Modelling Unit Interval COVID-19 Data: An Application of Unit Nadarajah-Haghighi Distribution

Hadiqa Basit[1], Shakila Bashir[2], Bushra Masood[3] and Nadia Mushtaq[4]

https://doi.org/10.62345/jads.2023.12.3.60

Abstract
During the year 2019 and onwards, people from all over the world faced a new deadly and fatal virus named COVID-19. It caused the death of millions of people from all around the world. Since then, scientists have been trying to develop antiviral research and finding appropriate effective medicines. In this research, a new unit interval probability distribution has been proposed to estimate and model the recovery rate of COVID-19. The proposed probability distribution is developed by transforming the variable and named as unit Nadarajah-Haghighi (UNH) distribution. Several statistical properties including reliability measures, quantile function, moments, some entropy measures, order statistics, stress-strength and stochastic ordering have been discussed. Estimation of parameters evaluated by numerical and simulation study. The UNH distribution plays an important role due to its flexibility and variety of shapes. Along with the recovery rate of COVID-19, milk production data is also used to check the usefulness of the UNH distribution. The proposed model is more competitive and flexible as compared to the other unit interval distributions available in the literature. It is necessary to think about strategies, diagnostics, and predicting future factors to lessen the epidemics. At this stage statisticians and policy makers can play an important role in preventing future viral epidemics by estimating and modeling.
Keywords: COVID-19; Unit Interval; Nadarajah-Haghighi; UNH; Entropies

Introduction

In 2020, with the widespread effect of COVID-19, there arose a need to model several aspects of the pandemic, one such aspect was the recovery rate. For those involved in controlling the spread of a disease it is imperative to understand the nature of how patients recover after being given a treatment. The development of distributions that allow modelling of the recovery rate allows for the required mathematical inference to take place.

Unit distributions can be used for modelling proportions and percentages, which is why a substantial body of literature has been published for the development of such distributions. Although in the presence of beta and Kumaraswamy distribution there is a need of development of new unit interval distributions due to certain reasons, for example firstly, beta distribution has unsolved expression for the cdf and quantile which ultimately create problems while modeling;

[1]*Department of Statistics, Forman Christian College (A Chartered University) Lahore, Pakistan*
[2]*Department of Statistics, Forman Christian College (A Chartered University) Lahore, Pakistan. Email: shakilabashir@fccollege.edu.pk.*
[3]*Department of Statistics, Forman Christian College (A Chartered University) Lahore, Pakistan*
[4]*Department of Statistics, Forman Christian College (A Chartered University) Lahore, Pakistan*

Compliance with ethical standards: There are no conflicts of interest (financial or non-financial). This study did not receive any funding.

secondly every distribution is not suitable to model all type of data sets due to different type of the data sets. Therefore, there is a need of the development of the new models always which could be modeled more efficiently on different type of data sets. In this regard some notable contributions in this field include the work of Mazucheli et al. (2018) on the unit Weibull and inference, Menezes et al. (2018) on the unit logistic, Mazucheli et al. (2019) on the unit Gompertz, Ghitany et al. (2019) on the unit inverse Gaussian, Mazucheli et al. (2020) on unit- Weibull distribution with quantile conditional, Bantan et al. (2020) on the unit Rayleigh, Korkmaz et al. (2021) on the unit Bur-XII, and Korkmaz et al. (2022) on the unit Chen. Haj et al. (2023) on unit exponential Pareto density for modelling COVID-19, Additionally, there have been significant efforts to apply unit distributions in modeling COVID-19 data. Notable contributions in this area include the work by Bantan et al. (2022) as well as the research conducted by Haj Ahmad, H. et al. (2023).
Nadarajah and Haghighi (2011) proposed an extension of the exponential distribution and named it Nadarajah Haghighi distribution (NHD). The probability density function (pdf) and cumulative distribution function (cdf) of the NHD is as follows:

$$f(x) = \alpha\beta(1+\beta x)^{\alpha-1}e^{1-(1+\beta x)^{\alpha}}, \qquad x>0,\ \alpha>0, \beta>0. \tag{1}$$
$$F(x) = 1 - e^{1-(1+\beta x)^{\alpha-1}}. \tag{2}$$

NHD is considered a very good alternative for the gamma and Weibull distributions. It has elegant properties such as variety of shapes for the pdf and hazard function as compared to the gamma, Weibull and exponentiated exponential distribution.
Later various work has been done on NHD including Nascimento et al (2019) proposed odd Nadarajah Haghighi family of Distributions, Pena-Ramirez et al. (2019) presented Nadarajah Haghighi Lindley distribution, Wu and Gui (2021) on estimation and prediction on NHD under censoring, Nagarjuna et al. (2022) proposed Nadarajah Haghighi Lomax distribution with applications, Newer et al (2023) suggested rank set sampling for NHD and inference, and many others have done work on NHD.
This paper considers the unit Nadarajah Haghighi distribution (UNHD) for the purpose of modelling the unit interval data sets such as recovery rates of COVID-19 and milk production. Although in the literature there are some well-known and classic unit interval distributions like beta, Kumaraswamy distributions etc. but in spite of these the proposed UNHD is still useful in the sense that some unit interval densities have complex cdf and quantile function to apply it for the life data analysis, and moreover all distributions are not always suitable for every type of data sets. The suitability of the distribution depends on many things such as the variety of density graphical shapes, hazard rate function (hrf) shapes and others. Therefore, the proposed density can be a very useful addition in the unit interval data applications.
This paper is divided into the following sections: section 2 formation of the model/density, reliability properties, density and hrf plots, section 3 represents basic properties of the proposed distribution including moments, quantile function, median etc. section 4 delves into different types of entropy and order statistics, section 5 elaborates on the estimation method used and Monte Carlo simulations, section 6 studies the performance of the distribution on real datasets, Finally, section 7 concludes this paper.

