Sains Malaysiana 52(2)(2023): 669-682
http://doi.org/10.17576/jsm-2023-5202-26

Statistical Properties and Estimation of the Three-Parameter Lindley Distribution with Application to COVID-19 Data

(Sifat Statistik dan Anggaran Taburan Lindley Tiga Parameter dengan Aplikasi pada Data COVID-19)

MATHIL KAMIL THAMER[1,2,*] & RAOUDHA ZINE[1]

[1]Laboratory of Probability and Statistics, Faculty of Sciences of Sfax, Sfax, Tunisia
[2]Department of Economics, College of Administration and Economics, University of Anbar, Iraq

Received: 14 July 2021/Accepted: 9 May 2022

ABSTRACT

In 2017, the three-parameter Lindley distribution has been studied. The present paper is a continuation of the investigation of the properties of this distribution because of its high flexibility for modeling lifetime data. We studied some statistical properties of this distribution as central tendency measures, dispersion measures, and shape measures. In addition, the mode and the quantile function of the distribution were derived by the authors. The three parameters were estimated by the Maximum Product of Spacing Method (MPS) due to its fame in estimating parameters. A simulation study is carried out to examine the consistency of estimators using mean square error (MSE). The estimators showed that they have the property of consistency because MSEs decrease with increasing the size of the sample. On the practical side, the MPS estimates were used to obtain statistical properties, probability density function (p.d.f), cumulative distribution function (c.d.f), survival function, and hazard function for real data which represents COVID-19 Data in Iraq/Al-Anbar Province. We found the flexibility of the distribution in representing life data and the possibility of getting the patients' probability of death and probability of survival for the time.

Keywords: COVID-19 data; mathematical model; maximum product of spacing method; three-parameter Lindley distribution

ABSTRAK

Pada tahun 2017, taburan Lindley tiga parameter telah dikaji. Makalah ini adalah kesinambungan daripada penyelidikan sifat pengedaran ini kerana kefleksibelannya yang tinggi untuk memodelkan data sepanjang hayat. Kami mengkaji beberapa sifat statistik taburan ini sebagai ukuran kecenderungan pusat, ukuran penyebaran dan ukuran bentuk. Di samping itu, mod dan fungsi kuantil taburan diperoleh oleh penulis. Ketiga-tiga parameter tersebut dianggarkan menggunakan Kaedah Maksimum Jarak Jauh (MPS) kerana kemasyhurannya dalam menganggar parameter. Suatu kajian simulasi dijalankan untuk mengkaji ketekalan penganggar menggunakan min ralat kuasa dua (MSE). Penganggar menunjukkan bahawa mereka memiliki sifat ketekalan kerana MSE menurun dengan peningkatan ukuran sampel. Dari segi praktikal, anggaran MPS digunakan untuk memperoleh sifat statistik, fungsi ketumpatan kebarangkalian (p.d.f), fungsi taburan kumulatif (c.d.f), fungsi kemandirian dan fungsi bahaya untuk data sebenar yang mewakili Data COVID-19 di Wilayah Iraq/Al-Anbar. Kami mendapati kefleksibelan penyebaran dalam mewakili data kehidupan dan kemungkinan mendapat kebarangkalian kematian pesakit dan kebarangkalian bertahan untuk masa ini.

Kata kunci: Data COVID-19; model matematik; produk maksimum kaedah jarak; taburan Lindley tiga parameter

INTRODUCTION

Survival analysis has become a very important statistical technique since the beginning of the twentieth century, and this theory relies on non-negative probability distributions that are often right-skewed distribution as Weibull, Rayleigh, Log-Normal, and many other probability distributions. Because most of the phenomena in real life are a mixture of several distributions, the so-called mixture distributions appeared, which means that a part of the population follows a distribution and the remaining follow others, and the most common mixed distributions are the t-Student, Laplace, Lomax distribution, and Lindley distribution.

Although the Lindley distribution was pioneered in 1958 by Lindley, but real interest began in 2008 when Ghitany, Atieh and Nadarajah studied its properties and applications, since then this distribution has begun to be developed, as generalized Lindley distribution by Zakerzadeh and Dolati (2009), a two-parameter weighted Lindley distribution by Ghitany et al. (2011), a discrete Lindley distribution by Déniz and Ojeda (2011), an extended Lindley distribution by Bakouch et al. (2012), a quasi Lindley distribution and a two-parameter Lindley distribution by Shanker and Mishra (2013a, 2013b), another two-parameter Lindley distribution by Shanker, Sharma and Shanker (2013), the truncated Lindley distribution by Singh, Singh and Sharma (2014), the Log-Lindley distribution by Déniz, Sordo and Ojeda (2014), Lindley distribution with a location parameter as a three-parameter distribution by Abd El-Monsef (2015), the inverse Lindley distribution by Sharma et al. (2015), a three-parameter weighted Lindley distribution by Shanker, Kamlesh and Fesshaye (2017), and another three-parameter Lindley distribution by Shanker et al. (2017), the unit-Lindley distribution by Mazucheli, Menezes and Chakraborty (2018), and the double Lindley distribution by Kumar and Jose (2019).

The Lindley distribution is a mixture of Exponential distribution (θ) with Gamma (2, θ) as given:

$$f(x;\theta) = wf_1(x;\theta) + (1-w)f_2(x;\theta),$$

where $f_1(x;\theta) = \theta e^{-\theta x}$ and $f_2(x;\theta) = \theta^2 x e^{-\theta x}$.

The respective probability density function (p.d.f) and cumulative distribution function (c.d.f) of the one-parameter Lindley distribution are given by:

$$f(x;\theta) = \frac{\theta^2}{\theta+1}(1+x)e^{-\theta x}\,; x,\theta > 0,$$

$$F(x;\theta) = 1 - \left[\frac{\theta+1+\theta x}{\theta+1}\right]e^{-\theta x}\,; x,\theta > 0.$$

Statistical properties and estimation of different Lindley distributions were studied by many authors as the quantile function by Jodrá (2010), reliability estimation in Lindley distribution with censored sample using Bayes method by Krishna and Kumar (2011), applications the two-parameter Lindley distribution by Shanker, Fesshaye and Sharma (2016), and estimation of the reliability function and the hazard function of the two-parameter Lindley distribution with fuzzy data by Al-Bayati (2018), estimation the probability density function and the cumulative distribution function of the one-parameter Lindley distribution by Maiti and Mukherjee (2018), but the three-parameter Lindley distribution has not been studied yet, so in this paper, we will derive the mode and the quantile function of the distribution, discuss some of the other statistical properties, and estimate the parameters by the MPS method with an application.

MATERIALS AND METHODS

THE THREE-PARAMETER LINDLEY DISTRIBUTION AND SURVIVAL MEASURES

The p.d.f of the three-parameter Lindley distribution (THPLD) introduced by Shanker et al. (2017), having parameters θ, α and β is given by:

$$f(x;\theta,\alpha,\beta) = \frac{\theta^2}{\alpha\theta+\beta}(\alpha+\beta x)e^{-\theta x};\ x,\theta,\beta > 0 \text{ and}$$

$$\alpha\theta+\beta > 0. \tag{1}$$

The corresponding c.d.f of (1) is given by:

$$F(x;\theta,\alpha,\theta) = 1 - \left[\frac{\alpha\theta+\beta+\beta\theta x}{\alpha\theta+\beta}\right]e^{-\theta x}\,;\ x,\theta,\beta > 0 \text{ and}$$

$$\alpha\theta+\beta > 0. \tag{2}$$

Special cases of the three-parameter Lindley distribution were shown in Table 1.

