

The New $r_T - X$ Family of Distributions: Some Properties with Applications

Clement Boateng Ampadu¹ and Abdulzeid Yen Anafo²

¹ 31 Carrolton Road, Boston, MA 02132-6303, USA; e-mail: drampadu@hotmail.com

² African Institute for Mathematical Sciences (AIMS), P. O. Box DL Adisadel - Cape Coast N1,
Accra-Cape Coast Road, Biriwa, Ghana, West Africa; e-mail: abdulzeid@aims.edu.gh

Abstract

The $r_T - X$ family of distributions induced by V which have been introduced in [1] is further explored in this paper. In particular, we have obtained some basic mathematical properties of this new family. The simulation study shows the method of maximum likelihood is adequate in estimating the unknown parameters in sub-models of this new class of statistical distributions. Further, the application shows that sub-models of this new family of distributions are useful in material science engineering and related disciplines that call for modeling and forecasting of related data sets. Finally, inspired by the Ampadu-G family of distributions [2], we propose a new class of distributions that have never appeared in the literature, and ask the reader to investigate some properties and applications of this new class of distributions.

1. Introduction and Preliminaries

We begin by recalling the $r_T - X$ family of distributions induced by V which have been introduced in [1]. In particular we start with the following

Definition 1.1. A random variable K is said to follow the generalized new $r_T - X$

Received: May 18, 2019; Accepted: August 1, 2019

2010 Mathematics Subject Classification: 62Exx.

Keywords and phrases: continuous distributions, $T - X(W)$ family of distributions, carbon fibers data.

Copyright © 2019 Clement Boateng Ampadu and Abdulzeid Yen Anafo. 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.

family of distributions of type I, if the CDF can be expressed as the following integral

$$J_1(x) = \int_0^{\left(\frac{F(x)-ProductLog(F(x))}{F(x)}\right)^{\frac{1}{\alpha}}} \frac{1}{(q_T \circ R_T)(t)} dt,$$

where the random variable X has CDF $F(x)$, and the random variable T with support $[0, \infty)$ has quantile density q_T and CDF R_T , $ProductLog[z]$ gives the principal solution for w in $z = we^w$, and $\alpha > 0$.

Remark 1.2. Note that $V(x) = \frac{x - ProductLog(x)}{x}$ is our weight in the above definition, and V^{-1} gives $\frac{-\log(1-x)}{1-x}$ which is the weight function introduced in [3].

Since $R' = \frac{1}{q \circ R}$, letting $\alpha = 1$ in the above definition implies the following

Theorem 1.3. *The CDF of the new $r_T - X$ family of distributions of type I is given by*

$$J_1(x) = R_T\left(\frac{F(x) - ProductLog(F(x))}{F(x)}\right),$$

where the random variable T with support $[0, \infty)$ has CDF R_T , $ProductLog[z]$ gives the principal solution for w in $z = we^w$, and the random variable X has CDF $F(x)$.

Remark 1.4. By differentiating the CDF above, the PDF of the new $r_T - X$ family of distributions of type I can be obtained.

2. The Approximation to the New Family

It is well known [4] that the principal solution for w in $z = we^w$, denoted $ProductLog[z]$, admit the following Taylor series expansion

$$\sum_{n=1}^{\infty} \frac{(-n)^{n-1}}{n!} z^n = z - z^2 + \frac{3}{2} z^3 - \frac{8}{3} z^4 + \frac{125}{24} z^5 - \dots.$$

Using the first two terms of this series to approximate $\text{ProductLog}[z]$, gives an approximation to the $r_T - X$ family of distributions, in particular we define the generic approximated distribution as follows:

Corollary 2.1. *The CDF of the approximated new $r_T - X$ family of distributions of type I is given by*

$$J_2(x) = R_T(F(x)),$$

where the random variable T with support $[0, \infty)$ has CDF R_T , and the random variable X has CDF $F(x)$.

By differentiating the CDF above, we get the following:

Corollary 2.2. *The PDF of the approximated new $r_T - X$ family of distributions of type I is given by*

$$j_2(x) = r_T(F(x))f(x),$$

where the random variable T with support $[0, \infty)$ has PDF r_T , and the random variable X has CDF $F(x)$ and PDF $f(x)$.

3. Some Sub-models of the New Family

Assuming the random variable T is Exponentially distributed so that its CDF is given by

$$R_T(t) = 1 - e^{-at}$$

for $t > 0$ and $a > 0$, and the random variable X is Logistic distributed, so that its CDF is given by

$$F_X(x) = \frac{1}{e^{\frac{x-m}{s}} + 1},$$

where $m \in \mathbb{R}$, $s > 0$, and $x \in \mathbb{R}$, then we have the following from Theorem 1.3.

Theorem 3.1. *The CDF of the new Exponential-Logistic distribution of type I is given by*

$$1 - \exp \left(a \left(- \left(e^{\frac{x-m}{s}} + 1 \right) \right) \left(\frac{1}{e^{\frac{x-m}{s}} + 1} - \text{ProductLog} \left(\frac{1}{e^{\frac{x-m}{s}} + 1} \right) \right) \right),$$

where $a, s > 0$, $m, x \in \mathbb{R}$, and $\text{ProductLog}(z)$ gives the principal solution for w in $z = we^w$.

Remark 3.2. If a random variable L has CDF given by the theorem immediately above, we write

$$L \sim \text{NELDI}(m, s, a).$$

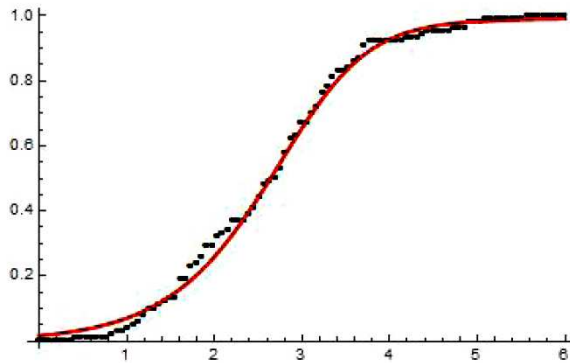


Figure 1. The CDF of $\text{NELDI}(4.4345312, 0.6843875, 11.2804404)$ fitted to the empirical distribution of the carbon fibers data, [5, Section 6].

Remark 3.3. The PDF of the NELDI distribution can be obtained by differentiating the CDF.

