2\sigma^2}\left[\ln\left(\frac{x_i^k}{\lambda^k-x_i^k}\right)-\mu\right]^2\\-\sum_{i=1}^{n}\ln x_i-\sum_{i=1}^{n}\ln\left(\lambda^k-x_i^k\right)\end{array}\right\} \quad (27)$$

From equation (24), we have

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n}\left[\ln\left(\frac{x_i^k}{\lambda^k - x_i^k}\right)\right]$$

From equation (25), we have

$$\hat{\sigma} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left\{\left[\ln\left(\frac{x_i^k}{\lambda^k - x_i^k}\right) - \mu\right]^2\right\}}$$

Parameter k is not in closed form, thus, Newton Rapson numerical method is used to estimate the parameter of k., developed by R package (maxLik or optim) (R Development Core Team 2009) [27].

The parameter λ cannot be estimated using maximum likelihood because it is an upper bound of the distribution. So, λ can be estimated directly from data as

$$\hat{\lambda} = \max\left(x_i\right) + \sigma_{\bar{x}}$$

where $\sigma_{\bar{x}}$ is the standard error of $\bar{x}$. Max(x_i) is the maximum of x and $\bar{x}$ is the arithmetic mean of x.

4.2. Simulation Studies

In this subsection, we present the simulation study for the NPLD to investigate its flexible behavior and consistency of the parameter estimates. We examined the "mean estimates (ME), variance, biases, and root mean square errors (RMSEs)" of the MLEs. Below are the procedures used to perform the simulation studies:

1. Uniform distribution is used to generate n quantiles, p.
2. The quantile function defined in equation (15) is used to generate NPLD random variates.
3. The sample sizes are drawn as $n = 50, 100, 200$, and 300.
4. The parameters values are set as $\lambda = k = \sigma = \mu = 1$, $\lambda = k = \sigma = \mu = 2$, and $\lambda = k = \sigma = \mu = 0.5$.

Table 1. Mean estimates, Variance, Biases, and RMSE of the MLEs for $\lambda = k = \sigma = \mu = 1$

n	Parameter values $\lambda = k = \sigma = \mu = 1$	Mean	Bias	Variance	RMSE
50		1.0388	0.0388	2e-04	0.0410
		1.0046	0.0046	2e-04	0.0158
		0.0019	-0.9981	0e+00	0.9981
		1.4142	0.4142	1e-04	0.4144
		1.0383	0.0383	1e-04	0.0391
		1.0034	0.0034	0e+00	0.0042
		0.0018	-0.9982	0e+00	0.9982
		1.4147	0.4037	0e+00	0.4147
200		1.0368	0.0368	0.0000	0.0371
		1.0030	0.0030	0.0000	0.0032
		0.0018	-0.9982	0.0000	0.9982
		1.4154	0.4154	0.0000	0.4154
300		1.0361	0.0361	0.0000	0.0361
		1.0028	0.0028	0.0000	0.0028
		0.0018	-0.9982	0.0000	0.9982
		1.4158	0.4158	0.0000	0.4158

Table 2. Mean estimates, Variance, Biases, and RMSE of the MLEs for $\lambda = k = \sigma = \mu = 2$

n	Parameter values $\lambda = k = \sigma = \mu = 2$	Mean	Bias	Variance	RMSE
50		2.0644	0.0644	0.0009	0.0710
		1.9986	-0.0014	0.0103	0.1017
		0.0035	-1.9965	0.0000	1.9965
		2.8415	0.8415	0.0060	0.8451
100		2.0596	0.0596	1e-04	0.0605
		1.9817	-0.0183	7e-04	0.0325
		0.0036	-1.9964	0.0000	1.9964
		2.8544	0.8544	5e-04	0.8547
200		2.0585	0.0585	1e-04	0.0591
		1.9792	-0.0208	3e-04	0.0264
		0.0036	-1.9964	0e+00	1.9963
		2.8566	0.8536	2e-04	0.8538
300		2.0579	0.0579	1e-04	0.0584
		1.9780	-0.0220	2e-04	0.0260
		0.0036	-1.9964	0e+00	1.9962
		2.8578	0.8508	2e-04	0.8529

