Statistics and Applications {ISSN 2454-7395 (online)}
Volume 22, No. 2, 2024 (New Series), pp 1–27
http://www.ssca.org.in/journal

Modelling Climate, COVID-19, and Reliability Data: A New Continuous Lifetime Model under Different Methods of Estimation

Arvind Pandey[1], Ravindra Pratap Singh[1], Shikhar Tyagi[2] and Abhishek Tyagi[3]
[1]*Department of Statistics, Central University of Rajasthan, Rajasthan-305817, India*
[2]*Department of Statistics and Data Science, Christ deemed to be University, Bangalore-560073, India*
[3]*Department of Statistics, Chaudhary Charan Singh University, Meerut-250004, India*

Received: 31 July 2023; Revised: 30 September 2023; Accepted: 15 October 2023

Abstract

In this article, a new continuous probability distribution called Arvind distribution is developed and studied. The proposed distribution has only one parameter but it exhibits a wide variety of shapes for density and hazard rate functions. A number of important distributional properties including mode, quantile function, moments, skewness, kurtosis, mean deviation, probability-weighted moments, stress-strength reliability, order statistics, reliability and hazard rate functions, Bonferroni Lorenz and Zenga curves, conditional moments, mean residual and mean past life functions, and stochastic ordering of the Arvind distribution are derived. For point estimation of the parameter of the proposed distribution, six estimation procedures including maximum likelihood, maximum product spacings, least squares, weighted least squares, Cramér-von Mises, and Anderson-Darling estimators are used. The interval estimation of the unknown parameter has also been discussed using observed Fisher's information. A vast simulation study has been conducted to examine the behaviour of different estimation procedures. Finally, the applicability of the proposed model is demonstrated by using three real-life datasets. The results of the real data analysis clearly announce that the Arvind distribution can be a better alternative to several existing models for modelling different types of data from various fields.

Key words: Arvind distribution; Maximum likelihood estimation; Maximum product spacings; Least squares estimation; Stress strength reliability

AMS Subject Classifications: 62K05, 05B05

1. Introduction

In today's competitive world, the data generated in numerous disciplines such as engineering, economics, biological sciences, actuarial sciences, *etc.* is becoming more difficult to analyze. As a consequence, for modelling such data, we require distributions that are

Corresponding Author: Ravindra Pratap Singh
Email: stats.rpsingh@gmail.com

best suited for analyzing these multi-features and complicated data. For these reasons, the invention of novel probability distributions has dominated statistical research during the last few decades. In this order, the well-known reference was Mudholkar *et al.* (1996), who described a particular generalization of the Weibull distribution and applied it to survival data. Gupta and Kundu (1999) introduced generalized exponential distribution to provide more flexibility over baseline exponential distribution. This model has decreasing and unimodal shapes of the density function and its hazard rate can take increasing, decreasing, and U-shapes. Nadarajah and Kotz (2006) proposed beta exponential distribution with decreasing and unimodal density whereas the hazard rate can exhibit decreasing and increasing shapes. Nadarajah and Haghighi (2011) developed the Nadarajah-Haghighi distribution to model increasing, decreasing, and constant hazard rate functions. Chaubey and Zhang (2015) pioneered exponentiated Chen distribution with bathtub rate hazard function. Yadav *et al.* (2021a) proposed Burr-Hatke exponential distribution to model decreasing density as well as decreasing hazard rate function. Bakouch *et al.* (2021) proposed a unit half-normal distribution with unimodal and asymmetric (left and right skewed) density and increasing hazard rate function. El-Morshedy *et al.* (2021) proposed a new generalization of the odd Weibull-G family by consolidating two notable families of distributions. Choudhary *et al.* (2021) enhanced the modified Weibull distribution with an additional parameter to provide its density and hazard rate function greater flexibility. This distribution is capable of modeling the bathtub-shaped, decreasing, increasing and the constant hazard rate function. Recently, Alsuhabi *et al.* (2022) pioneered a four-parameter distribution named the extended odd Weibull Lomax distribution. This model has increasing, and decreasing, bell shapes and unimodal shapes of the density function and its hazard rate can take increasing and decreasing shapes. They also show the applicability of this model to COVID-19 data. Promiscuous crucial literature includes Tyagi *et al.* (2022) and Agiwal *et al.* (2023).

Traditional continuous models and their modified or generalized counterparts (in the existing literature) sometimes become very restricted, for example, some models have a complex form of density and hazard functions that are difficult to handle, and a few models are limited to the model-specific type of failure rate, a non-existence of moments, a high number of parameters, an excessive amount of complexity in calculations of some characteristics, *etc.* Although some continuous distributions are less restrictive, there is still room for the construction of more flexible continuous models that may be suitable for the analysis of different types of data generated from distinct fields. With this motivation, we developed a more flexible and simpler continuous distribution called Arvind distribution. This distribution is extremely flexible compared to conventional and recently developed continuous models and we have noticed this in real data applications. Another advantage of Arvind distribution over other rival models is that it has just one parameter and therefore the expressions of this distribution are not too complicated in both analytical and computational handling.

The rest of the structure of this article is as follows. Section 2 introduced Arvind distribution and portrayed various shapes of its density. In Section 3, we have derived various imperative distributional and reliability properties of Arvind distribution with some numerical illustrations. In Section 4, different methods of estimation like maximum likelihood, maximum product spacings, ordinary and weighted least squares, Cramér-von-Mises, and Anderson-Darling have been used to estimate the unknown parameters of the proposed model. Section 4 also includes the asymptotic confidence interval (ACI) for the unknown parameter based on Fisher's information. A detailed simulation study is presented to inspect the performance of various estimation methods in Section 5. In Section 6, the applicability of Arvind distribution has been demonstrated using three case studies from different fields over other well-known continuous models. In the end, some concluding remarks are provided in Section 7.

2. Synthesis of the Arvind distribution

Before discussing the density function and form of the proposed model, we mention the following proposition.

Proposition 1: For a random variable X with domain $(0,\infty)$, the following function is a valid cumulative distribution function (CDF).

$$F(x,\theta) = P\left(X \le x\right) = \begin{cases} 1 - \frac{\exp(-\theta x^2)}{(1+\theta x)}; & x > 0, \theta > 0 \\ 0 \qquad ; & otherwise \end{cases}, \tag{1}$$

where $\theta \in (0,\infty)$ is an unknown parameter.

Proof: Since, $x > 0$ and $\theta > 0$, therefore, we can see that $F(x,\theta) \le 1$ and $F(x,\theta) \ge 0$. Furthermore $\lim\limits_{x\to 0} F(x,\theta) = 0 = F(0,\theta)$ implying that $F(x,\theta)$ is continuous at 0 and a fortiori in $\mathbb{R}$. It is clear that $\lim\limits_{x\to+\infty} F(x,\theta) = 1$. Now, for $x > 0$, we have

$$F'(x,\theta) = \frac{d}{dx}F(x,\theta) = \frac{\theta\left(1 + 2x + 2\theta x^2\right)}{\left(1+\theta x\right)^2}\ \exp(-\theta x^2)\ \ge 0,$$

implying that $F(x,\theta)$ is non-decreasing. The required properties for a valid CDF are satisfied, therefore $F(x,\theta)$ is a valid CDF. □

Based on Proposition 1, we can easily define the Arvind distribution as follows:
Definition 1: A continuous random variable X is said to follow Arvind distribution with parameter θ if its CDF is of the form (1) or it can be specified by the following probability density function (PDF)

$$f(x,\theta) = \begin{cases} \frac{\theta(1+2x+2\theta x^2)}{(1+\theta x)^2}\exp(-\theta x^2); & x > 0, \theta > 0 \\ 0 \qquad ; & otherwise \end{cases}, \tag{2}$$

here, it is clear that $f(x,\theta) \ge 0$ and $\int\limits_0^\infty f(x,\theta)dx = 1$.

Some of the possible shapes of the PDF of the Arvind distribution for a few arbitrary values of the parameter θ are portrayed in Figure 1. From this figure, we can easily see that the PDF of the Arvind distribution is versatile enough as it takes a variety of shapes for different values of θ. Also, the limiting behaviour of the PDF of Arvind distribution can be defined as

$$\lim_{x\to 0} f(x,\theta) = \theta \quad \text{and} \quad \lim_{x\to\infty} f(x,\theta) = 0.$$

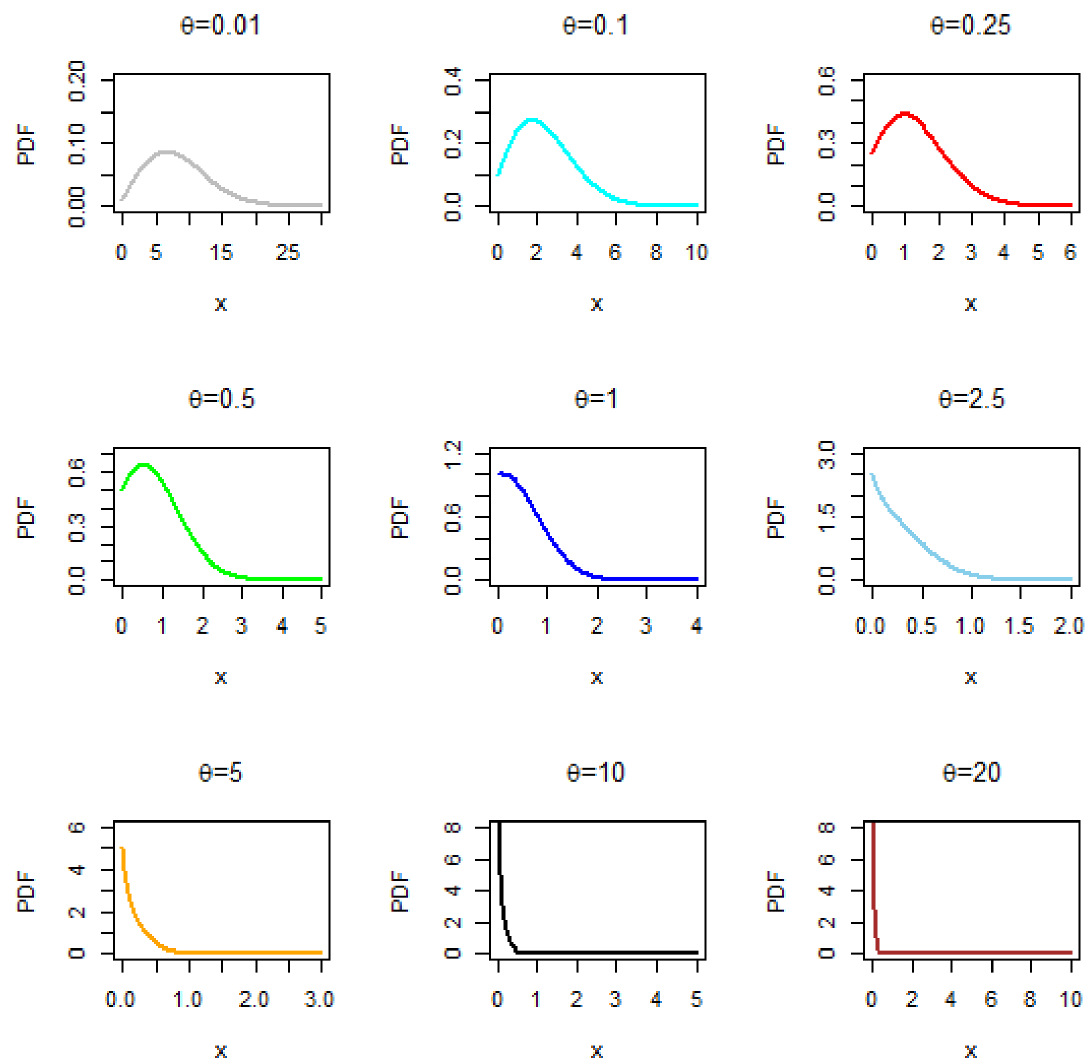


Figure 1: The PDF plots for Arvind distribution for different values of θ

3. Statistical properties of the Arvind distribution

The development of a probability distribution without discussing its statistical properties is not of much use. The Arvind distribution has many important distributional properties and some of them are presented below:

3.1. Mode

A value of a random variable that maximizes its PDF is known as a mode. In the case of Arvind distribution, the mode can be obtained by solving the following equation

$$\frac{\partial \log f(x,\theta)}{\partial x} = 0 \Rightarrow 2\theta^3x^4 + 4\theta^2x^3 + (2\theta+\theta^2)x^2 + \theta - 1 = 0.$$

The above equation cannot be solved analytically in closed form. Therefore, we have obtained the values of mode numerically for different values of θ, and these are listed in Table 1. From these tabulated values of mode, we have verified the conclusion of the PDF plot that the proposed distribution is unimodal.