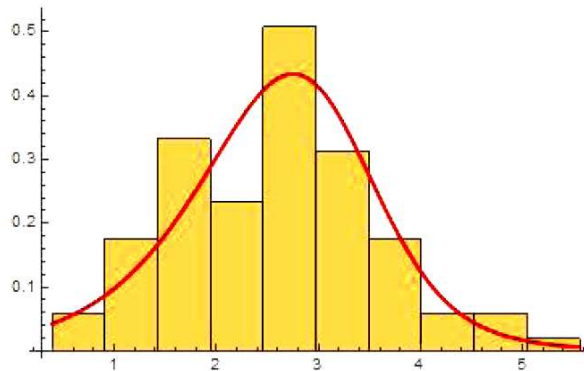


Figure 2. The PDF of $\text{NELDI}(4.4345312, 0.6843875, 11.2804404)$ fitted to the histogram of the carbon fibers data, [5, Section 6].

Assuming the random variable X is Weibull distributed, so that the CDF is given by

$$F_X(x) = 1 - e^{-\left(\frac{x}{c}\right)^b}$$

for $x, b, c > 0$, then we have the following from Theorem 1.3.

Theorem 3.4. *The CDF of the new Exponential-Weibull distribution of type I is given by*

$$1 - \exp \left[- \frac{a \left(- \text{ProductLog} \left(1 - e^{-\left(\frac{x}{c}\right)^b} \right) - e^{-\left(\frac{x}{c}\right)^b} + 1 \right)}{1 - e^{-\left(\frac{x}{c}\right)^b}} \right],$$

where $x, a, b, c > 0$, and $\text{ProductLog}(z)$ gives the principal solution for w in $z = we^w$.

Remark 3.5. If a random variable Y has CDF given by the new Exponential-Weibull distribution of type I, we write

$$Y \sim \text{NEWDI}(a, b, c).$$

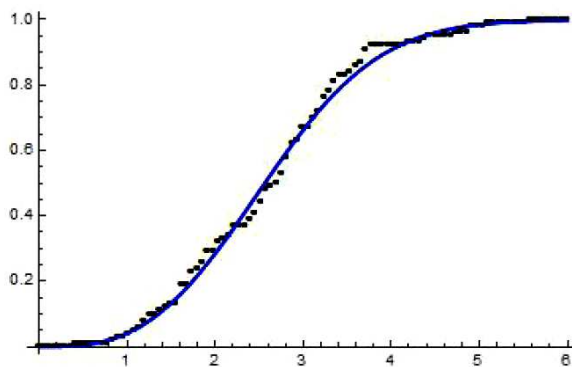


Figure 3. The CDF of $\text{NEWDI}(36.097941, 3.007914, 9.418255)$ fitted to the empirical distribution of the carbon fibers data, [5, Section 6].

Remark 3.6. The PDF of the NEWDI distribution can be obtained by differentiating the CDF.

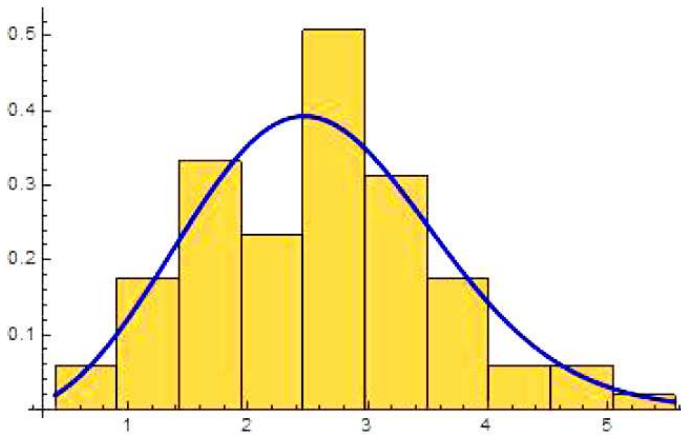


Figure 4. The PDF of $NEWDI(36.097941, 3.007914, 9.418255)$ fitted to the histogram of the carbon fibers data, [5, Section 6].

Assuming the random variable X is normally distributed, so that the CDF is given by

$$F_X(x) = \frac{1}{2} \operatorname{erfc}\left(\frac{b-x}{\sqrt{2c}}\right),$$

where $b, x \in \mathbb{R}$, $c > 0$, and $\operatorname{erfc}(x) = 1 - \operatorname{erf}(x) = 1 - \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$, then we have the following from Theorem 1.3.

Theorem 3.7. The CDF of the new Exponential-Normal distribution of type I is given by

$$1 - \exp\left(-\frac{2a\left(\frac{1}{2} \operatorname{erfc}\left(\frac{b-x}{\sqrt{2c}}\right) - \operatorname{ProductLog}\left(\frac{1}{2} \operatorname{erfc}\left(\frac{b-x}{\sqrt{2c}}\right)\right)\right)}{\operatorname{erfc}\left(\frac{b-x}{\sqrt{2c}}\right)}\right),$$

where $b, x \in \mathbb{R}$, $a, c > 0$, $\operatorname{erfc}(x) = 1 - \operatorname{erf}(x) = 1 - \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt$, and $\operatorname{ProductLog}(z)$ gives the principal solution for w in $z = we^w$.

Remark 3.8. If a random variable M has CDF given by the new Exponential-Normal distribution of type I, we write

$$H \sim NENDI(a, b, c).$$

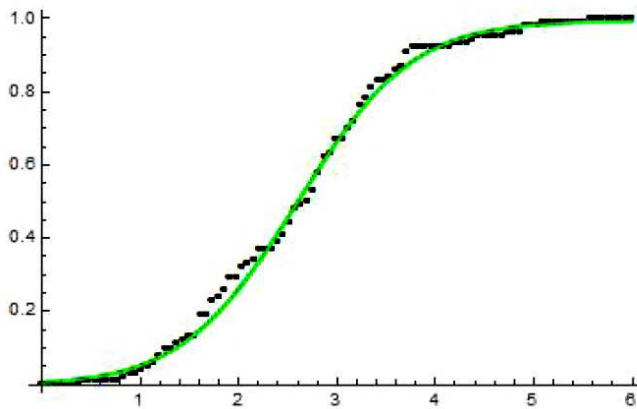


Figure 5. The CDF of $NENDI(13.865794, 5.048878, 1.522466)$ fitted to the empirical distribution of the carbon fibers data, [5, Section 6].

Remark 3.9. The PDF of the NEWDI distribution can be obtained by differentiating the CDF.