Table 3. Mean estimates, Variance, Biases, and RMSE of the MLEs for $\lambda = k = \sigma = \mu = 0.5$

n	Parameter values $\lambda = k = \sigma = \mu = 0.5$	Mean	Bias	Variance	RMSE
50		0.6175	0.1175	0.0012	0.1225
		0.5503	0.0503	0.0000	0.0503
		0.5005	0.5005	0.0000	0.5005
		0.1488	0.1488	0.0000	0.1488
100		0.6099	0.1099	3e-04	0.1111
		0.5503	0.0503	0e+00	0.0503
		0.0005	0.5005	0e+00	0.5005
		0.6488	0.1488	0e+00	0.1488
200		0.6081	0.1081	0.0000	0.1083
		0.5503	0.0503	0.0000	0.0503
		0.0005	0.5005	0.0000	0.5005
		0.6488	0.1488	0.0000	0.1488
300		0.6078	0.1078	0.0000	0.1078
		0.5503	0.0503	0.0000	0.0503
		0.0005	-0.5005	0.0000	0.5005
		0.6488	0.1488	0.0000	0.1488

Table 1 to Table 3 show that the maximum likelihood parameter estimates are consistent. The biases, variances, and RMSEs reduce as the sample size increases. This shows that the parameters are consistent.

4.3. Real Application of NPLD

In this subsection, NPLD is applied to two datasets, one in the engineering field and the other in the medical field. This will enable us to see the usefulness of the data in various fields of study.

Application 1

In this first application, we make use of the data set of gauge lengths of 10mm from Kundu and Raqab [28]. This data set consists of 63 observations:

1.901, 2.132, 2.203, 2.228, 2.257,2.350, 2.361, 2.396, 2.397, 2.445, 2.454, 2.474, 2.518, 2.522, 2.525, 2.532, 2.575, 2.614,2.616, 2.618, 2.624, 2.659, 2.675, 2.738, 2.740, 2.856, 2.917, 2.928, 2.937, 2.937, 2.977,2.996, 3.030, 3.125, 3.139, 3.145, 3.220, 3.223, 3.235, 3.243, 3.264, 3.272, 3.294, 3.332,3.346, 3.377, 3.408, 3.435, 3.493, 3.501, 3.537, 3.554, 3.562, 3.628, 3.852, 3.871, 3.886,3.971, 4.024, 4.027, 4.225, 4.395, 5.020.

The gauge lengths of 10mm data are fitted by using the NPLD, the Beta-Normal (BN) distribution defined and studied by [15] Kumaraswamy-Normal by [17]. Exponentiated Generalized-Normal (EGN) distribution by [19], Weibull-Normal (WN) distribution by [20], and Gamma-Normal (GN) distribution by [14]. The data is skewed to the right (skewness = 0.6328 and kurtosis = 3.2863). The results from the six distributions are presented in Table 4. The Log-Likelihood (LogL), Akaike Information Criterion (AIC), Consistent Akaike Information Criterion (CAIC), Bayesian Information Criterion (BIC), Hannan Quinn Information Criterion (HQIC), Cramer-von Mises statistic (W), Anderson-Darling statistic (A), the Kolmogorov-Smirnov (KS) goodness of fit statistic with its corresponding p-value are reported in Table 4. From Table 4, all six distributions provide an adequate fit to the data set using W, A, KS statistics, and p-value.

Table 4. Parameter estimates, Selection Criteria and Goodness of fit for gauge lengths data

Distributions	NPLD	BN	KuN	EGN	WeN	GN
Parameter Estimates	μ=-15.200	3.281	2.432	0.212	0.666	52.752
	σ=5.712	0.166	0.160	4.231	2.708	0.069
	k=28.693	1.940	2.116	1.369	2.741	1.316
	λ=5.080	0.389	0.355	0.447	0.301	
-LogL	55.864	55.966	55.904	56.657	56.359	56.510
AIC	117.727	119.932	119.407	121.315	120.719	119.020
CAIC	118.134	120.622	120.097	122.005	121.409	119.427
BIC	124.157	128.505	127.980	129.887	129.291	125.449
HQIC	120.256	123.304	122.779	124.687	124.091	121.549
W	0.055	0.049	0.048	0.060	0.060	0.059
A	0.322	0.270	0.275	0.366	0.326	0.336
KS Stat.	0.084	0.074	0.091	0.087	0.080	0.086
KS P-value	0.771	0.876	0.669	0.727	0.819	0.741

Table 4 shows that the NPLD performed best when compared with the other five distributions based on the model selection criteria using AIC, CAIC, BIC, and HQIC. This is because the smaller the values of the selection criteria, the better the model.