Methodology

The Nadarajah-Haghighi distribution (NHD) is a lifetime model whose domain is $(0,\infty)$, and in the methodology section the NHD is transformed into a unit interval model named as unit interval

Nadarajah-Haghighi distribution (UNHD). Let a random variable (RV) X follows the NHD and by using the transformation $X = \frac{Y}{1-Y}$, a new model is developed as UNHD as given below.

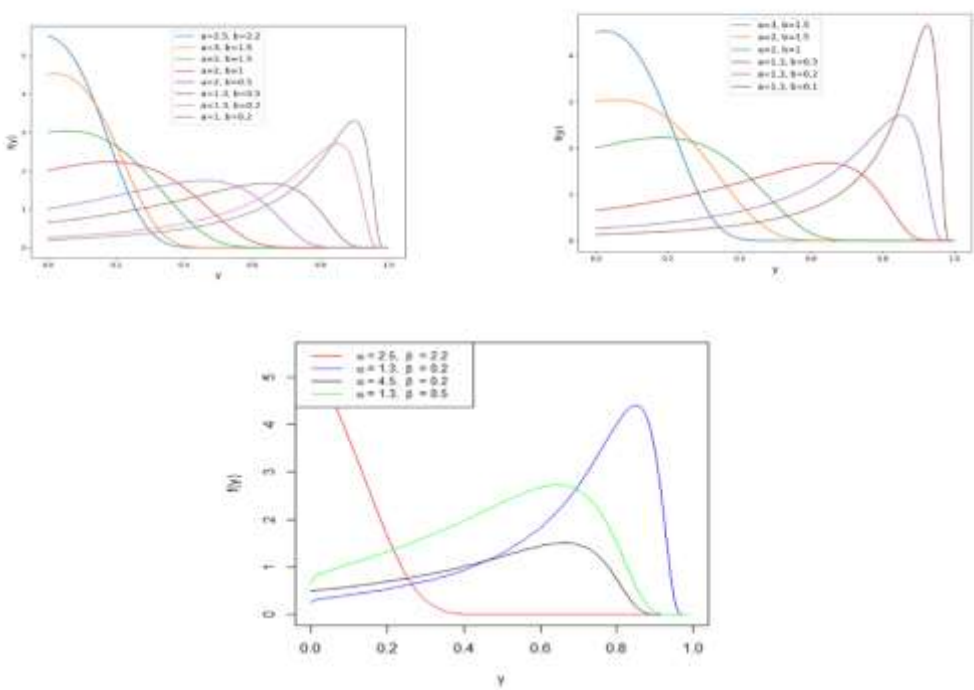
Unit Nadarajah-Haghighi Distribution (UNHD)

The newly proposed unit Nadarajah Haghighi distribution (UNHD) is obtained by applying a unique transformation as given $X = \frac{Y}{1-Y}$, which gives us the following form of the pdf and cdf of the UNHD.

$$f(y) = \frac{\alpha\beta(1+\beta\left(\frac{y}{1-y}\right))^{\alpha-1}}{(1-y)^2} e^{1-(1+\beta(\frac{y}{1-y}))^{\alpha}}, \quad 0 < y < 1, \ \alpha \,\&\, \beta > 0 \tag{3}$$

$$F(y) = 1 - e^{1-(1+\beta(\frac{y}{1-y}))^{\alpha}} \tag{4}$$

Figure 1: Density plot of UNHD for different parametric values



From figure 1, it can be seen that UNHD shows a variety of shapes such as right and left skewed, symmetrical, flat, and peaked. Therefore, UNHD can be appropriate for various types of data sets.

Results/Findings

Under this section, the results regarding the UNHD including reliability measures, statistical properties, entropies, order statistics, estimation of parameters, simulations, and applications are presented.

Reliability Measures

Some important functions for reliability analysis such as survival function (SF), hazard rate function (HRF), cumulative hazard rate function (CHRF), and reversed hazard rate function (RHRF) are given in the following discussion.

The SF for the UNHD is

$$S(y) = e^{1-(1+\beta(\frac{y}{1-y}))^{\alpha}} \quad (5)$$

The HRF for the UNHD is

$$h(y) = \frac{\alpha\beta(1+\beta\left(\frac{y}{1-y}\right))^{\alpha-1}}{(1-y)^2} \quad (6)$$

