[Delta University Scientific Journal](https://digitalcommons.aaru.edu.jo/dusj)

Volume 4
Number 1 *April (2021)*

Article 2

2021

Modelling of COVID-19 Data Using Discrete Distribution

Ehab M. Almetwally
Statistics, Delta University of Science and Technology, Egypt.

Sanku Dey
Department of Statistics, St. Anthony's College, Shillong, Meghalaya, India

DELTA UNIVERSITY
SCIENTIFIC JOURNAL
Journal home page: https://digitalcommons.aaru.edu.jo/dusj/

Modelling of COVID-19 Data Using Discrete Distribution

Ehab M. Almetwally,*[1] Sanku Dey[2]

1 **Statistics, Delta University of Science and Technology, Egypt.**

2 **Department of Statistics, St. Anthony's College, Shillong, Meghalaya, India.**

Keywords:

1- COVID-19
2- Hazard rate
3- Discrete distributions
4- Survival discretization
5- Maximum likelihood estimation

Abstract

In light of those current conditions that humanity is suffering from the outbreak of the Corona epidemic (COVID-19), which has caused an economic crisis for the entire world, and which also causes humanity, economic, and social losses. Which encouraged researchers in all fields to search and explore solutions to this epidemic. This is what prompted statisticians to provide probability distributions to describe this phenomenon, which is important in simulations and giving a certain probability of expected Incidence and deaths. Which helps in decision-making processes appropriate to the current situation. The purpose of this research is to find and classify the modeling of COVID-19 data by determining the optimal statistical modeling to evaluate the regular count of new COVID-19 fatalities, thus requiring discrete distributions. Some discrete models are checked and reviewed. A new discrete inverse Weibull distribution based on the discretization of survival has been reobtained. Probability mass function and the hazard rate is addressed. Discrete models are discussed based on the Maximum Likelihood estimate for parameters. A numerical analysis uses the regular count of new casualties in the countries of Angola, El Salvador, Estonia, and Greece. In-depth, the empirical findings are interpreted.

1. INTRODUCTION

Corona-Virus "COVID-19" was discovered in Wuhan, China in December 2019. The World Health Organization (WHO) described COVID-19 as a pandemic on March 11, 2020. Refer to Figures 1 and 2. Countries around the world have therefore increased their efforts to decrease the COVID-19 spread rate.

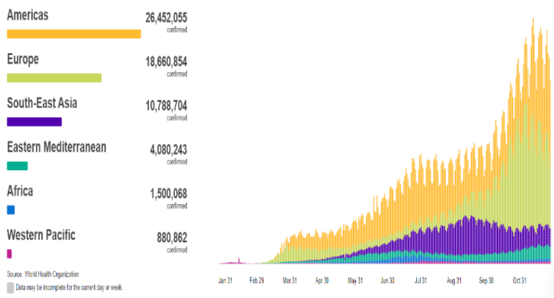

Figure 1: The situation for the daily new cases over the world by the WHO Region.

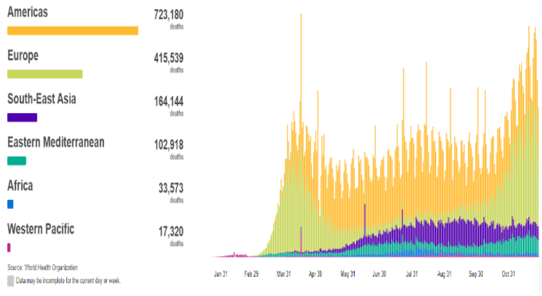

Figure 2: The situation for the daily new deaths over the world by the WHO Region.

A natural discrete Lindley distribution has been implemented by Al-Babtain et al. to model everyday cases and deaths in the world (2020). Almetwally et al. (2020) introduced a discrete Marshall Olkin generalized exponential distribution to discuss the recent Egyptian cases regularly. The research was performed by Hasab et al.

(2020) using the COVID-19 pandemic's Vulnerable Infected Recovered (SIR) epidemic dynamics to model the novel Coronavirus epidemic in Egypt. A new discrete distribution, called discrete generalized Lindley, was analyzed by El-Morshedy et al. (2020) to examine the counts of daily coronavirus cases in Hong Kong and new daily fatalities in Iran. Maleki et al. (2020) have used an autoregressive time series model based on the normal distribution of the two-piece scale mixture to estimate the recovered and reported cases of COVID-19. Nesteruk (2020) and Batista (2020b) have predicted regular new COVID-19 cases in China by using the mathematical model, named susceptible, infected, and recovered (SIR). The logistic growth regression model used by Batista (2020a) is used for estimating the final size of the coronavirus outbreak and its peak time.

Any researcher can come to mind the question: Why do we need discrete distributions? Since we see that most of the current continuous distributions do not set sufficient results for modeling the cases of COVID-19 in count data analysis. The explanation for this is that counts of deaths or new cases regularly indicate disproportionate dispersion.

This research aims to model the daily new fatalities of COVID-19 using a new statistical method to ensure that members of Egyptian society are aware of the risks resulting from the spread of Corona-Virus in the world. To accomplish this goal: First of all, we study separate models such as Poisson, geometric, negative binomial, discrete Burr, discrete Lindley, discrete alpha power inverse Lomax, discrete generalized exponential, discrete Marshall-Olkin Generalized exponential, discrete Gompertz-G-exponential, discrete Weibull, discrete inverse Weibull, exponentiated discrete Weibull, discrete Rayleigh, and new discrete Lindley. Secondly, we introduce a new discrete versatile model that can be denoted as a new discrete inverse Weibull distribution when $x = 0, 1, 2, \ldots$. Thirdly, in some countries such as Angola, El Salvador, Estonia, and Greece, we define the best discrete models that match different regular Coronavirus death datasets.

The need for mathematical and statistical modeling of the extent and spread of the Coronavirus, which tests the progress of medical solutions for drugs and vaccines in reducing the risk of virus spread, is an aspect of the value of science. In future studies, the authors propose that new and different applications, such as a censored sample and a competing risk model, will exist in this critical field. See Balakrishnan and Cramer (2014) for more details of this application and see Almetwally and Almongy (2018) and for more examples see Almetwally et al. (2019).

As follows, the remainder of the paper is structured. Analysis of discrete models in Section 2. Section 3 creates a new DIW. We address the parameter estimation of the discrete model in Section 4. In Section 5, the parameter is calculated using the maximum likelihood equation. Section 6 presents the new regular death of COVID-19 in the case of Angola, El Salvador, Estonia, and Greece to validate the use of models in suitable lifetime count results. Lastly, in Section 7, conclusions are made.