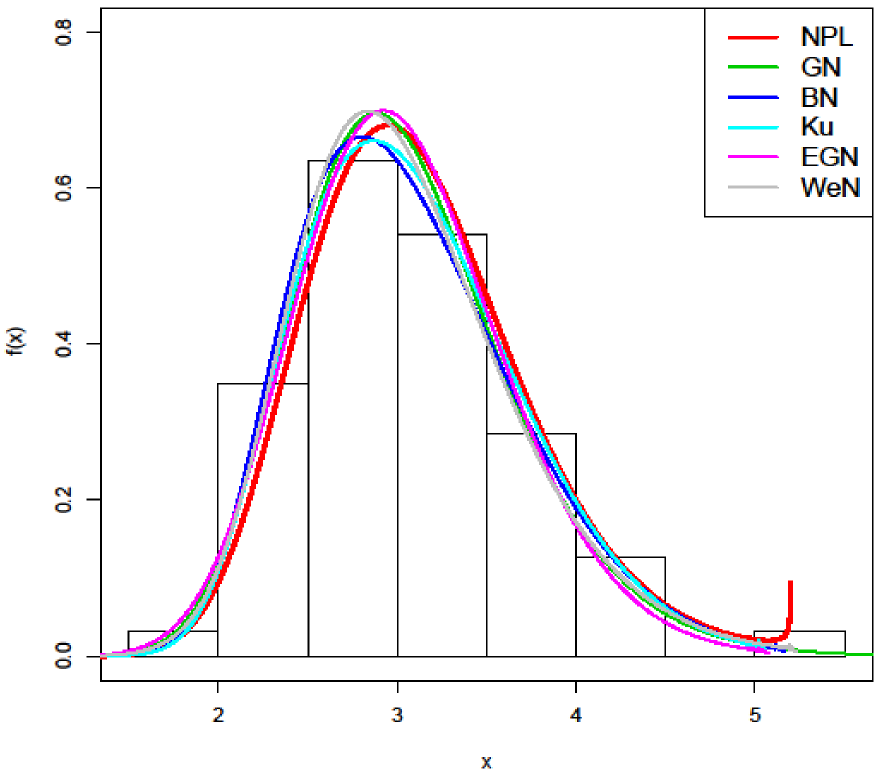

Figure 5. The PDF Curve of NPLD and other five distributions on Gauge Length

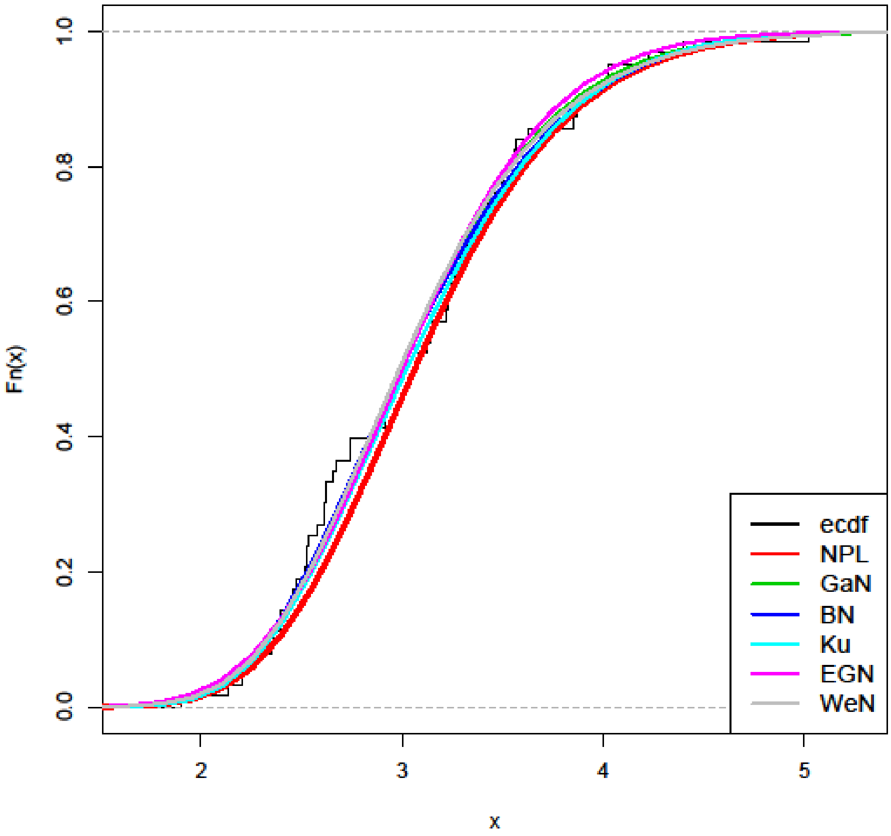

Figure 6. The CDF Curve of NPLD and other five distributions on Gauge Length

Figure 5 displayed the histogram of the gauge length data with the density curve of NPLD and the other five competing distributions. The NPLD appears to fit the data most and gives an account of the gap to the right of the histogram and rise to take care of the dispersed or extreme value. Figure 6 displayed the cumulative distribution of the data with the six distributions fitted on it. The NPLD shows almost a perfect fit.

Application 2

In this second application, we used Crude Mortality Rate (CMR) from Bradley [29]. The data set consists of 65 observations and it is given thus:

2.01, 6.32, 3.52, 2.15, 5.42, 2.04, 2.77, 2.26, 1.95, 1.00, 2.45, 0.74, 0.98, 1.27, 2.77, 3.68, 1.1, 1.09, 1.60, 0.57, 3.33, 0.91, 7.14, 2.08, 3.85, 1.99, 7.76, 2.52, 1.57, 4.67, 4.22, 1.92, 1.59, 4.08, 2.02, 0.84, 6.85, 2.18, 2.04, 1.05, 2.91, 1.37, 2.43, 2.28, 3.74, 1.30, 1.59, 1.83, 3.85, 6.30, 4.83, 0.50, 3.40, 2.33, 4.25, 3.49, 2.12, 0.83, 0.54, 3.23, 4.50, 0.71, 0.48, 2.30, 7.73

The second data is also fitted by using the same set of distributions. The data is positively skewed (skewness = 1.1235 and kurtosis = 3.6889). The results from the six distributions are presented in Table 5 using the same set of criteria.