3.2. Quantile function, random number generation and median

The quantile function is an important tool to specify a probability distribution. It is very useful in random number generation and computation of positional averages like median. The quantile (Q) of the Arvind distribution can be obtained by solving the following equation

$$\exp(-\theta Q^2) - (1+\theta Q)(1-u) = 0, \tag{3}$$

where u is the uniform random variable from $U(0,1)$. The median of the proposed distribution can be computed by putting $u = 0.5$ in Equation (3). We have numerically obtained the median for different values of θ and these are listed in Table 1, and it concludes that the median goes down as θ gets up. By solving Equation (3) for Q, we can generate random numbers from the proposed distribution for different values of u from $U(0,1)$.

3.3. Moments, skewness and kurtosis

For portraying different characteristics of a probability distribution like mean, variance, skewness (Sk), and kurtosis (Kur), moments are very useful in statistical theory. Suppose X is a random variable that follows Arvind distribution with parameter θ. Then, the r^{th} raw moment can be derived as

$$\begin{aligned}\mu'_r = E(X^r) &= \int_0^\infty x^r f(x,\theta)dx \\ &= \int_0^\infty x^r \frac{\theta(1+2x+2\theta x^2)}{(1+\theta x)^2}\exp(-\theta x^2)dx \\ &= r\int_0^\infty \frac{x^{r-1}\exp(-\theta x^2)}{(1+\theta x)}dx. \end{aligned} \tag{4}$$

In particular, the mean and variance of the proposed model can be presented as

$$\mu = \mu'_1 = E(X) = \int_0^\infty \frac{\exp(-\theta x^2)}{(1+\theta x)}dx$$

and

$$\begin{aligned}\mu_2 = Var(X) &= \mu'_2 - {\mu'_1}^2 \\ &= 2\int_0^\infty \tfrac{x\exp(-\theta x^2)}{(1+\theta x)}dx - \left[\int_0^\infty \tfrac{\exp(-\theta x^2)}{(1+\theta x)}dx\right]^2,\end{aligned}$$

respectively. Similarly, we can obtain other central moments using raw moments. From these raw moments, we can also calculate the Sk and Kur of the proposed model using the following formula:

$$Sk = \frac{E(X^4) - 3E(X^2)E(X) + 2(E(X))^3}{(Var(X))^{3/2}},$$

and

$$Kur = \frac{E(X^4) - 4E(X^2)E(X) + 6E(X^2)(E(X))^2 - 3(E(X))^4}{(Var(X))^2},$$

respectively. As we can easily observe the mean, variance, Sk, and Kur of the Arvind distribution cannot be found in closed expressions, therefore we compute them numerically for different values of θ, and these are listed in Table 1. From this table, we yield the following outcomes:

- The mean and variance of the Arvind distribution decrease as the value of θ increases.
- As the value of the coefficient of skewness based on moments is positive, the proposed model is positively skewed. Also, the Sk of the Arvind model increases as θ increases.
- From Table 1, since the value of the coefficient of kurtosis is less than 3, therefore, the proposed distribution is platykurtic and its peakedness increases as θ increases.

Table 1: Descriptive statistics for Arvind distribution for different values of θ

θ	Mode	Median	Mean	Variance	MD(μ)	MD(m)	Skewness	Kurtosis
0.1	1.77056	2.21965	2.40702	2.11585	1.17113	0.97464	0.69274	0.29880
0.5	0.52137	0.83134	0.93373	0.40649	0.51371	0.40521	0.81903	0.47350
1	0.00003	0.52237	0.60513	0.19600	0.35623	0.26777	0.91400	0.64761
1.5	0.11719	0.39208	0.46573	0.12693	0.28611	0.20687	0.98351	0.79551
2	0.26459	0.31746	0.38531	0.09288	0.24423	0.17082	1.03954	0.92660
2.5	0.42256	0.26827	0.33191	0.07271	0.21563	0.14647	1.08699	1.04560
3	0.10358	0.23306	0.29340	0.05943	0.19453	0.12869	1.12840	1.15534
4	0.24065	0.18560	0.24089	0.04308	0.16495	0.10420	1.19865	1.35382
5	0.71926	0.15482	0.20627	0.03347	0.14484	0.08796	1.25747	1.53157
10	0.07255	0.08580	0.12588	0.01503	0.09546	0.05024	1.46474	2.24050
20	0.65754	0.04588	0.07548	0.00657	0.06161	0.02741	1.71150	3.24746

3.4. Mean deviation

If we take the average absolute deviation about the mean (or median) it is known as the mean deviation about the mean (or median). Mean deviation about the mean (or median) is another important tool for measuring dispersion besides the variance. Suppose μ and m denote the mean and median, then the mean deviation about the mean (or median) can be defined as

$$MD(\zeta) = E\left|X-\zeta\right| = \int_0^\infty \left|x-\zeta\right| f(x,\theta)dx = 2\left\{\zeta F(\zeta,\theta) - \int_0^\zeta xf(x,\theta)dx\right\}, \quad (5)$$

where $\zeta = \mu$ or m . Using the above expression with some simplification, the mean deviation about mean (or median) for the Arvind distribution can be obtained as

$$MD(\zeta) = E\left|X-\zeta\right| = 2\left\{\zeta - \int_0^\zeta \frac{\exp(-\theta x^2)}{(1+\theta x)}dx\right\}. \quad (6)$$

The expression of mean deviation (6) cannot be bound up in closure form, so to measure the behaviour of mean deviation about mean (or median), we have calculated these average deviations numerically and they are listed in Table 1. This table announces that the mean deviation about the mean (or median) decreases as θ increases and the mean deviation about the median is smaller than the mean deviation about the mean as the theory claims.

3.5. Probability weighted moments

The generalization of the simple moments is known as probability-weighted moments (PWMs). They can be developed for a distribution whose ordinary moments can be derived. For the Arvind random variable X, the $(r,s)^{th}$ PWM is given by

$$\begin{aligned}
\varsigma_{r,s} &= E\left[X^r F^s(x,\theta)\right] \\
&= \int_0^\infty x^r F^s(x,\theta) f(x,\theta)dx \\
&= \int_0^\infty x^r \left(1 - \frac{\exp(-\theta x^2)}{1+\theta x}\right)^s \frac{\theta(1+2x+2\theta x^2)\exp(-\theta x^2)}{(1+\theta x)^2} dx \\
&= \sum_{j=0}^{s} (-1)^j \theta \binom{s}{j} \int_0^\infty x^r \left(\frac{1}{(1+\theta x)^{j+2}} + \frac{2x}{(1+\theta x)^{j+1}}\right) \exp(-(j+1)\theta x^2)dx.
\end{aligned}$$

After some simplification, the $(r,s)^{th}$ PWM of the Arvind distribution is given by

$$\varsigma_{r,s} = \sum_{j=0}^{s} (-1)^j \frac{r}{(j+1)} \binom{s}{j} \int_0^\infty \frac{x^{r-1}\exp(-(j+1)\theta x^2)}{(1+\theta x)^{j+1}} dx. \tag{7}$$

3.6. Stress-strength reliability

The probability $\varpi = P(X_2 < X_1)$ is referred to as stress-strength (S-S) reliability if the random variable X_1 represents the strength of a system under stress X_2, assuming that X_1 and X_2 are stochastically independent random variables. The S-S reliability is widely used in reliability theory, especially in engineering concepts like different structures, static fatigue of ceramic components, the aging of concrete pressure vessels, fatigue failure of aviation structures, *etc*. The research on S-S reliability models has received a lot of attention recently due to the expanded scope of S-S reliability. For more detail, see Goel and Singh (2020). In our case, suppose $X_1 \sim \text{Arvind}(\theta_1)$ and $X_2 \sim \text{Arvind}(\theta_2)$ distributions, then S-S reliability is given by

$$\begin{aligned}
\varpi = P(X_2 < X_1) &= \int_0^\infty P(X_2 < X_1 | X_1 = x) f_{X_1}(x,\theta_1)dx \\
&= \int_0^\infty F_{X_2}(x,\theta_2) f_{X_1}(x,\theta_1)dx \\
&= \int_0^\infty \left(1 - \frac{\exp(-\theta_2 x^2)}{(1+\theta_2 x)}\right)\left(\frac{\theta_1(1+2x+2\theta_1 x^2)\exp(-\theta_1 x^2)}{(1+\theta_1 x)^2}\right) dx \\
&= 1 - \theta_1 \int_0^\infty \frac{(1+2x+2\theta_1 x^2)\exp(-(\theta_1+\theta_2)x^2)}{(1+\theta_2 x)(1+\theta_1 x)^2} dx.
\end{aligned} \tag{8}$$

The expression (8) ϖ cannot be easily tractable in closed form. Therefore, to study the behaviour of ϖ for different values of θ_1 and θ_2, we have computed ϖ numerically. The outcomes of ϖ have been given in Table 2. From this table, we observe that

- For a fixed value of θ_1, the value of ϖ increases as θ_2 increases.

- For a fixed value of θ_2, the value of ϖ goes down as θ_1 increases.
- When the values of θ_1 and θ_2 are equal, the value of ϖ becomes 0.5.