The quantile function of the three-parameter Lindley distribution is given by:

$$F(x;\theta,\alpha,\beta) = 1 - \left[\frac{\alpha\theta+\beta+\beta\theta x}{\alpha\theta+\beta}\right]e^{-\theta x} = u$$

$$\Rightarrow 1-u = \left[\frac{\alpha\theta+\beta+\beta\theta x}{\alpha\theta+\beta}\right]e^{-\theta x}$$

$$\Rightarrow (\alpha\theta+\beta)(1-u) = [\alpha\theta+\beta+\beta\theta x]e^{-\theta x}$$

$$\Rightarrow \frac{(\alpha\theta+\beta)(1-u)}{\beta} = \left[\frac{\alpha\theta+\beta}{\beta}+\theta x\right]e^{-\theta x}$$

$$\Rightarrow -\frac{(\alpha\theta+\beta)(1-u)e^{-(\alpha\theta+\beta)/\beta}}{\beta} = -\left[\frac{\alpha\theta+\beta}{\beta}+\theta x\right]e^{-(\theta x+(\alpha\theta+\beta)/\beta)}$$

$$\Rightarrow W_{-1}\left[-\frac{(\alpha\theta+\beta)(1-u)e^{-(\alpha\theta+\beta)/\beta}}{\beta}\right] = -\left[\frac{\alpha\theta+\beta}{\beta}+\theta x\right]$$

$$\therefore x = Q(u;\theta,\alpha,\beta) = -\frac{\alpha}{\beta}-\frac{1}{\theta}-\frac{1}{\theta}W_{-1}\left[-\frac{(\alpha\theta+\beta)(1-u)e^{-(\alpha\theta+\beta)/\beta}}{\beta}\right];0<u<1. \quad (3)$$

where W_{-1} (.) denotes the negative branch of the Lambert W function.

TABLE 1. Special cases of the THPL distribution

Name of the distribution	Authors	***θ***	***α***	***β***
Three-Parameter Lindley	Shanker et al. (2017)	θ	α	β
Two-Parameter Lindley	Shanker, Kamlesh & Fesshaye (2017)	θ	α	1
Two-Parameter Lindley	Shanker & Mishra (2013b)	θ	1	α
One-Parameter Lindley	Lindley (1958)	θ	1	1
Gamma		θ	1	0
Exponential		θ	0	1

The survival function of the three-parameter Lindley distribution is given by:

$$S(t;\theta,\alpha,\beta) = 1-F(t;\theta,\alpha,\beta) = \left[\frac{\alpha\theta+\beta+\beta\theta t}{\alpha\theta+\beta}\right]e^{-\theta t}; \quad (4)$$

$$t,\theta,\beta>0 \; and \;\; \alpha\theta+\beta>0.$$

The hazard function of the three-parameter Lindley distribution is as follows:

$$h(t;\theta,\alpha,\beta) = \frac{f(t;\theta,\alpha,\beta)}{S(t;\theta,\alpha,\beta)} = \frac{\theta^2(\alpha+\beta t)}{\alpha\theta+\beta+\beta\theta t}; \quad (5)$$

$$t,\theta,\beta>0 \; and \;\; \alpha\theta+\beta>0.$$

The mean time to failure (MTTF) of the three-parameter Lindley distribution can be written as:

$$MTTF = \int_0^\infty S(t;\theta,\alpha,\beta)\,dt = \frac{(\alpha\theta+2\beta)}{(\alpha\theta+\beta)\theta}. \quad (6)$$

Figures 1-4 show the p.d.f, c.d.f, survival function, and hazard function for different values of the parameters.

STATISTICAL PROPERTIES AND RELATED MEASURES OF THPLD

Non-central moments about zero of the three-parameter Lindley distribution can be obtained as follows (Shanker et al. 2017):

$$\mu_r' = E[X^r] = \frac{\mathrm{r}!\,(\alpha\theta+(\mathrm{r}+1)\beta)}{(\alpha\theta+\beta)\theta^r}.$$

so, the first four moments of the three-parameter Lindley distribution about zero are:

$$\mu_1' = \frac{(\alpha\theta+2\beta)}{(\alpha\theta+\beta)\theta} \;\; \mu_2' = \frac{2(\alpha\theta+3\beta)}{(\alpha\theta+\beta)\theta^2}$$

$$\mu_3' = \frac{6(\alpha\theta+4\beta)}{(\alpha\theta+\beta)\theta^3} \;\; \mu_4' = \frac{24(\alpha\theta+5\beta)}{(\alpha\theta+\beta)\theta^4}$$

