Since January 2020 Elsevier has created a COVID-19 resource centre with free information in English and Mandarin on the novel coronavirus COVID-19. The COVID-19 resource centre is hosted on Elsevier Connect, the company’s public news and information website.

Elsevier hereby grants permission to make all its COVID-19-related research that is available on the COVID-19 resource centre - including this research content - immediately available in PubMed Central and other publicly funded repositories, such as the WHO COVID database with rights for unrestricted research re-use and analyses in any form or by any means with acknowledgement of the original source. These permissions are granted for free by Elsevier for as long as the COVID-19 resource centre remains active.
The new discrete distribution with application to COVID-19 Data

Ehab M. Almetwally\textsuperscript{a,b,*}, Doaa A. Abdou\textsuperscript{c}, E.H. Hafez\textsuperscript{d}, Taghreed M. Jawa\textsuperscript{e}, Neveen Sayed-Ahmed\textsuperscript{f}, Hisham M. Almongy\textsuperscript{c}

\textsuperscript{a} Faculty of Business Administration, Delta University of Science and Technology, Gamasa, 11152, Egypt
\textsuperscript{b} Faculty of Graduate Studies for Statistical Research, Cairo University, Giza, Egypt
\textsuperscript{c} Department of Mathematics, Faculty of Science, Helwan University, Cairo, Egypt
\textsuperscript{d} Faculty of Commerce, Mansoura University, Mansoura, 35516, Egypt
\textsuperscript{e} Department of Mathematics, College of Science, Taif University, P.O. Box 11099, Taif 21944, Saudi Arabia

\section*{A B S T R A C T}

This research aims to model the COVID-19 in different countries, including Italy, Puerto Rico, and Singapore. Due to the great applicability of the discrete distributions in analyzing count data, we model a new novel discrete distribution by using the survival discretization method. Because of importance Marshall–Olkin family and the inverse Toppe–Leone distribution, both of them were used to introduce a new discrete distribution called Marshall–Olkin inverse Toppe–Leone distribution, this new distribution namely the new discrete distribution called discrete Marshall–Olkin Inverse Toppe–Leone (DMOITL). This new model possesses only two parameters, also many properties have been obtained such as reliability measures and moment functions. The classical method as likelihood method and Bayesian estimation methods are applied to estimate the unknown parameters of DMOITL distributions. The Monte-Carlo simulation procedure is carried out to compare the maximum likelihood and Bayesian estimation methods. The highest posterior density (HPD) confidence intervals are used to discuss credible confidence intervals of parameters of new discrete distribution for the results of the Markov Chain Monte Carlo technique (MCMC).

\section*{Introduction}

Corona viruses are a huge family of viruses that can cause a variety of diseases varying from the common cold to much more serious conditions such as Middle East Respiratory Syndrome (MERS) and Severe Acute Respiratory Syndrome (SARS). In Wuhan, China, a new Coronavirus (COVID-19) was discovered in 2019. This is an extremely new coronavirus that has not been found in people before. The coronavirus disease 2019 (COVID-19) has been declared a pandemic by the World Health Organization (WHO). To stop the virus from spreading further, a concerted global effort is required. A pandemic affects a wide geographic area and affecting an exceptionally high proportion of the population. The H1N1 flu pandemic in 2009 is the last pandemic reported in the world.

There are numerous scientists that examined the pandemic Covid-19 and created models to match the data and offer predictions about the projected number of cases to aid the nations to make choices about prevention strategies. For example, see El-Morshedy et al. \cite{1} he presented a new discrete distribution, a discrete generalized Lindley, for analyzing everyday coronavirus infections in Hong Kong and daily new fatalities in Iran. Maleki et al. \cite{2} he predicted recovered and verified COVID19 cases using an autoregressive time series model based on the two-piece scale mixture normal distribution. Nesteruk \cite{3} and Batista \cite{4} they studied the daily new COVID-19 cases in China were anticipated using a mathematical model dubbed susceptible, infected, and recovered (SIR). Almongy et al. \cite{5} introduced a new modeling of the COVID-19 mortality rates in Italy, Mexico, and the Netherlands. Liu et al. \cite{6} discussed new modeling of the survival times for the COVID-19 mortality rates in China.