2. REVIEW FOR DISCRETE MODELS

In this section, the survival discretization method and some discrete distributions have been reviewed as discrete Burr, discrete Lindley, discrete alpha power inverse Lomax, discrete generalized exponential, discrete Marshall-Olkin Generalized exponential, discrete Gompertz-G-exponential, discrete Weibull, discrete inverse Weibull, exponentiated discrete Weibull, discrete Rayleigh, and new discrete Lindley.

2.1. Survival Discretization Method

In the statistics literature, sundry methods are available to obtain a discrete distribution from a continuous one. The most commonly used technique to generate discrete distribution is called a survival discretization method, it requires the existence of cumulative distribution function (CDF), survival function should be continuous and non-negative and times are divided into unit intervals. The PMF of discrete distribution is defined in Roy (2003) as

$P(X = x) = P(x \leq X \leq x + 1) = S(x) - S(x + 1)$ (1)

where $x = 0,1,2, \ldots$, $S(x) = P(X \geq x) = F(x;\,\Theta)$, where $F(x;\,\Theta)$ is a CDF of continuous distribution, and Θ is a vector of parameters. The random variable X is said to have the discrete distribution if its CDF is given by

$P(X < x) = F(x + 1;\Theta).$ (2)

The hazard rate is given by $hr(x) = \frac{P(X=x)}{S(x)}$.

The reversed failure rate of discrete distribution is given as

$rfr(x) = \frac{P(X=x)}{1-S(x)}$.

2.2. Discrete Burr Distribution

The PMF of the discrete Burr (DB) distribution has been defined by Krishna and Pundir (2009) is given as follows

$P(x;\theta,\alpha) = \theta^{\ln(1+x^{\alpha})} - \theta^{\ln(1+(x+1)^{\alpha})}$;
$x = 0,1,2,\ldots,\alpha > 0, 0 < \theta < 1$,

the CDF of the DBu distribution is

$F(x;\theta,\alpha) = \theta^{\ln(1+(x+1)^{\alpha})}$,

The hazard rate (hr) of the discrete Burr distribution is

$$hr(x;\theta,\alpha) = 1 - \theta^{\ln\left(\frac{1+(x+1)^{\alpha}}{1+x^{\alpha}}\right)}.$$

2.3. Discrete Lindley Distribution

The PMF of the discrete Lindley (DL) distribution has been defined by Gómez-Déniz and Calderín-Ojeda (2011) is given as follows

$$P(x;\theta) = \frac{\theta^{x}}{1-\ln(\theta)}[\theta\ln(\theta) + (1-\theta)(1-\ln(\theta^{x+1}))];$$
$x = 0,1,2,\ldots, 0 < \theta < 1.$

The CDF of the DLi distribution is

$$F(x;\theta) = \frac{1-\theta^{x+1} + [(2+x)\theta^{x+1} - 1]\ln(\theta)}{1-\ln(\theta)},$$

The hazard rate of the DLi distribution is

$$hr(x;\theta,\alpha) = \frac{\theta^{x}[\theta\ln(\theta) + (\theta-1)(\ln(\theta^{x+1}) - 1)]}{1-\theta^{x+1} + [(2+x)\theta^{x+1} - 1]\ln(\theta)}.$$

2.4. Discrete Alpha Power Inverse Lomax

The discrete alpha power inverse Lomax (DAPIL) distribution is introduced by Almetwally and Ibrahim (2020). The PMF and the CDF of the DAPIL distribution are respectively given by

$$P(x;\alpha,\vartheta,\delta) = \frac{\alpha^{\rho^{\ln\left(1+\frac{\vartheta}{x+1}\right)}} - \alpha^{\rho^{\ln\left(1+\frac{\vartheta}{x}\right)}}}{\alpha-1}; x = 0,1,2,$$

$$F(x;\alpha,\vartheta,\rho) = \frac{\alpha^{\rho^{\ln\left(1+\frac{\vartheta}{x+1}\right)}} - 1}{\alpha-1}, \quad x \in \mathbb{N}_0.$$

The *hr* function of the DAPIL distribution is given by

$$hr(x;\alpha,\vartheta,\rho) = \frac{\alpha^{\rho^{\ln\left(1+\frac{\vartheta}{x+1}\right)}} - \alpha^{\rho^{\ln\left(1+\frac{\vartheta}{x}\right)}}}{\alpha - \alpha^{\rho^{\ln\left(1+\frac{\vartheta}{x+1}\right)}}}, \quad x \in \mathbb{N}_0.$$

2.5. Discrete Generalized Exponential Distribution

The PMF of the discrete generalized exponential (DGE) distribution has been defined by Nekoukhou et al. (2013) is given as follows

$P(x;\theta,\alpha) = (1-\theta^{x+1})^{\alpha} - (1-\theta^{x})^{\alpha}$
$x = 0,1,2,\ldots,\alpha > 0, 0 < \theta < 1$, when $\theta = e^{-\lambda}; \lambda > 0$, the CDF of the DGEx distribution is

$F(x;\theta) = (1-\theta^{x+1})^{\alpha}$,

The hazard rate of the DGEx distribution is

$$hr(x;\theta,\alpha) = \frac{(1-\theta^{x+1})^{\alpha} - (1-\theta^{x})^{\alpha}}{1-(1-\theta^{x+1})^{\alpha}}.$$

2.6. The DMOGEx Distribution

The discrete Marshall-Olkin Generalized exponential (DMOGEx) distribution is introduced by Almetwally et al. (2020). The PMF and the CDF of the DMOGEx distribution are respectively given by

$$P(x,\alpha,\rho,\lambda)=\frac{\lambda[1-(1-\rho^{x})^{\alpha}]}{\lambda+(1-\lambda)(1-\rho^{x})^{\alpha}}-\frac{\lambda[1-(1-\rho^{x+1})^{\alpha}]}{\lambda+(1-\lambda)(1-\rho^{x+1})^{\alpha}}$$

and

$$F(x,\alpha,\rho,\lambda)=\frac{(1-\rho^{x+1})^{\alpha}}{\lambda+(1-\lambda)(1-\rho^{x+1})^{\alpha}},\quad x\in\mathbb{N}_0$$

where $0<\rho<1,\lambda,\theta>0$.