Table 2: The S-S reliability ϖ under different values of θ_1 and θ_2

$\theta_1 \rightarrow$ $\theta_2 \downarrow$	0.1	0.5	1	1.5	2	2.5	3	4	5	10	20
0.1	0.5	0.17715	0.10231	0.0734	0.05786	0.04808	0.04133	0.03255	0.02705	0.01526	0.00862
0.5	0.82285	0.5	0.34692	0.27031	0.22365	0.19196	0.1689	0.13737	0.11661	0.06916	0.04041
1	0.89769	0.65308	0.5	0.41108	0.35211	0.30975	0.27765	0.23187	0.2005	0.12452	0.07519
1.5	0.9266	0.72969	0.58892	0.5	0.43784	0.39151	0.35543	0.30244	0.26504	0.17049	0.1057
2	0.94214	0.77635	0.64789	0.56216	0.5	0.45243	0.41462	0.35785	0.31686	0.20963	0.13287
2.5	0.95192	0.80804	0.69025	0.60849	0.54757	0.5	0.46159	0.40289	0.35974	0.24359	0.15734
3	0.95867	0.8311	0.72235	0.64457	0.58538	0.53841	0.5	0.44045	0.39601	0.27348	0.17957
4	0.96745	0.86264	0.76813	0.69756	0.64215	0.59711	0.55955	0.5	0.45448	0.32402	0.21868
5	0.97295	0.88339	0.7995	0.73496	0.68314	0.64026	0.60399	0.54552	0.5	0.36549	0.25223
10	0.98474	0.93084	0.87548	0.82951	0.79037	0.75641	0.72652	0.67598	0.63451	0.5	0.37072
20	0.99138	0.95959	0.92481	0.8943	0.86713	0.84266	0.82043	0.78132	0.74777	0.62928	0.5

We have also portrayed a 3-D plot for ϖ under different values of θ_1 and θ_2 in Figure 2, this plot also announces that ϖ can take a variety of values from small to large for distinct values of θ_1 and θ_2.

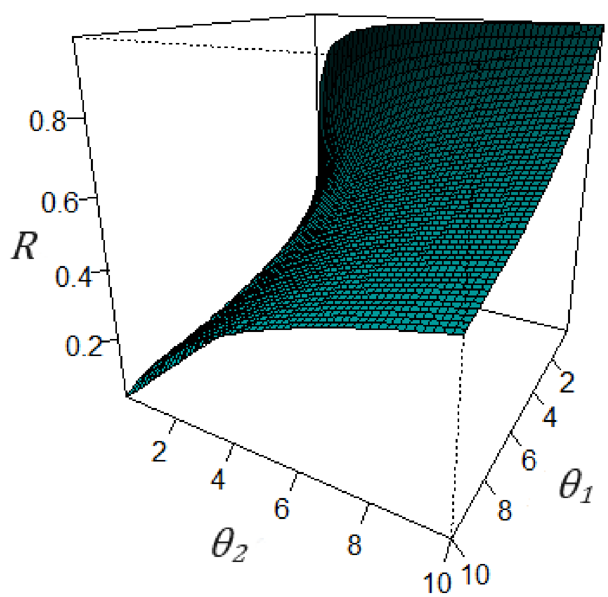


Figure 2: A 3-D plot for ϖ under different values of θ_1 and θ_2

3.7. Order statistics

Let $X_1, X_2, ..., X_n$ be a random sample of size n generated from Arvind(θ) distribution and $X_{1:n} \leq X_{2:n} \leq ..... \leq X_{n:n}$ denotes the corresponding order statistics. Then, the

PDF and CDF, respectively, of i^{th} order statistics are given as

$$f_{i:n}(x,\theta) = \frac{n!}{(i-1)!\,(n-i)!}[F(x,\theta)]^{i-1}[1-F(x,\theta)]^{n-i}f(x,\theta)$$
$$= \frac{n!\theta(1+2x+2\theta x^2)\exp(-\theta(n-i+1)x^2)(1+\theta x-\exp(-\theta x^2))^{i-1}}{(i-1)!\,(n-i)!(1+\theta x)^{n+1}}, \tag{9}$$

and

$$F_{i:n}(x,\theta) = \sum_{r=i}^{n}\binom{n}{r}[F(x,\theta)]^{r}[1-F(x,\theta)]^{n-r}$$
$$= \sum_{r=i}^{n}\sum_{j=0}^{n-r}(-1)^j\binom{n}{r}\binom{n-r}{j}[F(x,\theta)]^{j+r}$$
$$= \sum_{r=i}^{n}\sum_{j=0}^{n-r}(-1)^j\binom{n}{r}\binom{n-r}{j}\left(1-\frac{\exp(-\theta x^2)}{1+\theta x}\right)^{j+r}. \tag{10}$$

In particular, by putting $i = 1$ and $i = n$, respectively, we can find the PDF and CDF of minimum and maximum order statistics. For odd sample size n, we can obtain the PDF and CDF of the median order statistics by setting $i = \frac{n+1}{2}$.

3.8. Reliability and hazard rate functions

The reliability function (RF) $R(x,\theta)$ and hazard rate function (HRF) $h(x,\theta)$ of the Arvind(θ) distribution, respectively, are given by

$$R(x,\theta) = P(X > x) = \frac{\exp(-\theta x^2)}{(1+\theta x)}; \qquad x \geq 0, \theta > 0, \tag{11}$$

$$h(x,\theta) = \frac{f(x,\theta)}{R(x,\theta)} = \frac{\theta\,(1+2x+2\theta x^2)}{(1+\theta x)}; \qquad x \geq 0, \theta > 0. \tag{12}$$

We have plotted the hazard rate for different values of θ in Figure 3. From this figure, we can easily observe that the HRF of the Arvind distribution can increase, decrease, and U-shaped. Also, the limiting behaviour of the HRF can be stated as:

$$\lim_{x\to 0} h(x,\theta) = \theta \quad \text{and} \quad \lim_{x\to\infty} h(x,\theta) = \infty.$$

Also, the cumulative and reverse hazard rate (RHR) functions of the Arvind distribution, respectively, are given by

$$H(x,\theta) = \theta x^2 + \log(1+\theta x); \qquad x \geq 0, \theta > 0, \tag{13}$$

$$RHR(x,\theta) = \frac{\theta(1+2x+2\theta x^2)\exp(-\theta x^2)}{(1+\theta x)(1+\theta x-e^{-\theta x^2})}; \quad x \geq 0, \theta > 0. \tag{14}$$

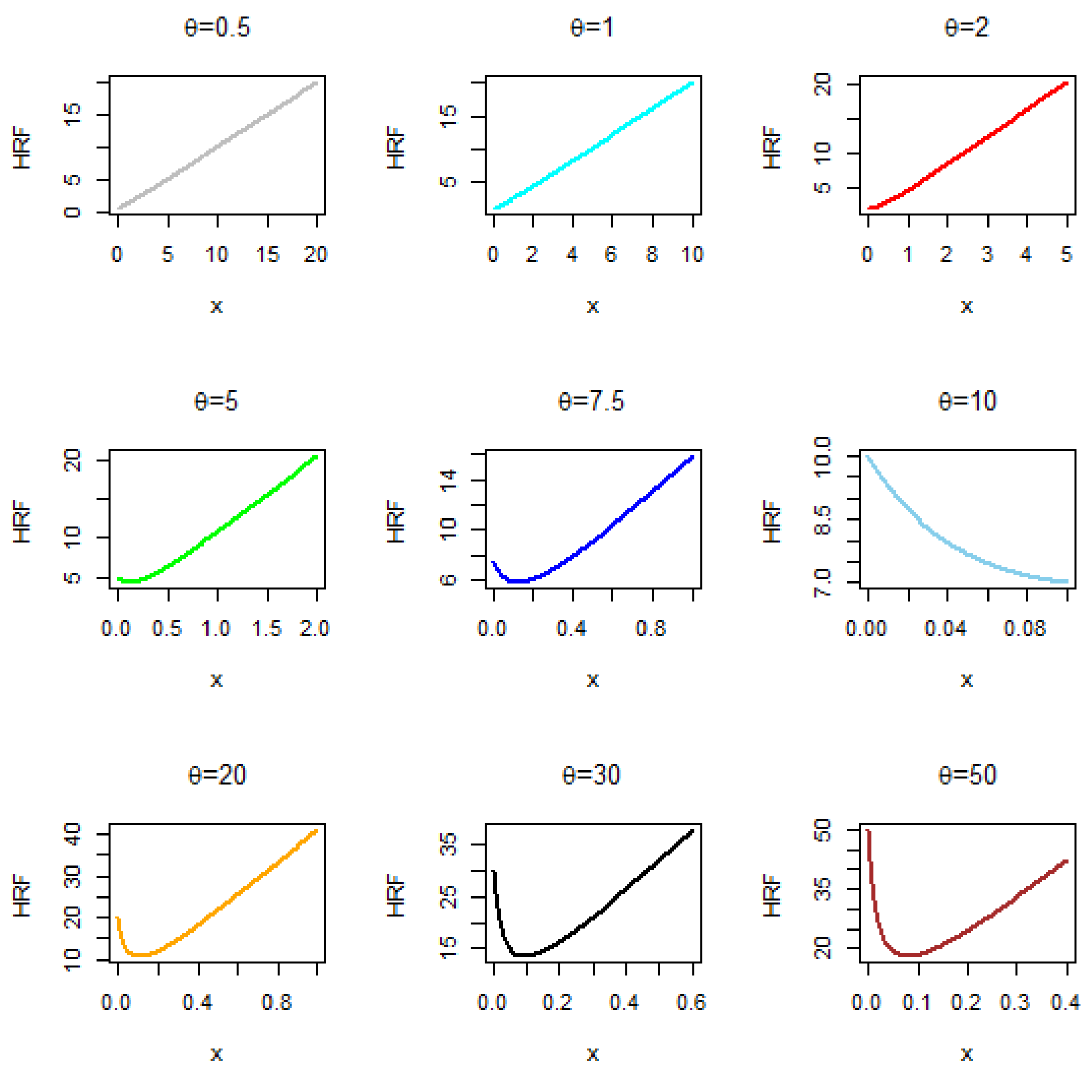


Figure 3: The HRF plot of Arvind distribution for different values of θ

3.9. Inequality measures: Lorenz, Bonferroni and Zenga curves

The Lorenz, Bonferroni, and Zenga curves are the most often used inequality measures in the literature (Lorenz (1905), Bonferroni (1930), and Zenga (2007)). These three curves can be defined using simply the population mean and the means of certain subgroups. Inequality curves are useful because they may be used to create a variety of orderings that allow for distribution comparisons based on inequality. Such comparisons within the same model make it possible to comprehend how distributional parameters influence inequality. For the Arvind distribution, the Lorenz, Bonferroni, and Zenga curves, respectively, are obtained as

$$L(p) = \frac{1}{\mu}\int_0^q xf(x)dx = \frac{1}{\mu}\left[-\frac{q\exp(-\theta q^2)}{1+\theta q} + \int_0^q \frac{\exp(-\theta x^2)}{(1+\theta x)}dx\right], \tag{15}$$

$$B(p) = \frac{L(p)}{p}, \tag{16}$$

$$Z(p) = \frac{p - L(p)}{p(1 - L(p))}, \tag{17}$$

where $q = F^{-1}(p)$ which can be computed numerically using Equation (3). Table 3 lists numerical values for Lorenz, Bonferroni, and Zenga curves of Arvind distribution for different values of $q = F^{-1}(p)$ and θ.