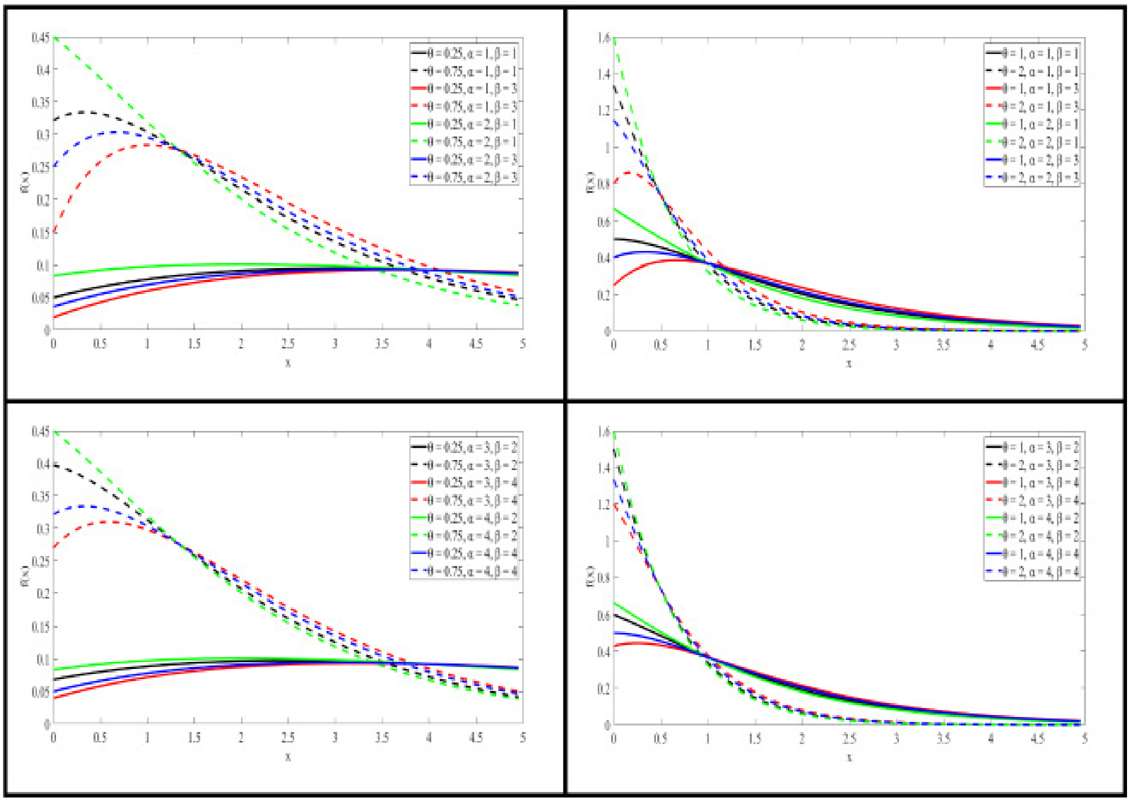


FIGURE 1. The p.d.f of the THPL distribution

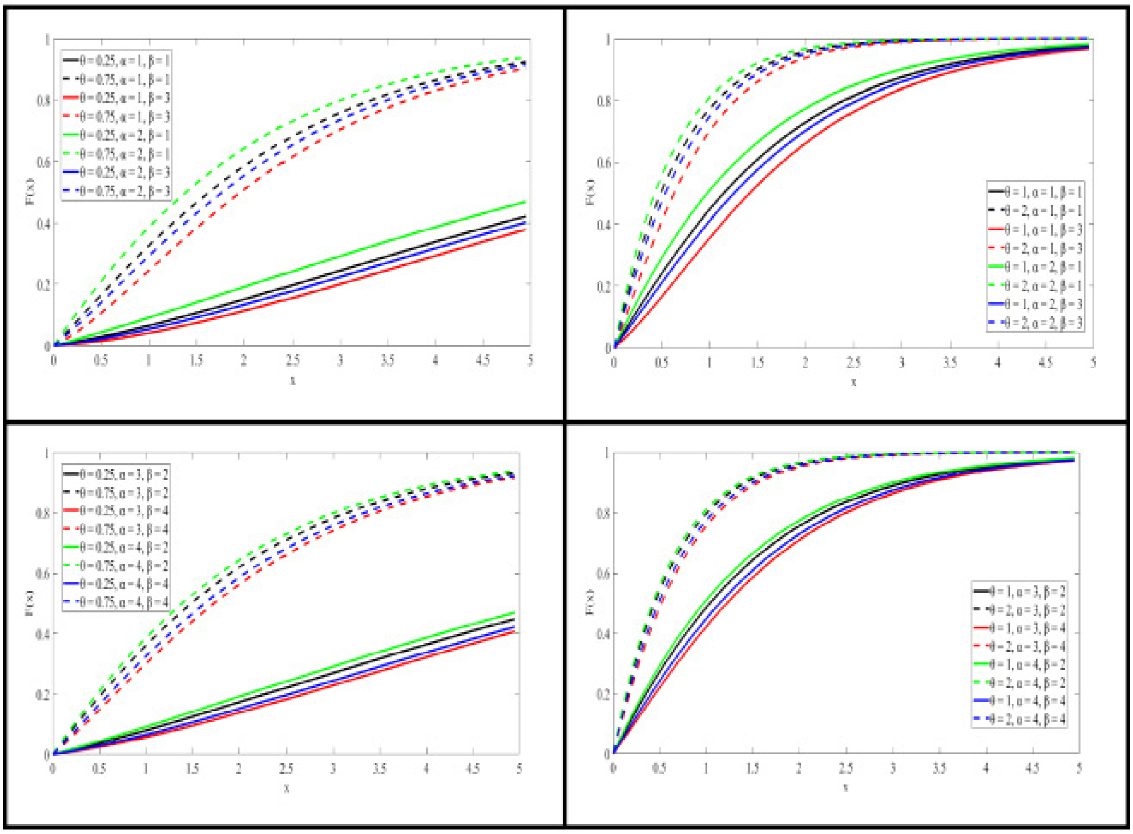


FIGURE 2. The c.d.f of the THPL distribution

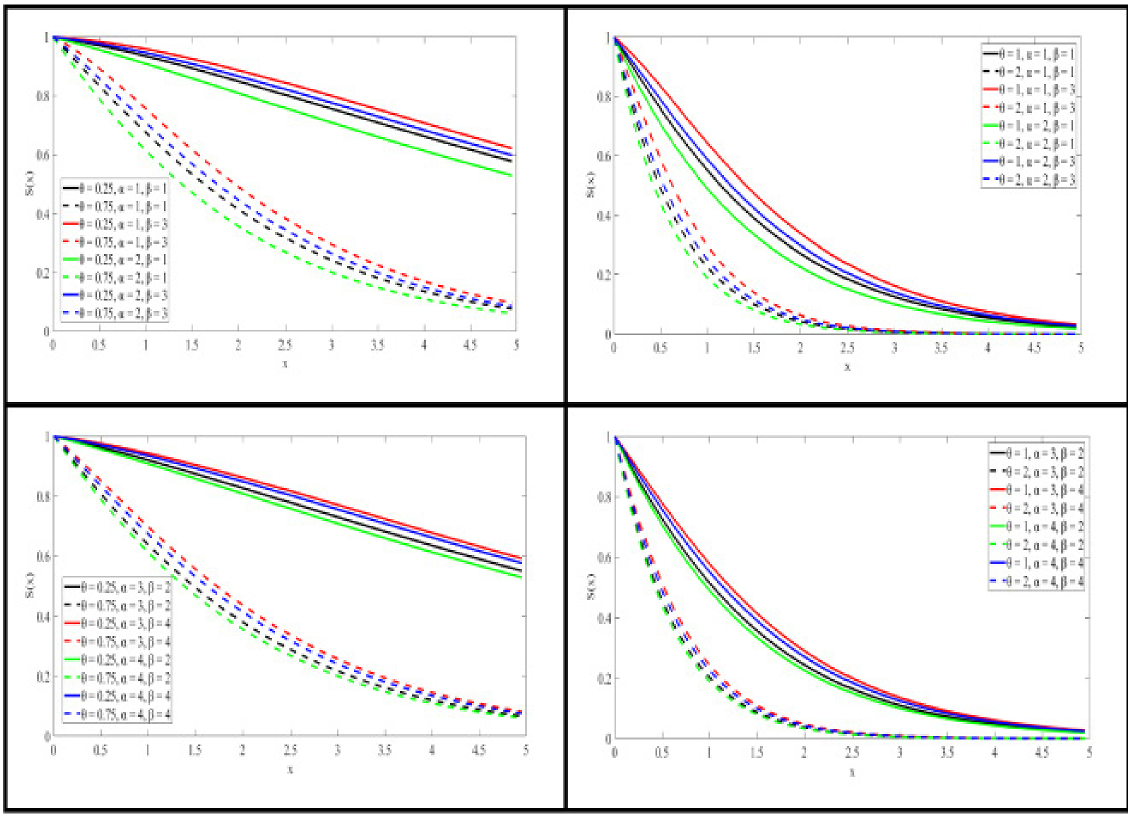


FIGURE 3. The survival function of the THPL distribution

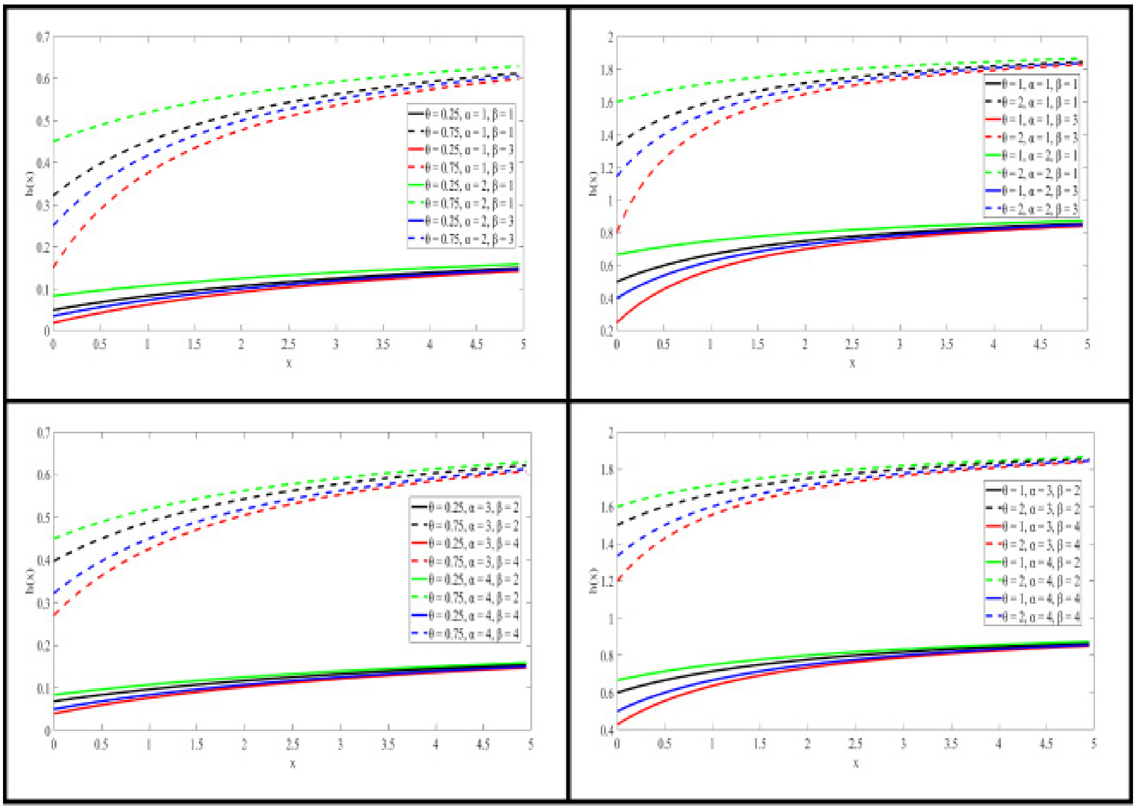


FIGURE 4. The hazard function of the THPL distribution

whereas r^{th} central moments of the three-parameter Lindley distribution are obtained as (Shanker et al. 2017):