The hr function of the DMOGEx distribution is given by

$$hr(x;\alpha,\rho,\lambda)=\frac{1-(1-\rho^{x})^{\alpha}}{1-(1-\rho^{x+1})^{\alpha}}\frac{\lambda+(1-\lambda)(1-\rho^{x+1})^{\alpha}}{\lambda+(1-\lambda)(1-\rho^{x})^{\alpha}}-1,\quad x\in\mathbb{N}_0.$$

2.7. Discrete Gompertz-G Exponential

The discrete Gompertz-G-exponential (DGzEx) distribution is introduced by Eliwa et al. (2020). The PMF and the CDF of the DGzEx distribution are respectively given by

$$P(x,\alpha,\rho,\lambda)=\rho^{\frac{-1}{\alpha}}\left(\rho^{\frac{1}{\alpha}e^{\lambda\alpha x}}-\rho^{\frac{1}{\alpha}e^{\lambda\alpha(x+1)}}\right);\ x\in\mathbb{N}_0,$$

and

$$F(x,\alpha,\rho,\lambda)=1-\rho^{\frac{1}{\alpha}e^{\lambda\alpha(x+1)}};\ x\in\mathbb{N}_0$$

The hr function of the DGzEx distribution is given by

$$hr(x;\alpha,\rho,\lambda)=\frac{\rho^{\frac{-1}{\alpha}}\left(\rho^{\frac{1}{\alpha}e^{\lambda\alpha x}}-\rho^{\frac{1}{\alpha}e^{\lambda\alpha(x+1)}}\right)}{\rho^{\frac{1}{\alpha}e^{\lambda\alpha(x+1)}}};\ x\in\mathbb{N}_0.$$

2.8. Discrete Weibull

A discrete Weibull (DW) distribution was introduced by Nakagawa and Osaki (1975), and is defined by the cumulative distribution function (CDF):

$$F(x;\alpha,\rho)=1-\rho^{(x+1)^{\alpha}};\ x\in\mathbb{N}_0,\alpha>0,0<\rho<1.$$

The DW distribution has PMF:

$$P(x;\alpha,\rho)=\rho^{x^{\alpha}}-\rho^{(x+1)^{\alpha}},$$

and the hazard rate of DW

$$h(x;\alpha,\rho)=\rho^{x^{\alpha}-(x+1)^{\alpha}}-1.$$

2.9. Discrete Inverse Weibull

A discrete inverse Weibull (DIW) distribution was introduced by Jazi et al. (2010), and is defined by the CDF:

$$F(x;\alpha,\rho)=\rho^{x^{-\alpha}};\ \alpha>0,0<\rho<1.$$

The DIW distribution has PMF:

$$P(x;\alpha,\rho)=\rho^{x^{-\alpha}}-\rho^{(x-1)^{-\alpha}},\qquad x=1,2,\ldots$$

and the hazard rate of DIW

$$h(x;\alpha,\rho)=\frac{1-\rho^{(x-1)^{-\alpha}}}{1-\rho^{x^{-\alpha}}}.$$

2.10. Exponentiated Discrete Weibull

The exponentiated discrete Weibull (EDW) distribution was introduced by Nekoukhou and Bidram (2015), and is defined by the CDF:

$$F(x;\alpha,\rho,\beta)=\left(1-\rho^{(x+1)^{\alpha}}\right)^{\beta};\ x\in\mathbb{N}_0,\alpha,\beta>0,0<\rho<1.$$

The DIW distribution has PMF:

$$P(x;\alpha,\rho,\beta)=\left(1-\rho^{(x+1)^{\alpha}}\right)^{\beta}-\left(1-\rho^{x^{\alpha}}\right)^{\beta},$$

and the hazard rate of DIW

$$h(x;\alpha,\rho,\beta)=\frac{\left(1-\rho^{(x+1)^{\alpha}}\right)^{\beta}-\left(1-\rho^{x^{\alpha}}\right)^{\beta}}{1-(1-\rho^{(x+1)^{\alpha}})^{\beta}}.$$

2.11. Discrete Rayleigh

The discrete Rayleigh (DR) distribution was introduced by Roy (2004), and is defined by the CDF:

$$F(x;\alpha,\rho,\beta)=1-\rho^{(x+1)^{2}};\ x\in\mathbb{N}_0,0<\rho<1.$$

The DR distribution has PMF:

$$P(x;\rho)=\rho^{x^{2}}-\rho^{(x+1)^{2}},$$

and the hazard rate of DR

$$h(x;\rho)=\frac{\rho^{x^{2}}-\rho^{(x+1)^{2}}}{\rho^{(x+1)^{2}}}.$$

2.12. New Discrete Lindley

The new discrete Lindley (NDL) distribution was introduced by Al-Babtain (2020), and is defined by the CDF:

$$F(x;\alpha,\rho,\beta)=1-\frac{1+\rho+\rho x}{1+\rho}(1-\rho)^{x};\ x\in\mathbb{N}_0,0<\rho<1.$$

The NDL distribution has PMF:

$$P(x;\rho)=\frac{\rho^{2}}{1+\rho}(2+x)(1-\rho)^{x},$$

and the hazard rate of NDL

$$h(x;\rho)=\frac{\rho^{2}(2+x)}{1+\rho+\rho x}.$$

3. NEW AND MODIFICATION OF DIW

In this section, we introduce a new flexible DIW model based on the survival discretization method, which can be donated as a new discrete inverse Weibull distribution. The survival of continuous inverse Weibull distribution is given by

$$S(x;\alpha,\beta)=1-e^{-\beta x^{\alpha}};x>0,\alpha,\beta>0.$$

By using the survival discretization method in Equations 1 and 2, the PMF of the new DIW can be defined as follows.

$$P(x;\alpha,\beta)=e^{-\beta(x+1)^{\alpha}}-e^{-\beta x^{\alpha}}.$$

Let $\rho=e^{-\beta};0<\rho<1$, then the PMF of the new DIW can be rewritten as following.

$$P(x;\alpha,\rho)=\rho^{(x+1)^{\alpha}}-\rho^{x^{\alpha}};x=0,1,\dots$$

The CDF of the new DIW is

$$F(x;\alpha,\rho)=1-\rho^{(x+1)^{\alpha}},$$

and the hazard rate of the new DIW is

$$h(x;\alpha,\rho)=1-\rho^{x^{\alpha}-(x+1)^{\alpha}}$$

4. PARAMETER ESTIMATION OF DISCRETE MODEL

In this section, we estimate the parameters of models using a maximum likelihood estimator (MLE). It is noted that the maximum likelihood method is also used to estimate unknown parameters of a statistical model because maximum likelihood estimates (MLEs) have several desirable properties; For example, they are asymptotically unbiased, symmetrical, consistent, asymptomatically normally distributed, etc. Let $x_1,x_2,\dots,x_n$ be a random sample of size n from the discrete distribution, then the log-likelihood function is given by

$$\mathcal{L}(\Theta|x)=\sum_{i=1}^{n}\log P(x_i;\Theta),$$

where $\Theta=(\Theta_1,\dots,\Theta_k)^T$, k is a length of Θ. The MLEs that can be obtained by derivatives of the log-likelihood function and equal zero

$$\frac{\partial\mathcal{L}(\Theta|x)}{\partial\Theta_1}=\sum_{i=1}^{n}\frac{\frac{\partial P(x_i;\Theta)}{\partial\Theta_1}}{P(x_i;\Theta)},\dots,\frac{\partial\mathcal{L}(\Theta|x)}{\partial\Theta_k}=\sum_{i=1}^{n}\frac{\frac{\partial P(x_i;\Theta)}{\partial\Theta_k}}{P(x_i;\Theta)}$$

provide the MLEs of Θ, say $\widehat{\Theta}=(\widehat{\Theta}_1,\dots,\widehat{\Theta}_k)^T$, then using a computational process such as the k variable Newton-Raphson Algorithm are given by the solutions of probability equations.