Table 5. Parameter estimates, Selection Criteria and Goodness of fit for CMR Data

Distributions	NPLD	BN	KuN	EGN	WeN	GN
Parameter Estimates	μ=-7.699	5.109	6.138	0.211	0.731	62.037
	σ=4.483	0.141	0.158	3.697	4.215	-5.692
	k=6.101	-0.601	-0.395	-1.713	1.491	3.594
	λ=7.987	1.021	1.029	1.236	0.777	
-LogL	114.967	122.164	122.608	126.654	123.791	124.617
AIC	235.934	252.328	253.217	261.307	255.582	255.234
CAIC	236.327	252.994	253.883	261.974	256.248	255.628
BIC	242.457	261.025	261.914	270.005	264.279	261.758
HQIC	238.507	255.760	256.649	264.739	259.013	257.808
W	0.048	0.117	0.115	0.212	0.141	0.155
A	0.272	0.781	0.766	1.358	0.926	1.003
KS Stat.	0.083	0.126	0.127	0.140	0.135	0.120
KS P-value	0.756	0.252	0.243	0.156	0.187	0.309

From Table 5, all six distributions provide an adequate fit to the data using the KS p-value, but NPLD has the best fit. Table 5 shows that the NPLD performed best when compared with the other five distributions based on the model selection criteria used.

Figure 7 displayed the histogram of the CMR data with the density curve of NPLD and the other five competing distributions. The NPLD appears to fit the data most showing the positive skewness on the histogram and rise to take care of the dispersed or extreme values. Figure 8 displayed the cumulative distribution of the CMR data with the six distributions fitted on it. The NPLD performed better on the curve.