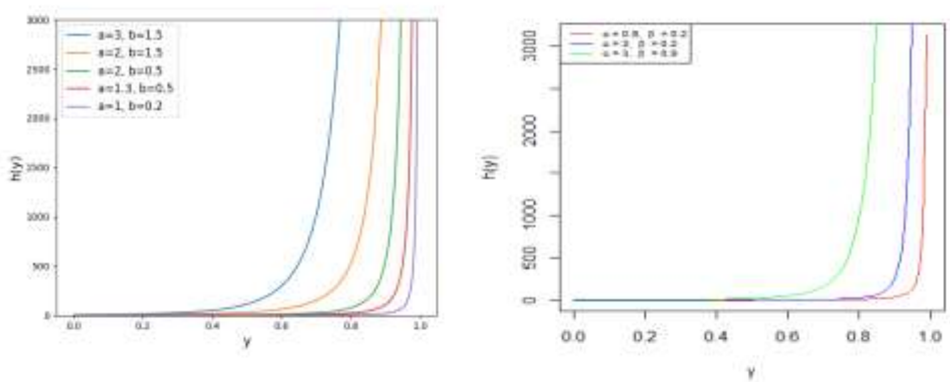
The CHRF for the UNHD is

$$H(y) = 1 - (1 + \beta(\frac{y}{1-y}))^{\alpha} \quad (7)$$

The RHRF for the UNHD is

$$r(y) = \frac{\frac{\alpha\beta(1+\beta\left(\frac{y}{1-y}\right))^{\alpha-1}}{(1-y)^2} e^{1-(1+\beta(\frac{y}{1-y}))^{\alpha}}}{1-e^{1-(1+\beta(\frac{y}{1-y}))^{\alpha}}} \quad (8)$$

Figure 2: HRF plot of UNHD for different parametric values



From figure 2, it is observed that the proposed density UNHD exhibits j shaped hazard rate, moreover the HRF for the UNHD is monotonically increasing which means that after a certain period of time the items/object are going to fail.

Some Statistical Properties

This section presents some basic properties such as moments, mean, variance, quantile function (QF) and median for the UNHD. QF is also used to generate the random number from UNHD.

The QF for the UNHD is

$$y_u = (1 + \beta\left((1 - \ln(1-u))^{1/\alpha} - 1\right)^{-1})^{-1} \quad (9)$$

From the above the median and the interquartile range can be found as follows, Median = $y_{0.5}$. IQR = $y_{0.75} - y_{0.25}$

The rth moments of the UNHD are defined by

$$\mu_r' = \sum_{i=0}^{\infty}\sum_{k=0}^{\infty}\binom{-r}{i}\binom{i}{k}(-1)^{i-k}\beta^i e\Gamma(\frac{\alpha+k}{\alpha}, 1) \quad (10)$$

The mean of the UNHD is

$$\mu_1' = \sum_{i=0}^{\infty}\sum_{k=0}^{\infty}\binom{-1}{i}\binom{i}{k}(-1)^{i-k}\beta^i e\Gamma(\frac{\alpha+k}{\alpha}, 1) \quad (11)$$

The variance of the UNHD is

$$\sigma^2 = \sum_{i=0}^{\infty}\sum_{k=0}^{\infty}\binom{-2}{i}\binom{i}{k}(-1)^{i-k}\beta^i e\Gamma(\frac{\alpha+k}{\alpha},1)$$
$$(\sum_{i=0}^{\infty}\sum_{k=0}^{\infty}\binom{-1}{i}\binom{i}{k}(-1)^{i-k}\beta^i e\Gamma(\frac{\alpha+k}{\alpha},1))^2 \quad (12)$$

Different Types of Entropy

Entropy functions provide a measure of uncertainty that is useful in reliability and risk analysis. Given below are some entropies and their numerical values for different parameter settings.

Rényi Entropy

The Rényi entropy is obtained using the following formula,

$$R_v = \frac{1}{1-v} log\left[\int_0^1 (f(x))^v dx\right], v > 0, v \neq 1$$

For UNHD, the Renyi entropy will be,

$$R_v = \frac{1}{1-v} log\left[\frac{(e^{1-(1+\beta(\frac{y}{1-y}))^{\alpha}} \alpha\beta(1+\beta(\frac{y}{1-y}))^{\alpha-1}}{(y-1)^2}\right]^v, \, , v > 0, v \neq 1 \quad (13)$$

Tsallis Entropy

The Tsallis entropy is obtained using the following formula,

$$T_v = \frac{1}{v-1}\left[1 - \int_0^1 (f(x))^v dx\right], v > 0, v \neq 1$$

For UNHD, the Tsallis entropy will be

$$T_v = \frac{1}{v-1}\left\{1 - \left[\frac{(e^{1-(1+\beta(\frac{y}{1-y}))^{\alpha}} \alpha\beta(1+\beta(\frac{y}{1-y}))^{\alpha-1}}{(y-1)^2}\right]^v\right\} \quad (14)$$

Arimoti Entropy

The Arimoto entropy is obtained using the following formula,

$$A_v = \frac{v}{1-v}\left[\left(\int_0^1 (f(x))^v dx\right)^{\frac{1}{v}} - 1\right], v > 0, v \neq 1$$

For UNHD, the Arimoto entropy will be,

$$A_v = \frac{v}{1-v}\left\{\left[\frac{(e^{1-(1+\beta(\frac{y}{1-y}))^{\alpha}} \alpha\beta(1+\beta(\frac{y}{1-y}))^{\alpha-1}}{(y-1)^2}\right]^v - 1\right\} \quad (15)$$

In Table 1 and 2, the numerical values for the entropies are mentioned.