For interval estimation and hypothesis tests on the model parameters, we require the information matrix. The $k\times k$ observed information matrix is

$$I_n(\widehat{\Theta})=\begin{bmatrix}\frac{\partial^2\mathcal{L}(\Theta|x)}{\partial\Theta_1^2} & \cdots & \frac{\partial^2\mathcal{L}(\Theta|x)}{\partial\Theta_k\partial\Theta_1}\\ \vdots & \ddots & \vdots\\ \frac{\partial^2\mathcal{L}(\Theta|x)}{\partial\Theta_1\partial\Theta_k} & \cdots & \frac{\partial^2\mathcal{L}(\Theta|x)}{\partial\Theta_k^2}\end{bmatrix}$$

One can use the normal distribution of $\widehat{\Theta}$ to construct approximate confidence interval regions for some parameters. Indeed, an asymptotic $100(1-\xi)$ confidence interval for each parameter $\Theta_j;j=1,\dots,k$, is given by

$$\left(\widehat{\Theta}_j-z_{\xi/2}\sqrt{\hat{\daleth}_{jj}},\quad \widehat{\Theta}_j+z_{\xi/2}\sqrt{\hat{\daleth}_{jj}}\right)$$

where $\hat{\daleth}_{jj}$ denotes the (i,i) diagonal element of $I_n^{-1}(\widehat{\Theta})$ and $z_{\xi/2}$ is the $\left(\mathbf{1}-\xi/2\right)_{\mathbf{th}}$ quantile of the standard normal distribution.

5. APPLICATION OF REAL DATA

In this section, we illustrate the empirical importance of the discrete distributions as DB, DLi, Binom, Pois, DR, DGE, Geometric, DW, DIW, discrete exponential (DE), NDL, DGzEx, DMOGE, DAPLo, and EDW distributions using four applications to real data. The fitted models are compared using some criteria; namely, Akaike information criterion (AIC), corrected AIC (CAIC), Hannan-Quinn information criterion (HQIC), Chi-square (χ2) with a degree of freedom and its p-value.

$$AIC=-2\mathcal{L}(\widehat{\Theta}|x)+2p,$$

$$CAIC=-2\mathcal{L}(\widehat{\Theta}|x)+2p+2\frac{p(p+1)}{n-p-1},$$

$$BIC=-2\mathcal{L}(\widehat{\Theta}|x)+p\log(n),$$

and

$$HQIC = -2\mathcal{L}(\widehat{\Theta}|x) + 2\,p\log[\log(n)].$$

5.1.Angola

This data represents the daily new deaths of 51 days from 10 October to 29 November 2020 belong to Angola country (see World Health Organization). The MLEs and the goodness of fit statistics are reported in Table 1.

Regarding Table 1, It is evident that all distributions are Fitted and work quite well for analyzing these data except for the DB distribution. However, we always search for the best model to get the best evaluation of the data, and therefore, concerning the AIC, BIC, CAIC, HQIC, X^2, and p-values, we can say that the DW model provides the best fit among all the tested models because it has the smallest values of AIC, CAIC, BIC, HQIC, X^2 and X^2 statistics, as well as having the highest p-value. Figure 1 supports the results of Table 1.

5.2.El Salvador

This data represents the daily new deaths of 81 days from 1 April to 20 June 2020 belong to El Salvador country (see World Health Organization). The MLEs and the goodness of fit statistics are reported in Table 2.

Regarding Table 2, It is evident that all distributions are Fitted and work immensely well for analyzing these data except for the DR distribution. However, we always search for the best model to get the best evaluation of the data, and therefore, concerning the AIC, BIC, CAIC, HQIC, X^2, and p-values, we can say that the DL model provides the best fit among all the tested models because it has the smallest values of AIC, CAIC, BIC, HQIC, X^2 and X^2 statistics, as well as having the highest p-value. Figure 2 supports the results of Table 2.

5.3.Estonia

This data represents the daily new deaths of 81 days from 1 April to 20 May 2020 belong to Estonia country (see World Health Organization). The MLEs and the goodness of fit statistics are reported in Table 3.

Regarding Table 3, It is evident that all distributions are Fitted and work quite well for analyzing these data except for the Binom, Pois, and DR distribution. However, we always search for the best model to get the best evaluation of the data, and therefore, concerning the AIC, BIC, CAIC, HQIC, X^2, and p-values, we can say that the DL model provides the best fit among all the tested models because it has the smallest values of AIC, CAIC, BIC, HQIC, X^2 and X^2 statistics, as well as having the highest p-value. Figure 3 supports the results of Table 3.

5.4.Greece

This data represents the daily new deaths of 111 days from 12 March to 30 June 2020 belong to Greece country (see World Health Organization). The MLEs and the goodness of fit statistics are reported in Table 4.

Regarding Table 4, It is evident that all distributions are Fitted and work quite well for analyzing these data except for the DB, Binom, Pois, DIW, and DR distribution. However, we always search for the best model to get the best evaluation of the data, and therefore, concerning the AIC, BIC, CAIC, HQIC, X^2, and p-values, we can say that the DE model provides the best fit among all the tested models because it has the smallest values of AIC, CAIC, BIC, HQIC, X^2 and X^2 statistics, as well as having the highest p-value. Figure 4 supports the results of Table 4.

6. CONCLUDING REMARKS

In this article, we use 15 discrete distribution to fit and determine the best model of daily Coronavirus deaths in some countries as Angola, El Salvador, Estonia, and Greece. We proposed and studied the new DIW distribution.