By using the inverse transformation to random variables, we proposed the inverse distributions. These distributions display different features in the behavior of the density and hazard rate shapes. Many authors discussed the inverted distributions and their applications. Some of the well-known inverted models are inverse Weibull distribution (Calabria and Pulcini \cite{7}, Muhammed and Almetwally \cite{8,9}), inverted Toppe–Leone (ITL) (Hassan et al. \cite{10}, Almetwally et al. \cite{11}, Hassan et al. \cite{12} and Almetwally \cite{13}) among others. Hassan et al. \cite{10} proposed the ITL with CDF given by

\begin{equation}
F(x; \theta) = 1 - \frac{(1+2x^\theta)}{(1+x^\theta)} ; x > 0, \theta > 0.
\end{equation}
where \( \theta > 0 \) is the shape parameter. The probability mass function (PMF) related to Eq. (2) is given by

\[
f(x; \theta) = 2\alpha x (1 + 2x)^{\alpha - 1} (1 + x)^{2\alpha + 1}.
\]

We utilize discrete distributions in countable data analysis since most existing continuous distributions do not produce appropriate results for modeling COVID-19 cases, and counts of deaths or daily new cases exhibit significant dispersion. The survival discretization method is the most often used method for generating discrete distributions, and it necessitates the presence of a cumulative distribution function (CDF). Time is divided into unit intervals, and the survival function should be continuous and non-negative. Roy [14] defines the discrete distribution PMF as follows:

\[
P(X = x) = P(x \leq X \leq x + 1) = S(x) - S(x + 1); \quad x = 0, 1, 2, \ldots \tag{3}
\]

Where \( S(x) = P(X \geq x) = 1 - F(x; \theta) \), where \( F(x; \theta) \) is a continuous distribution CDF and \( \theta \) is a parameter vector. If the CDF of the random variable \( X \) has \( P(X = x) = f(x + 1; \theta) \), it is considered to have a discrete distribution. The hazard rate is given by \( h_r(x) = P(X = x)/S(x) \). The discrete distribution’s reversed failure rate is given as \( r_f(x) = P(X = x)/(1 - S(x)) \).

Discrete Marshall–Olkin inverted Topp–Leone (DMOITL) distribution. We define DMOITL distribution. In Section ‘Parameter estimation’. While Section ‘Confidence intervals’ is concerned with the interval estimation methods. In Section ‘Simulation analysis’ we made a simulation study to compare the performance of the estimating approaches. Three real data sets from COVID-19 in different countries, including Italy, Puerto Rico, and Singapore, are used in Section ‘Data analysis’ to prove the efficiency of the DMOITL distribution with respect to other distributions. Finally, conclusions and major findings are given in Section ‘Conclusion’.

**DMOITL distribution**

In this part, we introduced the Marshall–Olkin inverted Topp–Leone (MOITL) distribution and converted this new continuous distribution to discrete distribution as discrete MOITL (DMOITL) distribution.

By using Eqs. (4), and Eq. (1), the survival function of MOITL distribution can obtained and written as follows:

\[
\bar{G}_r(x, \theta) = \frac{a (1 + 2x)^\theta}{(1 + x)^{2\theta} - \bar{F}(1 + 2x)\theta} \quad 0 < x < \infty, \quad a > 0, \tag{6}
\]

where \( \theta \) is defined as a vector parameters of MOITL distribution \( a \), and \( \delta \). The DMOITL distribution is obtained based on survival discretization method. Eq. (6) is used as the survival function of a baseline MOITL model using the parameter vector \( \theta \). As a result, the CDF of the DMOITL distribution is:

\[
F(x; \theta) = 1 - \frac{a (3 + 2x)^\delta}{(2 + x)^{2\delta} - \bar{F}(3 + 2x)^\delta}, \quad 0 < x < \infty, \quad a, \theta > 0, \tag{7}
\]

The corresponding PMF of Eq. (7) is defined by

\[
P(X = x; \theta) = \frac{a (1 + 2x)^\delta}{(1 + x)^{2\delta} - \bar{F}(1 + 2x)^\delta} - \frac{a (3 + 2x)^\delta}{(2 + x)^{2\delta} - \bar{F}(3 + 2x)^\delta}, \quad x \in \mathbb{R}, \tag{8}
\]

where \( \Theta \) is positive vector parameters. \( X \sim \text{DMOITL}(\theta) \) indicates the random variable with PMF (8).

Fig. 1 is a graphical representation for various shapes of the PMF of the DMOITL distribution. These figures show that the PMF of the DMOITL distribution can be right-skewed, symmetric, or decreasing curves. The DMOITL distribution, as seen in the application section, has a lot of versatility and can be used to simulate skewed data. Therefore it is extensively utilized in fields like biomedical studies, biology, dependability, physical engineering, and survival analysis.

Sub-models of the DMOITL model for selected values of the parameters are presented as: If \( a = 1 \), the DITL distribution with the PMF,

\[
P(X = x; \delta) = \frac{(1 + 2x)^\delta}{(1 + x)^{2\delta} - \bar{F}(3 + 2x)^\delta}, \quad x \in \mathbb{R},
\]

and the CDF of the DITL distribution is given by:

\[
F(x; \delta) = 1 - \frac{(3 + 2x)^\delta}{(2 + x)^{2\delta}},
\]

**Statistical properties**

The DMOITL distribution’s reliability measures, moments, and moment generating function (MGF) are shown here.