Table 1: Some numerical values of entropies of UNHD for different parametric values and $v = 0.5$ & 0.8

		$v = 0.5$			$v = 0.8$		
β	α	RE	TE	AE	RE	TE	AE
0.6	1.5	-0.2113	-0.2005	-0.3275	-0.2354	-0.2300	-0.2669
	2.0	-0.3256	-0.3005	-0.5029	-0.3555	-0.3432	-0.4007
	2.5	-0.4352	-0.3911	-0.7001	-0.4711	-0.4496	-0.5348
	3.5	-0.6308	-0.5410	-1.1201	-0.6766	-0.6328	-0.7837
	4.5	-0.7975	-0.6577	-1.5540	-0.8508	-0.7823	-1.0045
1.5	1.5	-0.5059	-0.4470	-0.9320	-0.5621	-0.5317	-0.6610
	2.0	-0.7244	-0.6077	-1.4738	-0.7917	-0.7322	-0.9478
	2.5	-0.9085	-0.7302	-2.0246	-0.9826	-0.8920	-1.1966
	3.5	-1.2033	-0.9042	-3.1370	-1.2846	-1.1329	-1.6125
	4.5	-1.4330	-1.0231	-4.2552	-1.5181	-1.3093	-1.9544
2.5	1.5	-0.7719	-0.6404	-1.7391	-0.8577	-0.7882	-1.0524
	2.0	-1.0441	-0.8134	-2.6528	-1.1371	-1.0171	-1.4246
	2.5	-1.2633	-0.9365	-3.5760	-1.3593	-1.1902	-1.7372
	3.5	-1.6004	-1.1015	-5.4346	-1.6985	-1.4401	-2.2467
	4.5	-1.8546	-1.2087	-7.3003	-1.9531	-1.6168	-2.6575

Table 2: Some numerical values of entropies of UNHD for different parametric values and $v = 0.5$ & 0.8

		$v = 1.6$			$v = 2.5$		
β	α	RE	TE	AE	RE	TE	AE
0.6	1.5	-0.2713	-0.2946	-0.2137	-0.2957	-0.3721	-0.1883
	2.0	-0.3955	-0.4464	-0.3193	-0.4189	-0.5829	-0.2835
	2.5	-0.5175	-0.6069	-0.4163	-0.5430	-0.8388	-0.3691
	3.5	-0.7352	-0.9240	-0.5792	-0.7664	-1.4380	-0.5092
	4.5	-0.9188	-1.2258	-0.7084	-0.9550	-2.1262	-0.6171
1.5	1.5	-0.6371	-0.7760	-0.4831	-0.6785	-1.1781	-0.4159
	2.0	-0.8805	-1.1601	-0.6590	-0.9292	-2.0201	-0.5657
	2.5	-1.0794	-1.5183	-0.7953	-1.1322	-2.9766	-0.6788
	3.5	-1.3902	-2.1713	-0.9934	-1.4475	-5.1798	-0.8370
	4.5	-1.6282	-2.7605	-1.1326	-1.6879	-7.7178	-0.9430
2.5	1.5	-0.9738	-1.3229	-0.7008	-1.0394	-2.5031	-0.5911
	2.0	-1.2602	-1.8834	-0.8941	-1.3284	-4.2233	-0.7509
	2.5	-1.4851	-2.3961	-1.0349	-1.5541	-6.1933	-0.8629
	3.5	-1.8258	-3.3176	-1.2293	-1.8950	-10.7730	-1.0095
	4.5	-2.0804	-4.1405	-1.3601	-2.1495	-16.0909	-1.1024

Order Statistics

Assuming that $X_1, X_2, \dots, X_n$ is a random sample derived from the UNHD and let $X_{(1)}, X_{(2)}, \dots, X_{(n)}$ be the corresponding order statistics. The pdf of the rth order statistic is given as follows,

$$f_r(x) = \frac{1}{B(r,n-r+1)} \frac{\alpha\beta(1+\beta(\frac{y}{1-y}))^{\alpha-1}}{(1-y)^2} e^{1-(1+\beta(\frac{y}{1-y}))^{\alpha}} \sum_{i=0}^{n-r} \binom{n-r}{i} (-1)^i [1-e^{1-(1+\beta(\frac{y}{1-y}))^{\alpha}}]^{i+r-1} \quad (16)$$

Maximum Likelihood Estimates

Maximum likelihood estimation has been used here to find the values of the parameters of UNH distribution, the log-likelihood function of the UNHD is

$$l(\theta) = nln\alpha + nln\beta + n\alpha + \alpha\beta \sum \frac{y}{1-y} - \beta \sum \frac{y}{1-y} - 2\sum \ln(1-y) - \sum(1+\beta(\frac{y}{1-y}))^{\alpha} \quad (17)$$

To find the values of the parameters we take partial derivatives of the likelihood function with respect to the parameters and get the following results.

$$\frac{dl(\theta)}{d\alpha} = \frac{n}{\alpha} + n + \beta \sum \frac{y}{1-y} - \alpha \sum (1+\beta(\frac{y}{1-y}))^{\alpha-1} \quad (18)$$

$$\frac{dl(\theta)}{d\beta} = \frac{n}{\beta} + \alpha \sum \frac{y}{1-y} - \sum \frac{y}{1-y} - \alpha \sum (1+\beta(\frac{y}{1-y}))^{\alpha-1} (\frac{y}{1-y}) \quad (19)$$

The above equations are non-linear, we will use the newton Raphson method to estimate the values of parameters by sung the R language.