$$\mu_r = E[(X-\mu)^r] = \sum_{i=0}^{r} C_i^r (-\mu)^{r-i} \mu_i'.$$

hence, the second, third and fourth central moments are:

$$\mu_2 = \frac{\alpha^2\theta^2 + 4\alpha\theta\beta + 2\beta^2}{(\alpha\theta+\beta)^2\theta^2}$$

$$\mu_3 = \frac{2(\alpha^3\theta^3 + 6\alpha^2\theta^2\beta + 6\alpha\theta\beta^2 + 2\beta^3)}{(\alpha\theta+\beta)^3\theta^3}$$

$$\mu_4 = \frac{3(3\alpha^4\theta^4 + 24\alpha^3\theta^3\beta + 44\alpha^2\theta^2\beta^2 + 32\alpha\theta\beta^3 + 8\beta^4)}{(\alpha\theta+\beta)^4\theta^4}$$

In statistics, the three most common measures of central tendency are the mean, median, and mode. For the three-parameter Lindley distribution, the mean and the median for the three-parameter Lindley distribution, respectively, can be obtained as (Shanker et al. 2017):

$$\mu = \frac{(\alpha\theta+2\beta)}{(\alpha\theta+\beta)\theta}$$

$$Me = F^{-1}(0.5) = Q(0.5)$$

In addition to the mode which we can get it as follows:

$$Mo = \underset{x}{Argmax}[f(x)]$$

$$f(x) = \frac{\theta^2}{\alpha\theta+\beta}(\alpha+\beta x)e^{-\theta x}$$

$$f'(x) = \frac{\theta^2}{\alpha\theta+\beta}\left[-\theta(\alpha+\beta x)e^{-\theta x} + \beta e^{-\theta x}\right] = 0$$

$$\Rightarrow \theta(\alpha+\beta x) = \beta \Rightarrow x = \frac{\beta-\alpha\theta}{\theta\beta}$$

$$\therefore Mo = \begin{cases} \dfrac{\beta-\alpha\theta}{\theta\beta}, & |\alpha\theta| < \beta; \\ 0, & otherwise. \end{cases}$$

The variance, standard deviation, coefficient of variation, and coefficient of dispersion for the three-parameter Lindley distribution are obtained as (Shanker et al. 2017):

$$\sigma^2 = E[(X-\mu)^2] = \frac{\alpha^2\theta^2 + 4\alpha\beta\theta + 2\beta^2}{(\alpha\theta+\beta)^2\theta^2}$$

$$\sigma = \sqrt{E[(X-\mu)^2]} = \frac{\sqrt{\alpha^2\theta^2 + 4\alpha\beta\theta + 2\beta^2}}{(\alpha\theta+\beta)\theta}$$

$$CV = \frac{\sigma}{\mu} = \frac{\sqrt{\alpha^2\theta^2 + 4\alpha\beta\theta + 2\beta^2}}{(\alpha\theta+2\beta)}$$

$$\gamma = \frac{\sigma^2}{\mu} = \frac{\alpha^2\theta^2 + 4\alpha\beta\theta + 2\beta^2}{(\alpha\theta+2\beta)(\alpha\theta+\beta)\theta}$$

Measures of shape give an idea of the distribution of the data such as skewness and kurtosis. The coefficient of skewness $\sqrt{\beta_1}$ and coefficient of kurtosis β_2 for the three-parameter Lindley distribution, respectively, are (Shanker et al. 2017):

$$\sqrt{\beta_1} = \frac{\mu_3}{\sigma^3} = \frac{2(\alpha^3\theta^3 + 6\alpha^2\theta^2\beta + 6\alpha\theta\beta^2 + 2\beta^3)}{(\alpha^2\theta^2 + 4\alpha\theta\beta + 2\beta^2)^{3/2}}$$

$$\beta_2 = \frac{\mu_4}{\sigma^4} = \frac{3(3\alpha^4\theta^4 + 24\alpha^3\theta^3\beta + 44\alpha^2\theta^2\beta^2 + 32\alpha\theta\beta^3 + 8\beta^4)}{(\alpha^2\theta^2 + 4\alpha\theta\beta + 2\beta^2)^2}$$

coefficients of variation, dispersion, skewness, and kurtosis of the three-parameter Lindley distribution are shown in Figures 5-8.

MPS ESTIMATORS OF THE THPLD PARAMETERS

In this paper, the Maximum Product of Spacing Method (MPS) was chosen for two reasons.

First, this method has not been used to estimate parameters of the three-parameter Lindley distribution. Second, this method is recommended by many researchers to estimate parameters of distributions such as Afify et al. (2020), Al-Mofleh, Afify and Ibrahim (2020), Arslan, Acitas & Senoglu (2017), Basu, Singh and Singh (2017), Do Espirito Santo and Mazucheli (2015), Kantar and Şenoğlu (2008), and Nassar et al. (2018).

MPS was derived by Cheng and Amin (1979), the idea of this method is to maximize the following function:

$$P = \frac{1}{n+1}\sum_{i=1}^{n+1}\log D_i \quad ; where\ D_i =$$

$$F\left(x_{(i)}\right) - F\left(x_{(i-1)}\right)\ and\ F\left(x_{(0)}\right) = 0; F\left(x_{(n+1)}\right) = 1$$