**Reliability measures**

The hazard rate function (HR) of the DMOITL distribution are given by

\[
h(x; \theta) = \frac{(1 + 2x)^\delta}{(1 + x)^{2\delta} - \bar{F}(3 + 2x)^\delta}.
\]
The survival functions of DMOITL is given as
\[ F(x, \theta) = \frac{\alpha (3 + 2x) \theta}{(2 + x)^{\theta} - \alpha (3 + 2x)^{\theta}}. \]

There are some important shapes of the HR of the DMOITL distribution in Figs. 2. The HR of the DMOITL distribution has some important shapes, containing decreasing, and upside down curve, which are appealing features for various count models.

The reverse hazard function of DMOITL is given as
\[ rh(x; \theta) = \frac{P(x, \theta)}{F(x, \theta)} = \frac{\alpha (1 + 2x)^{\theta} (2 + x)^{\theta} - \alpha (3 + 2x)^{\theta}}{(1 + x)^{2\theta} - (3 + 2x)^{\theta} - \alpha (1 + 2x)^{\theta}}. \]

\[ \mu_r' = \sum_{x=0}^{\infty} x^r P(x; \theta) = \alpha \sum_{x=0}^{\infty} x^r (1 + 2x)^{\theta} (2 + x)^{\theta} - \alpha \sum_{x=0}^{\infty} x^r (3 + 2x)^{\theta}. \]
In particular, the mean of DMOITL distribution is

$$\mu'_1 = \alpha \sum_{i=0}^{n} \frac{x(1 + 2x)^\theta}{(1 + x)^2(1 + 2x)^\theta} - \alpha \sum_{i=0}^{n} \frac{x(3 + 2x)^\theta}{(2 + x)^3 - \alpha (3 + 2x)^\theta}.$$  

(12)

The variance of DMOITL distribution is given as

$$\text{var}(x) = \alpha \sum_{i=0}^{n} \frac{x^2(1 + 2x)^\theta}{(1 + x)^2 - \pi (1 + 2x)^\theta} - \alpha \sum_{i=0}^{n} \frac{x^2(3 + 2x)^\theta}{(2 + x)^3 - \alpha (3 + 2x)^\theta} - \mu'^2.$$  

(13)

The dispersion index (DI) may be determined with the help of the following expression:

$$DI = \frac{V\text{ar}(x)}{\mu'_1}.$$  

(14)

The skewness value (SKV) for DMOITL distribution, can be positive, zero, negative, or undefined. It can be expressed in terms of the third raw moment:

$$SKV = \mu'_3 - 3\mu'_1\mu'_2 + 2\mu'_1^3$$  

(15)

The kurtosis value (KTV) for DMOITL distribution can be expressed in terms of the fourth raw moment:

$$KTV = \mu'_4 - 4\mu'_1\mu'_3 + 6\mu'_2\mu'_1^2 - 3\mu'_1^4$$  

(16)

Table 1 Different measures by moment function of DMOITL distribution.

| a    | Mean | \(\mu'_1\) | Var | DI   | \(\mu'_2\) | \(\mu'_3\) | SKV | KTV |
|------|------|------------|-----|------|------------|------------|-----|-----|
| 0.6  | 39.900 | 3.94E+04  | 3.77E+04 | 965.629 | 5.26E+07  | 7.22E+10  | 6.544 | 44.993 |
| 0.9  | 7.120  | 555.720   | 505.026 | 72.378  | 77002.960 | 1.16E+07  | 5.803 | 37.695 |
| 0.5  | 2.140  | 21.020    | 16.440 | 7.839   | 408.820   | 9692.300  | 4.402 | 24.816 |
| 3    | 0.840  | 2.000     | 1.294 | 1.572   | 7.320     | 34.640    | 2.353 | 10.157 |
| 5    | 0.460  | 0.660     | 0.448 | 0.995   | 1.180     | 2.580     | 1.545 | 5.533 |
| 0.6  | 233.460 | 1.46E+07  | 1.40E+06 | 6160.934 | 1.20E+10  | 1.01E+14  | 6.570 | 45.251 |
| 0.9  | 22.800 | 6193.040  | 5673.2  | 252.903 | 2908.060  | 1.47E+09  | 5.851 | 38.156 |
| 1.5  | 4.780  | 96.660    | 73.811 | 15.757  | 3905.260  | 194472.180 | 4.317 | 24.135 |
| 3    | 1.520  | 5.040     | 2.729 | 1.832   | 26.720    | 184.560   | 2.386 | 10.194 |
| 5    | 0.880  | 1.480     | 0.705 | 0.818   | 3.280     | 9.160     | 1.241 | 5.407 |

By differentiating Eq. (17), we can acquire the non-linear likelihood equations with respect to the parameters \(\alpha\) and \(\theta\), respectively:

$$\frac{\partial l(\Theta)}{\partial \alpha} = n \alpha \sum_{i=1}^{n} \frac{(1 + 2x_i)^\theta}{(1 + x_i)^2(1 + 2x_i)^\theta} - n \alpha \sum_{i=1}^{n} \frac{(3 + 2x_i)^\theta}{(2 + x_i)^3 - \sigma (3 + 2x_i)^\theta}.$$  