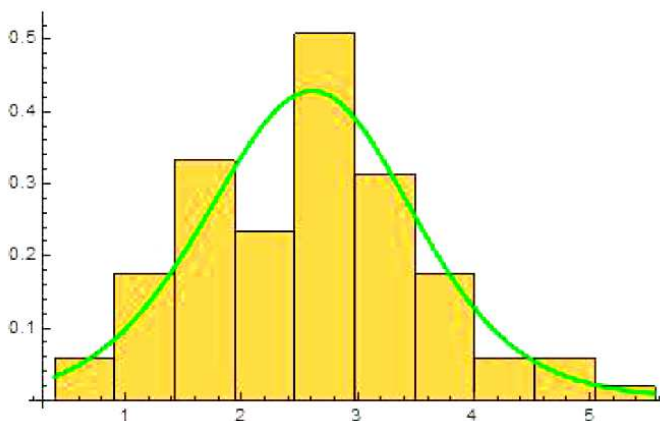


Figure 6. The PDF of $NENDI(13.865794, 5.048878, 1.522466)$ fitted to the histogram of the carbon fibers data, [5, Section 6].

4. Basic Mathematical Properties induced by the Approximated Family

Theorem 4.1. Suppose the random variable T with support $[0, \infty)$ has quantile Q_T , and the random variable X has quantile Q_X . If U is uniform on $(0, 1)$, then the random variable

$$Q_X[Q_T(U)]$$

follows the approximated new $r_T - X$ family of distributions of type I.

Theorem 4.2. The r th non-central moments of the approximated new $r_T - X$ family of distributions of type I can be expressed as

$$\mu'_r = \sum_{i,j=0}^{\infty} \delta_{r,i} \delta_{i,j} E[U^j],$$

where $\delta_{r,i} = (ih_0)^{-1} \sum_{s=1}^i [s(r+1) - i] h_s \delta_{r,i-s}$ with $\delta_{r,0} = h_0^r$ for $i = 1, 2, \dots$ [6], $\delta_{i,j} = (jh_0)^{-1} \sum_{t=1}^j [t(i+1) - j] h_t \delta_{i,j-t}$ with $\delta_{i,0} = h_0^i$ for $j = 1, 2, \dots$ [6], U is uniform on $[0, 1]$, and $E(\cdot)$ is an expectation.

Theorem 4.3. Suppose the random variable T with support $[0, \infty)$ has quantile Q_T , and the random variable X has quantile Q_X , then the quantile function of the approximated new $r_T - X$ family of distributions of type I is given by

$$Q_X[Q_T(p)]$$

for $0 < p < 1$.

Theorem 4.4. The moment generating function of the approximated new $r_T - X$ family of distributions of type I can be expressed as

$$M_J(z) = \sum_{r,i,j=0}^{\infty} \frac{z^r \delta_{r,i} \delta_{i,j} E[U^j]}{r!},$$

where $\delta_{r,i} = (ih_0)^{-1} \sum_{s=1}^i [s(r+1) - i] h_s \delta_{r,i-s}$ with $\delta_{r,0} = h_0^r$ for $i = 1, 2, \dots$ [6],

$\delta_{i,j} = (jh_0)^{-1} \sum_{t=1}^j [t(i+1) - j] h_t \delta_{i,j-t}$ with $\delta_{i,0} = h_0^i$ for $j = 1, 2, \dots$ [6], U is uniform on $[0, 1]$, and $E(\cdot)$ is an expectation.

Theorem 4.5. If the random variable V follows the approximated new $r_T - X$ family of distributions of type I, then the Shannon entropy, S_V , is given by

$$S_V = \eta_T - E[\log f(Q_X(T))],$$

where η_T is the Shannon entropy of the random variable T with support $[0, \infty)$, $E[\cdot]$ is an expectation, and the random variable X has PDF f and quantile Q_X .

Theorem 4.6. The CDF of the exponentiated approximated new standard exponential- X family of distributions of type I which is given by

$$(1 - e^{-F(x)})^\alpha$$

for some $\alpha > 0$, where $F(x)$ is the CDF of the random variable X , admit the following expansion

$$\sum_{k=0}^{\alpha} \sum_{q=0}^{\infty} \binom{\alpha}{k} \frac{(-1)^{k+q} (kF(x))^q}{q!}.$$

Theorem 4.7. The PDF of the exponentiated approximated new standard exponential- X family of distributions of type I which is given by

$$\alpha(1 - e^{-F(x)})^{\alpha-1} e^{-F(x)} f(x),$$

where $\alpha > 0$, $F(x)$ and $f(x)$ are the CDF and PDF, respectively, of the random variable X , admit the following expansion

$$\alpha \sum_{k=0}^{\alpha-1} \sum_{q=0}^{\infty} \binom{\alpha-1}{k} \frac{(-1)^{k+q} (k+1)^q F(x)^q f(x)}{q!}.$$

5. Parameter Estimation in a Sub-model of the New Family

In this section we discuss estimating the unknown parameters in the new Exponential- X family of distributions by using the methods of least squares estimation, weighted least squares estimation and the maximum likelihood estimation.

Observe we have the following from Theorem 1.3

Corollary 5.1. *The CDF of the new Exponential-X family of distributions of type I is given by*

$$S_1(x; a, \xi) = 1 - e^{-a \left(\frac{F(x; \xi) - \text{ProductLog}(F(x; \xi))}{F(x; \xi)} \right)},$$

where $a > 0$, $\text{ProductLog}[z]$ gives the principal solution for w in $z = we^w$, and the random variable X has CDF $F(x; \xi)$ with ξ being a vector of parameters in the distribution of the random variable X .

By differentiating the CDF above, we have the following

Corollary 5.2. *The PDF of the new Exponential-X family of distributions of type I is given by*

$$s_1(x; a, \xi) = \frac{af(x; \xi)e^{-a \left(\frac{F(x; \xi) - \text{ProductLog}[F(x; \xi)]}{F(x; \xi)} \right)} e^{-2\text{ProductLog}[F(x; \xi)]}}{1 + \text{ProductLog}[F(x; \xi)]},$$

where $a > 0$, $\text{ProductLog}[z]$ gives the principal solution for w in $z = we^w$, and the random variable X has CDF $F(x; \xi)$ and PDF $f(x; \xi)$, with ξ being a vector of parameters in the distribution of the random variable X .