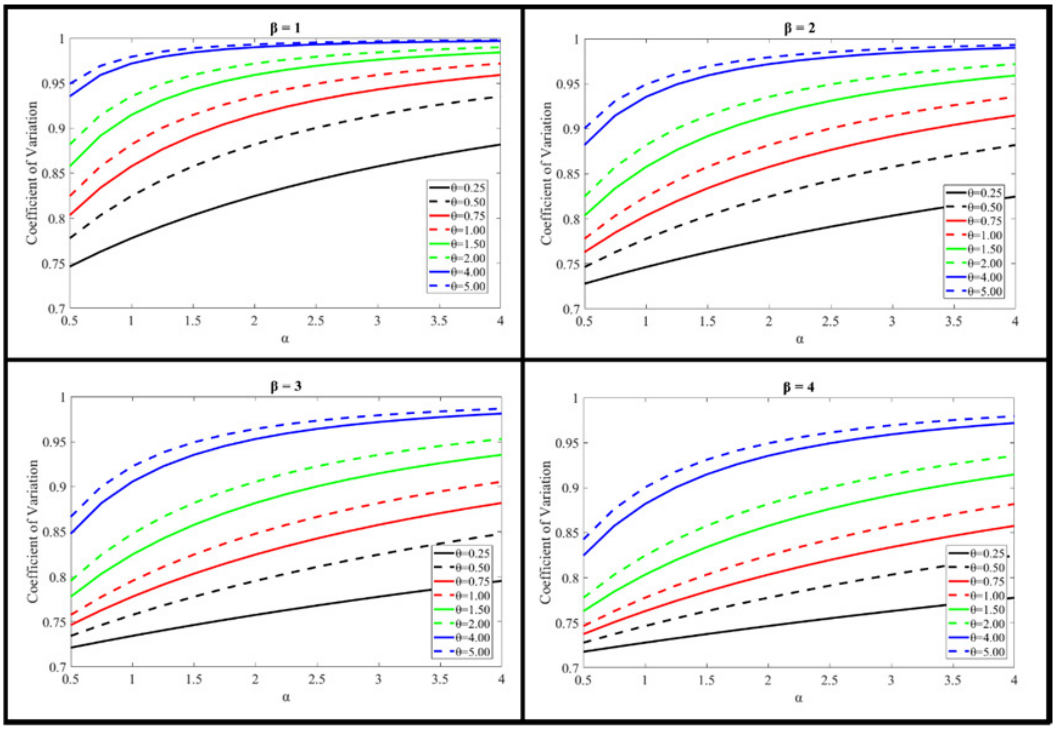


FIGURE 5. Coefficient of variation of the THPL distribution

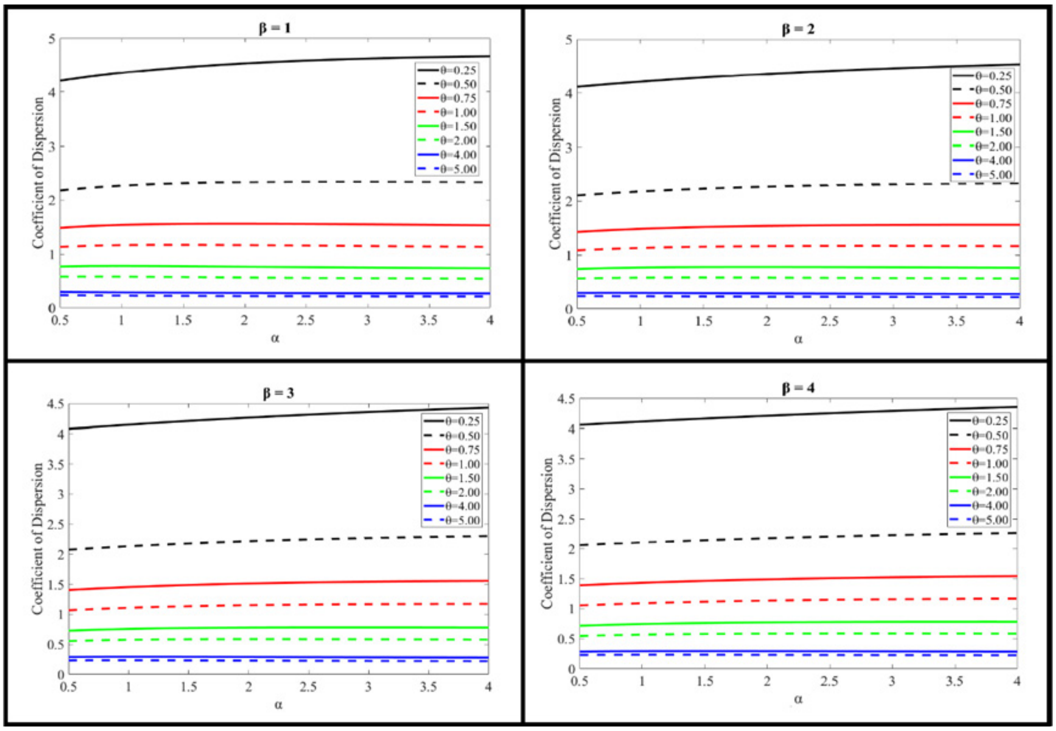


FIGURE 6. Coefficient of dispersion of the THPL distribution

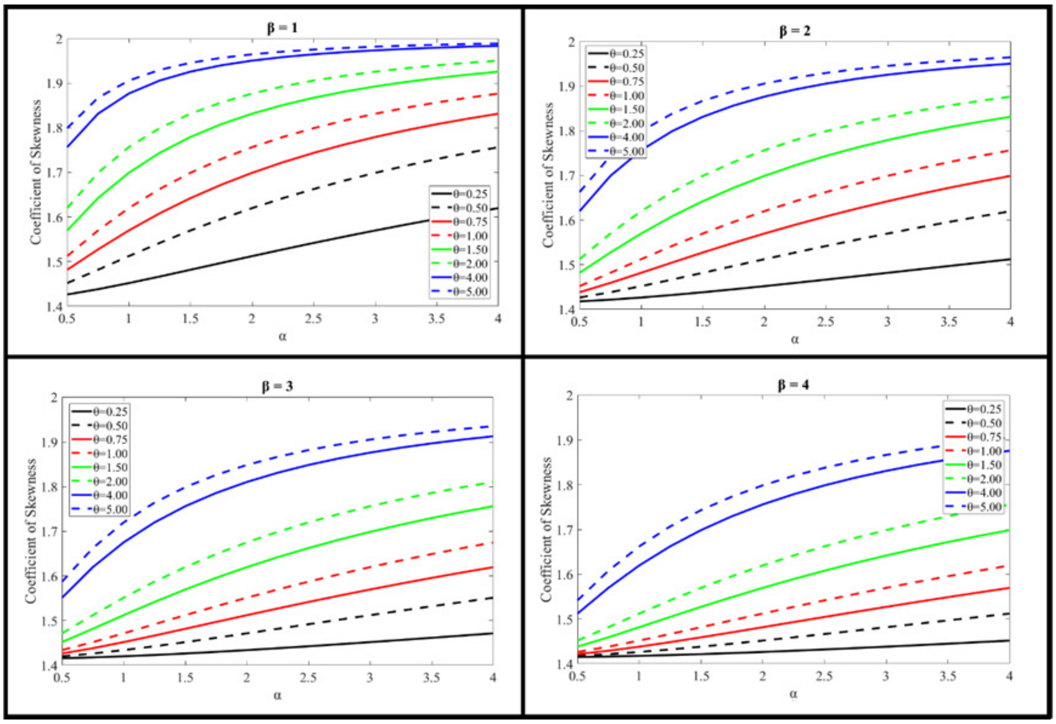


FIGURE 7. Coefficient of skewness of the THPL distribution

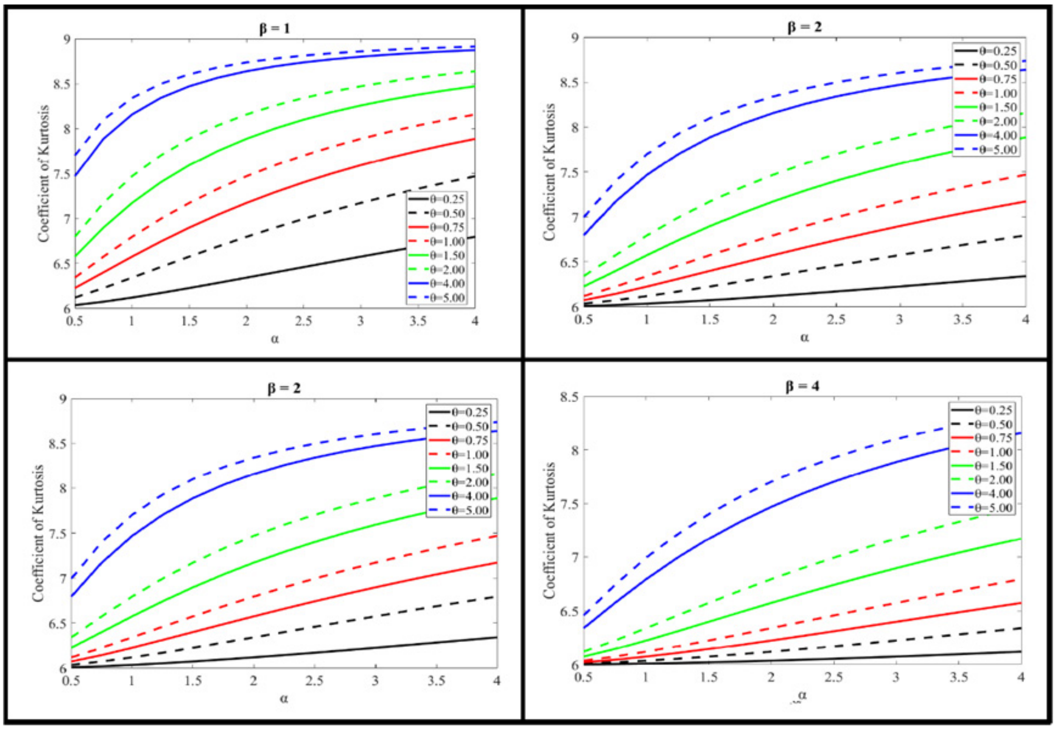


FIGURE 8. Coefficient of kurtosis of the THPL distribution

This function can be obtained for the three-parameter Lindley distribution after substituting for $F(x_{(i)})$ in the previous equation by its c.d.f defined in Equation (2) as follows:

$$P(\theta,\alpha,\beta)=\frac{1}{n+1}\sum_{i=1}^{n+1} log\ [Q_i]-log(\alpha\theta+\beta) \quad (7)$$

where

$$Q_i=\alpha\theta\left[e^{-\theta x_{(i-1)}}-e^{-\theta x_{(i)}}\right]+\beta\left[e^{-\theta x_{(i-1)}}-e^{-\theta x_{(i)}}\right]+$$
$$\beta\theta\left[x_{(i-1)}e^{-\theta x_{(i-1)}}-x_{(i)}e^{-\theta x_{(i)}}\right]$$

We can find MPS estimators by maximizing (10) via fminunc function or by solving the following equations:

$$\frac{\partial P(\theta,\alpha,\beta)}{\partial\theta}=\frac{1}{n+1}\sum_{i=1}^{n}\frac{1}{Q_i}\frac{\partial Q_i}{\partial\theta}-\frac{\alpha}{\alpha\theta+\beta}=0$$

$$\frac{\partial P(\theta,\alpha,\beta)}{\partial\alpha}=\frac{1}{n+1}\sum_{i=1}^{n}\frac{1}{Q_i}\frac{\partial Q_i}{\partial\alpha}-\frac{\theta}{\alpha\theta+\beta}=0$$

$$\frac{\partial P(\theta,\alpha,\beta)}{\partial\beta}=\frac{1}{n+1}\sum_{i=1}^{n}\frac{1}{Q_i}\frac{\partial Q_i}{\partial\beta}-\frac{1}{\alpha\theta+\beta}=0$$

where

$$\frac{\partial Q_i}{\partial\theta}=\alpha\left[\left(e^{-\theta x_{(i-1)}}-e^{-\theta x_{(i)}}\right)-\right.$$
$$\left.\theta\left(x_{(i-1)}e^{-\theta x_{(i-1)}}-x_{(i)}e^{-\theta x_{(i)}}\right)\right]$$
$$+\beta\left[\left(x_{(i-1)}e^{-\theta x_{(i-1)}}-x_{(i)}e^{-\theta x_{(i)}}\right)-\right.$$
$$\left.\theta\left(x_{(i-1)}^2e^{-\theta x_{(i-1)}}-x_{(i)}^2e^{-\theta x_{(i)}}\right)\right]$$

$$\frac{\partial Q_i}{\partial\alpha}=\theta\left(e^{-\theta x_{(i-1)}}-e^{-\theta x_{(i)}}\right)$$

$$\frac{\partial Q_i}{\partial\beta}=\left(e^{-\theta x_{(i-1)}}-e^{-\theta x_{(i)}}\right)+\theta\left(x_{(i-1)}e^{-\theta x_{(i-1)}}-x_{(i)}e^{-\theta x_{(i)}}\right)$$

Note that if there is a tie, we cannot find the natural logarithm of D_i for the corresponding observation, so we replace D_i with the observation's p.d.f, that is $D_i = f(x_{(i)})$.

Results and Discussion

SIMULATION

To test the consistency property of MPS estimators, data were generated from the three-parameter Lindley distribution based on the quantile function defined in Equation (3), which was derived by the researcher for this. Generation was performed for four different cases, as shown in Table 2, for each case, various sizes of samples were used (10, 30, 60, 80, 150, 250). The experiment will be repeated 10,000 times for each of the combinations, then the parameters will be estimated by the MPS method, and find the values Mean Square Error (MSE), which its formula is:

$$\frac{1}{10000}\sum_{j=1}^{10000}\left(\hat{\varphi}_j-\varphi\right)^2,$$

where $\varphi=\left(\hat{\theta},\ \hat{\alpha},\ \hat{\beta}\right)$ and $\hat{\varphi}=(\theta,\ \alpha,\ \beta)$. It is worth noting that all operations in this paper were done in Matlab 2020a.

TABLE 2. The cases of simulation

Cases	θ	α	β
Case 1	0.25	1	2
Case 2	0.50	2	2
Case 3	0.75	2	1
Case 4	1.00	3	2

TABLE 3. The simulated MSE values for the four cases

		Case 1	Case 2	Case 3	Case 4
n=10	θ	0.00842	0.03684	0.09316	0.14954
	α	7.77936	5.03375	0.36515	1.78596
	β	1.97869	5.00455	1.43334	4.13198
n=30	θ	0.0031	0.01557	0.04316	0.08264
	α	3.61839	1.55848	0.28803	1.49819
	β	0.91062	1.57176	1.12452	3.42073
n=60	θ	0.00116	0.00959	0.029	0.05632
	α	1.66567	1.04832	0.22678	1.12028
	β	0.41808	1.05479	0.89143	2.54612
n=80	θ	0.00116	0.00681	0.02502	0.04442
	α	1.66567	0.80042	0.18003	0.96882
	β	0.41808	0.8045	0.70862	2.20537
n=150	θ	0.00116	0.00681	0.02502	0.04442
	α	1.66567	0.80042	0.18003	0.96882
	β	0.41808	0.8045	0.70862	2.20537
n=250	θ	0.00038	0.00321	0.01717	0.03321
	α	0.57175	0.47251	0.12917	0.72412
	β	0.14333	0.47378	0.50980	1.64596

Tables 3 illustrates our simulation study. These results show that all the estimators have the property of consistency. Because MSEs decrease with increasing sample size.

APPLICATION

In this section, firstly we will verify the suitability and the high flexibility of the three-parameter Lindley distribution for real data by using the estimators in modelization, validation, and prediction. Secondly, we will try to give a clear vision of COVID-19 deaths probabilities to the medical sector in Iraq/Al-Anbar Province by taking a sample of COVID-19 patients, size 83, from the medical records of quarantine hospitals in the province by the researchers. Table 4 contains survival time to death for the patients.

TABLE 4. Survival time for 83 patients (in days)

1	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2	2	3
3	3	3	3	3	3	4	4	4	4	4	4	4	5	5	5	5	5
6	6	6	6	6	7	7	7	7	7	7	8	8	8	8	8	9	9
9	10	10	10	10	10	10	10	11	11	11	12	12	12	12	12	12	13
13	13	18	19	20	20	21	22	22	22	23							

The parameters of the three-parameter Lindley distribution for the COVID-19 data were estimated by the MPS method. The estimates were ($\hat{\theta} = 0.23503, \hat{\alpha} = 2.41325, \hat{\beta} = 2.87624$), and the Kolmogorov Smirnov (K-S) addition to its associated p-values were (K-S=0.07517, p-value = 0.70769), thus we can say the data follow the three-parameter Lindley distribution. Hence the MPS estimates will be used in estimating some statistical properties of these data as shown in Table 5.