(17)

and

$$\frac{\partial l(\Theta)}{\partial \theta} = \sum_{i=1}^{n} (1 + 2x_i)^\theta \log (1 + 2x_i) - \sum_{i=1}^{n} (3 + 2x_i)^\theta \log (3 + 2x_i).$$  

(19)

Bayesian estimation

Bayesian estimation is one of the most important and accurate methods of estimation. In Bayesian estimation the parameters is considered as a random variable that is distributed with a certain distribution. We assign a prior believe about the parameter by using a prior distribution for the two parameters. The capacity to integrate previous information into study helps make the Bayesian technique very valuable for reliability assessment, since one of the primary challenges involved with reliability analysis is data scarcity. For the \(a\) and \(\theta\) parameters of DMOITL distribution are distributed with gamma prior distributions,
where $\alpha$ and $\theta$ are non-negative values. The $\alpha$ and $\theta$ parameters as independent joint prior density functions can be expressed as follows:

$$\Pi(a, \theta) \propto a^{\alpha-1} \theta^{\theta-1} e^{-a \theta \alpha + b \theta \beta}.$$  

(20)

The joint posterior density function of $\Theta$ is derived from likelihood function of DMOITL distribution and joint prior density (20).

$$\pi(a, \theta|x) = \frac{L(x|a, \theta) \Pi(a, \theta)}{\int_a \int_\theta L(x|a, \theta) \Pi(a, \theta) \, da \, d\theta} \propto L(x|a, \theta) \Pi(a, \theta),$$

$$\propto \alpha^{\alpha+1-1} \theta^{\theta+1-1} e^{-a \theta \alpha + b \theta \beta} \times \sum_{i=1}^n \left[ \frac{(1 + 2x)^\theta}{(1 + x)^{2\theta} - \theta (1 + 2x)^\theta} - \frac{(3 + 2x)^\theta}{(2 + x)^{2\theta} - \theta (3 + 2x)^\theta} \right].$$

(21)

Under the symmetric loss functions, most of the Bayesian inference procedures have been developed squared-error loss function is commonly symmetric loss function. The Bayes estimators of $\Theta$, say ($\hat{\alpha}, \hat{\theta}$), based on squared error loss function is given by

$$\hat{\alpha} = E(\alpha|\theta, x) = \int_0^\infty \int_0^\infty a \times \pi(a, \theta|x) \, da \, d\theta,$$

and

$$\hat{\theta} = E(\theta|a, x) = \int_0^\infty \int_0^\infty \theta \times \pi(a, \theta|x) \, da \, d\theta.$$  

(22)

(23)

It is noticed that the integrals are given by (22) and (23) are not possible to derive explicitly. As a consequence, we estimate the value of integrals in (22) and (23) using the Markov Chain Monte Carlo (MCMC) approach. Many studies used MCMC techniques such Almetwally et al. [26,27], Basheer et al. [28], Almongy et al. [5,29], and Bantan et al. [30]. For more reading about Covid papers see [31].

Gibbs sampling and the more generic Metropolis within Gibbs samplers are significant sub classes of Markov chain Monte Carlo (MCMC) techniques. The Metropolis–Hastings (MH) and Gibbs sampling techniques are the two most often used instances of the MCMC method. The MH method, like acceptance–rejection sampling, believes that for each iteration of the algorithm, a candidate value from a proposal distribution can be produced for each iteration of the algorithm. The MH algorithm, similar to acceptance–rejection sampling, believes that for each iteration of the algorithm, a candidate value from a proposal distribution can be produced. To generate random samples of conditional posterior densities from the DMOITL distribution, we employ the MH within the Gibbs sampling steps:

$$\Pi(\alpha, \theta|x) \propto \alpha^{\alpha+1-1} e^{-a \theta \alpha + b \theta \beta} \sum_{i=1}^n \left[ \frac{(1 + 2x)^\theta}{(1 + x)^{2\theta} - \theta (1 + 2x)^\theta} - \frac{(3 + 2x)^\theta}{(2 + x)^{2\theta} - \theta (3 + 2x)^\theta} \right],$$

(24)

and

$$\Pi(\theta|\alpha, x) \propto \theta^{\theta+1-1} e^{-a \theta \alpha + b \theta \beta} \sum_{i=1}^n \left[ \frac{(1 + 2x)^\theta}{(1 + x)^{2\theta} - \theta (1 + 2x)^\theta} - \frac{(3 + 2x)^\theta}{(2 + x)^{2\theta} - \theta (3 + 2x)^\theta} \right].$$

(25)

Confidence intervals

In this section, we introduce the construction of confidence intervals with two different methods to estimate the unknown parameters of the DMOITL distribution, which are asymptotic confidence interval (ACI) in MLE and credible confidence interval in MCMC of $\alpha$, and $\theta$.