Least squares and weighted least squares estimators: Let X_1, X_2, \dots, X_n be a random sample of size n from the new Exponential-X family of distributions of type I, and suppose that $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ denotes the order statistics of the observed sample, then the following are well known

$$E(S_1(X_{(i)}, a, \xi)) = \frac{i}{n+1}, \quad i = 1, 2, \dots, n$$

and

$$V(S_1(X_{(i)}, a, \xi)) = \frac{i(n+1+i)}{(n+1)^2(n+2)}, \quad i = 1, 2, \dots, n,$$

where $E(\cdot)$ denotes an expectation and $V(\cdot)$ denotes variance. The least squares estimate

\hat{a}_{LS} and $\hat{\xi}_{LS}$ of a and ξ , respectively, are obtained by minimizing the equation

$$M(a, \xi) = \sum_{i=1}^n \left(1 - e^{-a \left(\frac{F(X_{(i)}; \xi) - \text{ProductLog}(F(X_{(i)}; \xi))}{F(X_{(i)}; \xi)} \right)} - \frac{i}{n+1} \right)^2$$

which amounts to solving the following system of equations for a, ξ :

$$\frac{\partial M(a, \xi)}{\partial a} = 0 \quad \text{and} \quad \frac{\partial M(a, \xi)}{\partial \xi} = 0.$$

Note that

$$\begin{aligned} \frac{\partial M(a, \xi)}{\partial a} &= 2 \sum_{i=1}^n \left(1 - e^{-a \left(\frac{F(X_{(i)}; \xi) - \text{ProductLog}(F(X_{(i)}; \xi))}{F(X_{(i)}; \xi)} \right)} - \frac{i}{n+1} \right) \\ &\quad \times e^{-a \left(\frac{F(X_{(i)}; \xi) - \text{ProductLog}(F(X_{(i)}; \xi))}{F(X_{(i)}; \xi)} \right)} \\ &\quad \times \left(\frac{F(X_{(i)}; \xi) - \text{ProductLog}(F(X_{(i)}; \xi))}{F(X_{(i)}; \xi)} \right) \end{aligned}$$

and

$$\begin{aligned} \frac{\partial M(a, \xi)}{\partial \xi} &= 2a \sum_{i=1}^n \left(1 - e^{-a \left(\frac{F(X_{(i)}; \xi) - \text{ProductLog}(F(X_{(i)}; \xi))}{F(X_{(i)}; \xi)} \right)} - \frac{i}{n+1} \right) \\ &\quad \times e^{-a \left(\frac{F(X_{(i)}; \xi) - \text{ProductLog}(F(X_{(i)}; \xi))}{F(X_{(i)}; \xi)} \right)} \\ &\quad \times \frac{\partial}{\partial \xi} \left(\frac{F(X_{(i)}; \xi) - \text{ProductLog}(F(X_{(i)}; \xi))}{F(X_{(i)}; \xi)} \right). \end{aligned}$$

The system of nonlinear equations

$$\frac{\partial M(a, \xi)}{\partial a} = 0 \quad \text{and} \quad \frac{\partial M(a, \xi)}{\partial \xi} = 0$$

can be solved using Newton's method or fixed point iteration techniques.

Note that the weighted least squares estimates \hat{a}_{WLS} and $\hat{\xi}_{WLS}$ of a and ξ , respectively, are obtained by minimizing the equation

$$N(a, \xi) = \sum_{i=1}^n \frac{(n+2)(n+1)^2}{i(n+i+1)} \left(1 - e^{-a \left(\frac{F(X_{(i)}; \xi) - \text{ProductLog}(F(X_{(i)}; \xi))}{F(X_{(i)}; \xi)} \right)} - \frac{i}{n+1} \right)^2.$$

Maximum likelihood estimators: For this, let X_1, X_2, \dots, X_n be a random sample of size n from the new Exponential-X family of distributions of type I. From Corollary 5.2, the likelihood function is given by

$$\begin{aligned} L &= \prod_{i=1}^n s_1(x_i; a, \xi) \\ &= a^n \prod_{i=1}^n f(x_i; \xi) e^{-a \sum_{i=1}^n \left(\frac{F(x_i; \xi) - \text{ProductLog}(F(x_i; \xi))}{F(x_i; \xi)} \right)} e^{-2 \sum_{i=1}^n \text{ProductLog}(F(x_i; \xi))} \\ &\quad \times \prod_{i=1}^n (1 + \text{ProductLog}(F(x_i; \xi)))^{-1}. \end{aligned}$$

From the above equation the log-likelihood function is given by

$$\begin{aligned} \ln L &= n \log(a) + \sum_{i=1}^n \log(f(x_i; \xi)) - a \sum_{i=1}^n \left(\frac{F(x_i; \xi) - \text{ProductLog}(F(x_i; \xi))}{F(x_i; \xi)} \right) \\ &\quad - 2 \sum_{i=1}^n \text{ProductLog}(F(x_i; \xi)) - \sum_{i=1}^n \log(1 + \text{ProductLog}(F(x_i; \xi))). \end{aligned}$$

The MLE's of a and ξ can be obtained by maximizing the equation immediately above. The derivatives of the equation immediately above with respect to the unknown

parameters are given as follows:

$$\frac{\partial \ln L}{\partial a} = \frac{n}{a} - \sum_{i=1}^n \left(\frac{F(x_i; \xi) - \text{ProductLog}(F(x_i; \xi))}{F(x_i; \xi)} \right)$$

and

$$\begin{aligned} \frac{\partial \ln L}{\partial \xi} &= \sum_{i=1}^n \frac{1}{f(x_i; \xi)} \frac{\partial f(x_i; \xi)}{\partial \xi} - a \sum_{i=1}^n \frac{e^{-2\text{ProductLog}(F(x_i; \xi))}}{\text{ProductLog}(F(x_i; \xi)) + 1} \frac{\partial F(x_i; \xi)}{\partial \xi} \\ &\quad - 2 \sum_{i=1}^n \frac{\text{ProductLog}(F(x_i; \xi))}{F(x_i; \xi) \text{ProductLog}(F(x_i; \xi)) + F(x_i; \xi)} \frac{\partial F(x_i; \xi)}{\partial \xi} \\ &\quad - \sum_{i=1}^n \frac{\text{ProductLog}(F(x_i; \xi))}{F(x_i; \xi) (\text{ProductLog}(F(x_i; \xi)) + 1)^2} \frac{\partial F(x_i; \xi)}{\partial \xi}. \end{aligned}$$

Now solving the system below for ξ and a gives the maximum likelihood estimators, $\hat{\xi}$ and \hat{a} of the unknown parameters:

$$\frac{\partial \ln L}{\partial a} = 0 \quad \text{and} \quad \frac{\partial \ln L}{\partial \xi} = 0.$$