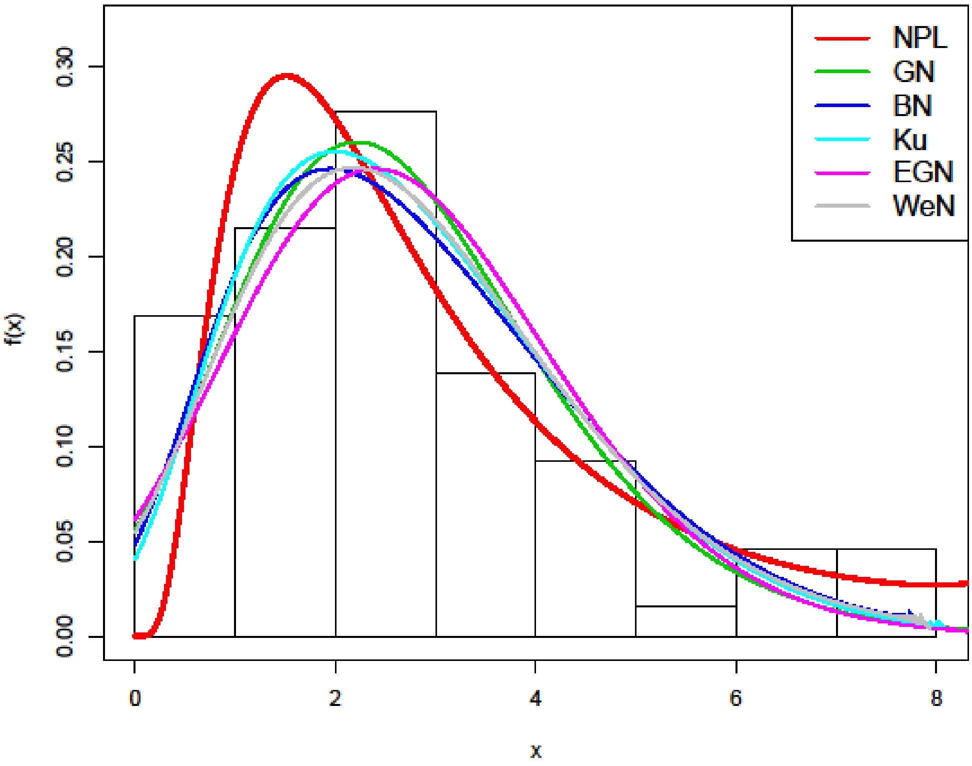

Figure 7. The PDF Curve of NPLD and other five distributions on CMR Data

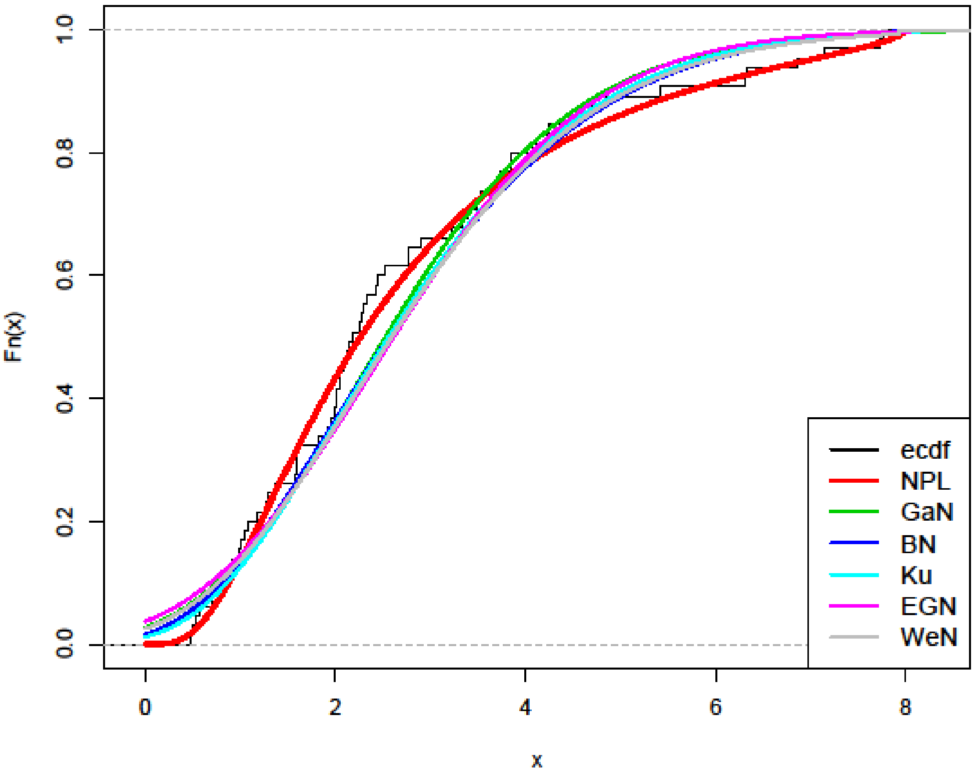

Figure 8. The CDF Curve of NPLD and other five distributions on CMR Data

5. Conclusion

A new generalization of the power function distribution is defined and studied. The Normal-Power function {log-logistic} distribution (NPLD) can be unimodal or bimodal, positively skewed, negatively skewed, symmetric, or bath up. Its pdf, cdf, survival function, hazard function, cumulative hazard function, and reversed hazard function are used to characterize the distribution. The quantile function is derived and displayed. With useful transformation, the four parameters GPLD reduces to the normal distribution with two parameters. The NPLD is not a heavy-tail function, its *r*-th moment and the mean inactive function exists. The method of maximum likelihood estimation was used to estimate the parameters of GPLD. The distribution is applied to two datasets, a gauge length of 10mm and a crude mortality rate. The analysis shows that the NPLD is capable of providing adequate fit to the two datasets that are about symmetric or skewed to the right, with little peakedness, and it is found to perform well in fitting the two datasets and compared favourably than other five convoluted distributions with normal distribution parameters. It is recommended that NPLD should be used to fit gauge length, crude mortality rate, and any similar data that slightly deviate from normal and cannot be fitted by the normal distribution. The deviation from normality can be handled by the shape parameter. It should be used in the generalized linear regression model whenever the ordinary linear regression model fails.