Simulation Study

In this section a simulation study is performed to observe the behavior of the distribution parameters. Then MLE estimates were calculated by sampling data from the cumulative distribution of the UNHD with varying sample sizes (30, 50, 75, 100, 300 and 500) from 10,000 of size.

Table 3: Estimated values of average biases, biases, MSE and MRE

		$\alpha = 0.5$				$\beta = 1.0$			
Methods		**20**	**50**	**100**	**300**	**20**	**50**	**100**	**300**
	Average bias	0.6524	0.5705	0.5459	0.5181	0.7300	0.8220	0.8637	0.9260
	Bias	0.1728	0.0863	0.0570	0.0258	0.2701	0.1780	0.1363	0.0741
	MSE	0.0591	0.0181	0.0075	0.0014	0.1519	0.0779	0.0459	0.0151
MLE	MRE	0.3457	0.1726	0.1139	0.0516	0.2701	0.1780	0.1363	0.0741

Table 4: Estimated values of average biases, biases, MSE and MRE

		$\alpha = 1.2$				$\beta = 0.8$			
Methods		**20**	**50**	**100**	**300**	**20**	**50**	**100**	**300**
	Average bias	0.9943	0.9956	0.9987	0.9997	0.9618	0.9788	0.9865	0.9965
	Bias	0.2057	0.2044	0.2013	0.2003	0.1707	0.1795	0.1865	0.1965
	MSE	0.0429	0.0421	0.0406	0.0401	0.0322	0.0341	0.0358	0.0388
MLE	MRE	0.1714	0.1704	0.1677	0.1669	0.2134	0.2244	0.2331	0.2456

Table 5: Estimated values of average biases, biases, MSE and MRE

		$\alpha = 0.1$				$\beta = 1.8$			
Methods		**20**	**50**	**100**	**300**	**20**	**50**	**100**	**300**
	Average bias	0.1151	0.1081	0.1071	0.1062	0.8241	0.9363	0.9699	0.9977
	Bias	0.0186	0.0100	0.0081	0.0063	0.9759	0.8638	0.8302	0.8023
	MSE	0.0007	0.0002	0.0001	0.0001	1.0291	0.7686	0.6982	0.6441
MLE	MRE	0.1858	0.1004	0.0805	0.0631	0.5422	0.4799	0.4612	0.4457

From table 3 to 5, as sample size is increases the MSE is decreases and Bias is also decrease with increasing sample size.

Applications

In this section, two real world data sets are modeled using UNHD to assess its performance against other distributions which are unit exponential Pareto distribution (UEPD), exponential Pareto distribution (EPD), the unit-Weibull (UW) distribution, Kumaraswamy, beta, Kumaraswamy (K), Marshall-Olkin Kumaraswamy (MOK), Marshall-Olkin extended Topp–Leone (MOETL), unit-Gompertz (UG), unit generalized log-Burr XII (UGLBXII), Topp–Leone (TL), and unit gamma/ Gompertz (UGG) distribution.

COVID-19 Recovery Rates in Turkey

As mentioned before, recovery rates are crucial information for professionals such as epidemiologists, health care workers and policy makers. The World Health Organization (WHO) reported the first recovery case in Turkey to be on 26 March 2020. The following are 25 observations on the daily recovery rates calculated between 27 March and 20 April.

Table 6: Daily recovery rates 25 observations

0.0074	0.0095	0.0113	0.015	0.018	0.0212	0.0229
0.0231	0.0328	0.0385	0.0439	0.0464	0.0483	0.0507
0.0515	0.0568	0.0605	0.0648	0.0737	0.0818	0.0955
0.1099	0.127	0.1388	0.1476			

