J. Nat. Acad. Math.
Vol. 34 (2020), pp.71-92

ISSN 0970-5228

ARCTAN GENERALIZED INVERTED EXPONENTIAL DISTRIBUTION: PROPERTIES AND APPLICATIONS

By

Ram Suresh Yadav and Vijay Kumar

Abstract

In this paper, a three-parameter new distribution called arctan generalized inverted exponential distribution is presented. Some mathematical properties of the distribution such as the shapes of the cumulative density, probability density, probability density, and hazard rate functions, survival function, quantile function, the kurtosis and skewness measures are established. To estimate the model parameters, we have employed three well-known estimation methods namely least-square estimation (LSE), maximum likelihood estimation (MLE), and Cramer-Von-Mises (CVM) methods. For the illustration purposes we have considered the two real data sets and goodness-of-fit statistics AIC, BIC, AICC and HQIC are calculated. It is found that the new distribution performs better as compared to some existing distribution.

Keywords : Arctan distribution, Generalized Inverted exponential, Estimation.
2010 AMS Subject Classification : 62F15, 65C05.

1. Introduction

Last few decades, most of the writers are paying attention towards the exponential distribution for its potential in modeling lifetime data, and it has been found that this distribution has performed remarkably in many applications due to the existence of closed form solutions to many reliability and survival analysis. It can easily be justified under the supposition of constant failure rate but in the practice, the failure rates are not always constant. Hence, haphazard use of exponential lifetime model seems to be inappropriate and unrealistic. In recent years, new classes of models have been introduced based on modifications of the existing classical probability models, Marshall and Olkin (2007). Recently, some attempts have been made to generate new distributions to extend well known distributions and at the same time provide great flexibility in modeling data in practice.

Received : July 26, 2019; Accepted : November 07, 2019

Several procedures could be employed to form a larger family from an existing distribution by incorporating extra parameters. So, several classes by adding one or more parameters to generate new models have been proposed in the statistical literature, Rinne (2009) and Pham and Lai (2007).

A random variable Y is said to follow exponential distribution with parameter λ if its cumulative probability distribution (CDF) function is given by

$$F(x;\theta) = 1 - e^{-\theta x}\,; x > 0, \theta > 0. \tag{1.1}$$

Many generalizations exponential distribution have been found in statistical literatures to generate more flexible life-time models. Some of them well known generalizations are as follows.

The modification of exponential distribution was introduced by Smith and Bain (1975) called exponential power distribution. The generalized exponential distribution was introduced by (Gupta & Kundu, 1999) which is flexible then exponential distribution, having increasing and decreasing failure rate hazard function. The probability density function of generalized exponential distribution is

$$f_{GE}\,(x;\,\alpha,\;\lambda) = \alpha\,\lambda\,e^{-\lambda\,x}\left\{1 - e^{-\lambda\,x}\right\}^{\alpha-1}\;;\;(\alpha,\;\lambda) > 0,\;x > 0. \tag{1.2}$$

Lan and Leemis (2008) has proposed the logistic-exponential distribution. It has increasing, de-creasing, bathtub (BT)-shaped, and upside-down bathtub (UBT)-shaped failure rates for various values of the parameters. Another generalization of exponential distribution was introduced by Nadarajah and Haghighi (2011) and called it as an extension of exponential distribution. Its density can have decreasing and unimodal shapes, and the hazard rate exhibits increasing and decreasing shapes. Joshi (2015) has proposed another extension of exponential distribution called new extended exponential (EEN) distribution having monotonically increasing and constant hazard rate shapes. The continuous random variable X follows EEN distribution with parameters α and λ if its CDF is given by

$$F(x) = 1 - \exp(-\alpha x e^{-\lambda/x})\;; x > 0, (\alpha, \lambda) > 0. \tag{1.3}$$

In this work, we have used the generalized inverted exponential distribution (GIED) (Dey & Dey, 2014) and this distribution was also used by (Abouammoh & Alshingiti, 2009) and (Krishna & Kumar, 2013) in reliability estimation. Dube et al., (2016) have also used GIED under the progressive first-failure censoring data. Similarly inverse generalized gompertz distribution has introduced by (Chaudhary & Kumar, 2017). Joshi and Kumar (2018) have introduced the new distribution called inverse upside down bathtub-shaped hazard function distribution.

Gomez-Doniz and Calderon-Ojeda (2015) have introduced the arctan distribution which was used to model Norwegian fire insurance data. This family of distribution has been proposed for an underlying Pareto distribution and the new distribution called Pareto arctan distribution and found that this distribution provide a good fit as compared to some well-known distributions. The CDF and PDF of arctan family of distribution with support $[a, b]$ are given by

$$F(x) = 1 - \frac{\arctan[\alpha\{1 - G(x)\}]}{\arctan(\alpha)}; x \geqslant 0, \alpha > 0; \alpha, x[a, b] \tag{1.4}$$

and

$$f(x) = \frac{1}{\arctan(\alpha)} \frac{\alpha g(x)}{1 + [\alpha\{1 - G(x)\}]^2}; x \geqslant 0, \alpha > 0, \tag{1.5}$$

respectively. Here $G(x)$ and $g(x)$ are the CDF and PDF of any base distribution.

The main objective of this study is to introduce a powerful distribution by adding just one extra parameter to the generalized inverted exponential distribution to attain a good fit to real data. We investigate the properties of the proposed distribution and illustrate its applicability. The different sections of the proposed study are arranged as follows. The new arctan generalized inverted exponential distribution is introduced and various distributional properties are discussed in Section 2. To estimate the model parameters, we have employed three well-known estimation methods namely maximum likelihood estimation (MLE), least-square estimation (LSE), and Cramer-Von-Mises (CVM) methods in Section 3. In Section 4 we have considered two real data sets to analyzed and explore the applications of the proposed distribution. In this section, we present the ML estimators of the parameters and approximate confidence intervals also for the above-mentioned method of estimations, AIC, BIC, AICC and HQIC are calculated to assess the validity of the arctan generalized exponential model. Finally, Section 5 ends up with some general concluding remarks.

2. The arctan generalized inverted exponential(ATGIE) distribution:

The generalized inverted exponential distribution was introduced by (Dey & Dey, 2014) and the CDF and PDF are

$$G(x; \beta, \lambda) = 1 - (1 - e^{-\lambda/x})^{\beta} \; ; x \geqslant 0, (\beta, \lambda) > 0. \tag{2.1}$$

$$g(x; \beta, \lambda) = \frac{\beta\lambda}{x^2} e^{-\lambda/x} (1 - e^{-\lambda/x})^{\beta - 1} \; ; x \geqslant 0, (\beta, \lambda) > 0. \tag{2.2}$$

The CDF and PDF of ATGIE distribution with parameters α, β and λ can be obtained by substituting the equations (2.1) and (2.2) in (1.4) and (1.5) as

$$F(x) = 1 - \frac{\arctan[\alpha(1-e^{-\lambda/x})^{\beta}]}{\arctan(\alpha)}; x \geqslant 0, (\alpha, \beta, \lambda) > 0 \tag{2.3}$$

$$f(x) = \frac{\alpha\beta\lambda}{\arctan(\alpha)} \frac{x^{-2}e^{-\lambda/x}(1-e^{-\lambda/x})^{\beta-1}}{1+[\alpha(1-e^{-\lambda/x})^{\beta}]^{2}} \; ; x \geqslant 0, (\alpha, \beta, \lambda) > 0. \tag{2.4}$$

We shall denote $X \sim ATGIE\,(\alpha, \beta, \lambda)$. Figure 1 demonstrates the graph for PDF and hazard function for ATGIE distribution for different values of parameters. From Figure 1 (left panel), the density function of the ATGIE distribution can bear different shapes according to the values of the parameters.