Table 1. The goodness of fit statistics for Dataset of Angola

value	count	DB	DL	Binom	pois	DR	DGE	Geo	DW	DIW	DE	NDL	DGzEx	DMOGE	DAPL	EDW
0	5	6.781	10.385	3.720	3.475	3.641	4.703	13.835	4.967	3.390	13.835	11.173	7.024	5.394	4.196	6.075
1	13	18.900	10.383	9.493	9.334	9.436	11.921	10.082	10.218	16.933	10.082	10.219	8.392	9.718	12.654	9.045
2	8	8.340	8.661	12.351	12.537	11.737	11.638	7.347	11.220	10.953	7.347	8.308	9.291	11.037	11.702	9.977
3	7	4.236	6.620	10.919	11.226	10.594	8.590	5.354	9.527	6.046	5.354	6.333	9.231	9.603	7.865	9.450
4	8	2.547	4.805	7.377	7.539	7.586	5.633	3.901	6.775	3.586	3.902	4.634	7.865	6.724	4.865	7.605
5	6	1.699	3.371	4.061	4.050	4.461	3.484	2.843	4.172	2.294	2.843	3.296	5.394	4.037	2.998	5.013
6	3	1.214	2.308	1.897	1.813	2.191	2.089	2.072	2.265	1.559	2.072	2.297	2.726	2.205	1.893	2.573
7	1	0.910	1.552	0.773	0.696	0.908	1.232	1.510	1.096	1.111	1.510	1.576	0.899	1.142	1.236	0.967
theta		4.860	0.593	0.950	2.686	0.929	0.577	0.271	0.903	0.066	0.729	0.390	0.894	1.657	0.001	3.021
alpha		0.814					2.773		1.786	1.559			1.105	0.490	2.531	0.441
lambda													0.323	4.117	0.031	0.992
X-squared		24.861	7.473	6.309	7.169	5.075	4.582	13.899	3.206	13.327	13.898	8.561	3.927	4.060	6.391	3.213
P-Value		0.001	0.381	0.504	0.411	0.651	0.711	0.053	0.869	0.065	0.053	0.286	0.788	0.773	0.495	0.865
AIC		233.040	212.202	207.594	209.245	207.438	208.278	221.781	205.509	222.517	221.781	214.016	206.435	208.733	215.744	206.090
CAIC		233.290	212.283	207.676	209.327	207.520	208.528	221.863	205.759	222.767	221.863	214.097	206.945	209.244	216.254	206.601
BIC		236.904	214.133	212.526	211.977	211.870	212.142	223.713	209.372	226.381	223.713	215.948	211.230	214.529	221.539	210.886
HQIC		234.517	212.940	208.332	209.983	208.976	209.755	222.519	206.985	223.994	222.519	214.754	208.649	210.948	217.958	208.305

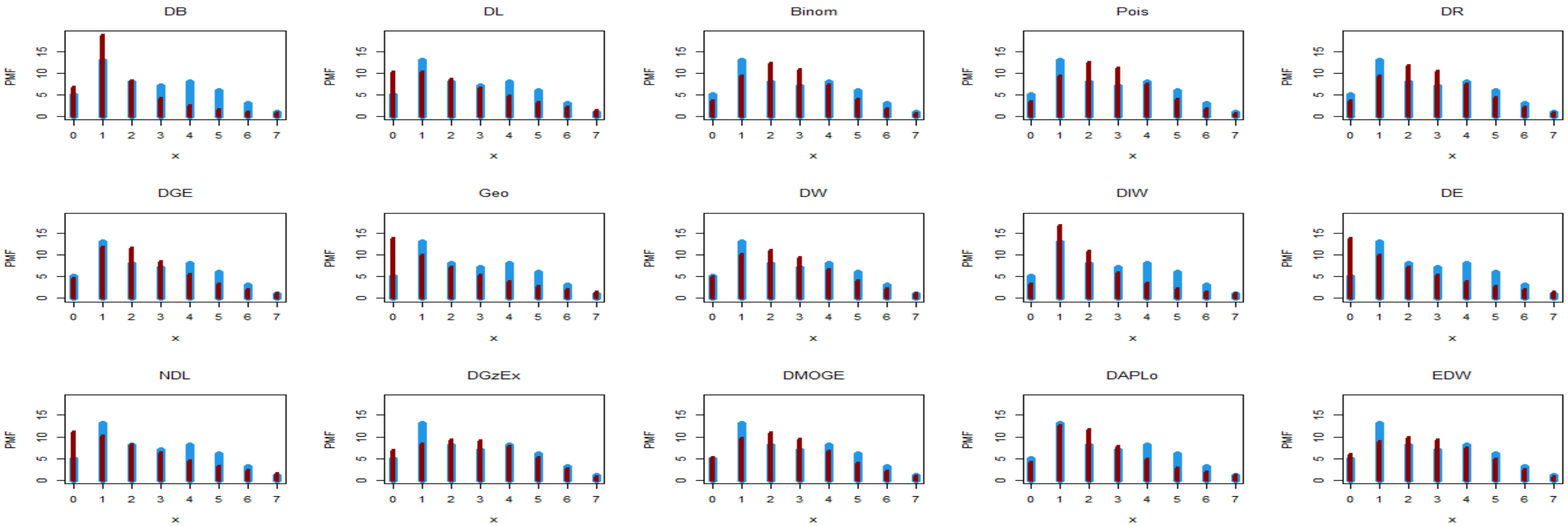

Figure 1. The fitted PMFs for Dataset of Angola.

Table 2. The goodness of fit statistics for Dataset of El Salvador

value	count	DB	DL	binom	pois	DR	DGE	Geo	DW	DIW	DE	NDL	DGzEx	DMOGE	DAPL	EDW
0	34	34.404	35.894	28.208	28.015	18.501	33.839	39.287	33.710	33.080	39.287	36.898	34.454	34.042	35.691	34.327
1	25	28.005	22.759	29.562	29.744	33.789	25.230	20.232	24.660	29.502	20.232	21.888	22.818	24.167	25.218	22.983
2	11	9.339	11.964	15.682	15.790	20.858	12.254	10.419	12.831	8.932	10.419	11.541	13.271	12.957	9.957	13.358
3	6	3.791	5.751	5.613	5.588	6.574	5.487	5.365	5.852	3.686	5.365	5.705	6.597	5.830	4.271	6.494
4	4	1.867	2.623	1.525	1.483	1.155	2.390	2.763	2.446	1.868	2.763	2.708	2.711	2.414	2.088	2.629
5	1	1.050	1.156	0.335	0.315	0.117	1.029	1.423	0.957	1.079	1.423	1.249	0.886	0.965	1.142	0.886
theta		2.414	0.372	0.987	1.062	0.772	0.427	0.485	0.584	0.408	0.515	0.605	0.661	0.940	0.002	1.710
alpha		0.450					1.566		1.245	1.795			1.212	0.390	2.171	0.542
lambda													0.162	2.329	0.124	0.795
X-squared		4.134	1.132	8.642	9.271	33.669	1.246	2.545	1.249	4.805	2.545	1.348	1.280	1.357	2.488	1.357
P-Value		0.530	0.951	0.124	0.099	0.000	0.940	0.770	0.940	0.440	0.770	0.930	0.937	0.929	0.778	0.929
AIC		238.453	230.577	235.033	235.447	253.201	232.446	233.361	232.019	239.779	233.361	231.128	233.570	234.283	239.291	233.750
CAIC		238.607	230.628	235.084	235.498	253.251	232.599	233.412	232.173	239.933	233.412	231.178	233.882	234.594	239.603	234.062
BIC		243.242	232.972	237.428	237.842	255.595	237.234	235.756	236.808	244.568	235.756	233.522	240.754	241.466	246.475	240.934
HQIC		240.374	231.538	235.994	236.408	254.161	234.367	234.322	233.940	241.700	234.322	232.089	236.452	237.165	242.174	236.632