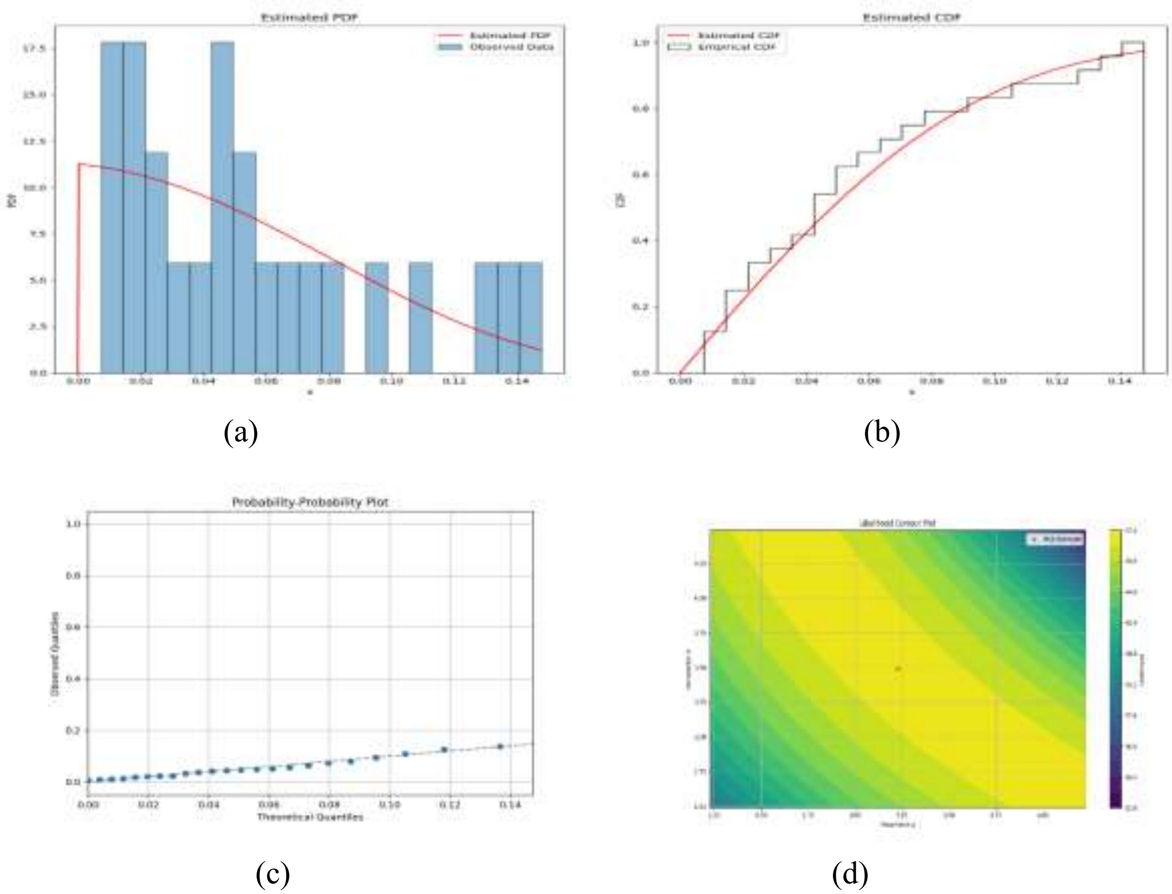
To assess the performance of the distribution the following goodness-of-fit tests were conducted: the Kolmogorov-Smirnov test (KS) and the Cramer-von Mises (CVM) test. The p-value of the KS test (PVKS) has also been listed.
Table 7 compares the performance of UNHD with 9 other distributions where it is observed that it obtains the minimum value for the KS test and has the maximum PVKS value. The CVM statistic is also small and competitive with other distributions.

Table 7: Estimates, Standard Errors (SE), test statistics for recovery of COVID-19 in Turkey

Models	Esti and SE	α	β	λ	KS	PVKS	CVM
UNHD	estimates	3.228	3.489		0.097	0.961	0.035
	SE	2.509	1.944				
UEPD	estimates	1.3419	0.1147	2.0628	0.1005	0.9407	0.0298
	SE	0.2088	0.054	1.0694			
UW	estimates	0.0054	4.1597		0.1362	0.6923	0.0652
	SE	0.0031	0.4182				

Kumaraswamy	estimates	1.4164	50.9406		0.1022	0.9329	0.0310
	SE	0.2303	31.3225				
MOK	estimates	0.1377	1.8755	47.5473	0.1129	0.8722	0.0359
	SE	0.1588	0.3459	52.9895			
UG	estimates	0.0166	1.1455		0.1602	0.4929	0.1242
	SE	0.0125	0.1801				
MOETL	estimates	0.0062	2.0660		0.1058	0.9147	0.0559
	SE	0.0046	0.2976				
UGLBXII	estimates	1.8385	2.7239	3.6048	0.0989	0.9471	0.0375
	SE	3.0788	1.1344	1.6640			
UGG	estimates	1.2880	29.9588	0.8056	0.2189	0.1563	0.0393
	SE	0.2831	14.0703	0.3997			
EPD	estimates	1.4316	0.1171	2.5012	0.1027	0.9304	0.0303
	SE	0.2244	0.9087	27.8069			

Figure 3: Histogram(a), cdf plot (b), pp plot (c) and contour plot (d) of UNHD for recovery rate of COVID-19 in Turkey



Milk Production Data

The following data is obtained for the total milk output in the first birth from 107 SNDI race cows. This data was used by Cordeiro et al. (2012) in their work on beta distribution. Similar to the previous application, the MLEs were calculated along with their standard deviations.

The KS test and the CVM test were also performed to check the goodness-of-fit. The values for the tests show that UNHD is competitive with other distributions.

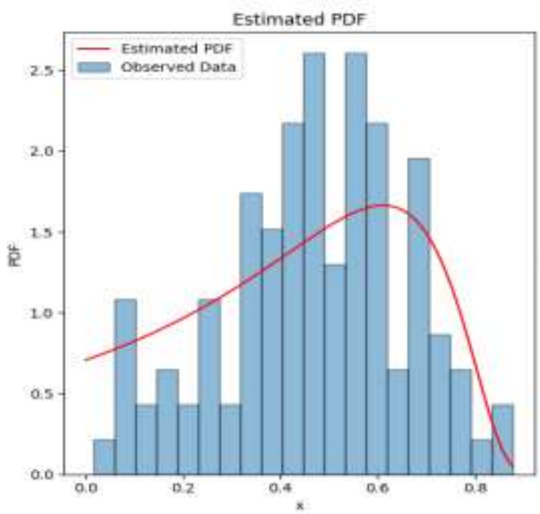
The estimated pdf and cdf plots are also given below along with the PP-plot and the MLE likelihood contour plot.