TABLE 5. Measures of central tendency, dispersion, and shape for the THPL distribution

N	Min	Max	μ	Me	Mo	σ^2	σ	γ	CV	$\sqrt{\beta_1}$	β_2
83	1	23	7.809	6.418	3.416	35.715	5.976	4.574	0.765	1.440	6.082

Mean time to failure mentioned in equation (6) was estimated as $MTTF = 7.809$, which means the average survival time until death is almost 8 days. The remaining of this section is structured in three steps as follows:

Step 1: 'Modelization'
In this step, we will construct models for the p.d.f, c.d.f, survival function, and hazard function of the three-parameter Lindley distribution based on the MPS estimates as the following:

$$\hat{f}(t) = 0.01604(2.41325 + 2.87624t)\, e^{-0.23503t}\,;\ t > 0, \quad (8)$$

$$\hat{F}(t) = 1 - \left[\frac{3.44343 + 0.676t}{3.44343}\right] e^{-0.23503t}\,;\ t > 0, \quad (9)$$

$$\hat{S}(t) = \left[\frac{3.44343 + 0.676t}{3.44343}\right] e^{-0.23503t}\,;\ t > 0, \quad (10)$$

$$\hat{h}(t) = \frac{0.05524\,(2.41325 + 2.87624t)}{3.44343 + 0.676t}\,;\ t > 0. \quad (11)$$

According to Equations 8-11, the p.d.f, c.d.f, survival function, and hazard function will estimate as shown in Table 6.

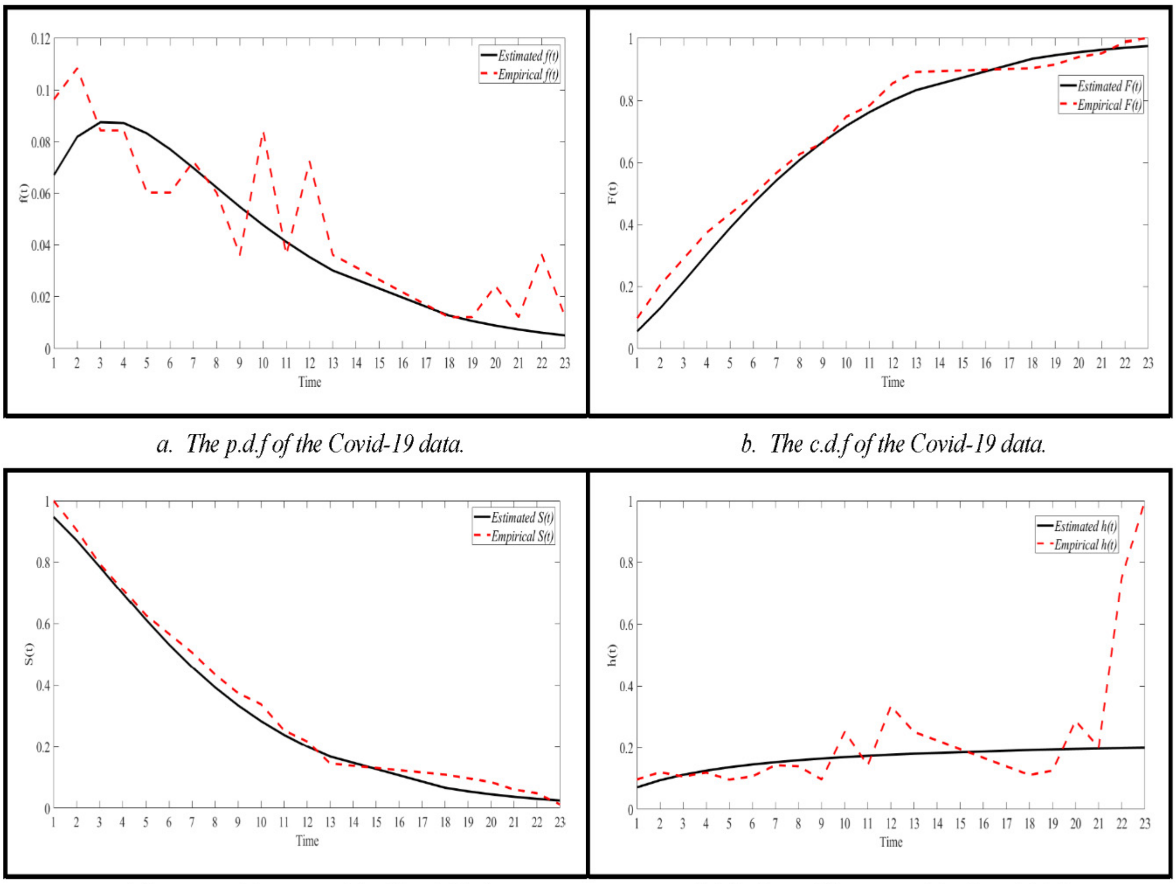
TABLE 6. The p.d.f, c.d.f, survival function, and hazard function estimators for the COVID-19 data

Time	$\hat{f}(t)$	$\hat{F}(t)$	$\hat{S}(t)$	$\hat{h}(t)$
1	0.06707	0.05424	0.94576	0.07091
2	0.08185	0.12963	0.87037	0.09404
3	0.08750	0.21492	0.78508	0.11146
4	0.08720	0.30266	0.69734	0.12504
5	0.08318	0.38808	0.61192	0.13593
6	0.07702	0.46831	0.53169	0.14486
7	0.06980	0.54177	0.45823	0.15232
8	0.06222	0.60779	0.39221	0.15863
9	0.05475	0.66625	0.33375	0.16405
10	0.04768	0.71742	0.28258	0.16875
11	0.04118	0.76180	0.23820	0.17287
12	0.03530	0.79999	0.20001	0.17650
13	0.03008	0.83263	0.16737	0.17973
18	0.01265	0.93403	0.06597	0.19170
19	0.01053	0.94559	0.05441	0.19350
20	0.00874	0.95520	0.04480	0.19515
21	0.00724	0.96317	0.03683	0.19668
22	0.00599	0.96977	0.03023	0.19809
23	0.00494	0.97522	0.02478	0.19941

From Table 6, we note that the probability of death on the first day is about 7%, on the second day is about 8%, on the third day is about 9%, ..., about 0.5% on the 23^{rd} day. That means that the p.d.f is not a monotonous function. The cumulative probability of death to the first day is about 5%, to the second day about 13%, to the third day about 21%, …, about 98% to the 23^{rd} day. So, the c.d.f is an increasing function according to its values. We see that the probability of survival for one day is about 95%, for two days is about 87%, for three days is about 79%, for four days is about 78%, and so on. That means that the number of those who died has increased and thus the survival function decreased and became close to 3% on the 23^{rd} day. This indicates that the survival function is inversely proportional to time. The rate of people who are at risk on the first day is about 7%, on the second day about 9%, on the third day about 11%, …, about 20% on the 23^{rd} day, hence the hazard function is directly proportional to time, so the survival function is inversely proportional to the hazard function.

Step 2: ‘Validation’

The estimated p.d.f, c.d.f, survival function, and hazard function based on the three-parameter Lindley distribution and their corresponding empirical functions for the COVID-19 data were illustrated in Figure 9.



a. The p.d.f of the Covid-19 data.

b. The c.d.f of the Covid-19 data.

c. The survival function of the Covid-19 data.

d. The hazard function of the Covid-19 data.

FIGURE 9. The p.d.f, c.d.f, survival function, and hazard function of the COVID-19 real dat

When we see Figure 9, we note that the behavior of the estimated functions is somewhat close to the behavior of the empirical functions, and this is a good indication that the estimated models can represent the COVID-19 data.