Asymptotic confidence intervals

Using the asymptotic normal distribution of the MLE is the most popular method to set confidence bounds for the parameters. Fisher information matrix $I(\Theta)$ is constructed of the negative second derivatives of the natural logarithm of the likelihood function evaluated at $\Theta = (\alpha, \theta)$ in connection to the asymptotic variance–covariance matrix of the MLE of the parameters. Suppose the asymptotic variance–covariance matrix of the parameter vector $\Theta$ is

$${I(\Theta)}^{-1} =$$

where $V(\Theta) = {I(\Theta)}^{-1}$

Highest posterior density

This method is similar to the ACI for more information see Chen and Shao [32]. The HPD intervals: Chen and Shao [32] discussed this technique to generate the HPD intervals of unknown parameters of the benefit distribution. In this study, samples drawn with the proposed MH algorithm should be used to generate time-lapse estimates. For example, using the MCMC sampling outputs and the percentile tail points, a $(1 - \gamma)\%$ HPD interval with two points for 2th parameters of the DMOITL distribution can be generated. According to [32], the BCIs of the parameters of DMOITL distribution $\alpha$, $\theta$ can be obtained through the following steps:

1. Arrange $\hat{\alpha}$, and $\hat{\theta}$ as $(\hat{\alpha}[1] \leq \hat{\alpha}[2] \leq \ldots \leq \hat{\alpha}[M])$ and $(\hat{\theta}[1] \leq \hat{\theta}[2] \leq \ldots \leq \hat{\theta}[M])$, where $A$ denotes the length of the generated of MH algorithm.

2. The 100$(1 - \gamma)\%$ symmetric credible intervals of $\alpha$, $\theta$ are obtained as: \(\hat{\alpha}[M/2] \leq \hat{\alpha} \leq \hat{\alpha}[M(1 - \gamma)/2]\) and \(\hat{\theta}[M/2] \leq \hat{\theta} \leq \hat{\theta}[M(1 - \gamma)/2]\).

Simulation analysis

In this part of the paper, we made a simulation study to assess the performance of the distribution by varying the values of the actual values for both parameters and observing the effect of this change on the accuracy of estimation for both methods. The Monte-Carlo simulation process is used in this section to compare the conventional estimation methods: MLE, and Bayesian estimation methods under the square error loss function based on MCMC, for the estimation of DMOITL distribution parameters by R software. Monte-Carlo experiments are carried out on the basis of 10000 randomly generated DMOITL distribution samples, where $x$ represents the DMOITL for various parameter actual values such as: Case 1: $\alpha = 0.5$ with different $\theta = 0.3, 0.6, 1.5$. Case 2: $\alpha = 1.5$ with different $\theta = 0.3, 0.6, 1.5$. Case 3: $\alpha = 3$ with different $\theta = 0.3, 0.6, 1.5$, and different sample sizes $n$: (20, 50, and 100).

Concluding remarks on the simulation results

The Tables 2–4 summarize the simulation findings for the methodologies provided in this work for estimating parameters of the DMOITL distribution using point estimate and interval estimation. It is necessary to compare the different point estimating approaches by calculating the Bias, MSE, and lower and higher confidence intervals. These tables let the following conclusions to be drawn:

1. As $n$ rises, the Bias and MSE of the DMOITL distribution drop.
2. Bias and MSE for $\alpha$ and $\theta$ parameters grow as $\theta$ increases.
3. As the value of $\alpha$ grows, the Bias and MSE values for the $\alpha$ and $\theta$ parameters decrease.
4. Bayesian estimation is the best approach for estimating the parameters as it provides the smallest MSE and Bias and also has the shortest confidence interval.

E.M. Almetwally et al. Results in Physics 32 (2022) 104987
Darling (AD), Kolmogorov–Smirnov (KS) and Akaike information criterion (AIC) statistics for the all models fitted based on three real data sets. These tables also include the MLE of the parameters for the models under consideration. Figs. 3, 5 and 7 show the fitted DMOITL, PMF, CDF, PP-plot, and QQ-plot of the three data sets, respectively. These statistics show that among all fitted models, the DMOITL distribution has the lowest CvM, AD, KS, and AIC values. Using alternative data, Table 8 presented MLE and Bayesian estimation methods for parameters of the DMOITL distribution. Figs. 4, 6, 8 show convergence plots of MCMC for parameter estimates of DMOITL distribution for different parameters under consideration. Figs. 3, 5 and 7 show the fitted DMOITL, PMF, CDF, PP-plot, and QQ-plot of the three data sets, respectively. These statistics show that among all fitted models, the DMOITL distribution has the lowest CvM, AD, KS, and AIC values. Using alternative data, Table 8 presented MLE and Bayesian estimation methods for parameters of the DMOITL distribution. Figs. 4, 6, 8 show convergence plots of MCMC for parameter estimates of DMOITL distribution for different data set.