Survival function: The survival function $R\,(t)$, which is the probability of an item surviving up to time t, is defined by $R\,(t) = 1 - F\,(t)$. The survival /reliability function of a ATGIE distribution is given by

$$R(x) = 1 - F(x) = \frac{\arctan[\alpha\{(1-e^{-\lambda/x})^{\beta}\}]}{\arctan(\alpha)}; x \geqslant 0, (\alpha, \beta, \lambda) > 0. \tag{2.5}$$

The hazard rate function (HRF): Let t be survival time of a component or item and the probability that it will not survive for an additional time Δt then, hazard rate function is,

$$h\,(t) = \frac{f\,(t)}{R\,(t)},$$

where $R(t)$ is a survival function. Hence let, $X \sim ATGIE\,(\alpha, \beta, \lambda)$ then its hazard rate function is

$$h(x) = \frac{\alpha\beta\lambda}{\arctan[\alpha\{(1-e^{-\lambda/x})^{\beta}\}]} \frac{x^{-2}e^{-\lambda/x}(1-e^{-\lambda/x})^{\beta-1}}{1+[\alpha(1-e^{-\lambda/x})^{\beta}]^{2}}. \tag{2.6}$$

The Reversed hazard rate function: The reversed hazard rate function is given by

$$r(x) = \frac{f(x)}{F(x)} = \frac{\alpha\beta\lambda}{\arctan(\alpha) - \arctan[\alpha\{(1-e^{-\lambda/x})^{\beta}\}]} \frac{x^{-2}e^{-\lambda/x}(1-e^{-\lambda/x})^{\beta-1}}{1+[\alpha(1-e^{-\lambda/x})^{\beta}]^{2}}. \tag{2.7}$$

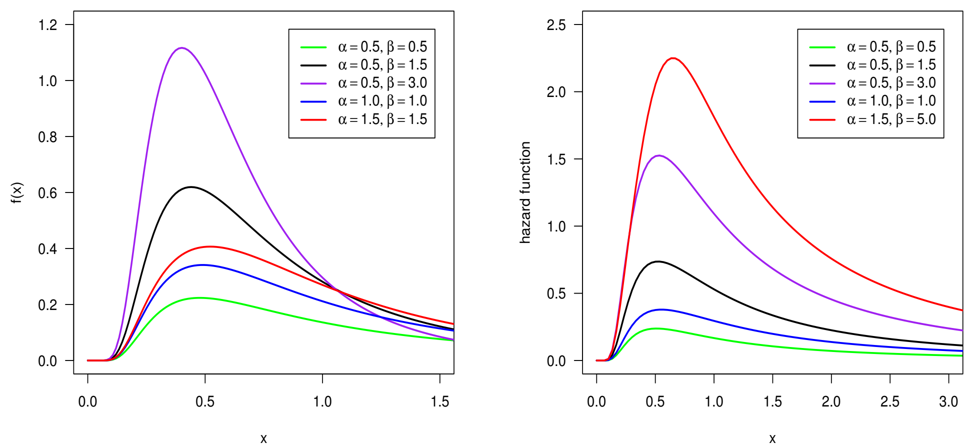

Figure 1. Plots of the probability density function(left panel) and hazard function (right panel), for λ=1 and different values of α and β.

Quantile function: The value of the p_{th} quantile can be obtained by solving the following equation,

$$Q\left(p\right)=F^{-1}\left(p\right),$$

and we get quantile function by inverting (2.3) as

$$Q(p)=-\lambda\left[\log\left\{1-\left[\frac{1}{\alpha}\tan\left\{(1-u)\arctan\alpha\right\}\right]^{1/\beta}\right\}\right]^{-1};0<p<1. \qquad (2.8)$$

For the generation of the random numbers of the ATGIE distribution, we suppose simulating values of random variable X with the CDF (2.3). Let U denote a uniform random variable in (0,1), then the simulated values of X can be obtained by

$$x=-\lambda\left[\log\left\{1-\left[\frac{1}{\alpha}\tan\left\{(1-u)\arctan\alpha\right\}\right]^{1/\beta}\right\}\right]^{-1};0<u<1. \qquad (2.9)$$

Skewness and Kurtosis: The skewness and kurtosis measures are used in statistical analyses to characterize a distribution or a data set. The Bowley's skewness measure based on quartiles is given by

$$S_k=\frac{Q\left(0.75\right)+Q\left(0.25\right)-2Q\left(0.5\right)}{Q\left(0.75\right)-Q\left(0.25\right)}, \qquad (2.10)$$

and the Moors's kurtosis measure based on octiles (Moors, 1988)) is given by

$$K_u = \frac{Q\,(0.875) - Q\,(0.625) + Q\,(0.375) - Q\,(0.125)}{Q\,(0.75) - Q\,(0.25)}, \qquad (2.11)$$

where the $Q\,(.)$ is the quantile function.The skewness and kurtosis measures based on quantiles like Bowley's skewness and Moors's kurtosis have a number of advantages compared to the classical measures of skewness and kurtosis, e.g. they are less sensitive to outliers and they exist for the distributions even without defined the moments.

Some useful expansion of ATGIE distribution:

Here we have expand the CDF and PDF of ATGIE distribution by using the following series expansions

$$(1-a)^{-n} = \sum_{j=0}^{\infty} \binom{n+j-1}{n-1} \mathrm{a}^j,$$

$$Arc\tan x = \frac{\pi}{2} - \sum_{i=0}^{\infty} \frac{(-1)^i}{2i+1} x^{-(2i+1)}, \text{ and}$$

$$e^{-x} = \sum_{n=0}^{\infty} \frac{(-1)^n x^n}{n!}$$

The CDF of ATGIE distribution defined in (2.3) is

$$\begin{aligned} F(x) &= 1 - \frac{\arctan[\alpha(1-e^{-\lambda/x})^{\beta}]}{\arctan(\alpha)}; x \geqslant 0, (\alpha, \beta, \lambda) > 0 \\ &= 1 - \frac{1}{\arctan(\alpha)} \left[\frac{\pi}{2} - \sum_{i=0}^{\infty} \frac{(-1)^i}{2i+1} \left\{ \alpha(1-e^{-\lambda/x})^{\beta} \right\}^{-(2i+1)} \right] \qquad (2.12) \\ &= \omega + \sum_{i=0}^{\infty} \sum_{k=0}^{\infty} d_{ik} e^{-\lambda k/x} \end{aligned}$$

where

$$d_{ik} = \frac{(-1)^i}{(2i+1)} \frac{\alpha^{-(2i+1)}}{\arctan(\alpha)} \binom{k+\beta(2i+1)-1}{\beta(2i+1)-1}$$

and

$$\omega = 1 - \frac{\pi}{2\arctan(\alpha)}.$$

Differentiating (2.12) with respect to x we get PDF of ATGIE distribution as,

$$f(x) = \sum_{i=0}^{\infty} \sum_{k=0}^{\infty} d_{ik} e^{-\lambda k/x} \frac{\lambda k}{x^2} = \sum_{i=0}^{\infty} \sum_{k=0}^{\infty} d^*_{ik} x^{-2} e^{-\lambda k/x} \qquad (2.13)$$

where $d^*_{ik} = d_{ik} \times \lambda k$.

Moments : The r^{th} moment about origin of ATGIE distribution can be written as

$$\begin{aligned}
\mu'_r &= E\left(X^r\right) \\
&= \int_0^\infty x^r f\left(x\right) dx \\
&= \int_0^\infty \sum_{i=0}^\infty \sum_{k=0}^\infty d^*_{ik}\; x^{r-2} e^{-\lambda k/x} dx \\
&= \sum_{i=0}^\infty \sum_{k=0}^\infty d^*_{ik} \frac{\Gamma(1-r)}{[\lambda k]^{1-r}}.
\end{aligned}$$

where $\int\limits_0^\infty x^n e^{-ax} dx = \frac{\Gamma(n+1)}{a^{n+1}}$ is a standard gamma integral.