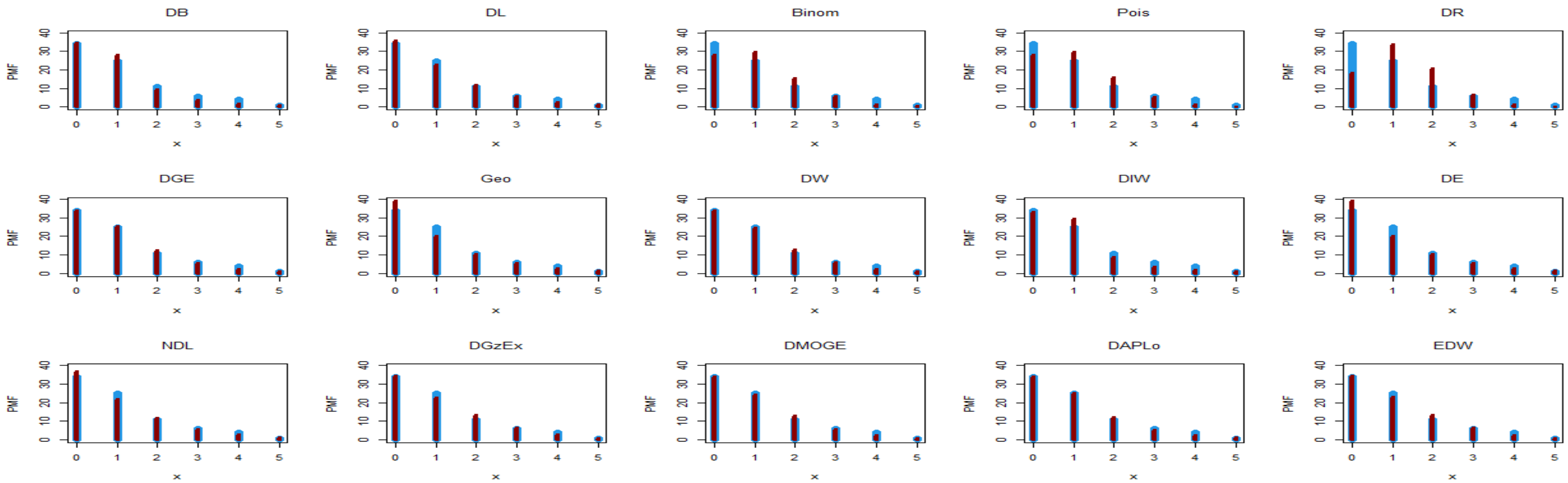

Figure 2. The fitted PMFs for Dataset of El Salvador.

Table 3. The goodness of fit statistics for Dataset of Estonia

value	count	DB	DL	binom	pois	DR	DGE	Geo	DW	DIW	DE	NDL	DGzEx	DMOGE	DAPL	EDW
0	20	20.314	20.413	15.268	15.058	9.516	19.946	22.727	19.891	19.378	22.727	21.090	20.373	20.050	19.588	20.197
1	15	16.759	13.910	17.899	18.071	18.994	14.684	12.397	14.364	17.749	12.397	13.393	13.367	14.148	15.653	13.602
2	6	6.051	7.866	10.701	10.844	14.011	7.805	6.762	8.042	5.858	6.762	7.560	8.086	8.085	7.750	8.177
3	5	2.613	4.069	4.349	4.338	5.772	3.897	3.688	4.108	2.553	3.688	4.001	4.458	4.102	3.574	4.390
4	3	1.348	1.998	1.351	1.301	1.451	1.901	2.012	1.978	1.344	2.012	2.033	2.211	1.956	1.668	2.124
5	0	0.786	0.947	0.342	0.312	0.230	0.918	1.097	0.912	0.800	1.097	1.004	0.971	0.905	0.809	0.932
6	1	0.499	0.438	0.074	0.062	0.023	0.442	0.599	0.406	0.517	0.599	0.486	0.371	0.413	0.412	0.373
theta		2.334	0.400	0.977	1.200	0.810	0.479	0.455	0.602	0.388	0.545	0.577	0.676	1.023	0.000	1.487
alpha		0.471					1.410		1.188	1.671			1.246	0.451	8.577	0.644
lambda													0.110	1.763	0.220	0.755
X-squared		5.445	2.896	18.123	21.001	59.692	2.969	3.221	3.032	5.799	3.221	2.801	3.119	3.059	3.669	3.141
P-Value		0.488	0.822	0.006	0.002	0.000	0.813	0.781	0.805	0.446	0.781	0.833	0.794	0.801	0.721	0.791
AIC		158.573	152.290	157.901	158.584	170.826	154.399	153.582	154.216	159.305	153.582	152.457	156.043	156.336	157.083	156.116
CAIC		158.828	152.373	157.985	158.667	170.910	154.655	153.665	154.472	159.560	153.665	152.541	156.565	156.858	157.605	156.638
BIC		162.397	154.202	159.814	160.496	172.738	158.223	155.494	158.040	163.129	155.494	154.369	161.779	162.072	162.819	161.852
HQIC		160.029	153.018	158.630	159.312	171.554	155.855	154.310	155.673	160.761	154.310	153.185	158.228	158.521	159.267	158.301