Competing Interests

The authors declare that they have no competing interests.

Authors' Contributions

All Authors drafted the manuscript together. All authors read and approved the final manuscript.

References

[1] Okorie, I.E., Akpanta, A.C., Ohakwe, J. and Chikezie, D.C. The modified Power function distribution. *Cogent Mathematics*, 4: 2017, 1319592.

[2] Ekum, M.I., Adeleke, I.A. and Akarawak, E.E.. Lambda Upper Bound Distribution: Some Properties and Applications. *Benin Journal of Statistics*, 3: 2020a, 12-40.

[3] De Pascoa, M.A.R., Ortega, E.M.M. and Cordeiro, G.M. The Kumaraswamy generalized gamma distribution with application in survival analysis. *Statistical Methodology*, 8, 2011, 411-433.

[4] Aljarrah, M.A., Lee, C., Famoye, F. On generating T-X family of distributions using quantile functions, *Journal of Statistical Distributions and Applications*, 1(2), 2014.

[5] Alzaatreh, A., Lee, C., Famoye, F. A new method for generating families of continuous distributions, *METRON 71*, 2013, 63-79.

[6] Alzaatreh, A., Famoye, F. and Lee, C. The gamma-normal distribution: Properties and application. *Computational Statistics & Data Analysis*, 69, 2014a, 67-80.

[7] Ekum, M. I., Adamu, M. O., Adeleke, I. A., Akarawak, E. E. and Arowolo, O. T. Class of Generalized Power Function Distributions: Properties and Applications. 2020b.

[8] Arshad, M. A., Iqbal, M. Z., and Ahmadm M. Exponentiated Power Function Distribution: Properties and Applications. *Journal of Statistical Theory and Applications*, 19 (2), 2020, 297-313.

[9] Dallas, A.C. Characterization of Pareto and power function distribution. *Annals of Mathematical Statistics*, 28, 1976, 491-497.