Moment Generating Function (mgf): Let X be a random variable, then the mgf of X can be defined as

$$\begin{aligned}
M_X\left(t\right) = E\left(e^{tX}\right) &= \sum_{l=0}^\infty \frac{t^l}{l!} \mu'_r \\
&= \sum_{l=0}^\infty \sum_{i=0}^\infty \sum_{k=0}^\infty d^*_{ik} \frac{t^l}{l!} \frac{\Gamma(1-r)}{[\lambda k]^{1-r}}.
\end{aligned}$$

Conditional Moments: The conditional moments of random variable X that follows ATGIE distribution can be expressed as

$$\begin{aligned}
E\left(X^n | X > x\right) &= \frac{1}{S(x)} \int_x^\infty x^n\; f(x)\; dx \\
&= \frac{1}{S(x)} \sum_{i=0}^\infty \sum_{k=0}^\infty d^*_{ik} \int_x^\infty x^{n-2}\; e^{-\lambda k/x}\; dx \\
&= \frac{1}{S(x)} \sum_{i=0}^\infty \sum_{k=0}^n d^*_{ik} \frac{\gamma\left(1-n, \lambda k x\right)}{[\lambda k]^{1-n}}
\end{aligned}$$

where $S(x)$ is survival function and $\int\limits_t^\infty x^a e^{-bx} dx = \frac{\gamma(a+1,bt)}{b^{a+1}}$ is lower incomplete gamma function.

Order Statistics for ATGIE distribution:

Let $X_{k:n}$ represents the k^{th} order statistic of $X_1, ..., X_n$ and $f_{k:n}$ indicates PDF of k^{th} then

$$f_{k:n}(x) = \frac{n!}{(k-1)!\,(n-k)!} f(x)\,[F(x)]^{k-1}[1-F(x)]^{n-k}$$
$$= \frac{n!}{(k-1)!\,(n-k)!} f(x) \sum_{j=1}^{n-k} \binom{n-k}{j} [F(x)]^{j+k-1}$$

Using CDF and PDF defined in (2.12) and (2.13) we get

$$f_{k:n}(x) = \sum_{i=0}^{\infty}\sum_{k=0}^{n} D_{ik} x^{-2} e^{-\lambda k/x} \sum_{j=1}^{n-k} \binom{n-k}{j} \left[\omega + \sum_{i=0}^{\infty}\sum_{k=0}^{\infty} d_{ik} e^{-\lambda k/x}\right]^{j+k-1}$$

where $D_{ik} = \frac{n!}{(k-1)!(n-k)!} \times d^*_{ik}$.

3. METHODS OF ESTIMATION

The object of estimation is to evaluate a model parameter value based on sample information. The estimation theory deals with the basic problem of inferring some relevant features of a chance experiment centered on the observation of the experiment outcomes. There are so many methods which are used to evaluate values of parameters. Three kinds of parameter estimation methods have been considered, such as MLE, LSE, and the Cramer-von Mises (CVM) methods.

(a) Maximum Likelihood Estimation:

In this section, we have illustrated the maximum likelihood estimators (MLE's) of the ATGIE(α, β, λ) distribution. Let $\underline{x} = (x_1, \ldots, x_n)$ be the observed values of size n from ATGIE(α, β, λ) then the likelihood function for the parameter vector $\Psi = (\alpha, \beta, \lambda)^T$ can be written as,

$$L(\Psi) = \left(\frac{\alpha\beta\lambda}{\arctan(\alpha)}\right)^n \prod_{i=1}^{n} \frac{x_i^{-2} e^{-\lambda/x_i} \left(1 - e^{-\lambda/x_i}\right)^{\beta-1}}{1 + \left[\alpha(1 - e^{-\lambda/x_i})^{\beta}\right]^2}$$

It is easy to deal with log-likelihood function as,

$$\ell(\alpha, \beta, \lambda|\underline{x}) = n\ln(\alpha\beta\lambda) - n\ln\{arctan(\alpha)\} - 2\sum_{i=1}^{n} x_i - \lambda\sum_{i=1}^{n} x_i^{-1}$$
$$-(\beta-1)\sum_{i=1}^{n}\ln(1 - e^{-\lambda/x_i}) - \sum_{i=1}^{n}\ln\left\{1 + \left[\alpha(1 - e^{-\lambda/x_i})^{\beta}\right]^2\right\} \tag{3.1}$$

The elements of the score function $B(\Psi) = (B_\alpha, B_\beta, B_\lambda)$ are obtained as

$$\frac{\partial \ell}{\partial \alpha} = \frac{n}{\alpha} - \frac{n}{\arctan(\alpha)[1+\alpha^2]} - 2\alpha \sum_{i=1}^{n} \frac{\left(1-e^{-\lambda/x_i}\right)^{2\beta}}{1+\left[\alpha(1-e^{-\lambda/x_i})^{\beta}\right]^2]}$$

$$\begin{aligned}\frac{\partial \ell}{\partial \beta} &= \frac{n}{\beta} + \sum_{i=1}^{n} \ln(1-e^{-\lambda/x_i}) - \alpha^2 \sum_{i=1}^{n} \frac{\left(1-e^{-\lambda/x_i}\right)^{2\beta} \ln(1-e^{-\lambda/x_i})}{1+\left[\alpha(1-e^{-\lambda/x_i})^{\beta}\right]^2} \\ \frac{\partial \ell}{\partial \lambda} &= \frac{n}{\lambda} - \sum_{i=1}^{n} x_i^{-1} - (\beta-1) \sum_{i=1}^{n} \frac{e^{-\lambda/x_i}}{x_i(1-e^{-\lambda/x_i})} \\ &\quad -2\alpha^2\beta \sum_{i=1'}^{n} \frac{x_i^{-1}(1-e^{-\lambda/x_i})^{2\beta-1} e^{-\lambda/x_i}}{1+\left[\alpha(1-e^{-\lambda/x_i})^{\beta}\right]^2}\end{aligned} \tag{3.2}$$

Equating B_α, B_β and B_λ to zero and solving these non-linear equations simultaneously gives the MLE $\hat{\Psi} = \left(\hat{\alpha}, \hat{\beta}, \hat{\lambda}\right)$ of $\Psi = (\alpha, \beta, \lambda)^T$. These equations cannot be solved analytically and by using the computer software R, Mathematica, Matlab, or any other programs and Newton-Raphson's iteration method, one can solve these equations. Using the asymptotic normality of MLEs, approximate $100(1-\gamma)\%$ confidence intervals for α, β and λ can be constructed as,

$$\hat{\alpha} \pm z_{\gamma/2}\sqrt{\operatorname{var}(\hat{\alpha})},\ \hat{\beta} \pm z_{\gamma/2}\sqrt{\operatorname{var}(\hat{\beta})} \text{ and } \hat{\lambda} \pm z_{\gamma/2}\sqrt{\operatorname{var}(\hat{\lambda})}, \tag{3.3}$$

where $z_{\gamma/2}$ is the upper percentile of standard normal variate.

(b) Least-Square Estimation (LSE)Method:

The least-square estimators of the unknown parameters α, β and λ of ATGIE distribution can be obtained by minimizing

$$M(X; \alpha, \beta, \lambda) = \sum_{i=1}^{n} \left[F(X_{(i)}) - \frac{i}{n+1}\right]^2, \tag{3.4}$$

with respect to (w.r.t.) α, β and λ, (Swain et al., 1988).