Step 3: ‘Prediction’

Based on the previous steps, we can use the models to predict the probabilities of death and survival for days in which we did not have information. Table 7 shows that the predicted values of the p.d.f, c.d.f, survival function, and hazard function.

TABLE 7. The predicted p.d.f, c.d.f, survival function, and hazard function for the COVID-19 data

Time	$\hat{f}(t)$	$\hat{F}(t)$	$\hat{S}(t)$	$\hat{h}(t)$
14	0.02550	0.86037	0.13963	0.18263
15	0.02152	0.88383	0.11617	0.18523
16	0.01809	0.90359	0.09641	0.18759
17	0.01515	0.92017	0.07983	0.18974
24	0.00407	0.97971	0.02029	0.20063
25	0.00335	0.98341	0.01659	0.20177
26	0.00275	0.98645	0.01355	0.20284
27	0.00225	0.98894	0.01106	0.20384
28	0.00185	0.99098	0.00902	0.20478
29	0.00151	0.99266	0.00734	0.20567
30	0.00123	0.99402	0.00598	0.20651

CONCLUSION

In this paper, the statistical properties of the three-parameter Lindley distribution (THPLD) were studied, the mode and the quantile function were derived. the three parameters were estimated by the MPS method, and the property of consistency was proven of the three estimators based on a simulation study, because the MSEs decline when the sample size is raising.

In the application section, the COVID-19 data is distributed as the three-parameter Lindley distribution according to the K-S value. the MPS estimates are used to find some statistical properties, the p.d.f, c.d.f, survival function, hazard function, and mean time to failure. The mean time to failure for the data under study is approximately 8 days, p.d.f is a non-monotonous function, the c.d.f is an increasing function, the survival function is a decreasing function and the hazard function for it is an increasing function. Models were constructed based on the data and their validation was verified, then used for prediction. The limitation is that the distribution can take more time than others in estimation the parameters.

ACKNOWLEDGEMENTS

The authors are grateful to Adnan M. Hussein for his programming contribution.

REFERENCES

Abd El-Monsef, M.M.E. 2015. A new Lindley distribution with location parameter. *Communications in Statistics-Theory and Methods* 45(17): 5204-5219.

Afify, A.Z., Nassar, M., Cordeiro, G.M. & Kumar, D. 2020. The Weibull Marshall-Olkin Lindley distribution: Properties and estimation. *Journal of Taibah University for Science* 14(1): 192-204.

Al-Bayati, R.S.S. 2018. Some estimation methods for lindley distribution. A Master Thesis, College of Science, Mustansiriyah University, Baghdad, Iraq (Unpublished).

Al-Mofleh, H., Afify, A.Z. & Ibrahim, N.A. 2020. A new extended two-parameter distribution: Properties, estimation methods, and applications in medicine and geology. *Mathematics* 8(9): 1578.

Arslan, T., Acitas, S. & Senoglu, B. 2017. Generalized Lindley and power Lindley distributions for modeling the wind speed data. *Energy Conversion and Management* 152: 300-311.

Bakouch, H.S., Al-Zahrani, B.M., Al-Shomrani, A.A., Marchi, V.A. & Louzada, F. 2012. An extended Lindley distribution. *Journal of the Korean Statistical Society* 41: 75-85.

Basu, S., Singh, S.K. & Singh, U. 2017. Parameter estimation of inverse Lindley distribution for Type-I censored data. *Computational Statistics* 32(1): 367-385.

Cheng, R.C.H. & Amin, N.A.K. 1979. Maximum product-of-spacings estimation with application to the lognormal distribution. University of Wales IST, *Math. Report.* p. 79-1.

Déniz, G.E. & Ojeda, C.E. 2011. The discrete Lindley distribution: Properties and applications. *Journal of Statistical Computation and Simulation* 81(11): 1405-1416.

Déniz, G.E., Sordo, M.A. & Ojeda, C.E. 2014. The Log-Lindley distribution as an alternative to the beta regression model with applications in insurance. *Insurance: Mathematics and Economics* 54: 49-57.

Do Espirito Santo, A.P.J. & Mazucheli, J. 2015. Comparison of estimation methods for the Marshall–Olkin extended Lindley distribution. *Journal of Statistical Computation and Simulation* 85(17): 3437-3450.

Ghitany, M.E., Atieh, B. & Nadarajah, S. 2008. Lindley distribution and its application. *Mathematics Computing and Simulation* 78: 493-506.

Ghitany, M.E., Alqallaf, F., Al-Mutairi, D.K. & Husain, H.A. 2011. A two-parameter weighted Lindley distribution and its applications to survival data. *Mathematics and Computers in Simulation* 81(6): 1190-1201.

Jodrá, P. 2010. Computer generation of random variables with Lindley or Poisson-Lindley distribution via the Lambert W function. *Mathematics and Computers in Simulation* 81(4): 851-859.

Kantar, Y.M. & Şenoğlu, B. 2008. A comparative study for the location and scale parameters of the Weibull distribution with given shape parameter. *Computers & Geosciences* 34(12): 1900-1909.

Krishna, H. & Kumar, K. 2011. Reliability estimation in Lindley distribution with progressively type II right censored sample. *Mathematics and Computers in Simulation* 82(2): 281-294.

Kumar, S.C. & Jose, R. 2019. On double Lindley distribution and some of its properties. *American Journal of Mathematical and Management Sciences* 38(1): 23-43.

Lindley, D.V. 1958. Fiducial distribution and Bayes theorem. *Journal of the Royal Statistical Society, Series B* 20(1): 102-107.

Maiti, S.S. & Mukherjee, I. 2018. On estimation of the PDF and CDF of the Lindley distribution. *Communications in Statistics-Simulation and Computation* 47(5): 1370-1381.

Mazucheli, J., Menezes, A.F.B. & Chakraborty, S. 2018. On the one parameter unit-Lindley distribution and its associated regression model for proportion data. *Journal of Applied Statistics* 46(4): 700-714.

Nassar, M., Afify, A.Z., Dey, S. & Kumar, D. 2018. A new extension of Weibull distribution: Properties and different methods of estimation. *Journal of Computational and Applied Mathematics* 336: 439-457.

Shanker, R. & Mishra, A. 2013a. A quasi Lindley distribution. *African Journal of Mathematics and Computer Science Research* 6(4): 64-71.

Shanker, R. & Mishra, A. 2013b. A two-parameter Lindley distribution. *Statistics in Transition-New Series* 14(1): 45-56.

Shanker, R., Kamlesh, K.K. & Fesshaye, H. 2017. A two parameter Lindley distribution: Its properties and applications. *Biostatistics and Biometrics Open Access Journal* 1(4): 85-90.

Shanker, R., Fesshaye, H. & Sharma, S. 2016. On two parameter Lindley distribution and its applications to model lifetime data. *Biom. Biostat. Int. J.* 3(1): 9-15.

Shanker, R., Sharma, S. & Shanker, R. 2013. A two-parameter Lindley distribution for modeling waiting and survival times data. *Applied Mathematics* 4: 363-368.

Shanker, R., Shukla, K.K., Shanker, R. & Tekie, A.L. 2017. A three-parameter Lindley distribution. *American Journal of Mathematics and Statistics* 7(1): 15-26.

Sharma, V.K., Singh, S.K., Singh, U. & 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(3): 162-173.

Singh, S.K., Singh, U. & Sharma, V.K. 2014. The truncated Lindley distribution: Inference and application. *Journal of Statistics Applications and Probability* 3(2): 219-228.

Zakerzadeh, H. & Dolati, A. 2009. Generalized Lindley distribution. *Journal of Mathematical Extension* 3(2): 13-25.

*Corresponding author; email: mathil.thamir@uoanbar.edu.iq