[10] Zaka, A., and Akhter, A.S. Methods for estimating the parameters of power function distribution. *Pakistan Journal of Statistics and Operations Research* 9, 2013, 213-224.

[11] Meniconi, M., and Barry, D.M. The power function distribution: A useful and simple distribution to assess electrical component reliability. *Microelectronics Reliability*, 36: 1996, 1207-1212.

[12] Tahir, M. H., Zubair, M., Cordeiro, G. M., Alzaatreh, A., and Mansoor, M. The Weibull-power Cauchy distribution Model, properties and applications. *Hacettepe Journal of Mathematics and Statistics*, *46* (4), 2016.

[13] Ogunsanya, A. S., Sanni, O.O. M., and Yahya, W. B. Exploring Some Properties of Odd Lomax-Exponential Distribution. A*nnals of Statistical Theory and Applications* (ASTA) 1: 2019, 21-30.

[14] Gupta, R. C., Gupta, P. L., and Gupta, R. D. Modeling failure time data by Lehmann alternatives. *Communications in Statistics - Theory and Methods* 27, 1998, 887-904.

[15] Eugene, N., Lee, C., and Famoye, F. Beta-Normal distribution and its applications, *Communications in Statistics - Theory and Methods*, 31(4), 2002, 497-512.

[16] Zografos, K., and Balakrishnan, N. On families of beta- and generalized gamma-generated distributions and associated inference. *Stat. Methodol*. 6: 2009, 344-362.

[17] Cordeiro, G. M., and de Castro, M. A new family of generalized distributions. *J. Stat. Comput. Simul*. 81: 2011, 883-898.

[18] Alzaatreh, A., Lee, C., Famoye, F. *T*-normal family of distributions: a new approach to generalize the normal distribution. *Journal of Statistical Distributions and Applications.* 1(16), 2014b, 1-18.

[19] Cordeiro, G. M., Ortega, E. M. M., and da Cunha, D. C. C. The Exponentiated Generalized Class of Distributions. *Journal of Data Science* 11, 2013, 1-27.

[20] Famoye, F., Akarawak, E., and Ekum, M. Weibull-Normal distribution and its Applications. *Journal of Statistical Theory and Applications*. Vol. 17(4), December 2018, 719-727.

[21] Amalare, A. A., Ogunsanya, A. S. Ekum, M. I. and Owolabi, T. Lomax-Cauchy {Uniform} Distribution: Properties and Application to Exceedances of Flood Peaks of Wheaton River, *Benin Journal of Statistics*, 3, 2020, 66-81.

[22] Ogunsanya, A. S., Akarawak, E. E., and Ekum, M. I. On some properties of Rayleigh-Cauchy distribution", Journal of Statistics and Management Systems: 2021a.

[23] Ogunsanya, A.S., Yahya, W. B., Adegoke, T. M., Iluno, C., Aderele, O. R. and Ekum M. I. A New Three-Parameter Weibull Inverse Rayleigh Distribution: Theoretical Development and Applications. *Mathematics and Statistics*, *9(3),* 2021b, 249-272.

[24] Akarawak, E. E. E., Adeleke, I. A. and Okafor, R. O. (2013). The Weibull-Rayleigh Distribution and its Properties". *Journal of Engineering Research*, 18(1): 56-67.

[25] Shaked, M., and Shantikumar, J. G. (1994). *Stochastic Orders and Their Applications*, Academic Press, New York.

[26] Yu, Y. 2009. Stochastic Ordering of Exponential Family Distributions and Their Mixtures. Department of Statistics, University of California, Irvine, CA 92697, USA.

[27] R Development Core Team R- A Language and Environment for Statistical Computing (R Foundation for Statistical Computing, Austria, Vienna, 2009).

[28] D. Kundu and M. Z. Raqab, (2005). Generalized Rayleigh Distribution: Different Methods of Estimation. Computational Statistics and Data Analysis, 49(1), 187-200.

[29] Bradley, E. H, Webster, T. R., Baker, D., Schlesinger, M., Inouye, S. K., Barth, M. C, Lapane, K. L, Lipson, D., Stone, R., and Koren M. J. (2004). Translating research into practice: speeding the adoption of innovative health care programs. Issue Brief (Commonwealth Fund). 724: 1-12.