From a distribution function $F(.)$, consider random sample be denoted by $\{X_1, \ldots, X_n\}$ with sample size is n where $F(X_{(i)})$ represents the distribution function of the random variables ordered $X_{(1)}, X_{(2)}, \ldots, X_{(n)}$. Then LSE ($\tilde{\alpha}, \tilde{\beta}$ and $\tilde{\lambda}$) is acquired with minimization of

$$M(X; \alpha, \beta, \lambda) = \sum_{i=1}^{n} \left[1 - \frac{\arctan[\alpha(1-e^{-\lambda/x_{(i)}})^{\beta}]}{\arctan(\alpha)} - \frac{i}{n+1}\right]^2 \tag{3.5}$$

with respect to α, β and λ. Differentiation of (3.5) with respect to α, β and λ yields,

$$\frac{\partial M}{\partial \alpha} = -2\sum_{i=1}^{n}\left[1-\frac{\arctan[\alpha T\left(x_i\right)]}{\arctan(\alpha)} \quad -\frac{i}{n+1}\right]\times$$
$$\left[\frac{T\left(x_i\right)}{(\alpha^2\{T\left(x_i\right)\}^2+1)}-\frac{\arctan[\alpha T\left(x_i\right)]}{(\alpha^2+1)\arctan^2(\alpha)}\right]$$
$$\frac{\partial M}{\partial \beta} = \frac{2\alpha}{\arctan(\alpha)}\sum_{i=1}^{n}\left[1-\frac{\arctan[\alpha T\left(x_{(i)}\right)]}{\arctan(\alpha)} \quad -\frac{i}{n+1}\right]\left[\frac{T\left(x_{(i)}\right)\ln(1-e^{-\lambda/x_{(i)}})}{(\alpha^2\{T\left(x_{(i)}\right)\}^2+1)}\right]$$
$$\frac{\partial M}{\partial \lambda} = -\frac{2\alpha\beta}{\arctan(\alpha)}\sum_{i=1}^{n}\left[1-\frac{\arctan[\alpha T\left(x_{(i)}\right)]}{\arctan(\alpha)} \quad -\frac{i}{n+1}\right]\left[\frac{\left(1-e^{-\lambda/x_{(i)}}\right)^{\beta-1}e^{-\lambda/x_{(i)}}}{x_{(i)}(\alpha^2\{T\left(x_{(i)}\right)\}^2+1)}\right].$$

where

$$T\left(x_{(i)}\right) = (1-e^{-\lambda/x_{(i)}})^{\beta}.$$

The LSE estimators can be obtained by solving

$$\frac{\partial M}{\partial \alpha}=0, \frac{\partial M}{\partial \beta}=0 \; and \; \frac{\partial M}{\partial \lambda}=0. \tag{3.7}$$

simultaneously.

Likewise, the weighted LSEs can be found with minimization w.r.t. α, β and λ.

$$M\left(X;\alpha,\beta,\lambda\right) = \sum_{i=1}^{n} w_i\left[F(X_{(i)})-\frac{i}{n+1}\right]^2.$$

The weights w_i are

$$w_i = \frac{1}{Var(X_{(i)})} = \frac{\left(n+1\right)^2\left(n+2\right)}{i\left(n-i+1\right)}.$$

Hence, the weighted LSEs of α, β and λ can be found respectively by minimizing following function w.r.t. α, β and λ.

$$M\left(X;\alpha,\beta,\lambda\right) = \sum_{i=1}^{n}\frac{\left(n+1\right)^2\left(n+2\right)}{i\left(n-i+1\right)}\left[1-\frac{\arctan[\alpha(1-e^{-\lambda/x_{(i)}})^{\beta}]}{\arctan(\alpha)}-\frac{i}{n+1}\right]^2$$

(c) Cramer-Von-Mises estimation (CVME):

The CVM estimators for α, β and λ are obtained by minimization of

$$
\begin{aligned}
A\left(X;\alpha,\beta,\lambda\right) &= \frac{1}{12n}+\sum_{i=1}^{n}\left[F\left(x_{(i)}|\alpha,\beta,\lambda\right)-\frac{2i-1}{2n}\right]^{2} \\
&= \frac{1}{12n}+\sum_{i=1}^{n}\left[1-\frac{\arctan[\alpha(1-e^{-\lambda/x_{(i)}})^{\beta}]}{\arctan(\alpha)}-\frac{2i-1}{2n}\right]^{2}.
\end{aligned} \tag{3.6}
$$

Differentiating (3.6) with respect to α, β and λ we get,

$$
\frac{\partial A}{\partial \alpha} = -2\sum_{i=1}^{n}\left[1-\frac{\arctan[\alpha T\left(x_{(i)}\right)]}{\arctan(\alpha)}-\frac{2i-1}{2n}\right]\left[\frac{T\left(x_{(i)}\right)}{(\alpha^{2}\{T\left(x_{(i)}\right)\}^{2}+1)}-\frac{\arctan[\alpha T\left(x_{(i)}\right)]}{(\alpha^{2}+1)\arctan^{2}(\alpha)}\right]
$$

$$
\frac{\partial A}{\partial \beta} = \frac{2\alpha}{\arctan(\alpha)}\sum_{i=1}^{n}\left[1-\frac{\arctan[\alpha T\left(x_{(i)}\right)]}{\arctan(\alpha)}-\frac{2i-1}{2n}\right]\left[\frac{T\left(x_{(i)}\right)\ln(1-e^{-\lambda/x_{(i)}})}{(\alpha^{2}\{T\left(x_{(i)}\right)\}^{2}+1)}\right]
$$

$$
\frac{\partial A}{\partial \lambda} = -\frac{2\alpha\beta}{\arctan(\alpha)}\sum_{i=1}^{n}\left[1-\frac{\arctan[\alpha T\left(x_{(i)}\right)]}{\arctan(\alpha)}-\frac{2i-1}{2n}\right]\left[\frac{(1-e^{-\lambda/x_{(i)}})^{\beta-1}e^{-\lambda/x_{(i)}}}{x_{(i)}(\alpha^{2}\{T\left(x_{(i)}\right)\}^{2}+1)}\right]
$$

where $T\left(x_{(i)}\right) = (1-e^{-\lambda/x_{(i)}})^{\beta}$. The CVM estimators can be found by solving

$$
\frac{\partial A}{\partial \alpha}=0, \frac{\partial A}{\partial \beta}=0 \; and \; \frac{\partial A}{\partial \lambda}=0. \tag{3.7}
$$

simultaneously.

4. Data Analysis: Application

In this section, we illustrate the applicability of arctan generalized inverted exponential distribution using two real datasets used by earlier researchers. To illustrate the goodness of fit of the ATGIE distribution, we have select some well known distribution for comparison purpose which are listed below:

(i) Generalized Exponential Extension (GEE) distribution: The probability density function of GEE introduced by (Lemonte, 2013) having upside down bathtub-shaped hazard function distribution with parameters (α, β, λ) is

$$
f_{GEE}\left(x;\alpha,\beta,\lambda\right) = \alpha\beta\lambda\left(1+\lambda x\right)^{\alpha-1}\exp\left\{1-(1+\lambda x)^{\alpha}\right\}\left[1-\exp\left\{1-(1+\lambda x)^{\alpha}\right\}\right]^{\beta-1} \quad ; \; x \geqslant 0
$$

(ii) Generalized Exponential (GE) distribution: The probability density function of generalized exponential distribution (Gupta & Kundu, 1999) is.

$$f_{GE}\,(x;\,\alpha,\;\lambda) = \alpha\;\lambda\;e^{-\lambda\,x}\Big\{1 - e^{-\lambda\,x}\Big\}^{\alpha-1}\;;\;(\alpha,\;\lambda) > 0,\;x > 0.$$

(iii) Generalized Rayleigh distribution The probability density function of Generalized Rayleigh (GR) distribution (Kundu & Raqab, 2005) is

$$f_{GR}\;(\mathrm{x};\,\alpha,\,\lambda) = 2\;\alpha\;\lambda^2\;\mathrm{x}\;\mathrm{e}^{-(\lambda x)^2}\Big\{1\;-\;\mathrm{e}^{-(\lambda x)^2}\Big\}^{\alpha-1}\quad;\;\;x > 0$$

Here $\alpha > 0$ and $\lambda > 0$ are the shape and scale parameters respectively.

(iv) Modified Weibull (MW) The modified Weibull (MW) distribution was introduced by (Lai et al., 2003) with three parameters $\alpha > 0, \beta > 0$ and $\lambda > 0$. The probability density function (pdf) is

$$f_{MW}(x) = \alpha\,(\lambda + \beta x)\,x^{\lambda-1}\exp(\beta x - \alpha x^{\lambda}e^{\beta x}\;; x > 0$$

(v) Weibull Extension (WE) Model: The probability density function of Weibull extension (WE) distribution (Tang et al., 2003) with three parameters $(\alpha > 0, \beta > 0, \lambda > 0)$ is

$$f_{WE}(x;\alpha,\beta,\lambda) = \lambda\beta\,\left(\frac{x}{\alpha}\right)^{\beta-1}\exp\left(\frac{x}{\alpha}\right)^{\beta}$$
$$\exp\left\{-\lambda\alpha\;\left(\exp\left(\frac{x}{\alpha}\right)^{\beta} - 1\right)\right\}\quad;\;\;x > 0.$$

Dataset I: The data below are from the tensile strength of 69 observations of failure stresses of single carbon fibers of length 50 mm,(Bader & Priest, 1982).