Firstly: This is a COVID-19 data set from Puerto Rico that spans 38 days, from February 26 to April 4, 2021. This data set is comprised of newly reported instances on a daily basis. The data are as follows: 100, 160, 235, 193, 216, 67, 69, 332, 212, 330, 295, 227, 145, 78, 260, 399, 311, 114, 253, 287, 151, 30, 102, 199, 261, 305, 185, 120, 68, 46, 356, 160, 235, 193, 216, 67, 69, 332, 212, 330, 295, 227, 145, 78, 260, 399, 268, 595, 447, 170, 365, 510, 881.

Secondly: This is a 61-day COVID-19 data set from Italy, recorded between 13 June and 12 August 2021. This data set is comprised of newly reported instances on a daily basis. The data are as follows: 268, 595, 447, 170, 365, 510, 881.
Table 4
This table contains the simulation results when $\alpha = 3$.

| $\theta$ | $n$ | $a$ | $\beta$ | $\gamma$ | MSE | CP |
|----------|-----|-----|---------|---------|-----|----|
| 0.3      |     |     |         |         |     |    |
| 20       |     |     | 0.1802  | 0.6961  | 1.5780 | 4.7825 | 93.06% |
|          |     |     | 0.0134  | 0.0051  | 0.1750 | 0.4519 | 96.53% |
| 50       |     |     | 0.0123  | 0.3528  | 1.8278 | 4.1968 | 93.10% |
|          |     |     | 0.0117  | 0.0017  | 0.2324 | 0.3910 | 93.10% |
| 100      |     |     | 0.1360  | 0.1126  | 2.5247 | 3.7473 | 93.55% |

Table 5
MLE, CvM, AD, KS and AIC for different alternative models of DMOITL distribution: Puerto Rico.

| $\alpha$ | $\beta$ | $\lambda$ | CvM | AD | KS | AIC |
|----------|---------|-----------|-----|----|----|-----|
| DMOITL   | 116654.6968 | 2.5126 | 0.0711 | 0.4019 | 0.0758 | 487.0636 |
| DBuur    | 16.5248   | 0.9886  | 0.1237 | 0.6821 | 0.4977 | 607.7753 |
| DW       | 0.9999    | 1.5876  | 12.5759 | 76.0974 | 0.9937 | 487.1050 |
| DIW      | 0.0000    | 0.8312  | 0.1745 | 0.9684 | 1.0000 | 510.3790 |
| NB       | 0.1339    | 0.5969  | 0.3400 | 0.3400 | 884.7561 |
| Poisson  | 245.8474  | 0.3611  | 0.9437 | 0.4713 | 4067.0956 |
| DAPL     | 1.3874    | 1.4894  | 3.145E-25 | 0.1771 | 0.9839 | 2.0548 |
| DITL     | 0.2175    | 0.1305  | 1.3804 | 2.2066 | 94.50% | 1.4226 |

Fig. 3. Plots of estimated pmfs of distributions for Data set of Puerto Rico.
Fig. 4. Convergence plots of MCMC for parameter estimates of DMOITL distribution for data set of Puerto Rico.

Fig. 5. Plots of estimated pmfs of distributions for Data set of Italy.
Fig. 6. Convergence plots of MCMC for parameter estimates of DMOITL distribution for data set of Italy.

Fig. 7. Plots of estimated pmfs of distributions for Data set of Singapore.
Fig. 8. Convergence plots of MCMC for parameter estimates of DMOITL distribution for data set of Singapore.

Table 6
MLE, CvM, AD, KS and AIC for different alternative models of DMOITL distribution: Italy.

| Model  | $\alpha$  | $\delta$ | $\lambda$ | CvM  | AD   | KS   | AIC   |
|--------|-----------|----------|-----------|------|------|------|-------|
| DMOITL | 3065.8285 | 3.3396   |           | 0.0681 | 0.3680 | 0.0715 | 475.1201 |
| DBuur  | 16.2005   | 0.9795   |           | 0.1821 | 1.0424 | 0.4526 | 623.6174 |
| DW     | 0.9983    | 1.9497   |           | 19.9308 | 121.8355 | 0.9900 | 476.0156 |
| DIW    | 0.0000    | 1.4558   |           | 0.2993 | 1.7367 | 1.0000 | 495.8519 |
| NB     | 0.0846    | 0.3810   |           | 0.9983 | 1.9497 | 0.2351 | 608.7627 |
| Poisson| 22.6230   | 22.6230  |           | 0.0918 | 0.4535 | 0.2674 | 700.5536 |
| DAPL   | 0.0020    | 1.4589   | 1.08E−13  | 0.2188 | 0.2188 | 0.1265 | 486.1749 |
| Dil    | 0.9202    | 0.0940   |           | 0.0940 | 0.4254 | 0.1464 | 481.7783 |
| DITL   | 0.4203    | 0.4203   |           | 0.1592 | 0.9048 | 0.4320 | 604.7508 |