Table 3: Values for Lorenz, Bonferroni, and Zenga curves of Arvind distribution for a variety of θ

$\theta \rightarrow$	0.5			1			2			5		
$p \downarrow$	$L(p)$	$B(p)$	$Z(p)$	$L(p)$	$B(p)$	$Z(p)$	$L(p)$	$B(p)$	$Z(p)$	$L(p)$	$B(p)$	$Z(p)$
0.05	0.00260	0.05191	0.95055	0.00207	0.04132	0.96066	0.00165	0.03298	0.96862	0.00124	0.02487	0.97635
0.1	0.01013	0.10126	0.90793	0.00827	0.08267	0.92497	0.00670	0.06704	0.93926	0.00512	0.05118	0.95370
0.15	0.02229	0.14863	0.87079	0.01860	0.12398	0.89262	0.01533	0.10219	0.91179	0.01184	0.07893	0.93211
0.2	0.03891	0.19453	0.83808	0.03310	0.16548	0.86309	0.02769	0.13843	0.88610	0.02165	0.10824	0.91149
0.25	0.05983	0.23933	0.80908	0.05178	0.20710	0.83619	0.04395	0.17579	0.86209	0.03481	0.13923	0.89181
0.3	0.08500	0.28334	0.78323	0.07469	0.24895	0.81167	0.06429	0.21430	0.83968	0.05167	0.17225	0.87286
0.35	0.11439	0.32683	0.76012	0.10190	0.29113	0.78929	0.08885	0.25385	0.81891	0.07243	0.20694	0.85499
0.4	0.14801	0.37002	0.73942	0.13349	0.33374	0.76891	0.11791	0.29478	0.79949	0.09748	0.24371	0.83798
0.45	0.18591	0.41314	0.72088	0.16960	0.37689	0.75037	0.15163	0.33697	0.78154	0.12723	0.28272	0.82183
0.5	0.22820	0.45640	0.70433	0.21037	0.42074	0.73358	0.19028	0.38055	0.76501	0.16206	0.32412	0.80660
0.55	0.27501	0.50002	0.68964	0.25600	0.46545	0.71848	0.23413	0.42569	0.74988	0.20248	0.36814	0.79228
0.6	0.32652	0.54419	0.67679	0.30674	0.51124	0.70502	0.28353	0.47255	0.73618	0.24892	0.41487	0.77906
0.65	0.38306	0.58932	0.66567	0.36293	0.55836	0.69324	0.33891	0.52140	0.72396	0.30212	0.46479	0.76690
0.7	0.44497	0.63567	0.65641	0.42500	0.60714	0.68323	0.40079	0.57256	0.71334	0.36278	0.51826	0.75601
0.75	0.51277	0.68369	0.64920	0.49353	0.65804	0.67518	0.46988	0.62650	0.70455	0.43187	0.57582	0.74662
0.8	0.58718	0.73397	0.64441	0.56936	0.71170	0.66947	0.54713	0.68392	0.69796	0.51064	0.63830	0.73913
0.85	0.66931	0.78742	0.64283	0.65371	0.76907	0.66687	0.63398	0.74586	0.69433	0.60094	0.70699	0.73426
0.9	0.76101	0.84556	0.64620	0.74860	0.83178	0.66914	0.73278	0.81420	0.69531	0.70570	0.78411	0.73357
0.95	0.86601	0.91159	0.65984	0.85826	0.90343	0.68131	0.84824	0.89289	0.70581	0.83075	0.87447	0.74167

3.10. Conditional moments

In the context of lifetime models, it is also useful to have a knowledge of the expression $E(X^r|X > x)$. This expression is called the r^{th} conditional moment of the random variable 'X'. The computation of mean deviations around the mean and the median, as well as the mean residual life function (See, Section 3.11) are all areas in which the conditional moments find widespread usage. The r^{th} conditional moment of a random variable following Arvind(θ) distribution can be obtained as

$$E(X^r|X > x) = \frac{1}{1 - F(x, \theta)} \Lambda_r(x, \theta), \tag{18}$$

where

$$\Lambda_r(x, \theta) = \int_x^\infty v^r f(v, \theta) dv = \frac{x^r \exp(-\theta x^2)}{(1 + \theta x)} + r \int_x^\infty \frac{v^{r-1} \exp(-\theta v^2)}{1 + \theta v} dv.$$

3.11. Mean residual life

The expected value of the remaining lifetimes after a fixed time point x, is called the mean residual life (MRL) function. Since it is representative of the aging mechanism, the MRL function is put to considerable use in a broad range of fields, including reliability engineering, survival analysis, and biological research. For Arvind(θ) distribution, it can be derived as

$$\begin{aligned} MRL(x,\theta) = E(X-x|X>x) &= \frac{1}{1-F(x,\theta)}\int_x^\infty vf(v,\theta)dv - x \\ &= \frac{1+\theta x}{\exp(-\theta x^2)}\int_x^\infty \frac{\exp(-\theta v^2)}{1+\theta v}dv. \end{aligned} \tag{19}$$

From the above expression of MRL, we can easily observe that the MRL is an application of conditional moments and it can be obtained by putting $r = 1$ in Equation (18).

3.12. Mean past life

The expected time elapsed from the failure of a system given that its lifetime is less than or equal to a time point $x (x \geq 0)$ is referred to as the mean past life (MPL) function. Similar to the MRL function, the MPL function has applications in a vast array of fields, such as actuarial research, forensic science, reliability theory, and survival analysis. The expression of the MPL function for Arvind(θ) distribution can be developed as

$$\begin{aligned} MPL(x,\theta) = E(x-X|X\leq x) &= x - \frac{1}{F(x,\theta)}\int_0^x vf(v,\theta)dv \\ &= \frac{1}{F(x,\theta)}\left[x - \int_0^x \frac{\exp(-\theta v^2)}{1+\theta v}dv\right]. \end{aligned} \tag{20}$$

3.13. Stochastic ordering

It is crucial to compare two or more random variables indicating the state of things in two or more circumstances. In the situation of two random variables that are independent, stochastic orderings are extremely advantageous. For two independent random variables Y and Z if $F_Y(y) \geq F_Z(y)$ for all y, Y is said to be stochastically smaller than Z *i.e.* $Y \leq_{st} Z$. Similarly, we can define stochastic ordering in terms of hazard rate, mean residual life, and likelihood ratio functions as

- hazard rate order $(Y \leq_{hr} Z)$ if $h_Y(y) \geq h_Z(y)$ for all y.
- mean residual life order $(Y \leq_{mrl} Z)$ if $MRL_Y(y) \leq MRL_Z(y)$ for all y.
- likelihood ratio order $(Y \leq_{lr} Z)$ if $f_Y(y)/f_Z(y)$ decreases in y.

The following implications Shaked and Shanthikumar (2007) are well-known

$$\begin{array}{c} Y \leq_{lr} Z \Rightarrow Y \leq_{hr} Z \Rightarrow Y \leq_{mrl} Z \\ \Downarrow \\ Y \leq_{st} Z. \end{array}$$

The Arvind distributions are ordered with respect to the strongest "likelihood ratio" ordering as we can easily observe from the following theorem.

Theorem 1: Let Y and Z be two independent random variables form Arvind(θ_1) and Arvind(θ_2) distributions, respectively. If $\theta_1 > \theta_2$ then $Y \leq_{lr} Z$ and hence $Y \leq_{hr} Z$, $Y \leq_{mrl} Z$ and $Y \leq_{st} Z$.

Proof: Firstly, we observe that

$$\frac{f_Y(y,\theta_1)}{f_Z(y,\theta_2)} = \frac{\theta_1(1+2y+2\theta_1 y^2)(1+\theta_2 y)^2 \exp(-\theta_1 y^2)}{\theta_2(1+2y+2\theta_2 y^2)(1+\theta_1 y)^2 \exp(-\theta_2 y^2)}, \quad y > 0.$$

Since, for $\theta_1 > \theta_2$,

$$\frac{d}{dx}\log\left(\frac{f_Y(y,\theta_1)}{f_Z(y,\theta_2)}\right) = 2(\theta_2-\theta_1)\left[\frac{\begin{array}{l} 2+(6+\theta_1+\theta_2)y+(6(1+\theta_2)+\theta_1(6+\theta_2))y^2 \\ +2(\theta_1^2+\theta_2(5+\theta_2)+\theta_1(5+3\theta_2))y^3 \\ +2(2\theta_2^2+\theta_1^2(2+\theta_2)+\theta_1\theta_2(9+\theta_2))y^4 \\ +8\theta_1\theta_2(\theta_1+\theta_2)y^5+4\theta_1^2\theta_2^2 y^6 \end{array}}{(1+2y+2\theta_1 y^2)(1+2y+2\theta_2 y^2)(1+\theta_1 y)(1+\theta_2 y)}\right] < 0,$$

i.e. $f_Y(y)/f_Z(y)$ is decreasing in y. It implies that $Y \leq_{lr} Z$. The rest of the ordering is a direct consequence of the results provided by Shaked and Shanthikumar (2007). □

4. Parameter estimation of the Arvind distribution

Under this section, the estimation of the unknown parameter of the proposed model has been discussed using six different classical approaches, namely, the method of maximum likelihood, maximum product spacings, ordinary and weighted least squares, Cramér-von-Mises, and Anderson Darling method of estimation. These methods are briefly discussed as follows,

4.1. Maximum likelihood estimation

Suppose $\underset{\sim}{X} \equiv X_1, X_2, ..., X_n$ be a random sample of size n from the Arvind distribution. Then, the log-likelihood (logL) function can be written as

$$\log L = n\log(\theta) - \theta\sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} log\left(1+2x_i+2\theta x_i^2\right) - 2\sum_{i=1}^{n} log\left(1+\theta x_i\right). \tag{21}$$

To find out the maximum likelihood estimator (MLE) of θ, the normal equation is given by

$$\frac{\partial \log L}{\partial\theta} = \frac{n}{\theta} - \sum_{i=1}^{n}\left[x_i^2 + \frac{2}{(1+\theta x_i)} - \frac{2x_i^2}{(1+2x_i+2\theta x_i^2)}\right] = 0. \tag{22}$$

The solution of Equation (22) yields the MLE of θ. Unfortunately, the above normal equation cannot be solved analytically. Therefore, we can use numerical iteration procedures such as Newton-Raphson (NR) through the open-source programming language R.

4.2. Observed Fisher's information and asymptotic confidence interval

The observed Fisher's information for Arvind(θ) distribution is specified by

$$I_o(\hat{\theta}) = -\frac{\partial^2 \log L}{\partial \theta^2}|\theta = \hat{\theta},$$

where the second-order derivative of the log-likelihood function (21) with respect to θ is given by

$$\frac{\partial^2 \log L}{\partial \theta^2} = -\frac{n}{\theta^2} - \sum_{i=1}^{n} \left[-\frac{2x_i}{(1+\theta x_i)^2} + \frac{4x_i^4}{(1+2x_i+2\theta x_i^2)^2} \right].$$

Using this Fisher's information, the asymptotic variance of $\hat{\theta}$ can be obtained as

$$Var_o(\hat{\theta}) = \frac{1}{I_o(\hat{\theta})}.$$

Under some regularity conditions, the sampling distribution of $(\hat{\theta}-\theta)/\sqrt{Var_o(\hat{\theta})}$ can be approximated by a standard normal distribution. The large-sample $100 \times (1-\alpha)\,\%$ confidence interval (also called ACI) for θ is given by

$$[\hat{\theta}_L, \hat{\theta}_U] = \hat{\theta} \mp z_{\alpha/2} \sqrt{Var_o(\hat{\theta})}.$$

Using simulation, we can estimate the coverage probability $P\left[\left|\frac{(\theta-\hat{\theta})}{\sqrt{Var_o(\hat{\theta})}}\right| \le z_{\alpha/2}\right]$, here z_p is such that $p = \int\limits_{z_p}^{\infty} (1/\sqrt{2\pi}) e^{-z^2/2} dz$.