Let $\Delta = (a; \xi)$, for the purposes of constructing confidence intervals for the parameters in the new Exponential-X family of distributions of type I, the observed information matrix, call it $J(\Delta)$, can be used due to the complex nature of the expected information matrix. The observed information matrix is given by

$$J(\Delta) = - \begin{bmatrix} \frac{\partial^2 \ln L}{\partial^2 a} & \frac{\partial^2 \ln L}{\partial a \partial \xi} \\ \frac{\partial^2 \ln L}{\partial^2 \xi} \end{bmatrix}.$$

The elements of the observed information matrix are given below

$$\frac{\partial^2 \ln L}{\partial^2 a} = \frac{-n}{a^2},$$

$$\begin{aligned}
\frac{\partial^2 \ln L}{\partial a \partial \xi} &= - \sum_{i=1}^n \frac{e^{-2 \text{ProductLog}(F(x_i; \xi))}}{\text{ProductLog}(F(x_i; \xi)) + 1} \frac{\partial F(x_i; \xi)}{\partial \xi} \\
\frac{\partial^2 \ln L}{\partial^2 \xi} &= \sum_{i=1}^n \left[\frac{1}{f(x_i; \xi)} \frac{\partial^2 f(x_i; \xi)}{\partial^2 \xi} - \frac{1}{f(x_i; \xi)^2} \left(\frac{\partial f(x_i; \xi)}{\partial \xi} \right)^2 \right] \\
&\quad - a \sum_{i=1}^n \left[\frac{e^{-2 \text{ProductLog}(F(x_i; \xi))}}{\text{ProductLog}(F(x_i; \xi)) + 1} \frac{\partial^2 F(x_i; \xi)}{\partial^2 \xi} \right. \\
&\quad \left. - \frac{e^{-3 \text{ProductLog}(F(x_i; \xi))} (2 \text{ProductLog}(F(x_i; \xi)) + 3) \left(\frac{\partial F(x_i; \xi)}{\partial \xi} \right)^2}{(\text{ProductLog}(F(x_i; \xi)) + 1)^3} \right] \\
&\quad - 2 \sum_{i=1}^n \left[\frac{\text{ProductLog}(F(x_i; \xi))}{F(x_i; \xi) \text{ProductLog}(F(x_i; \xi)) + F(x_i; \xi)} \frac{\partial^2 F(x_i; \xi)}{\partial^2 \xi} \right. \\
&\quad \left. - \frac{e^{-2 \text{ProductLog}(F(x_i; \xi))} (\text{ProductLog}(F(x_i; \xi)) + 2) \left(\frac{\partial F(x_i; \xi)}{\partial \xi} \right)^2}{(\text{ProductLog}(F(x_i; \xi)) + 1)^3} \right] \\
&\quad - \sum_{i=1}^n \left[\frac{\text{ProductLog}(F(x_i; \xi))}{F(x_i; \xi) (\text{ProductLog}(F(x_i; \xi)) + 1)^2} \frac{\partial^2 F(x_i; \xi)}{\partial^2 \xi} \right. \\
&\quad \left. - \frac{e^{-2 \text{ProductLog}(F(x_i; \xi))} (\text{ProductLog}(F(x_i; \xi)) + 3) \left(\frac{\partial F(x_i; \xi)}{\partial \xi} \right)^2}{(\text{ProductLog}(F(x_i; \xi)) + 1)^4} \right].
\end{aligned}$$

When the usual regularity conditions are satisfied and that the parameters are within the interior of the parameter space, but not on the boundary, the distribution of $\sqrt{n}(\hat{\Delta} - \Delta)$ converges to the multivariate normal distribution $N_{n+1}(0, I^{-1}(\Delta))$, where $I(\Delta)$ is the expected information matrix, and it is assumed that $\xi = (\xi_1, \dots, \xi_n)$. The asymptotic behavior remains valid when $I(\Delta)$ is replaced by the observed information matrix evaluated at $J(\hat{\Delta})$. The asymptotic multivariate normal distribution $N_{n+1}(0, J^{-1}(\hat{\Delta}))$ is a very useful tool for constructing an approximate $100(1 - \psi)\%$ two-sided confidence intervals for the model parameters, where ψ is the significance level.

6. Simulation Study Induced by the Approximated Family

In this section, a Monte Carlo simulation study is carried out to assess the performance of the maximum likelihood estimation method in the new $r_T - X$ family of distributions of type I by using its approximated family. At first we demonstrate the equivalence of the approximated new $r_T - X$ family of distributions of type I and the new $r_T - X$ family of distributions of type I. It implies the adequacy of the maximum likelihood estimation method in the new $r_T - X$ family of distributions of type I is equivalent to the adequacy of the maximum likelihood estimation method in the approximated new $r_T - X$ family of distributions of type I.

Recall the CDF of the new Exponential-Logistic distribution of type I is given by Theorem 3.1, it follows from Corollary 2.1, that its approximation is given by the following:

Theorem 6.1. *The CDF of the approximated new Exponential-Logistic distribution of type I is given by*

$$1 - e^{-e^{\frac{a}{\frac{x-m}{s} + 1}}},$$

where $a, s > 0, m, x \in \mathbb{R}$.

Remark 6.2. If a random variable L^* has CDF given by the theorem immediately above, we write

$$L^* \sim \text{ANELDI}(m, s, a).$$

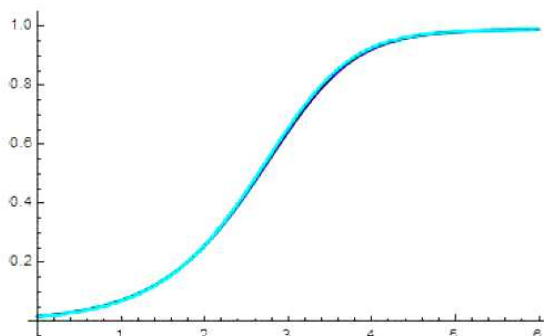


Figure 7. A comparison of the CDF's $\text{NELDI}(4.4345312, 0.6843875, 11.2804404)$ (Cyan) and $\text{ANELDI}(3.9, 0.7, 4.8)$ (Purple).

From the above we can see that approximating the true distribution, NELDI, by ANELDI is appropriate, as the CDF's appear identical.