1.312, 1.314, 1.479, 1.552, 1.700, 1.803, 1.861, 1.865, 1.944, 1.958,
1.966, 1.997, 2.006, 2.021, 2.027, 2.055, 2.063, 2.098, 2.140, 2.179,
2.224, 2.240, 2.253, 2.270, 2.272, 2.274, 2.301, 2.301, 2.359, 2.382,
2.382, 2.426, 2.434, 2.435, 2.478, 2.490, 2.511, 2.514, 2.535, 2.554,
2.566, 2.570, 2.586, 2.629, 2.633, 2.642, 2.648, 2.684, 2.697, 2.726,
2.770, 2.773, 2.800, 2.809, 2.818, 2.821, 2.848, 2.880, 2.954, 3.012,
3.067, 3.084, 3.090, 3.096, 3.128, 3.233, 3.433, 3.585, 3.585

We have presented the MLEs directly by using optim() function (R Core Team, 2015) and Rizzo (2008) by maximizing the likelihood function (3.1). We have obtained $\hat{\alpha} = 1.3544(0.9132)$, $\hat{\beta} = 124.7936(4.1417)$, $\hat{\lambda} = 11.8788(0.6331)$ and corresponding value of log-likelihood is -49.5653. Using the method described in Section 3, we can construct the approximate confidence intervals(ACI) based on MLE's. Table 1 shows the MLE's with their standard errors(SE) and 95% confidence intervals for α, β and λ. In Table 1 we have presented the MLE's with their standard errors (SE) and 95% asymptotic confidence intervals(ACI) for α, β and λ.

In Figure 2 we have displayed the graph of profile log-likelihood functions of ML estimates of α, β and λ. We have noticed that ML estimates of α, β and λ exist and can be obtained uniquely.

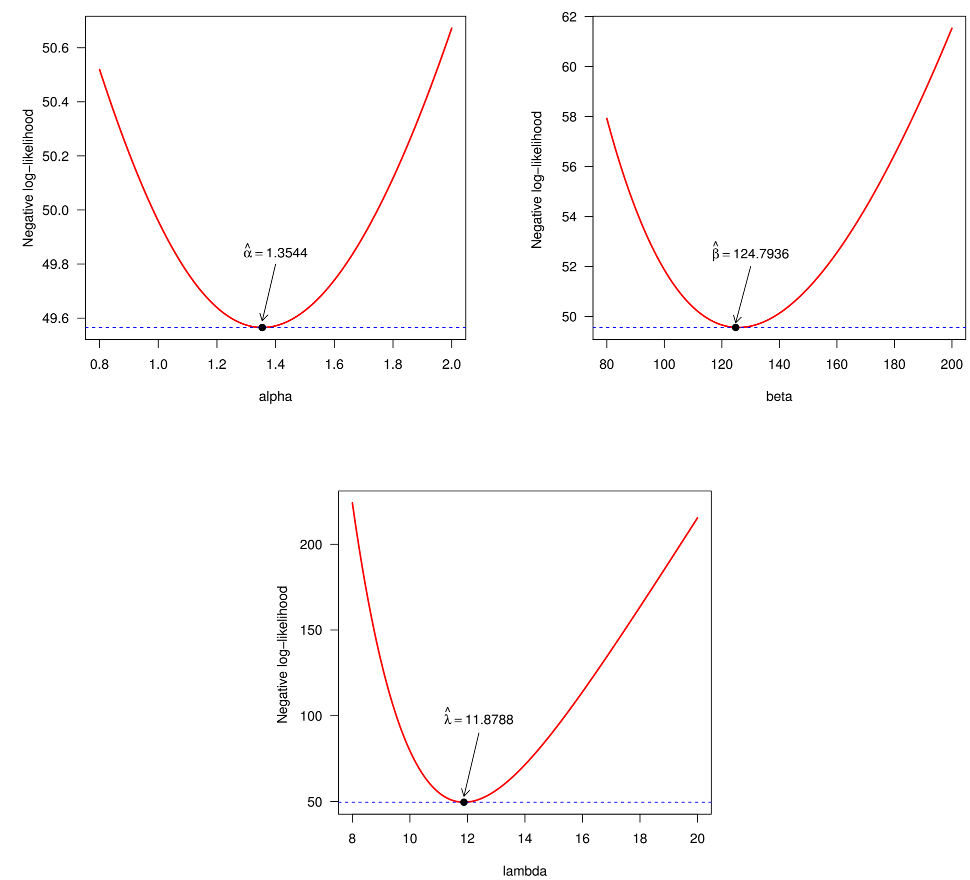

Figure 2. Profile log-likelihood functions of α, β and λ.

To evaluate the goodness of fit of a given distribution we generally use the PDF and CDF plot. To get the additional information we have to plot Q-Q and P-P plots. In particular, the Q-Q plot may provide information about the lack-of-fit at the tails of the distribution, whereas the P-P plot emphasizes the lack-of-fit, (Kumar & Ligges 2011). From Figure 3 we have shown that the ATGIE model fits the data very well.

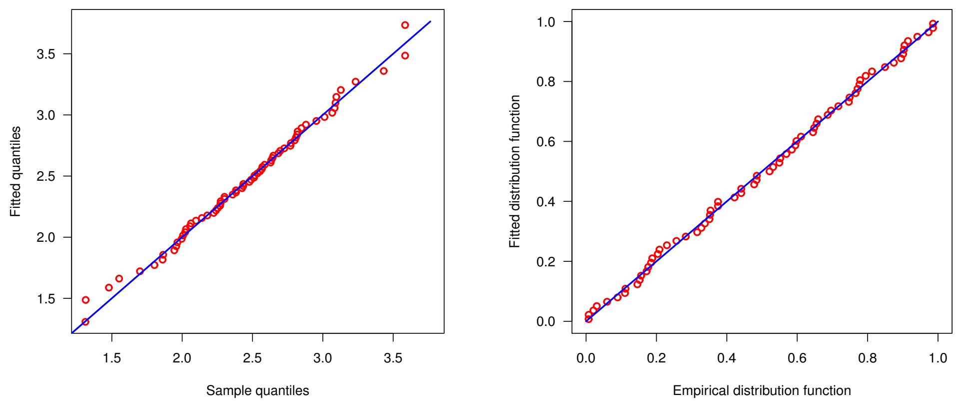

Figure 3. PP plot(left panel) and QQ plot(right panel).