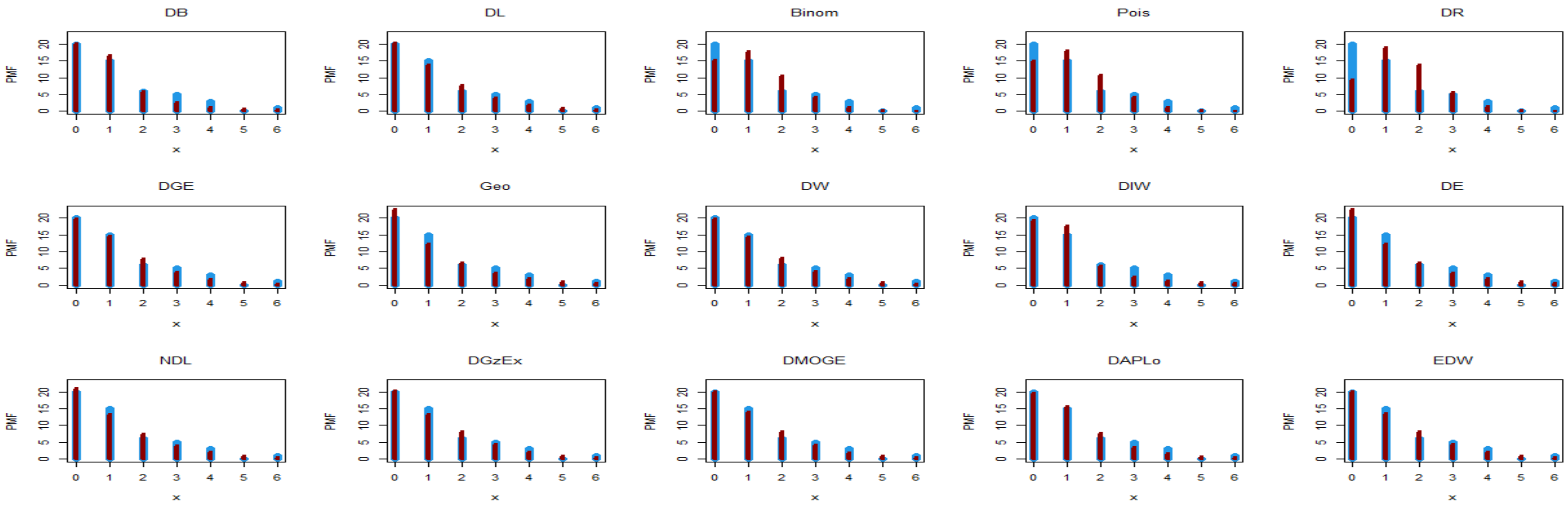

Figure 3. The fitted PMFs for Dataset of Estonia

Table 4. The goodness of fit statistics for Dataset of Greece

value	count	DB	DL	binom	pois	DR	DGE	Geo	DW	DIW	DE	NDL	DGzEx	DMOGE	DAPL	EDW
0	39	39.973	34.353	20.121	19.860	12.202	38.434	40.798	37.670	36.667	40.798	36.079	37.142	39.100	38.256	39.647
1	26	33.012	28.330	34.099	34.176	29.132	27.135	25.803	27.174	35.507	25.803	27.465	26.213	24.518	28.750	23.258
2	17	14.343	19.440	29.154	29.405	30.749	17.284	16.319	17.685	14.478	16.319	18.585	17.918	17.657	17.394	16.952
3	9	7.182	12.213	16.766	16.867	21.694	10.791	10.321	11.146	7.250	10.321	11.790	11.833	11.777	10.234	12.058
4	6	4.137	7.283	7.295	7.256	11.184	6.681	6.527	6.891	4.213	6.527	7.180	7.531	7.385	6.057	8.119
5	7	2.631	4.197	2.562	2.497	4.361	4.118	4.128	4.203	2.701	4.128	4.251	4.606	4.445	3.648	5.101
6	6	1.795	2.360	0.756	0.716	1.308	2.533	2.611	2.536	1.856	2.611	2.466	2.699	2.608	2.247	2.963
7	0	1.291	1.303	0.193	0.176	0.305	1.556	1.651	1.517	1.342	1.651	1.408	1.510	1.507	1.417	1.580
8	0	0.966	0.709	0.043	0.038	0.055	0.955	1.044	0.901	1.008	1.044	0.794	0.804	0.864	0.914	0.769
9	1	0.746	0.382	0.009	0.007	0.008	0.586	0.661	0.532	0.781	0.661	0.443	0.406	0.492	0.604	0.341
theta		2.097	0.487	0.985	1.721	0.890	0.613	0.368	0.661	0.330	0.632	0.492	0.756	0.562	0.000	1.893
alpha		0.525					1.117		1.082	1.364			1.406	0.566	16.773	0.378
lambda													0.052	3.070	0.293	0.935
X-squared		21.525	12.646	184.826	212.928	218.335	9.902	9.479	10.003	21.504	9.478	10.992	9.568	9.742	12.190	9.585
P-Value		0.011	0.179	0.000	0.000	0.000	0.358	0.393	0.350	0.011	0.394	0.276	0.387	0.372	0.203	0.389
AIC		416.943	399.550	440.250	442.207	462.050	400.857	399.219	400.506	418.319	399.214	401.673	401.406	402.214	405.199	400.974
CAIC		417.055	399.587	440.287	442.244	462.087	400.968	399.259	400.617	418.430	399.250	401.710	401.631	402.438	405.423	401.199
BIC		422.362	402.259	442.960	444.916	464.760	406.276	401.926	405.925	423.738	401.923	404.382	409.535	410.342	413.327	409.103
HQIC		419.142	400.649	441.349	443.306	463.150	403.055	400.316	402.704	420.517	400.313	402.772	404.704	405.511	408.496	404.272