Now we demonstrate adequacy of the maximum likelihood estimation method in the approximated new $r_T - X$ family of distributions of type I via a brief simulation study which implies adequacy of the maximum likelihood estimation method in the new $r_T - X$ family of distributions of type I. For this, a Monte Carlo simulation study is carried out to assess the performance of the estimation method in the ANELDI distribution. Samples of sizes 200, 350, 500, and 700, are drawn from the ANELDI distribution, and the samples have been drawn for the following set of parameters

- (a) Set I: $(m, s, a) = (3.9, 0.7, 4.8)$.
- (b) Set II: $(m, s, a) = (3.9, 3.9, 3.9)$.
- (c) Set III: $(m, s, a) = (0.7, 0.7, 4.8)$.

The maximum likelihood estimators for the parameters m , s , and a are obtained. The procedure has been repeated 1000 times, and the mean and mean square error for the estimates are computed, and the results are summarized in Tables 1-3 below for each of sets I, II, and III, respectively, considered above.

Table 1. Result of simulation study for Set I.

Parameter m		
Sample Size	Average Estimate	MSE
200	3.949551	0.03355402
350	3.90274	0.0104246
500	3.884419	0.00199555
700	3.881701	0.0007860998
Parameter s		
Sample Size	Average Estimate	MSE
200	0.6989851	0.0007454304
350	0.69232	0.0002275311
500	0.691042	0.0001474762
700	0.6901352	0.000106438

Parameter a		
Sample Size	Average Estimate	MSE
200	5.238702	1.398739
350	4.9495	0.2121618
500	4.888973	0.07475331
700	4.857556	0.01783414

From Table 1, we find that the simulated estimates are close to the true values of the parameters and hence the estimation method is adequate. We have also observed that the estimated mean square errors (MSEs) consistently decrease with increasing sample size as seen in Figures 8-10.

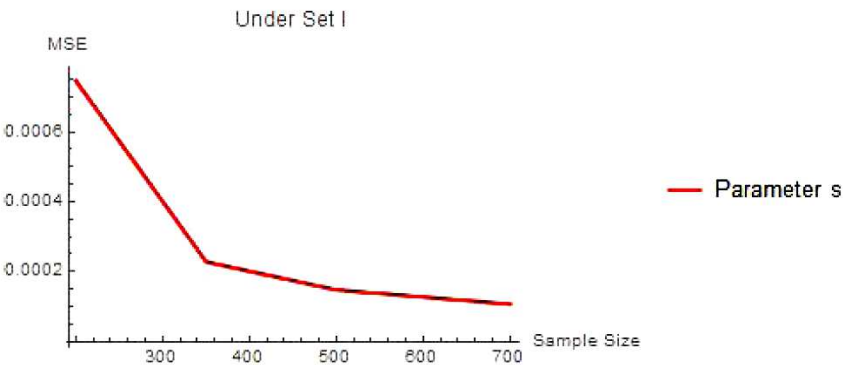


Figure 8. Decreasing MSE for increasing sample size.

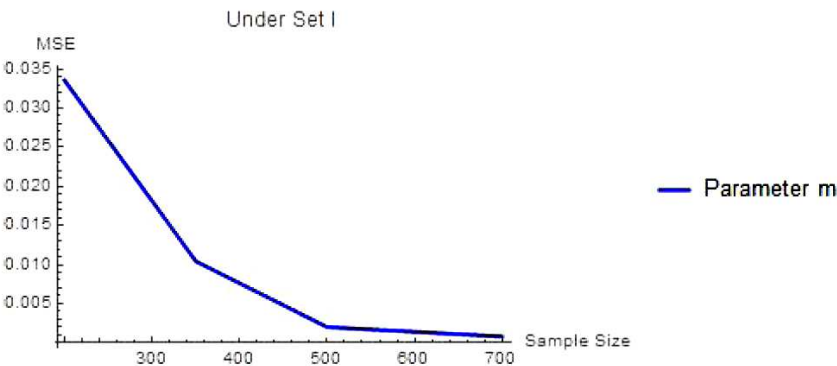


Figure 9. Decreasing MSE for increasing sample size.

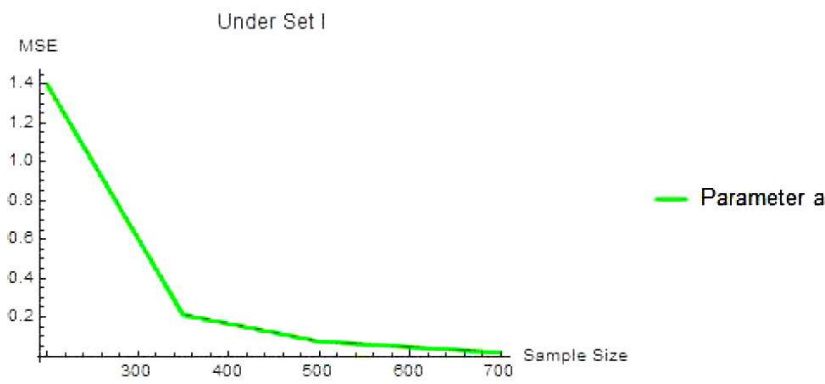


Figure 10. Decreasing MSE for increasing sample size.

Table 2. Result of simulation study for Set II

Parameter <i>m</i>		
Sample Size	Average Estimate	MSE
200	3.937262	0.183125
350	3.882132	0.004860528
500	3.88	0.0004
700	3.88	0.0004

Parameter <i>s</i>		
Sample Size	Average Estimate	MSE
200	3.886086	0.003141356
350	3.881337	0.001435791
500	3.88	0.0004
700	3.88	0.0004

Parameter <i>a</i>		
Sample Size	Average Estimate	MSE
200	3.923786	0.1808325
350	3.883832	0.007635531
500	3.88	0.0004
700	3.88	0.0004

From Table 2, we find that the simulated estimates are close to the true values of the parameters and hence the estimation method is adequate. We have also observed that the estimated mean square errors (MSEs) consistently decrease with increasing sample size as seen in Figures 11-13.