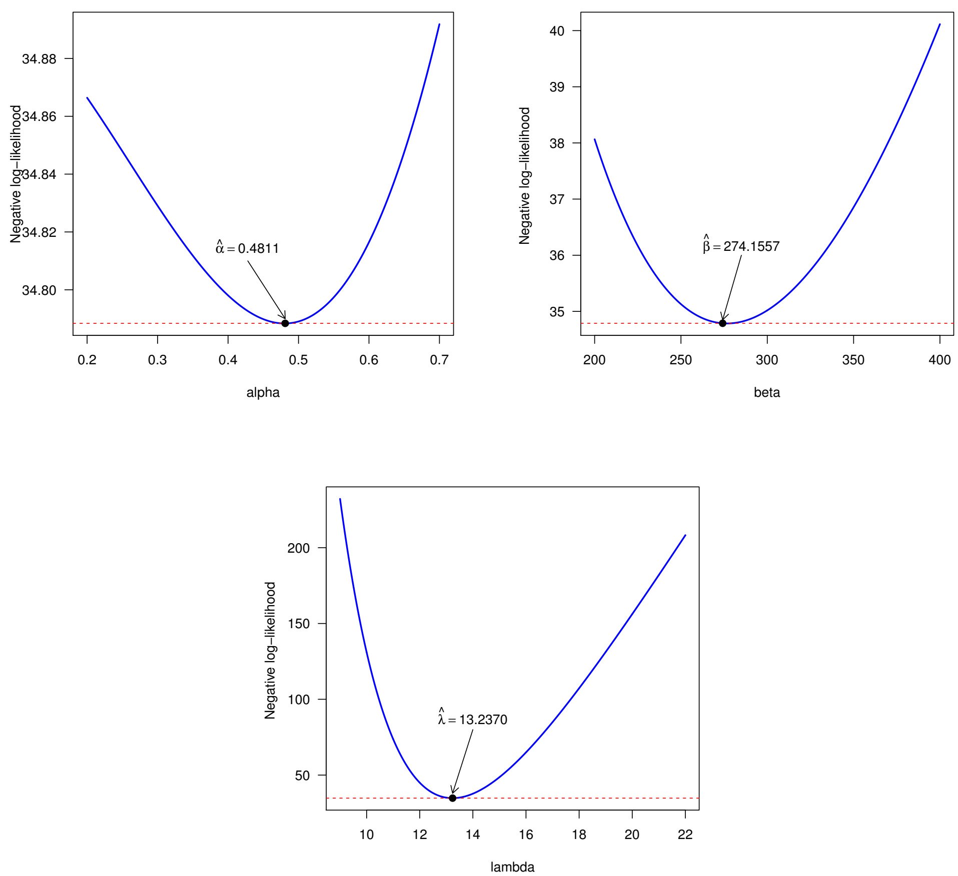

Figure 4. Profile log-likelihood functions of α, β and λ.

Dataset II: For the illustration, we consider the another data set to fit our model on the tensile strength of 65 observations of failure stresses of single carbon fibers of length 50 mm (Bader & Priest, 1982). The data set is also used by Muhammad & Liu (2019). The data is as follows:

1.339, 1.434, 1.549, 1.574, 1.589, 1.613, 1.746, 1.753, 1.764, 1.807,
1.812, 1.840, 1.852, 1.852, 1.862, 1.864, 1.931, 1.952, 1.974, 2.019,
2.051, 2.055, 2.058, 2.088, 2.125, 2.162, 2.171, 2.172, 2.180, 2.194,
2.211, 2.270, 2.272, 2.280, 2.299, 2.308, 2.335, 2.349, 2.356, 2.386,
2.390, 2.410, 2.430, 2.431, 2.458, 2.471, 2.497, 2.514, 2.558, 2.577,
2.593, 2.601, 2.604, 2.620, 2.633, 2.670, 2.682, 2.699, 2.705, 2.735,
2.785, 3.020, 3.042, 3.116, 3.174

We have obtained the MLEs with their standard errors (SE) in parenthesis of ATGIE distribution as $\hat{\alpha} = 0.4811(1.1763), \hat{\beta} = 274.1557(3.0760), \hat{\lambda} = 13.2370(0.6467)$ and corresponding value of log-likelihood is -34.7884. In Figure 5 we have displayed the graph of profile log-likelihood functions of ML estimates of α, β and λ. We have noticed that ML estimates of α, β and λ exist and can be obtained uniquely.

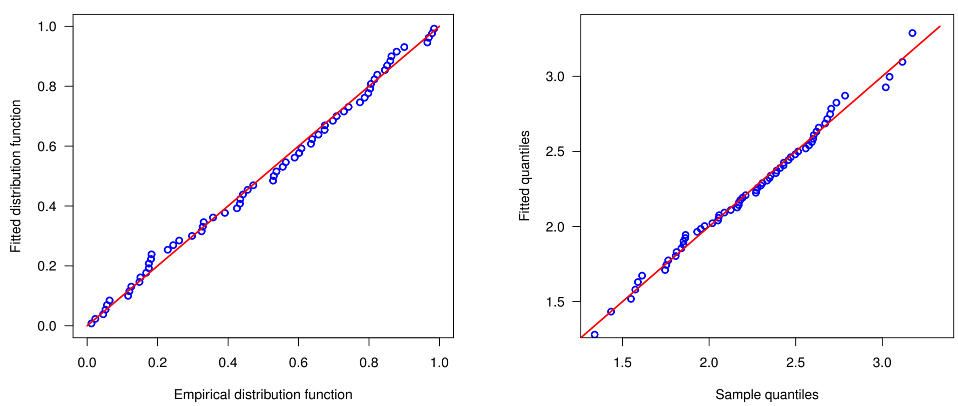

Figure 5. The P-P plot (left panel) and Q-Q plot (right panel).

By using MLE method we estimate the parameter of each of these distributions. For the goodness of fit purpose we use log-likelihood $(l(\hat{\theta}))$ where $\hat{\theta} = (\hat{\alpha}, \hat{\beta}, \hat{\lambda})$ to compute Akaike information criterion (AIC), Bayesian information criterion (BIC), Corrected Akaike Information criterion (CAIC) and Hannan-Quinn information criterion (HQIC), statistic to select the best model among selected models. The expressions to calculate AIC, BIC, CAIC and HQIC are listed below:

$$AIC = -2l(\hat{\theta}) + 2k\,,$$
$$BIC = -2l(\hat{\theta}) + k\log\left(n\right)\,,$$
$$CAIC = AIC + \frac{2k\left(k+1\right)}{n-k-1}\,,$$
$$HQIC = -2l(\hat{\theta}) + 2k\log\left[\log\left(n\right)\right]\,,$$

where k is the number of parameters and n is the size of the sample in the model under consideration. Further, in order to evaluate the fits of the AT-GIE distribution with some selected distributions we have taken the Kolmogorov-Smirnov (KS), the Anderson-Darling (W) and the Cramer-Von Mises (A^2) statistic. These statistics are widely used to compare non-nested models and to illustrate how closely a specific CDF fits the empirical distribution to the given data set. These statistics are calculated as

$$KS = \max_{1\leqslant i\leqslant n}\left(d_i - \frac{i-1}{n}, \frac{i}{n} - d_i\right)\,,$$
$$W = -n - \frac{1}{n}\sum_{i=1}^{n}\left(2i-1\right)\left[\ln d_i + \ln\left(1 - d_{n+1-i}\right)\right],$$
$$A^2 = \frac{1}{12n} + \sum_{i=1}^{n}\left[\frac{\left(2i-1\right)}{2n} - d_i\right]^2,$$

where $d_i = CDF\left(x_i\right)$; the x_i's being the ordered observations, (D'Agostino and Stephens,1986).

For the assessment of potentiality of the proposed model we have calculated the Akaike information criterion (AIC), Bayesian information criterion (BIC), Corrected Akaike information criterion (CAIC) and Hannan-Quinn information criterion (HQIC) and these are presented in Table 2.

Table 1

Estimated parameters, log-likelihood AIC and KS(Data set-I

Method	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	ll	AIC	KS(p-value)
MLE	1.3544	124.7936	11.8788	-49.5653	105.1306	0.0378(0.9999)
LSE	4.3231	47.1571	8.2149	-51.8238	109.6476	0.0377(0.9999)
CVME	2.3575	82.2636	10.2617	-50.0926	106.1853	0.0388(0.9999)

The Histogram and the density function of fitted distributions and Empirical distribution function with estimated distribution function of ATGIE, generalized

Gompertz (GG), generalized exponential extension (GEE), exponential extension (EE), Weibull and EEP distributions are presented in Figure 4.