0.0168	0.1546	0.3188	0.3751	0.4332	0.4612	0.515	0.5553	0.6012	0.6768	0.7471
0.0609	0.216	0.3259	0.3821	0.4365	0.4675	0.5232	0.5627	0.6058	0.6789	0.7629
0.065	0.2303	0.3323	0.3891	0.4371	0.4694	0.5285	0.5629	0.6114	0.6844	0.7687
0.0671	0.2356	0.3383	0.3906	0.4438	0.4741	0.5349	0.5707	0.6174	0.686	0.7804
0.0776	0.2361	0.3406	0.3945	0.447	0.4752	0.535	0.5744	0.6196	0.6891	0.8147
0.0854	0.2605	0.3413	0.4049	0.4517	0.48	0.5394	0.577	0.622	0.6907	0.8492
0.1131	0.2681	0.348	0.4111	0.453	0.4823	0.5447	0.5853	0.6465	0.6927	0.8781
0.1167	0.2747	0.3598	0.4143	0.4553	0.499	0.5481	0.5878	0.6488	0.7131	
0.1479	0.3134	0.3627	0.4151	0.4564	0.5113	0.5483	0.5912	0.6707	0.7261	
0.1525	0.3175	0.3635	0.426	0.4576	0.514	0.5529	0.5941	0.675	0.729	

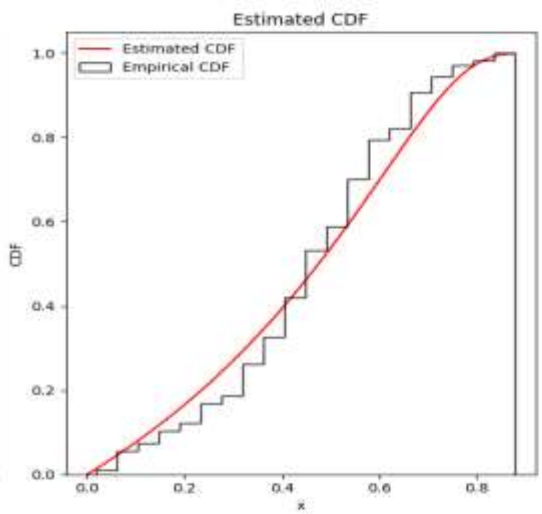
Table 8: estimates, standard errors (SE), and test statistics for milk production data

Models	Estimates and SE	α	β	λ	θ	KS	PVKS	CVM
UNHD	estimates	1.391	0.508			0.117	0.096	0.269
	SE	12.411	22.943					
UEPD	estimates	1.2087	1.1238	0.8427		0.0787	0.5213	0.1088
	SE	0.0867	1.0180	0.4151				
UW	estimates	0.9846	1.5619			0.1206	0.0890	0.3963
	SE	0.1015	0.1064					
Kumaraswamy	estimates	2.1949	3.4366			0.0796	0.5162	0.1561
	SE	0.2224	0.5821					
Beta	estimates	2.4125	2.8297			0.0910	0.3384	0.2083
	SE	0.3145	0.3744					
KK	estimates	0.3781	5.1808	3.8954	1.4834	0.0790	0.5165	0.1120
	SE	0.2337	3.9516	9.2215	2.8696			
UG	estimates	2.1191	0.3878			0.1835	0.0015	0.5206
	SE	0.8684	0.1145					
MOETL	estimates	1.0535	2.0230			0.0968	0.2682	0.2246
	SE	0.3475	0.4118					
UGG	estimates	4.1576	5.2380	0.4268		0.1067	0.1747	0.2195
	SE	1.0592	1.6032	0.1393				
EPD	estimates	2.6012	0.6366	1.6623		0.0832	0.4487	0.2318
	SE	0.2098	6.2038	42.1400				

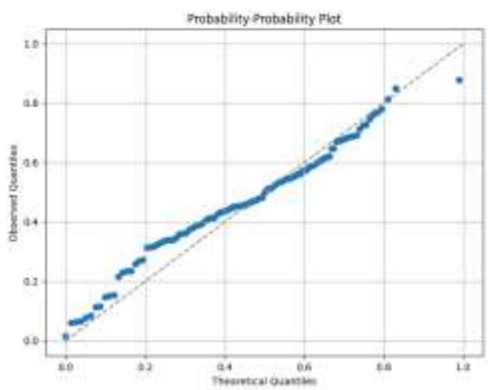
Figure 4: Histogram (a), cdf plot (b), pp plot (c) and contour plot (d) of UNHD for milk production data



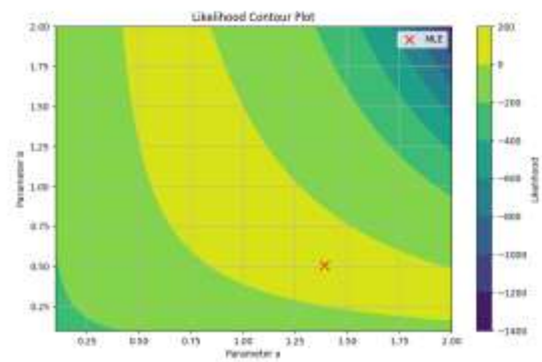
(a)



(b)



(c)



(d)