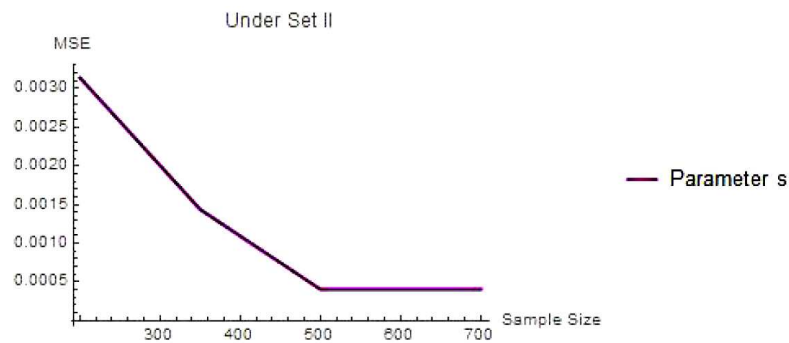


Figure 11. Decreasing MSE for increasing sample size.

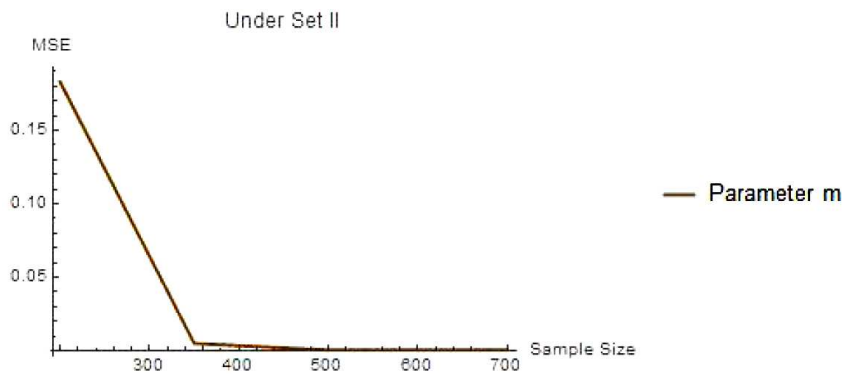


Figure 12. Decreasing MSE for increasing sample size.

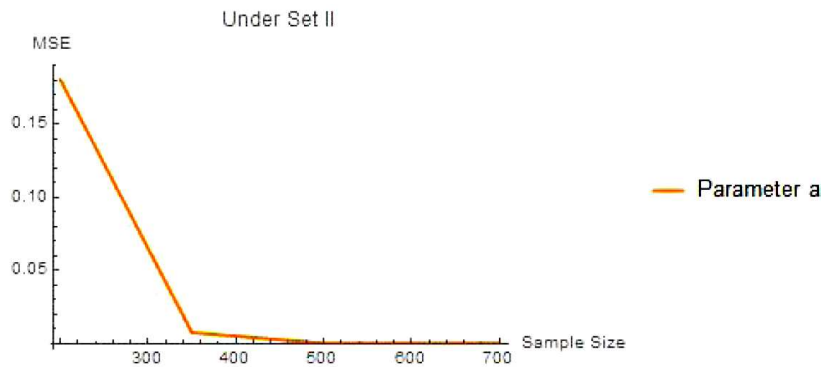


Figure 13. Decreasing MSE for increasing sample size.

Table 3. Result of simulation study for Set III

Parameter m		
Sample Size	Average Estimate	MSE
200	0.7544334	0.03499849
350	0.7067013	0.007117539
500	0.6946308	0.00157214
700	0.6910952	0.0004869904
Parameter s		
Sample Size	Average Estimate	MSE
200	0.6991688	0.0008378809
350	0.6924615	0.000256121
500	0.6908999	0.0001496388
700	0.6902325	0.0001096653
Parameter a		
Sample Size	Average Estimate	MSE
200	5.208542	1.139824
350	4.94128	0.186333
500	4.872674	0.04494567
700	4.856782	0.01662355

From Table 3, we find that the simulated estimates are close to the true values of the parameters and hence the estimation method is adequate. We have also observed that the estimated mean square errors (MSEs) consistently decrease with increasing sample size as seen in Figures 14-16.

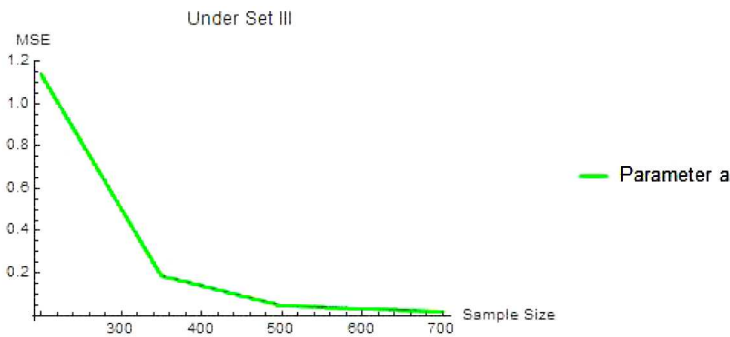


Figure 14. Decreasing MSE for increasing sample size.

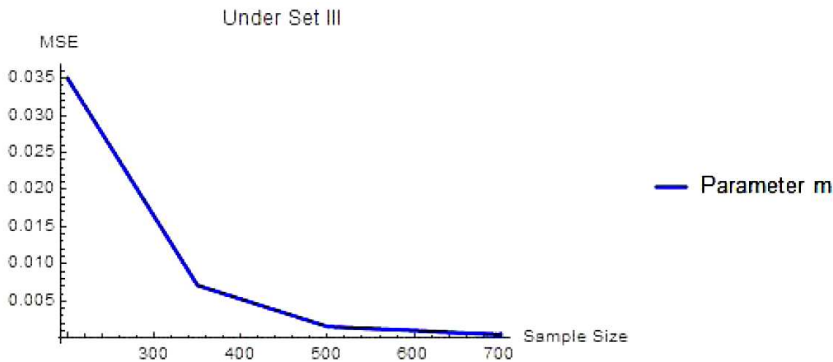


Figure 15. Decreasing MSE for increasing sample size.

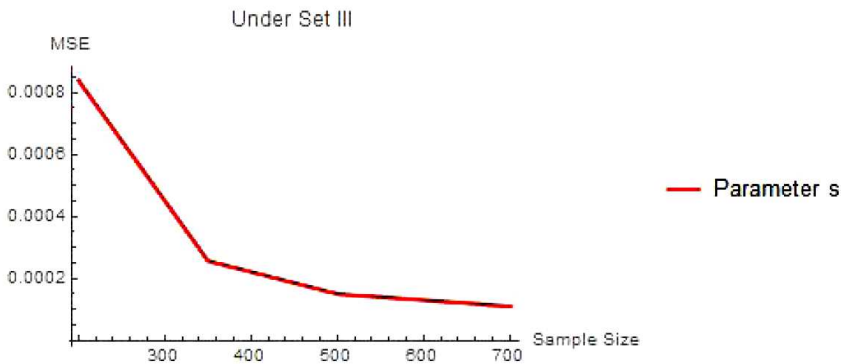


Figure 16. Decreasing MSE for increasing sample size.