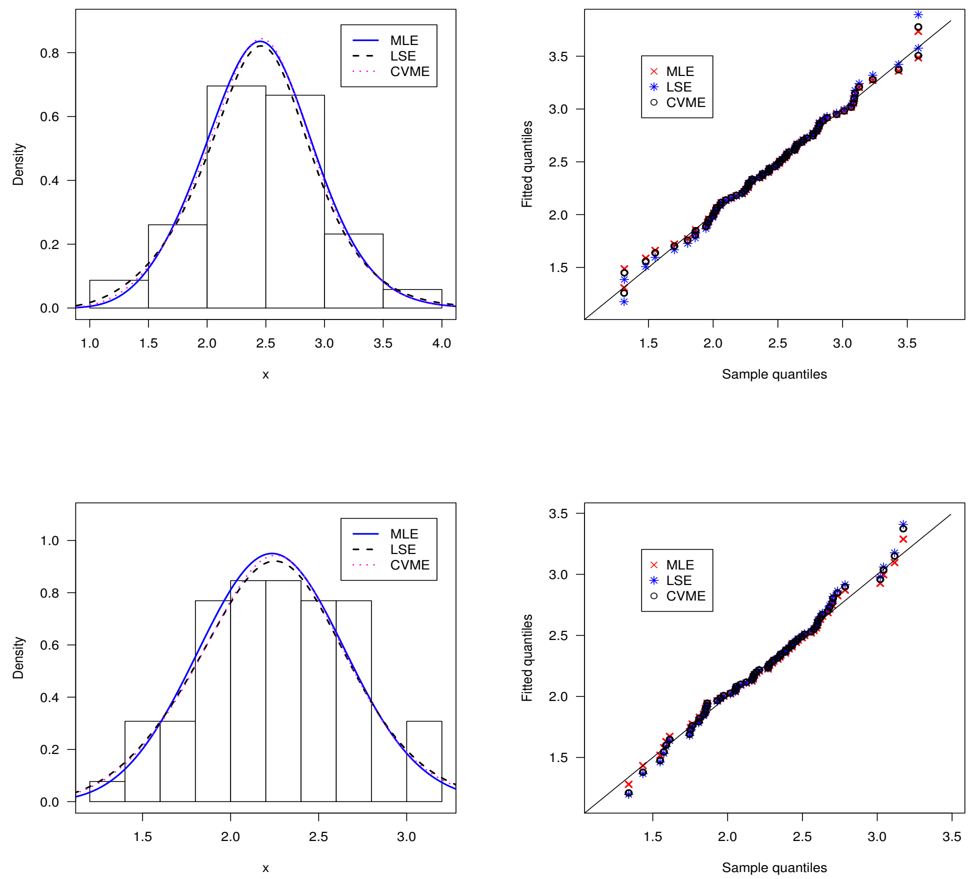

Figure 6. The Histogram and the PDF of fitted distributions(first column) and the fitted qunatiles and sample quantiles(second column) of datasets I and II, respectively for estimation methods(MLE, LSE and CVM).

To compare the goodness-of-fit of the ATGIE distribution with other competing distributions we have presented the value of Kolmogorov-Smirnov (KS), the Anderson-Darling (AD) and the Cramer-Von Mises (CVM) statistics. These three statistics are widely used to compare non-nested models and to illustrate how closely a specific CDF fits the empirical distribution of a given data set. From Table 3 the result shows that the ATGIE distribution has the minimum value of the test statistic and higher p-value hence we conclude that the ATGIE distribution gets quite better fit and more consistent and reliable results from others taken for comparison.

Table 2

Estimated parameters, log-likelihood AIC and KS(Data set-II

Method	$\hat{\alpha}$	$\hat{\beta}$	$\hat{\lambda}$	ll	AIC	KS(p-value)
MLE	0.4811	274.1557	13.2370	-34.7884	75.5767	0.0634(0.9565)
LSE	1.5978	114.4953	10.5397	-35.8880	77.7760	0.0580(0.9809)
CVME	1.6213	126.4087	10.7511	-35.6623	77.3246	0.0604(0.9716)

To evaluate the goodness of fit of a given distribution we generally use the PDF and CDF plot. To get the additional information we have to plot Q-Q and P-P plots. In particular, the Q-Q plot may provide information about the lack-of-fit at the tails of the distribution, whereas the P-P plot emphasizes the lack-of-fit. From Figure 6 we have shown that the ATGIE model fits the data very well.

For the assessment of potentiality of the proposed model we have calculated the Akaike information criterion (AIC), Bayesian information criterion (BIC), Corrected Akaike information criterion (CAIC) and Hannan-Quinn information criterion (HQIC) which are presented in Table 3 and Table 4 for date sets I and II.

Table 3

Log-likelihood(LL), AIC, BIC, CAIC and HQIC(Dataset-I)

Model	-LL	AIC	BIC	CAIC	HQIC
ATGIE	-49.5653	105.1306	111.8329	105.4999	107.7897
MW	-49.6017	105.2033	111.9056	105.5725	107.8623
GEE	-49.6465	105.2930	111.9954	105.6623	107.9521
GR	-50.6292	105.2584	109.7266	105.4402	107.0311
WE	-50.7239	107.4479	114.1502	107.8171	110.1069
GE	-54.6205	113.2409	117.7091	113.4227	115.0136

For the both datasets we have presented the Histogram and the density function of fitted distributions and Empirical distribution function with the estimated distribution function of ATGIE and some selected distributions are presented in Figure 7.

Table 4
Log-likelihood(LL), AIC, BIC, CAIC and HQIC

Model	-LL	AIC	BIC	CAIC	HQIC
ATGIE	-34.7884	75.5767	82.0999	75.9702	78.1505
GEE	-35.0445	76.0889	82.6121	76.4824	78.6627
MW	-35.4552	76.9103	83.4335	77.3038	79.4841
WE	-35.4760	76.9521	83.4753	77.3455	79.5259
GR	-35.7674	75.5349	79.8836	75.7284	77.2507
GE	-38.3657	80.7315	85.0803	80.9250	82.4474

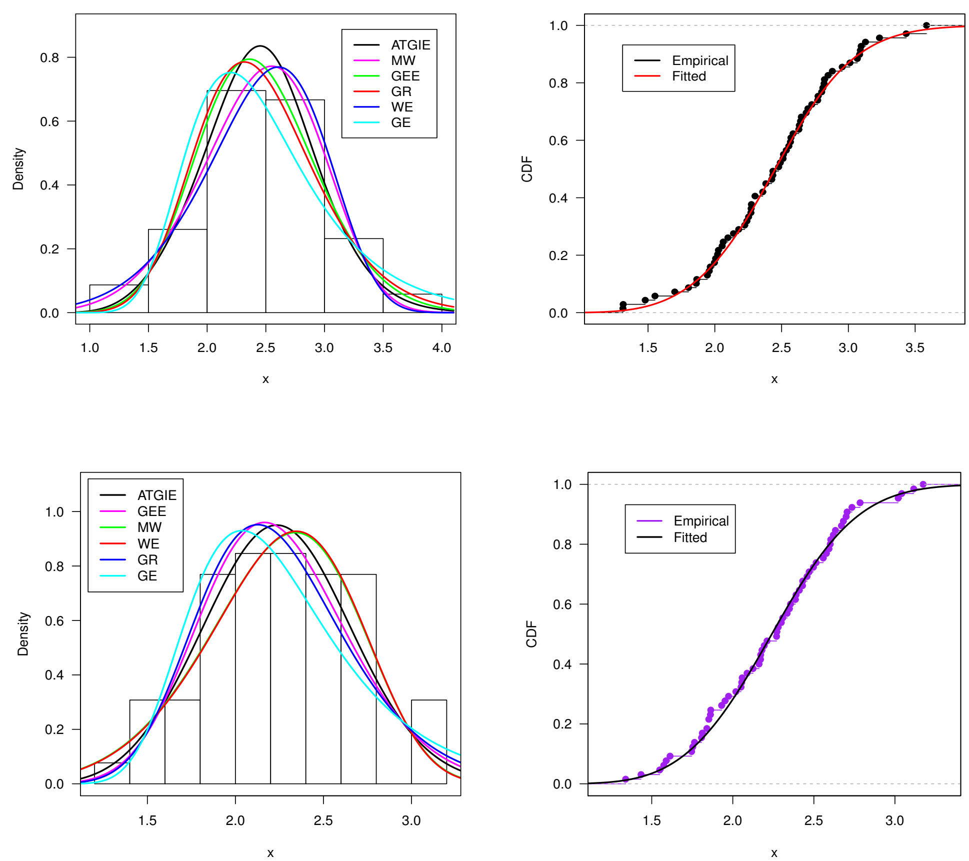

Figure 7. The Histogram and the density function of fitted distributions and Empirical distribution function with estimated distribution function for Dataset-I(left panel) and for Dataset-II(right panel).