Conclusion

In this research article, a new unit interval distribution named unit Nadarajah-Haghighi distribution is used to model accurate data in the domain from zero to one. Many vital characteristics were studied, such as the shape of density and hazard rate functions: the quantile function, moments, and order statistics were obtained. Unknown parameters are estimated using the MLE method, and a Monte Carlo simulation study is also conducted to see the performance of the estimated parameters. The proposed density shows a variety of shapes in the context of PDF and HRF, which offers flexibility over different types of unit interval data sets. From the entropy computation in tables 1 and 2, it is concluded that the entropy computes the disorder of a system. If this order is higher, then the probability is lower. Entropy measures the amount of information delivered by discovering the outcome of a random trial. For the proposed density, the entropy values are lower, which means the order is less, and probabilities will be higher. In other words, it is observed that entropies for UNHD are less. That's why UNHD has less uncertainty. Finally, the UNHD is applied to two real-life data sets: the recovery rate of COVID-19 in Turkey and the milk production rate. It is observed that UNHD is showing efficient results on both data sets. UNHD is more flexible as compared to the other competitive unit interval models. By applying the UNHD to these data sets, we can predict future pandemics and take proper actions to avoid disasters. In the case of production, we can estimate the shows and take appropriate measures for it.

References

- Bantan, R.A.R., Chesneau, C., Jamal, F., Elgarhy, M., Tahir, M.H., Ali, A., Zubair, M., Anam, S. (2020). Some New Facts about the Unit-Rayleigh Distribution with Applications. *Mathematics, 8*. 1954.
- Bantan, R.A.R., Shafiq, S., Tahir, M.H., Elhassanein, A., Jamal, F., Almutiry, W, Elgarhy, M. (2022). Statistical Analysis of COVID-19 Data: Using A New Univariate and Bivariate Statistical Model. *J. Funct. Spaces.*
- Cordeiro, G.M., & Santos B. R. (2012). The beta power distribution. *Braz. J. Probab. Stat. 26*, 88–112.
- Haj A, H., Almetwally, E.M., Elgarhy, M., Ramadan, D.A. (2023). On Unit Exponential Pareto Distribution for Modeling the Recovery Rate of COVID-19. *Processes*, *11*, 232.
- Korkmaz, M.Ç., Chesneau, C. (2021). On the unit Burr-XII distribution with the quantile regression modeling and applications. *Comp. Appl. Math, 40.*
- Korkmaz, M.Ç., Emrah, A., Chesneau, C., Yousof, H.M. (2022). On the Unit-Chen distribution with associated quantile regression and applications. *Math. Slovaca*, *72*, 765–786.
- M. E. Ghitany, J. Mazucheli, A. F. B. Menezes & F. Alqallaf (2019). The unit-inverse Gaussian distribution: A new alternative to two-parameter distributions on the unit interval, Communications in Statistics. *Theory and Methods, 48*(14), 3423-3438, DOI: 10.1080/03610926.2018.1476717.
- Mazucheli, J., Menezes, A.F., Dey, S. (2019). Unit-Gompertz distribution with applications. *Statistica*, *79*, 25–43.
- Mazucheli, J., Menezes, A.F.B., Fernandes, L.B., de Oliveira, R.P, Ghitany, M.E. (2020). The unit-Weibull distribution as an alternative to the Kumaraswamy distribution for the modeling of quantiles conditional on covariates. *J. Appl. Stat. 47*, 954–974
- Mazucheli, J., Menezes, A.F.B., Ghitany, M.E. (2018). The unit-Weibull distribution and associated inference. *J. Appl. Probably. Stat. 13*, 1–22
- Menezes, André & Mazucheli, Josmar & Dey, Sanku. (2018). The Unit-Logistic Distribution: Different Methods of Estimation. *Pesquisa Operacional Statistica*, *79*, 25–43.
- Nadarajah, S.; Haghighi, F. An extension of the exponential distribution. Statistics, 45, 543–558.
- Nagarjuna, V.B.V., Vardhan, R.V., and Chesneau, C. (2022). Nadarajah–Haghighi Lomax Distribution and Its Applications. *Mathematical and Computational Applications*, *27*(2), 30; https://doi.org/10.3390/mca27020030.
- Nascimento, Abraão D.C., Silva, Kássio F., Cordeiro, Gauss M., Alizadeh, Morad, Yousof, Haitham M., and Hamedani, G. G. (2-19). The Odd Nadarajah-Haghighi Family of Distributions: Properties and Applications. *Mathematical and Statistical Science Faculty Research and Publications. 95*. https://epublications.marquette.edu/math_fac/95

- Newer, H.A., Mohie El-Din, M.M., Ali, H.S., Al-Shbeil, I., and Emam, W. (2023). Statistical inference for the Nadarajah-Haghighi distribution based on ranked set sampling with applications. *AIMS Mathematics*, *8*(9), 21572-21590. doi: 10.3934/math.20231099.
- Pena-Ramirez, F. A., Guerra, R. R., & Cordeiro, G. M., (2019). The Nadarajah-Haghighi Lindley distribution. *Annals of the Brazilian Academy of Sciences 91*(1): e20170856. DOI 10.1590/10.1590/0001-3765201920170856.
- Wu, M. and Gui, W. (2021). Estimation and Prediction for Nadarajah-Haghighi Distribution under Progressive Type-II Censoring. *Symmetry*, *13*(6). https://doi.org/10.3390/sym13060999