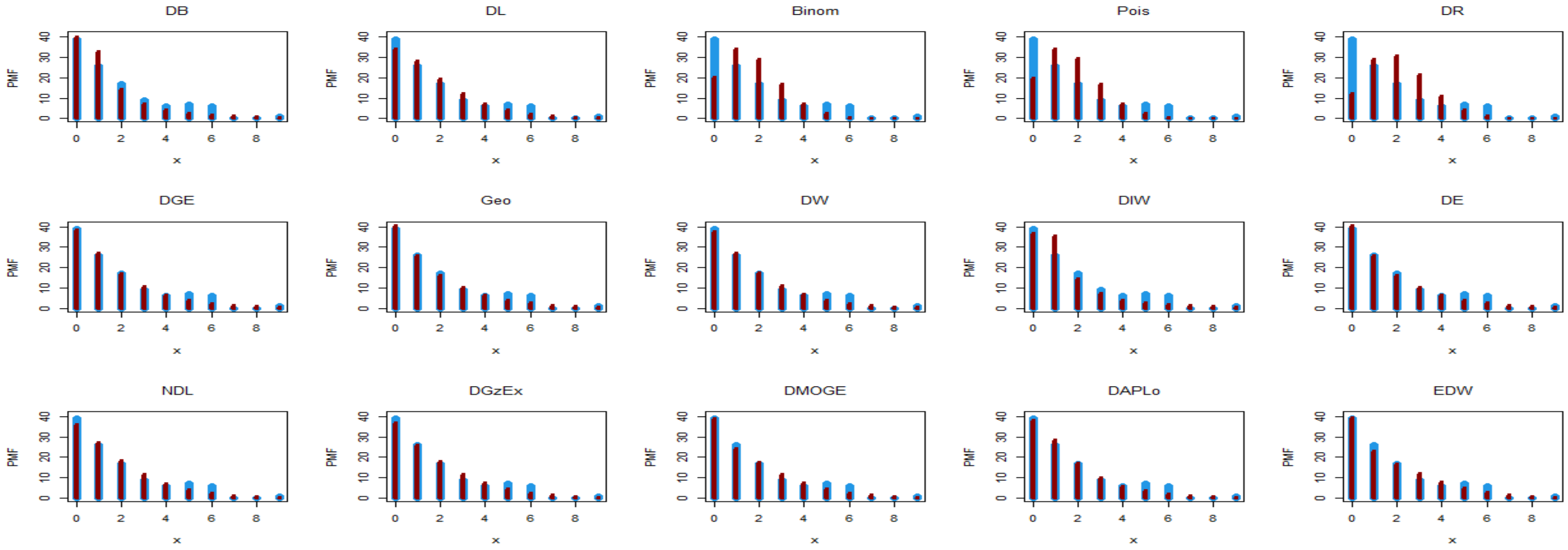

Figure 4. The fitted PMFs for Dataset of Greece

A review of some important discrete distributions has been provided as DB, DL, DMOGE, DGE, DAPL, DR, DE, Geometric, Binomial, NDL, DGzEx, and EDW distribution. The maximum likelihood estimation method is discussed to estimate the parameter of the discrete distribution. We prove empirically that the discrete models fit different datasets of daily Coronavirus deaths in some countries as Angola, El Salvador, Estonia, and Greece. DW and DB reveal its superiority over other competitive models for the analysis of daily deaths of the COVID-19 in the case of Angola, El Salvador, Estonia, and Greece.

REFERENCES

[1] Al-Babtain, A. A., Ahmed, A. H. N., Afify, A. Z. (2020). A New Discrete Analog of the Continuous Lindley Distribution, with Reliability Applications. Entropy, 22(6), 603.

[2] Almetwally, E. M., Ibrahim, G. M. Discrete Alpha Power Inverse Lomax Distribution with Application of COVID-19 Data. International Journal of Applied Mathematics, 9(6), 11-22.

[3] Almetwally, E. M., Almongy, H. M., & ElSherpieny, E. A. (2019). Adaptive type-II progressive censoring schemes based on maximum product spacing with the application of generalized Rayleigh distribution. Journal of Data Science, 17(4), 802-831.

[4] Almetwally, E. M., Almongy, H. M., & Saleh, H. A. (2020). Managing risk of spreading "COVID-19" in Egypt: Modelling using a discrete Marshall-Olkin generalized exponential distribution. International Journal of Probability and Statistics, 9(2), 33-41.

[5] Almetwaly, E. M., Almongy, H. M. (2018). Estimation of the generalized power Weibull distribution parameters using progressive censoring schemes. International Journal of Probability and Statistics, 7(2), 51-61.

[6] Balakrishnan, N., Cramer, E. (2014). The art of progressive censoring. Statistics for industry and technology.

[7] Batista, M. (2020a). Estimation of the final size of the coronavirus epidemic by the logistic model. doi.org/10.1101/2020.02.16.20023606.

[8] Batista, M. (2020b). Estimation of the final size of the coronavirus epidemic by the SIR model. Online paper, ResearchGate.

[9] Eliwa, M. S., Alhussain, Z. A., El-Morshedy, M. (2020). Discrete Gompertz-G family of distributions for over-and under-dispersed data with properties, estimation, and applications. Mathematics, 8(3), 358.

[10] El-Morshedy M., Altun E, Eliwa MS. (2020) A new statistical approach to model the counts of novel coronavirus cases. Research Square. DOI: 10.21203/rs.3.rs-31163/v1.

[11] Gómez-Déniz, E., Calderín-Ojeda, E. (2011). The discrete Lindley distribution: properties and applications. *Journal of Statistical Computation and Simulation*, *81*(11), 1405-1416.

[12] Hasab, A. A., El-Ghitany, E. M., Ahmed, N. N. (2020). Situational Analysis and Epidemic Modeling of COVID-19 in Egypt. *Journal of High Institute of Public Health*, *50*(1), 46-51.

[13] https://covid19.who.int/

[14] https://www.worldometers.info/coronavirus/

[15] Jazi, M. A., Lai, C. D., & Alamatsaz, M. H. (2010). A discrete inverse Weibull distribution and estimation of its parameters. Statistical Methodology, 7(2), 121-132.

[16] Krishna, H., & Pundir, P. S. (2009). Discrete Burr and discrete Pareto distributions. *Statistical Methodology*, *6*(2), 177-188.

[17] Maleki, M., Mahmoudi, M. R., Wraith, D., & Pho, K. H. (2020). Time series modeling to forecast the confirmed and recovered cases of COVID-19. Travel Medicine and Infectious Disease, 101742.

[18] Nakagawa, T., & Osaki, S. (1975). The discrete Weibull distribution. IEEE transactions on reliability, 24(5), 300-301.

[19] Nekoukhou, V., & Bidram, H. (2015). The exponentiated discrete Weibull distribution. Sort, 39, 127-146.

[20] Nekoukhou, V., Alamatsaz, M. H., & Bidram, H. (2013). Discrete generalized exponential distribution of a second type. *Statistics*, *47*(4), 876-887.

[21] Nesteruk, I. (2020). Statistics-based predictions of coronavirus epidemic spreading in mainland China, Innovative Biosystems and Bioengineering, 4(1), 13-18.

[22] Roy, D. (2003). The discrete normal distribution. *Communications in Statistics-theory and Methods*, *32*(10), 1871-1883.

[23] Roy, D. (2004). Discrete rayleigh distribution. *IEEE Transactions on Reliability*, *53*(2), 255-260.

* Corresponding author: Ehab M. Almetwally
Tel.: +0201274450338
E-mail address: ehabxp_2009@hotmail.com
Received: 30 November 2020
Received in revised form
Accepted 6 August 2021
Available online