To compare the goodness-of-fit of the ATGIE distribution with other competing distributions we have presented the value of Kolmogorov-Smirnov (KS), the Anderson-Darling (W) and the Cramer-Von Mises (A^2) statistics in Table 5 and Table 6. It is observed that the ATGIE distribution has the minimum value of the test statistic and higher p-value for the both datasets I and II, thus we conclude that the ATGIE distribution gets quite better fit and more consistent and reliable results from others taken for comparison.

Table 5

The goodness-of-fit statistics and their corresponding $p-$value (Dataset-I)

Model	*KS(p-value)*	*AD(p-value)*	*CVM(p-value)*
ATGIE	0.0378(0.9999)	0.0142(0.9998)	0.1406(0.9992)
MW	0.0542(0.9873)	0.0326(0.9677)	0.2717(0.9577)
GEE	0.0559(0.9823)	0.0413(0.9279)	0.2924(0.9436)
GR	0.0658(0.9264)	0.0625(0.8000)	0.4417(0.8061)
WE	0.0647(0.9348)	0.0568(0.8357)	0.4431(0.8046)
GE	0.0949(0.5629)	0.1603(0.3603)	1.1235(0.2983)

Table 6

The goodness-of-fit statistics and their corresponding $p-$value (Dataset-II)

Model	*KS(p-value)*	*AD(p-value)*	*CVM(p-value)*
ATGIE	0.0634(0.9565)	0.0284(0.9819)	0.2031(0.9896)
GEE	0.0721(0.8876)	0.0537(0.8550)	0.3239(0.9189)
MW	0.0538(0.9918)	0.0243(0.9915)	0.2696(0.9590)
WE	0.0560(0.9869)	0.0260(0.9880)	0.2787(0.9531)
GR	0.0814(0.7824)	0.0714(0.7446)	0.4282(0.8199)
GE	0.0966(0.5783)	0.1290(0.4614)	0.8200(0.4663)

5. Conclusion

In this paper, a continuous probability distribution called arctan generalized inverted exponential distribution. Some mathematical and statistical properties of the ATGIE distribution are presented such as the shapes of the probability density, cumulative density and hazard rate functions, survival function, quantile function, the skewness, and kurtosis measures are derived and established and found that the proposed model has flexible hazard rate function. The model parameters are estimated by using three well-known estimation methods namely maximum likelihood estimation (MLE), least-square estimation (LSE), and Cramer-Von-Mises (CVM) methods and it is found that MLEs are quite good than LSEs and CVMs. Two real datasets is considered to explore the applicability and suitability potentiality of the proposed distribution and found that the proposed model is quite better than other lifetime model taken into consideration. We hope this model may be an alternative in the field of reliability analysis, applied statistics and probability theory.

Ram Suresh Yadav
email : yadav.ram2345@gmail.com
Vijay Kumar
email : vkgkp@rediffmail.com
Department of Mathematics & Statistics
DDU Gorakhpur University, Gorakhpur

References

1. **Abouammoh, A.M. & Alshingiti, A.M. :** Reliability estimation of generalized inverted exponential distribution. Journal of Statistical Computation and Simulation, **79(11)** (2009), 1301-1315.
2. **Bader, M. G., & Priest, A. M. :** Statistical aspects of fibre and bundle strength in hybrid composites. Progress in science and engineering of composites, 1129-1136, (1982).
3. **Birnbaum, Z.W., & Saunders, S.C. :** Estimation for a family of life distributions with applications to fatigue, Journal of Applied Probability, **6** (1969), 328-347.
4. **Chaudhary, A.K.& Kumar, V. :** Inverse Generalized Gompertz Distribution with Properties and Applications, Journal of National Academy of Mathematics India, **31** (2017), 01-15.
5. **Dey, S. & Dey, T. :** Generalized inverted exponential distribution: different methods of estimation. American Journal of Mathematical and Management Sciences, **33(3)** (2014), 194-215.
6. **Dube, M., Krishna, H. & Garg, R. :** Generalized inverted exponential distribution under progressive first-failure censoring. Journal of Statistical Computation and Simulation, **86(6)** (2016), 1095-1114.
7. **El-Gohary, A., Alshamrani, A., & Al-Otaibi, A. N. :** The generalized Gompertz distribution. Applied Mathematical Modelling, **37(1-2)** (2013), 13-24.
8. **Gomez-Doniz, E., & Calderon-Ojeda, E. :** Modelling insurance data with the Pareto ArcTan distribution. ASTIN Bulletin: The Journal of the IAA, **45(3)** (2015), 639-660.
9. **Gupta, R. D., & Kundu, D. :** Theory & methods: Generalized exponential distributions. Australian & New Zealand Journal of Statistics, **41(2)** (1999), 173-188.
10. **Joshi, R. K & Kumar, V. :** Inverse Upside Down Bathtub-shaped Hazard function Distribution: Theory and Applications, Journal of National Academy of Mathematics India, **32** (2018), 06-20.

11. **Krishna, H. & Kumar, K. :** Reliability estimation in generalized inverted exponential distribution with progressive type II censored sample. Journal of Statistical Computation and Simulation, **83(6)** (2013), 1007-1019.
12. **Kumar, V. and Ligges, U. :** (2011). reliaR: A package for some probability distributions, http://cran.r-project.org/web/packages/reliaR/index.html.
13. **Kundu, D., and Raqab, M.Z. :** Generalized Rayleigh Distribution: Different Methods of Estimation, Computational Statistics and Data Analysis, **49** (2005), 187-200.
14. **Lan, Y., & Leemis, L. M. :** The logistic exponential survival distribution. Naval Research Logistics (NRL), **55(3)** (2008), 252-264.
15. **Lai, C., Xie, M., Murthy, D. :** A modified weibull distribution. IEEE Trans Reliab **52** (2003), 33-37.
16. **Lemonte, A. J. :** A new exponential-type distribution with constant, decreasing, increasing, upside-down bathtub and bathtub-shaped failure rate function. Computational Statistics & Data Analysis, **62** (2013), 149-170.
17. **Marshall, A. W. and Olkin, I. :** Life Distributions: Structure of Nonparametric, Semiparametric and Parametric Families, Springer, New York (2007).
18. **Ming Hui, E. G. :** Learn R for applied statistics. Springer, New York (2019).
19. **Moors, J. J. A. :** A quantile alternative for kurtosis. Journal of the Royal Statistical Society: Series D (The Statistician), **37(1)** (1988), 25-32.
20. **Murthy, D.N.P., Xie, M. and Jiang, R. :** Weibull Models, Wiley, New York (2003).
21. **Nadarajah, S., & Haghighi, F. :** An extension of the exponential distribution. Statistics, **45(6)** (2011), 543-558.
22. **Nelson, W., & Doganaksoy, N. :** Statistical analysis of life or strength data from specimens of various sizes using the power-(log) normal model. Recent Advances in Life-Testing and Reliability, 377-408 (1995).
23. **Pham, H., & Lai, C. D. :** On recent generalizations of the Weibull distribution. IEEE transactions on reliability, **56(3)** (2007), 454-458.
24. **R Core Team :** (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.
25. **Rinne, H. :** The Weibull distribution: A Handbook CRC Press. Boca Raton, (2009).
26. **Smith, R. M., & Bain, L. J. :** An exponential power life-testing distribution. Communications in Statistics-Theory and Methods, **4(5)** (1975), 469-481.
27. **Swain, J. J., Venkatraman, S. & Wilson, J. R. :** Least-squares estimation of distribution functions in Johnson?s translation system. Journal of Statistical Computation and Simulation, **29(4)** (1988), 271-297.
28. **Tang, Y., Xie, M., & Goh, T. N. :** Statistical analysis of a Weibull extension model. Communications in Statistics-Theory and Methods, **32(5)** (2003), 913-928.