7. Application based on the New Family

In this section, we illustrate the usefulness of the new $r_T - X$ family of distributions of type I. We compare the fits of three sub-models of this family of distributions discussed in Section 3 to the carbon fibers data, Section 6 [5]. The estimates of the unknown parameters of the three sub-models of the new $r_T - X$ family of distributions of type I are obtained by the maximum likelihood method using the R language software. The measures of goodness of fit considered included Akaike information criterion (AIC), consistent Akaike information criterion (CAIC), Bayesian information criterion (BIC) and Hannan-Quinn information criterion (HQIC) statistics and they are defined as follows:

$$AIC = -2l + 2k$$

$$BIC = -2l + k \log(n)$$

$$CAIC = AIC + \frac{2k(k+1)}{n-k-1}$$

$$HQIC = 2 \log[\log(n)(k-2l)],$$

where k is the number of parameters in the statistical model, n is the sample size, and $l(\cdot)$ is the maximized value of the log-likelihood function under the considered model.

Table 4. Estimated Parameters for the carbons fiber data

Model	Parameter Estimate	Standard Error
$NELDI(m, s, a)$	$(\hat{m}, \hat{s}, \hat{a}) = (4.4345312, 0.6843875, 11.2804404)$	$(0.31919564, 0.06906468, 2.43779722)$
$NEWDI(a, b, c)$	$(\hat{a}, \hat{b}, \hat{c}) = (36.097941, 3.007914, 9.418255)$	$(50.1500545, 0.3833479, 5.8907485)$
$NENDI(a, b, c)$	$(\hat{a}, \hat{b}, \hat{c}) = (13.865794, 5.048878, 1.522466)$	$(4.098553, 0.552467, 0.199181)$

Table 5. Criteria for comparison

Model	AIC	BIC	CAIC	HQIC
$NELDI(m, s, a)$	297.1061	304.9216	297.3561	14.42224
$NEWDI(a, b, c)$	288.5277	296.3432	288.7777	14.36304
$NENDI(a, b, c)$	291.8597	299.6753	292.1097	14.38624

Since the $NEWDI(a, b, c)$ distribution has the smallest AIC, BIC, CAIC, and HQIC values compared to the other distributions in Table 5, it can be considered a better fit to the carbon fibers data, Section 6 [5].

The asymptotic variance-covariance matrix of the MLEs under the $NELDI(m, s, a)$ $NEWDI(a, b, c)$ and $NENDI(a, b, c)$ distributions, respectively, are given by

$$\begin{bmatrix} 0.10188585 & 0.01798639 & 0.64958252 \\ 0.01798639 & 0.00476993 & 0.08563882 \\ 0.64958252 & 0.08563882 & 5.94285527 \end{bmatrix}$$

$$\begin{bmatrix} 2515.02797 & -15.0434655 & 292.125036 \\ -15.04347 & 0.1469556 & -1.939528 \\ 292.12504 & -1.9395279 & 34.700918 \end{bmatrix}$$

$$\begin{bmatrix} 16.7981229 & 2.0626902 & 0.63823575 \\ 2.0626902 & 0.3052194 & 0.10385962 \\ 0.6382357 & 0.1038596 & 0.03967303 \end{bmatrix}.$$

Hence the approximate 95% confidence intervals for the parameters under the $NELDI(m, s, a)$, $NEWDI(a, b, c)$, and $NENDI(a, b, c)$ distributions, respectively, are given in Tables 6-8.

Table 6. $NELDI(m, s, a)$ distribution

Parameter	CI
m	(4.414515, 4.454547)
s	(0.6800567, 0.6887183)
a	(11.12757, 11.43331)

Table 7. $NEWDI(a, b, c)$ distribution

Parameter	CI
a	(32.95319, 39.24269)
b	(2.983876, 3.031953)
c	(9.048866, 9.787645)

Table 8. $NENDI(a, b, c)$ distribution

Parameter	CI
a	(13.60879, 14.12280)
b	(5.014235, 5.083522)
c	(1.509976, 1.534956)

8. Further Developments

The Ampadu-G family of distributions appeared in [2] as a further development we ask the reader to investigate properties and applications of a so-called Ampadu-new $r_T - X$ family of distributions of type I. We leave the reader with the following.

Definition 8.1. A random variable M will be called *Ampadu-new $r_T - X$ distributed of type I* if the CDF is given by

$$\frac{1 - e^{-\lambda \left(R_T \left(\frac{F(x; \xi) - \text{ProductLog}(F(x; \xi))}{F(x; \xi)} \right) \right)^2}}{1 - e^{-\lambda}},$$

where the random variable T with support $[0, \infty)$ has CDF R_T , $\text{ProductLog}[z]$ gives the principal solution for w in $z = we^w$, and the random variable X has CDF $F(x; \xi)$, $\lambda > 0$, and $\xi > 0$ is a vector of parameters all of whose entries are positive.

References

- [1] Clement Boateng Ampadu, The $r_T - X$ family of distributions induced by V: A new method for generating continuous distributions with illustration to cancer patients data, *Annals of Biostatistics & Biometric Applications* 2(2) (2019). ABBA.MS.ID.000531. <https://doi.org/10.33552/ABBA.2019.02.000531>
- [2] Clement Boateng Ampadu, The Ampadu-G family of distributions with application to the $T-X(W)$ class of distributions, *Annals of Biostatistics & Biometric Applications* 1(4) (2019). ABBA.MS.ID.000518. <https://doi.org/10.33552/ABBA.2019.01.000518>
- [3] Zubair Ahmad, M. Elgarhy and G. G. Hamedani, A new Weibull-X family of distributions: properties, characterizations and applications, *Journal of Statistical Distributions and Applications* 5 (2018), Article No. 5. <https://doi.org/10.1186/s40488-018-0087-6>
- [4] Wikipedia, Lambert W function. https://en.wikipedia.org/wiki/Lambert_W_function
- [5] Suleman Nasiru, Peter N. Mwita and Oscar Ngesa, Exponentiated generalized transformed-transformer family of distributions, *Journal of Statistical and Econometric Methods* 6(4) (2017), 1-17.
- [6] I. S. Gradshteyn and I. M. Ryzhik, *Tables of Integrals, Series, and Products*, New York: Academic Press, 2007.