as follows: 4, 4, 5, 12, 5, 18, 7, 5, 4, 6, 8, 5, 10, 2, 9, 3, 13, 5, 13, 12, 14, 6, 6, 8, 7, 5, 16, 12, 24, 9, 17, 19, 10, 29, 21, 13, 14, 10, 5, 5, 13, 12, 14, 13, 12, 27, 30, 31, 33, 35, 24, 28, 31, 33, 23, 29, 42, 22, 17, 38, 45, 30, 24, 30, 14, 30, 40, 38, 15, 10, 48, 44, 14, 25, 34, 24, 58, 29, 29, 19, 18, 22, 25, 26, 24, 22, 11, 15, 12, 18, 9, 14, 9, 1, 11, 11, 14, 12, 11, 10, 4, 7, 10, 13, 12, 11, 8, 23, 19, 9, 13, 13, 13, 6, 10, 8, 10, 8, 17, 12, 11, 9, 15, 15, 17, 12, 12, 13, 15, 17, 12, 23, 12, 21, 26, 34, 26, 43, 18, 10, 17, 24, 35, 21, 26, 32, 20, 25, 14, 27, 16, 34, 39, 23, 20, 14, 15, 24, 39, 23, 40, 12, 45, 23, 35, 24, 34, 39, 17, 17, 16, 18, 25, 20, 28, 19, 25, 16, 34, 52, 31, 49, 28, 38, 38, 41, 40, 29, 25, 36, 30, 26, 24, 30, 33, 25, 23, 18, 31, 45, 13, 18, 20, 14, 9, 4, 13, 9, 18, 13, 25, 14, 24, 27, 16, 21, 11, 16, 18, 22, 23, 20, 17, 14, 9, 10, 16, 10, 10, 7, 11, 13, 10, 12, 16, 10, 6, 8, 26, 26, 60, 48, 61, 68, 92.

Concluding remarks on the data analysis

After applying the three data sets on the proposed distribution and observing the results in Tables 5–7 that provide values of Gramercy–von Mises (CvM), Anderson–Darling (AD), Kolmogorov–Smirnov (KS) and Akaike information criterion (AIC) statistics for all models fitted based on three real data sets, we found that our proposed distribution is the best model as it has the lowest value of AIC and KS values.
Fig. 9. Existence and uniqueness for the log-likelihood for data set of Puerto Rico.

Fig. 10. Existence and uniqueness for the log-likelihood for data set of Italy.

Fig. 11. Existence and uniqueness for the log-likelihood for data set of Singapore.

By referring to these values, we can make sure that our proposed distribution is superior among all its competitors.

Existence and uniqueness for the log-likelihood

We sketched the log-likelihood for each parameter as shown in Figs. 9–11 by fixing one parameter and varying the other. The figures show that the two roots of the parameters are global maximum, and also by differentiating the log-likelihood with respect to each parameters, we found that the function is a decreasing function and it intersects the x-axis in a single point which is the root of the parameter, and that assures that the roots are unique.

Conclusion

In this paper, we introduce Discrete Marshall–Olkin Inverted Topp-Leone distribution which is called DMOITL. We derived its statistical properties. We made the point and interval estimation by classical and Bayesian estimation methods for the DMOITL unknown parameters \( a \) and \( \theta \). We conducted simulation analysis using the R package to differentiate the performance of different estimation methods. We deduced that the Bayesian method is very efficient than the classical method as it gets more efficient results through the values of the MSE and the length of the confidence interval as it is always shorter and the MSE
is always smaller. In order to prove the superiority and applicability of the proposed distribution, we made a data analysis through the COVID-19 data. We used three data sets in three different countries thought different intervals of time, and by referring to the results in Tables 5–7 that provide values of Cramér–von Mises (CvM), Anderson–Darling (AD), Kolmogorov-Smirnov (KS) and Akaike information criterion (AIC) statistics for the all models fitted based on three real data sets we found that our proposed distribution is the best model as it has the lowest value of AIC and KS values.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The paper includes the data used to support the study’s results.

Acknowledgments

Taif University Researchers Supporting Project number (TURSP-2020/318), Taif University, Taif, Saudi Arabia.

References

[1] El-Morshedy M, Altun E, Eliwa MS. A new statistical approach to model the counts of novel coronavirus cases. Math Sci 2021;1:1–14. http://dx.doi.org/10.21203/rs.3.rs-3163/v1.

[2] Maleki M, Mahmoudi MR, Wraith D, Pho KH. Time series modelling to forecast the confirmed and recovered cases of COVID-19. Travel Med Infect Dis 2020;37:101742.

[3] Nesteruk I. Statistics-based predictions of coronavirus epidemic spreading in mainland China. Innov Bioiyst Bioeng 2020;4(1):13–8.

[4] Batista M. Estimation of the final size of the coronavirus epidemic by the SIR model. 2020, Online paper, ResearchGate, 2.