4.3. Maximum product of spacings method of estimation

As an alternative to the approach of maximum likelihood, the maximum product spacing (MPS) method was developed by Cheng and Amin (1979) for estimating the unknown parameters of continuous univariate distributions. Cheng and Amin (1983) proved that this technique is just as efficient as the maximum likelihood estimation and that it is consistent under more general conditions. Suppose $X_1, X_2, ..., X_n$ be a random sample from Arvind distribution $F(x,\theta)$, and $X_{1:n} \le X_{2:n} \le ... \le X_{n:n}$ be the corresponding ordered values. Based on this random sample, let us define the uniform spacings as

$$D_i(\theta) = F(x_{i:n}, \theta) - F(x_{i-1:n}, \theta), \qquad i = 1, 2, ..., n,$$

where $F(x_{0:n}, \theta) = 0$, $F(x_{n+1:n}, \theta) = 1$, and $\sum\limits_{i=1}^{n+1} D_i(\theta) = 1$. The MPS estimator $\hat{\theta}_{MPS}$ of the parameter θ is determined by maximizing the geometric mean of the spacings with respect to θ, or, evenly, by maximizing the following function

$$H(\theta) = \frac{1}{n+1} \sum_{i=1}^{n+1} \log(D_i(\theta)).$$

The estimator $\hat{\theta}_{MPS}$ can also be obtained by solving the following non-linear equation,

$$\frac{1}{n+1}\sum_{i=1}^{n+1}\frac{1}{D_i(\theta)}\left[\xi(x_{i:n},\theta)-\xi(x_{i-1:n},\theta)\right]=0, \tag{23}$$

where $\xi(x_{i:n},\theta)=\left(\frac{x_i e^{-\theta x_i^2}}{1+\theta x_i}\right)\left(x_i+\frac{1}{1+\theta x_i}\right)$.

4.4. Ordinary and weighted least squares estimation

Swain *et al.* (1988) firstly introduced regression-based estimators called ordinary least squares (OLS) and weighted least squares (WLS) estimators for estimating the unknown parameters of the beta distribution. These two methods are based on the combination of the non-parametric and parametric distribution functions. Suppose $X_1, X_2, ..., X_n$ be a random sample from Arvind distribution, and $X_{1:n} \leq X_{2:n} \leq ... \leq X_{n:n}$ be the corresponding ordered values. Then, the OLS estimator of θ, say $\hat{\theta}_{OLS}$ can be derived by minimizing the following function with respect to θ

$$V(\theta)=\sum_{i=1}^{n}\left[F\left(x_{i:n},\theta\right)-\frac{i}{n+1}\right]^2.$$

Alternatively, we can obtain the OLS estimator of θ by solving the following expression for θ,

$$\sum_{i=1}^{n}\left[F\left(x_{i:n},\theta\right)-\frac{i}{n+1}\right]\xi(x_{i:n},\theta)=0. \tag{24}$$

The WLS estimator of θ, say $\hat{\theta}_{WLS}$ can be found by minimizing the following equation,

$$W(\theta)=\sum_{i=1}^{n}\frac{(n+1)^2(n+2)}{i(n-i+1)}\left[F\left(x_i,\theta\right)-\frac{i}{n+1}\right]^2.$$

The WLS estimator $\hat{\theta}_{WLS}$ can also be obtained by solving the following non-linear equation with respect to θ,

$$\sum_{i=1}^{n}\frac{(n+1)^2(n+2)}{i(n-i+1)}\left[F\left(x_{i:n},\theta\right)-\frac{i}{n+1}\right]\xi(x_{i:n},\theta)=0, \tag{25}$$

where $\xi(x_{i:n},\theta)$ is defined in Section 4.3.

4.5. Cramér-von-Mises estimation

Cramer-von-Mises type minimum distance estimator is a widely used minimum distance estimator since the empirical data suggests that the bias of this estimator is lower than that of the other minimum distance estimators. In our case, the Cramér-von Mises (CVM) minimum distance estimator of θ can be obtained by minimizing, the following function:

$$C(\theta)=\frac{1}{12n}\left[F\left(x_i,\theta\right)-\frac{2i-1}{2n}\right]^2.$$

Moreover, we can obtain the CVM estimator of θ by solving the following equation for θ,

$$\sum_{i=1}^{n}\left[F\left(x_{i:n},\theta\right)-\frac{2i-1}{2n}\right]\xi(x_{i:n},\theta)=0, \tag{26}$$

where $\xi(x_{i:n},\theta)$ is already defined in Section 4.3.

4.6. Anderson-Darling method of estimation

The Anderson-Darling estimator (ADE) is another sort of minimum distance estimator that utilizes Anderson–Darling statistics. The ADE of θ, say $\hat{\theta}_{ADE}$, can be obtained by minimizing the following function with respect to θ,

$$A\left(\theta\right)=-n-\frac{1}{n}\sum_{i=1}^{n}\left(2i-1\right)\left\{\log F\left(x_{i:n},\theta\right)+\log\bar{F}\left(x_{n+1-i:n},\theta\right)\right\}.$$

The estimator $\hat{\theta}_{ADE}$ can also be achieved by simplifying the following nonlinear equation

$$\sum_{i=1}^{n+1}\left(2i-1\right)\left[\frac{\xi(x_{i:n},\theta)}{F(x_{i:n},\theta)}-\frac{\xi(x_{n+1-i:n},\theta)}{\bar{F}(x_{n+1-i:n},\theta)}\right]=0, \tag{27}$$

where $\xi(x_{i:n},\theta)$ is given in Section 4.3.

5. A Monte Carlo simulation study

This section showcases the behaviour of different estimation procedures for estimating the unknown parameter of the Arvind distribution. For this purpose, we have performed an empirical experiment which utilizes the following steps:

1. Generate 2,500 samples of size n =10, 20, 40, 60, 80, and 100 from Arvind distribution with $\theta=0.5,1.0,2.0,$ and 4.0. For sample generation, Equation (3) has been used.
2. Calculate the MLE, MPS, OLS, WLS, CVM, and AD estimators for the 2,500 samples, say $\hat{\theta}_j; j=1,2,...,2,500$. Also, compute the 95% ACI for the above-generated samples.
3. Determine the expected value (EV), mean-squared error (MSE), and average bias (AB) for all point estimators, whereas, for 95% ACI, we compute the average lower confidence limit (ALCL), average upper confidence limit (AUCL), average width (AW), and coverage probability, *i.e.*,

$$EV=\frac{1}{2500}\sum_{j=1}^{2500}\hat{\theta}_j,\quad MSE=\frac{1}{2500}\sum_{j=1}^{2500}\left(\hat{\theta}_j-\theta\right)^2,\quad AB=\frac{1}{2500}\sum_{j=1}^{2500}\left(\hat{\theta}_j-\theta\right),$$

$$ALCL=\frac{1}{2500}\sum_{j=1}^{2500}LCL_j,\quad AUCL=\frac{1}{2500}\sum_{j=1}^{2500}UCL_j,$$

$$AW=\frac{1}{2500}\sum_{j=1}^{2500}(UCL_j-LCL_j),\quad CP=\frac{1}{2500}\sum_{j=1}^{2500}I_j(LCL_j<\theta_j<UCL_j),$$

where LCL_j and UCL_j denotes the upper and lower confidence limit for the j^{th} sample, respectively and $I_j(\bullet)$ is the indicator function takes value 1 if $LCL_j < \theta < UCL_j$ otherwise 0.

The simulation study was conducted using the R software and the codes are available upon request. Various classical estimates of θ with their MSE, AB are listed in Table 4. On the other hand, Table 5 contains the ALCL, AUCL, AW, and CP for 95% ACI.
From this empirical study, the following outcomes have been noted:

- We found that the average bias and MSE of all estimators approach zero for large n, indicating that the parameter estimates are consistent and asymptotically unbiased.
- The performance of all of the estimating techniques is satisfactory. However, in the overall comparison, for the proposed model, MPS is the most favourable estimation procedure while CVM is the least favourable estimation method.
- Additionally, the hierarchy of the best estimation technique among the numerous methods taken into consideration for estimating the parameter of the proposed distribution, as determined by the MSE, is as follows:
$$MPS \rightarrow MLE \rightarrow WLSE \rightarrow ADE \rightarrow LSE \rightarrow CVM$$
$$(HighlyPreferable \rightarrow Less\ Preferable)$$
- Except for MPS, all classical point estimators overestimate the parameter of the proposed model.
- From Table 5, we can simply conclude that ACI performed well. Even with a small sample size, for all values of θ, the asymptotic intervals computed here are able to sustain nominal levels of coverage probability. Furthermore, when we increase the sample size n, the AW of the ACI diminishes.

6. Application of Arvind distribution

The fitting capabilities of the Arvind distribution are shown in this section using three real datasets. We have used three distinct datasets from different areas. The detailed summary and graphical representation of these datasets can be found in Table 6 and Figure 4, respectively. The fitting of the proposed model has been compared with that of numerous well-known conventional and recently developed models. A list of the rival models can be found in Table 7. The fitted models' parameters have been estimated using MLE estimation for comparison's sake. Based on -logL, the Akaike information criterion (AIC), the corrected Akaike information criterion (CAIC), the Bayesian information criterion (BIC), and Kolmogorov-Smirnov (KS) statistic with the related P-value, the model comparison has been carried out. The open-source program R has been used to do the necessary calculations. The datasets along with their fitting summary are as follows:

Table 4: Various classical point estimates for different values of n

θ	n	MLE			MPS			OLS			WLS			CVM			ADE		
		EV	MSE	AB	EV	MSE	AB	EV	MSE	AB	EV	MSE	AB	EV	MSE	AB	EV	MSE	AB
0.5	10	0.5663	0.0550	0.0663	0.4990	0.0390	-0.0010	0.5510	0.0770	0.0510	0.5470	0.0710	0.0470	0.5600	0.0780	0.0600	0.5420	0.0540	0.0420
	20	0.5313	0.0199	0.0313	0.4900	0.0160	-0.0100	0.5240	0.0260	0.0240	0.5220	0.0240	0.0220	0.5290	0.0270	0.0290	0.5200	0.0210	0.0200
	40	0.5148	0.0085	0.0148	0.4897	0.0076	-0.0103	0.5109	0.0112	0.0109	0.5102	0.0101	0.0102	0.5133	0.0113	0.0133	0.5095	0.0096	0.0095
	60	0.5099	0.0053	0.0099	0.4913	0.0049	-0.0087	0.5068	0.0070	0.0068	0.5065	0.0063	0.0065	0.5085	0.0071	0.0085	0.5060	0.0061	0.0060
	80	0.5081	0.0040	0.0081	0.4931	0.0037	-0.0069	0.5057	0.0053	0.0057	0.5057	0.0048	0.0057	0.5070	0.0053	0.0070	0.5052	0.0046	0.0052
	100	0.5062	0.0031	0.0062	0.4936	0.0030	-0.0064	0.5043	0.0041	0.0043	0.5043	0.0037	0.0043	0.5052	0.0041	0.0052	0.5038	0.0036	0.0038
1	10	1.1423	0.2458	0.1423	1.0025	0.1722	0.0025	1.1112	0.3376	0.1112	1.1027	0.3136	0.1027	1.1285	0.3398	0.1285	1.0899	0.2397	0.0899
	20	1.0668	0.0872	0.0668	0.9812	0.0701	-0.0188	1.0520	0.1151	0.0520	1.0476	0.1051	0.0476	1.0614	0.1171	0.0614	1.0438	0.0935	0.0438
	40	1.0315	0.0368	0.0315	0.9794	0.0326	-0.0206	1.0235	0.0485	0.0235	1.0219	0.0438	0.0219	1.0283	0.0490	0.0283	1.0204	0.0415	0.0204
	60	1.0210	0.0228	0.0210	0.9825	0.0210	-0.0175	1.0150	0.0303	0.0150	1.0143	0.0272	0.0143	1.0182	0.0306	0.0182	1.0131	0.0263	0.0131
	80	1.0171	0.0172	0.0171	0.9860	0.0161	-0.0140	1.0123	0.0227	0.0123	1.0121	0.0204	0.0121	1.0147	0.0229	0.0147	1.0111	0.0199	0.0111
	100	1.0131	0.0133	0.0131	0.9870	0.0127	-0.0130	1.0093	0.0177	0.0093	1.0093	0.0158	0.0093	1.0112	0.0178	0.0112	1.0083	0.0155	0.0083
2	10	2.3080	1.1200	0.3080	2.0184	0.7785	0.0184	2.2417	1.4897	0.2417	2.2242	1.3867	0.2242	2.2752	1.4966	0.2752	2.1962	1.0748	0.1962
	20	2.1441	0.3889	0.1441	1.9662	0.3100	-0.0338	2.1137	0.5094	0.1137	2.1044	0.4659	0.1044	2.1321	0.5172	0.1321	2.0953	0.4139	0.0953
	40	2.0680	0.1629	0.0680	1.9594	0.1433	-0.0406	2.0518	0.2140	0.0518	2.0484	0.1934	0.0484	2.0613	0.2162	0.0613	2.0447	0.1827	0.0447
	60	2.0453	0.1008	0.0453	1.9649	0.0921	-0.0351	2.0328	0.1332	0.0328	2.0313	0.1196	0.0313	2.0392	0.1342	0.0392	2.0286	0.1154	0.0286
	80	2.0367	0.0755	0.0367	1.9716	0.0702	-0.0284	2.0269	0.0999	0.0269	2.0264	0.0899	0.0264	2.0317	0.1005	0.0317	2.0242	0.0873	0.0242
	100	2.0281	0.0584	0.0281	1.9733	0.0552	-0.0267	2.0204	0.0776	0.0204	2.0202	0.0694	0.0202	2.0242	0.0780	0.0242	2.0181	0.0680	0.0181
4	10	4.6605	5.0300	0.6605	4.0686	3.5644	0.0686	4.5206	6.5489	0.5206	4.4866	6.1365	0.4866	4.5832	6.5622	0.5832	4.4271	4.8272	0.4271
	20	4.3121	1.7657	0.3121	3.9425	1.3947	-0.0575	4.2468	2.2763	0.2468	4.2273	2.0916	0.2273	4.2815	2.3042	0.2815	4.2081	1.8653	0.2081
	40	4.1480	0.7350	0.1480	3.9208	0.6414	-0.0792	4.1137	0.9582	0.1137	4.1068	0.8691	0.1068	4.1318	0.9661	0.1318	4.0983	0.8203	0.0983
	60	4.0986	0.4527	0.0986	3.9296	0.4110	-0.0704	4.0721	0.5914	0.0721	4.0688	0.5324	0.0688	4.0843	0.5950	0.0843	4.0629	0.5147	0.0629
	80	4.0790	0.3363	0.0790	3.9421	0.3108	-0.0579	4.0590	0.4435	0.0590	4.0578	0.3997	0.0578	4.0683	0.4457	0.0683	4.0530	0.3886	0.0530
	100	4.0603	0.2581	0.0603	3.9448	0.2427	-0.0552	4.0447	0.3428	0.0447	4.0442	0.3070	0.0442	4.0521	0.3441	0.0521	4.0393	0.3004	0.0393

Table 5: Classical confidence intervals for different values of n

θ	n	ALCL	AUCL	AW	CP	θ	n	ALCL	Upper	Width	CP
0.5	10	0.1829	0.9497	0.7668	95.52	2.0	10	0.5978	4.0183	3.4205	95.33
0.5	20	0.2787	0.7839	0.5052	95.57	2.0	20	1.0360	3.2530	2.2170	95.65
0.5	40	0.3423	0.6874	0.3451	95.19	2.0	40	1.3160	2.8200	1.5030	95.25
0.5	60	0.3705	0.6493	0.2788	95.41	2.0	60	1.4400	2.6510	1.2120	95.37
0.5	80	0.3878	0.6283	0.2405	95.43	2.0	80	1.5150	2.5590	1.0440	95.36
0.5	100	0.3991	0.6133	0.2142	95.25	2.0	100	1.5634	2.4928	0.9294	95.21
1	10	0.3371	1.9475	1.6104	95.48	4.0	10	1.0020	8.3190	7.3160	95.02
1	20	0.5403	1.5933	1.0531	95.61	4.0	20	1.9600	6.6650	4.7050	95.47
1	40	0.6731	1.3899	0.7168	95.29	4.0	40	2.5600	5.7360	3.1760	95.13
1	60	0.7317	1.3103	0.5786	95.41	4.0	60	2.8200	5.3770	2.5560	95.37
1	80	0.7677	1.2665	0.4989	95.52	4.0	80	2.9780	5.1790	2.2010	95.35
1	100	0.7910	1.2352	0.4443	95.21	4.0	100	3.0810	5.0390	1.9580	95.29

Table 6: Summary of dataset I, II, and III

Dataset No.	Minimum	1st Quartile	Median	Mean	3rd Quartile	Maximum	SD
I	0.3200	0.9150	1.4700	1.6750	2.0870	4.7500	1.0006
II	2.8870	5.3290	8.1440	9.8870	13.8360	23.3940	5.8567
III	0.1100	0.7175	1.2350	1.5427	1.9425	4.7300	1.1276

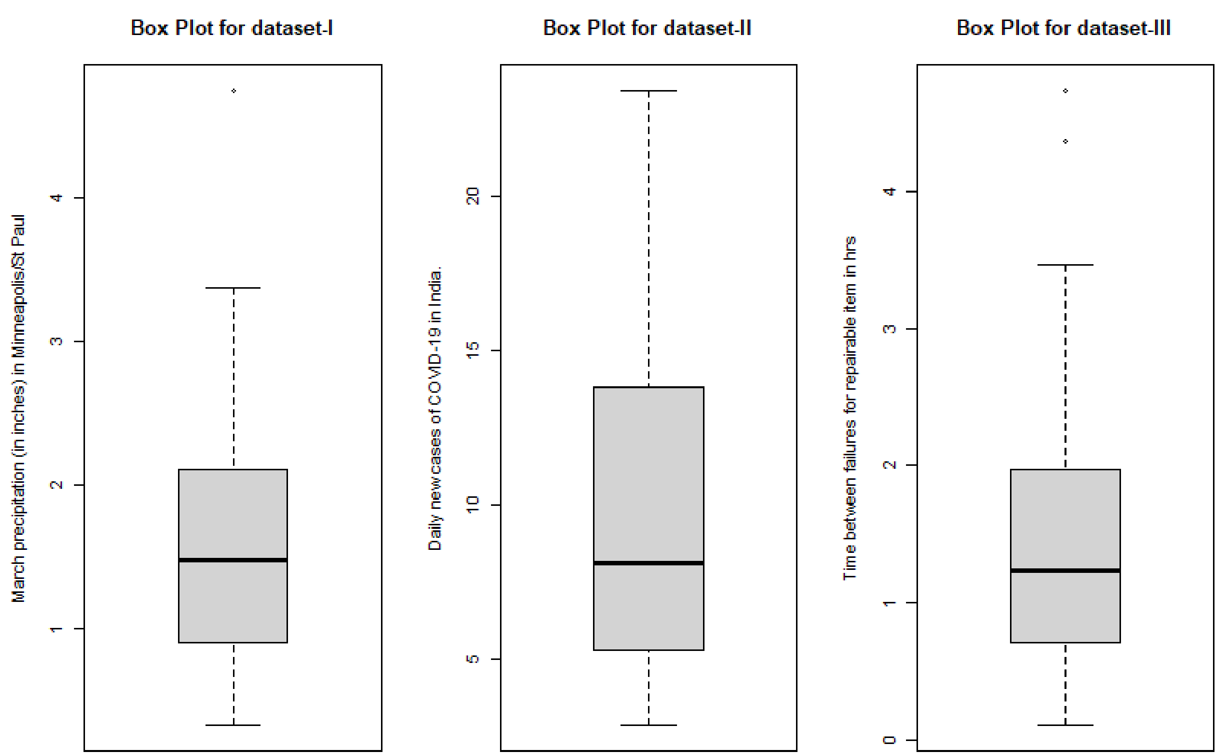


Figure 4: Graphical representation of dataset I, II, and III using boxplot

Table 7: The Competitive models

Model	Abbreviation	Parameter(s)	Author(s)
Lindley	L	θ	Lindley (1958)
Inverse Lindley	IL	θ	Sharma *et al.* (2015)
Inverted Modified Lindley	IML	θ	Chesneau *et al.* (2020)
Exponential	E	λ	-
Inverted Exponential	IE	β	Lin *et al.* (1989)
Inverse Rayleigh	IR	σ	Voda (1972)
Inverse Xgamma	IXG	θ	Yadav *et al.* (2021b)
Inverted Gamma	IG	α, β	Lin *et al.* (1989)
Inverse Weibull	IW	η, β	Khan *et al.* (2008)
Inverted Nadarajah–Haghighi	INH	λ, α	Tahir *et al.* (2018)
Inverted Topp-Leone	ITL	θ	Hassan *et al.* (2020)
Burr-Hatke Exponential	BHE	λ	Yadav *et al.* (2021a)
Maxwell Distribution	M	θ	Bekker and Roux (2005)
Laplace Distribution	La	μ, b	Kotz *et al.* (2001)
Inverse Lomax Distribution	ILo	α, β	Kleiber (2004)
Exponential Poisson Distribution	EP	λ, β	Kuş (2007)
Rayleigh	R	σ	Siddiqui (1962)

Dataset (I): The first real dataset represents thirty successive values of March precipitation (in inches) in Minneapolis/St Paul Yousef *et al.* (2023). The data values are: 0.77, 1.74, 0.81, 1.2, 1.95, 1.2, 0.47, 1.43, 3.37, 2.2, 3, 3.09, 1.51, 2.1, 0.52, 1.62, 1.31, 0.32, 0.59, 0.81, 2.81, 1.87, 1.18, 1.35, 4.75, 2.48, 0.96, 1.89, 0.9, 2.05.

Under dataset I, the fitting of the Arvind distribution is compared with L, IL, IML, E, IE, IR, IXG, IG, IW, and INH models. Table 8 lists the MLEs of the unknown parameters (standard errors (SEs) between parentheses) with the values of the –logL, AIC, BIC, CAIC, and KS statistic with associated P-value. Table 8 demonstrates that the suggested model has the lowest -logL, AIC, BIC, CAIC, and KS statistic as well as the highest P-value, hence, the Arvind distribution is superior to a number of competing models for this dataset. Figure 5 depicts the density and empirical vs fitted CDF plots for the proposed model with respect to dataset I. This graph also indicates that the Arvind distribution closely resembles the pattern of this real data.