[5] Almongy HM, Almetwally EM, Aljohani HM, Alghamdi AS, Hafez EH. A new statistical model. 2020, Online paper, ResearchGate, 2.

[6] Liu X, Ahmad Z, Gmez AM, Abdulrahman AT, Hafez EH, Khalil N. Modeling the survival times of the COVID-19 patients with a new statistical model: A case study from China. Plos One 2021;16(7):e0254999.

[7] Calabria R, Pulcini G. On the maximum likelihood and least-squares estimation in the inverse Weibull distribution. Statist Appl 1990;2(1):53–66.

[8] Almongy HM, Almetwally EM, Aljohani HM, Alghamdi AS, Hafez EH. A new statistical model. 2020, Online paper, ResearchGate, 2.

[9] Hassan AS, Almetwally EM, Ibrahim GM. Kumaraswamy inverted Topp-Leone distribution with applications to COVID-19 data. CMC-Comput Mater Continua 2021;68(1):337–58.

[10] Almetwally EM, The odd Weibull inverse Topp-Leone distribution with applications to COVID-19 data. Ann Data Sci 2021;1–20.

[11] Roy D. Discrete rayleigh distribution. IEEE Trans Reliaib 2004;53(2):255–60.

[12] Hassan AS, Almetwally EM, Ibrahim GM. Kumaraswamy inverted Topp-Leone distribution with applications to COVID-19 data. CMC-Comput Mater Continua 2021;68(1):337–58.

[13] Almetwally EM. The odd Weibull inverse Topp-Leone distribution with applications to COVID-19 data. Ann Data Sci 2021;1–20.

[14] Roy D. Discrete rayleigh distribution. IEEE Trans Reliaib 2004;53(2):255–60.

[15] Para BA, Jan TR. On discrete three parameter Burr type XII and discrete lomax distributions and their applications to model count data from medical science. Biom Biostat Int J 2016;4(2):1–15.

[16] Nakagawa T, Osaki S. The discrete Weibull distribution. IEEE Trans Reliaib 1975;24(5):300–1.

[17] Krishna H, Pandur PS. Discrete Burr and discrete Pareto distributions. Stat Methodol 2009;6(2):177–88.

[18] Aslam SA, Nekoukhou V, Alamatsaz MH. A discrete generalized exponential distribution of the second type. Statistics 2013;47(4):876–87.

[19] Al-Babtain AA, Ahmed AS, Afify A. A new discrete analog of the continuous lindley distribution, with reliability applications. Entropy 2020;22(6):603.

[20] Eliaawa MS, Alhussain ZA, El-Morshedy M. Discrete Gompertz-G family of distributions for over-and under-dispersed data with properties, estimation, and applications. Mathematics 2020;8(3):558.

[21] Gil-road J, Balogan OS, Almetwally EM, Sherwani RAK, Jamal F, Joseph J. A new model for modeling the COVID-19 epidemic in Egypt: Modelling using a discrete Marshall-Olkin generalized exponential distribution. Int J Probab Statist 2020;9(2):33–41.

[22] Marshall AW, Olikin I. A new method for adding a parameter to a family of distributions with an application to the exponential and Weibull families. Biometrika 1997;84(3):641–52.

[23] Sankaran PG, Jayakumar K. On proportional odds models. Statist Papers 2008;49(4):779–89.

[24] Almetwally EM, Almongy HM, El sayed Mubarak A. Bayesian and maximum likelihood estimation for the Weibull generalized exponential distribution parameters using progressive censoring schemes. Pakistan J Statist Oper Res 2018;14(4):853–68.

[25] Almetwally EM, Almogny HM, Rastogi MK, Ibrahim M. Maximum product spacing estimation of Weibull distribution under adaptive type-II progressive censoring schemes. Ann Data Sci 2020;7(2):257–79.

[26] Basheer AM, Almetwally EM, Okasha HM, Marshall-Olkin alpha power inverse Weibull distribution: Non Bayesian and Bayesian estimations. J Statist Appl Probab 2021;10(2):329–45.

[27] Almongy HM, Almetwally EM, Mubarak AE, Marshall-Olkin alpha power log- distribution: Estimation methods, applications on physics and economics. Pakistan J Statist Oper Res 2021;17(1):137–53.

[28] Bantan R, Hassan AS, Almetwally EM, Elgarhy M, Jamal F, Cheneau C, et al. Bayesian analysis in partially accelerated life tests for weighted lomax distributions and their applications to COVID-19 data. Results Phys 2021;23:104012.

[29] Chen MH, Shao QM. Monte Carlo estimation of Bayesian credible and HPD intervals. J Comput Graph Statist 1999;8(1):69–92.

[30] Almetwally EM, Almongy HM, Ibrahim GM. Discrete alpha power inverse lomax distribution with application of COVID-19 data. J Appl Math 2020;9(6):11–22.