Dataset (II): The second application takes into account the daily new cases of COVID-19 that have been reported in India. The data can be accessed at `https://www.worldometers.info/coronavirus/country/india/` and describes the daily new cases (in thousands) that occurred between the 16th of March 2021 and the 16th of April 2021. The data values are as follows:
28869, 35838, 39643, 40950, 43815, 40611, 47264, 53419, 59069, 62291, 62631, 68206, 56119, 53158, 72182, 81441, 89019, 92998, 103793, 96557, 115269, 126315, 131893, 144829, 152682, 169914, 160694, 185248, 199509, 216850, 233943.

To facilitate fitting, this dataset has been divided by 10000. The Arvind distribution's fit to this COVID data is compared to the L, IL, ITL, IXG, BHE, M, La, ILo, EP, and INH models. Table 9 summarizes the MLEs of the parameters (SEs between parentheses) as well

Table 8: The goodness-of-fit statistics for various models under dataset I

Model	MLE (SEs)	-logL	AIC	BIC	CAIC	KS	P-value
Arvind	0.1928 (0.0367)	39.7202	81.4403	82.8415	81.5832	0.0899	0.9685
L	0.9096 (0.1247)	43.1437	88.2875	89.6886	88.4303	0.1882	0.2383
IL	1.5835 (0.2267)	45.2212	92.4424	93.8436	92.5852	0.2279	0.0886
IML	1.247 (0.1906)	43.8683	89.7366	92.5390	90.1810	0.1975	0.1925
E	0.5971 (0.1090)	45.4744	92.9488	94.3500	93.0917	0.2352	0.0723
IE	1.1405 (0.2083)	46.2726	94.5452	95.9464	94.6881	0.2538	0.0420
IR	0.9267 (0.0846)	44.1365	90.2730	91.6740	90.4160	0.9360	0.0640
IXG	1.9440 (0.2680)	46.9850	95.9701	97.3713	96.1129	0.2632	0.0313
IG	2.5928 (0.6306), 2.9599 (0.7944)	40.3072	84.6144	87.4168	85.0589	0.1380	0.6174
IW	1.0163 (0.1273), 1.5495 (0.2026)	41.9170	87.8340	90.6364	88.2785	0.1523	0.4896
INH	3.0625 (2.8279), 0.2647 (0.2975)	44.5344	93.0689	95.8713	93.5133	0.1961	0.1989

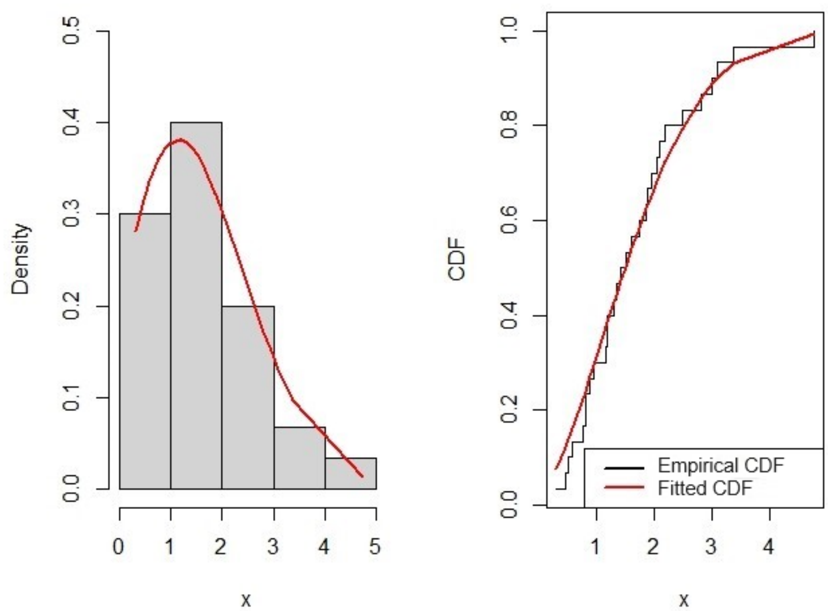
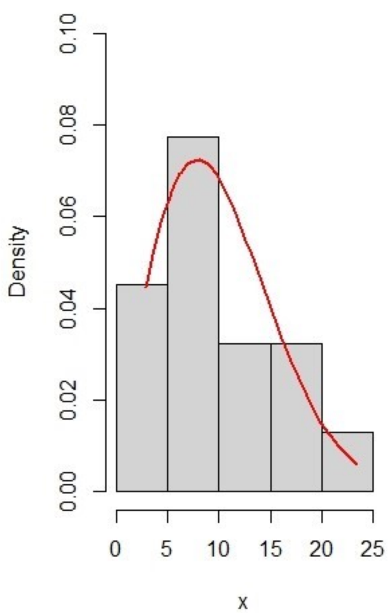


Figure 5: Histogram and the empirical vs fitted CDF under datasets I

as the -logL, AIC, BIC, CAIC, and KS statistic with its P-value. The developed model has the lowest -logL, AIC, BIC, CAIC, and KS statistic as well as the maximum P-value, as shown in Table 9; as a consequence, the Arvind distribution surpasses other competing models for this dataset. Figure 6 depicts the density and empirical vs fitted CDF for the proposed model under dataset II. This figure also reveals that the Arvind distribution closely follows the actual data pattern.

Table 9: The goodness-of-fit statistics for various models under dataset II

Model	MLE (SEs)	-logL	AIC	BIC	CAIC	KS	P-value
Arvind	0.0071 (0.0013)	95.0043	192.0085	193.4425	192.1465	0.1432	0.5032
L	0.1863 (0.0238)	96.7713	195.5426	196.9766	195.6805	0.1662	0.3216
IL	7.8937 (1.2790)	101.7100	205.4199	206.8539	205.5578	0.2593	0.0250
ITL	0.6169 (0.1108)	117.8740	237.7481	239.1820	237.8860	0.4094	<0.0001
IXG	8.4800 (1.3980)	102.7350	207.4706	208.9046	207.6086	0.2726	0.0158
BHE	0.0552 (0.1108)	103.7590	209.5174	210.9514	209.6554	0.2831	0.0108
M	0.0229 (0.0034)	97.5042	197.0083	198.4423	197.1463	0.2450	0.0401
La	8.1441 (0.0020), 4.6943 (0.8431)	100.4290	204.8579	207.7259	205.2865	0.1632	0.3434
ILo	0.6935 (0.2864), 10.6855 (3.7119)	102.8453	209.6907	212.5587	210.1193	0.2681	0.0186
EP	0.1010 (0.0182), 2.904e-07(0.01461)	102.0284	208.0567	210.9247	208.4853	0.2715	0.0165
INH	14.7742 (6.8420), 0.3258 (0.1251)	96.9953	197.9905	200.8585	198.4191	0.2028	0.1353



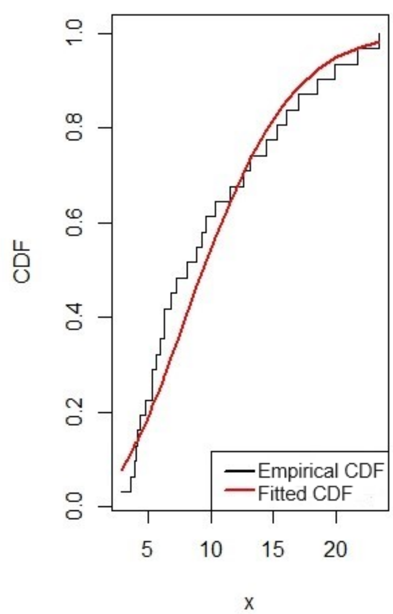


Figure 6: Histogram and the empirical vs fitted CDF under datasets II

Dataset (III): The third dataset includes the time between failures for repairable items, Murthy *et al.* (2004). The data values are as follows:
1.43, 0.11, 0.71, 0.77, 2.63, 1.49, 3.46, 2.46, 0.59, 0.74, 1.23, 0.94, 4.36, 0.40, 1.74, 4.73, 2.23, 0.45, 0.70, 1.06, 1.46, 0.30, 1.82, 2.37, 0.63, 1.23, 1.24, 1.97, 1.86, 1.17.

The Arvind distribution is fitted to this data and compared to the L, IL, E, R, IXG, La, ILo, IG, IW, and INH models. Table 10 depicts the MLEs of the parameters (SEs between parentheses) with the different fitting measures. The suggested model has the lowest -logL, AIC, BIC, CAIC, and KS statistics, as well as the greatest P-value, as shown in Table 10; as a result, the Arvind distribution excels other competing models for this dataset. Figure 7 shows the density and empirical vs fitted CDF for the Arvind model under dataset II. This graphic also demonstrates how well the Arvind distribution fits the actual data pattern.

Table 10: The goodness-of-fit statistics for various models under dataset III

Model	MLE (SEs)	-logL	AIC	BIC	CAIC	KS	P-value
Arvind	0.2042 (0.0388)	40.9532	83.9063	85.3075	84.0492	0.1205	0.7763
L	0.9761 (0.1345)	41.5473	85.0946	86.4958	85.2374	0.1406	0.5931
IL	1.1605 (0.1619)	46.9329	95.8658	97.2670	96.0087	0.1412	0.5885
E	0.6484 (0.1184)	43.0054	88.0108	89.4120	88.1536	0.1845	0.2586
R	1.3434 (0.1226)	42.9183	87.8366	89.2378	87.9794	0.1865	0.2479
IXG	1.4160 (0.1892)	48.9037	99.8073	101.2085	99.9502	0.1556	0.4615
La	1.2374 (8.3889), 0.8074 (0.1474)	44.37386	92.74771	95.55011	93.19216	0.12375	0.7478
ILo	0.11873 (0.05041), 7.73475 (2.62162)	46.01338	96.02677	98.82916	96.47121	0.18931	0.2325
IG	1.4209 (0.3325), 1.1271 (0.3152)	45.5074	95.0147	97.8171	95.4591	0.1576	0.4452
IW	0.7665 (0.1388), 1.0730 (0.1314)	46.3756	96.7512	99.5536	97.1957	0.1338	0.6557
INH	0.8517 (0.2348), 1.0347 (0.5133)	46.3701	96.7402	99.5426	97.1846	0.1786	0.2942

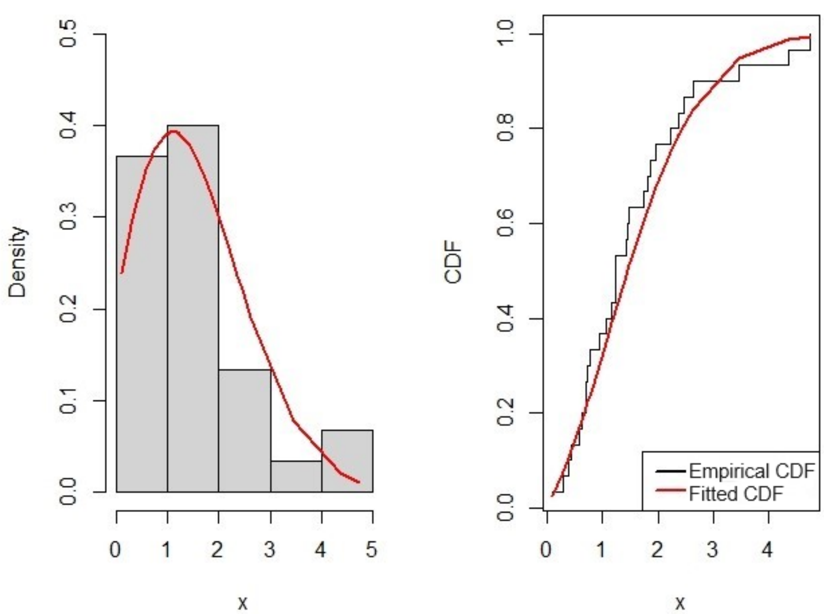


Figure 7: Histogram and the empirical vs fitted CDF under datasets III

6.1. Other classical estimates for datasets I, II, and III

Here, we estimate the Arvind distribution's unknown parameter using several different approaches that have been used in this article. Under I, II, and III datasets, we also obtain the 95% ACI for the unknown parameter θ. Table 11 includes estimates for θ from MLE, MPS, OLS, WLS, CVM, and AD estimation methods along with respective SEs, and 95% ACI. In order to compare various approaches, Table 11 also contains the KS statistic and corresponding P-value for all approaches. From Table 11, we can easily see the reverse trend from the simulation section, as MPS is the least favourable estimation method for datasets I, II, and III.

Table 11: Classical estimates for dataset I, II, and III

Dataset	Methods	Estimate	SEs	KS	P-value	ACI (Width)
I	MLE	0.1928	0.0367	0.0899	0.9685	[0.1209, 0.2647] (0.1439)
	MPS	0.1824	0.1941	0.0952	0.9484	
	OLS	0.1920	0.0039	0.0892	0.9707	
	WLS	0.1878	0.0051	0.0851	0.9816	
	CVM	0.1927	0.0048	0.0899	0.9684	
	ADE	0.1904	0.0401	0.0877	0.9751	
II	MLE	0.0071	0.0013	0.1432	0.5032	[0.0046, 0.0096] (0.0050)
	MPS	0.0067	0.0068	0.1571	0.3877	
	OLS	0.0081	0.0003	0.1114	0.7960	
	WLS	0.0075	0.0004	0.1304	0.6209	
	CVM	0.0081	0.0003	0.1107	0.8026	
	ADE	0.0075	0.0015	0.1303	0.6220	
III	MLE	0.2042	0.0388	0.1205	0.7763	[0.1281, 0.2803] (0.1522)
	MPS	0.1910	0.2029	0.1427	0.5745	
	OLS	0.2417	0.0050	0.0698	0.9986	
	WLS	0.2364	0.0066	0.0709	0.9982	
	CVM	0.2425	0.0049	0.0702	0.9985	
	ADE	0.2292	0.0499	0.0815	0.9886	

7. Concluding remarks

A new lifetime model named Arvind distribution has been developed for modelling different types of data. The suggested model's PDF and HRF have a variety of forms that make it possible to analyze a broad range of real data. Its impressive statistical properties have been derived. Six different estimation methods namely the maximum likelihood, maximum product spacings, ordinary and weighted least square, Cramér-von Mises, and Anderson-Darling are discussed for estimating the unknown parameter. The asymptotic confidence interval has also been provided for the unknown parameter. An extensive simulation study has been performed to study the performance of the considered methods of estimations. This study suggests that methods of maximum product spacings and maximum likelihood are highly preferable whereas Cramér-von Mises is the least preferable method of estimation for the proposed model.

The goodness-of-fit of the proposed distribution has been explained with three real datasets from different fields and the fits of the proposed model have been found quite satisfactory over other existing lifetime models like Lindley, inverse Lindley, inverted modified Lindley, inverse Xgamma, inverse gamma, inverse Weibull, inverted Nadarajah-Haghighi, Burr-Hatke Exponential, *etc.* As a result, we may draw the conclusion that the proposed model may be utilized as a substitute for several well-known current models to analyze data produced from diverse fields. In the future, we will extend this work by implementing censoring and different stress-strength models in the Arvind Distribution under various classical and non-classical estimation procedures.

Funding

No funds, grants, or other support were received.

Data availability

The authors confirm that the data supporting the study's findings are presented in the manuscript.

Conflict of interest

The authors state that they do not have any conflicts of interest.

Acknowledgements

The authors express their deep gratitude to the respected editor and reviewer for the valuable time and effort put into enhancing the overall quality of this manuscript.

References

Agiwal, V., Tyagi, S., and Chesneau, C. (2023). Bayesian and frequentist estimation of stress-strength reliability from a new extended Burr XII distribution: Accepted: March 2023. *REVSTAT-Statistical Journal*.

Alsuhabi, H., Alkhairy, I., Almetwally, E. M., Almongy, H. M., Gemeay, A. M., Hafez, E., Aldallal, R., and Sabry, M. (2022). A superior extension for the Lomax distribution with application to covid-19 infections real data. *Alexandria Engineering Journal*, **61**, 11077–11090.

Bakouch, H. S., Nik, A. S., Asgharzadeh, A., and Salinas, H. S. (2021). A flexible probability model for proportion data: Unit-half-normal distribution. *Communications in Statistics: Case Studies, Data Analysis and Applications*, **7**, 271–288.

Bekker, A. and Roux, J. (2005). Reliability characteristics of the maxwell distribution: A Bayes estimation study. *Communications in Statistics-Theory and Methods*, **34**, 2169–2178.

Bonferroni, C. (1930). *Elmenti Di Statistica Generale*. Libreria Seber, Firenze.

Chaubey, Y. P. and Zhang, R. (2015). An extension of Chen's family of survival distributions with bathtub shape or increasing hazard rate function. *Communications in Statistics-Theory and Methods*, **44**, 4049–4064.

Cheng, R. and Amin, N. (1979). Maximum product-of-spacings estimation with applications to the lognormal distribution. *Math Report*, **791**.

Cheng, R. and Amin, N. (1983). Estimating parameters in continuous univariate distributions with a shifted origin. *Journal of the Royal Statistical Society: Series B (Methodological)*, **45**, 394–403.

Chesneau, C., Tomy, L., Gillariose, J., and Jamal, F. (2020). The inverted modified Lindley distribution. *Journal of Statistical Theory and Practice*, **14**, 46.

Choudhary, N., Tyagi, A., and Singh, B. (2021). A flexible bathtub-shaped failure time model: Properties and associated inference. *Statistica*, **81**, 65–92.

El-Morshedy, M., Alshammari, F. S., Tyagi, A., Elbatal, I., Hamed, Y. S., and Eliwa, M. S. (2021). Bayesian and frequentist inferences on a type I half-logistic odd Weibull generator with applications in engineering. *Entropy*, **23**, 446.

Goel, R. and Singh, B. (2020). Estimation of p (y < x) for modified weibull distribution under progressive type-II censoring. *Life Cycle Reliability and Safety Engineering*, **9**, 227–240.

Gupta, R. D. and Kundu, D. (1999). Generalized exponential distributions. *Australian & New Zealand Journal of Statistics*, **41**, 173–188.

Hassan, A. S., Elgarhy, M., and Ragab, R. (2020). Statistical properties and estimation of inverted Topp-Leone distribution. *Journal of Statistics, Applications and Probability*, **9**, 319–331.

Khan, M. S., Pasha, G., and Pasha, A. H. (2008). Theoretical analysis of inverse weibull distribution. *WSEAS Transactions on Mathematics*, **7**, 30–38.

Kleiber, C. (2004). Lorenz ordering of order statistics from log-logistic and related distributions. *Journal of Statistical Planning and Inference*, **120**, 13–19.

Kotz, S., Kozubowski, T., and Podgórski, K. (2001). *The Laplace Distribution and Generalizations: A Revisit with Applications to Communications, Economics, Engineering, and Finance*. Number 183. Springer Science & Business Media.

Kuş, C. (2007). A new lifetime distribution. *Computational Statistics & Data Analysis*, **51**, 4497–4509.

Lin, C., Duran, B., and Lewis, T. (1989). Inverted gamma as a life distribution. *Microelectronics Reliability*, **29**, 619–626.

Lindley, D. V. (1958). Fiducial distributions and Bayes' theorem. *Journal of the Royal Statistical Society. Series B (Methodological)*, 102–107.

Lorenz, M. O. (1905). Methods of measuring the concentration of wealth. *Publications of the American statistical association*, **9**, 209–219.

Mudholkar, G. S., Srivastava, D. K., and Kollia, G. D. (1996). A generalization of the Weibull distribution with application to the analysis of survival data. *Journal of the American Statistical Association*, **91**, 1575–1583.

Murthy, D. P., Xie, M., and Jiang, R. (2004). *Weibull Models*. John Wiley & Sons.

Nadarajah, S. and Haghighi, F. (2011). An extension of the exponential distribution. *Statistics*, **45**, 543–558.

Nadarajah, S. and Kotz, S. (2006). The beta exponential distribution. *Reliability Engineering & System safety*, **91**, 689–697.

Shaked, M. and Shanthikumar, J. G. (2007). *Stochastic Orders*. Springer.

Sharma, V. K., Singh, S. K., Singh, U., and Agiwal, V. (2015). The inverse Lindley distribution: A stress-strength reliability model with application to head and neck cancer data. *Journal of Industrial and Production Engineering*, **32**, 162–173.

Siddiqui, M. M. (1962). Some problems connected with Rayleigh distributions. *Journal of Research of the National Bureau of Standards D*, **66**, 167–174.

Swain, J. J., Venkatraman, S., and Wilson, J. R. (1988). Least-squares estimation of distribution functions in Johnson's translation system. *Journal of Statistical Computation and Simulation*, **29**, 271–297.

Tahir, M., Cordeiro, G. M., Ali, S., Dey, S., and Manzoor, A. (2018). The inverted Nadarajah–Haghighi distribution: Estimation methods and applications. *Journal of Statistical Computation and Simulation*, **88**, 2775–2798.

Tyagi, S., Kumar, S., Pandey, A., Saha, M., and Bagariya, H. (2022). Power xgamma distribution: Properties and its applications to cancer data. *International Journal of Statistics and Reliability Engineering*, **9**, 51–60.

Voda, V. G. (1972). On the inverse Rayleigh distributed random variable. *Reports of Statistical Application Research*, **19**, 13–21.

Yadav, A. S., Altun, E., and Yousof, H. M. (2021a). Burr–Hatke exponential distribution: A decreasing failure rate model, statistical inference and applications. *Annals of Data Science*, **8**, 241–260.

Yadav, A. S., Maiti, S. S., and Saha, M. (2021b). The inverse xgamma distribution: statistical properties and different methods of estimation. *Annals of Data Science*, **8**, 275–293.

Yousef, M. M., Alsultan, R., and Nassr, S. G. (2023). Parametric inference on partially accelerated life testing for the inverted kumaraswamy distribution based on type-ii progressive censoring data. *Mathematical Biosciences and Engineering*, **20**, 1674–1694.

Zenga, M. (2007). Inequality curve and inequality index based on the ratios between lower and upper arithmetic means. *Statistica & Applicazioni*, **5**, 3–27.