International Journal of Mathematics Trends and Technology
ISSN: 2231-5373/ https://doi.org/10.14445/22315373/IJMTT-V68I12P512

Volume 68 Issue 12, 106-111, December 2022

Original Article

The Modelling Number of Daily Death Covid 19 Data Using Some of Two and One Parameter Distributions in Indonesia

Anisha Putri Ramadhani[1], Rado Yendra[2], Muhammad Marizal[3], Rahmadeni[4]

[1,2,3,4] *Department of Mathematics, Faculty of Science and Technology, Universitas Islan Negeri Sultan Syarif Kasim Riau, Indonesia*

Received: 30 October 2022 Revised: 05 December 2022 Accepted: 17 December 2022 Published: 31 December 2022

Abstract - *The outbreak of COVID-19 pandemic has caused many deaths. The number of deaths that occur randomly per day has caused various problems, including reduced population, increased land for graves, and increased fear of the situation that has arisen which indirectly impacts decreased performance which ultimately weakens a country's economy. However, it is the concern of governments and others' responsibilities to provide the correct statistics and figures to take any practicable necessary steps. Where the statistical literature supposes that a model governs every real phenomenon, once we know the model, we can evaluate the dilemma. Therefore, in this article, we compare and analyze the frequency of a number of death from COVID-19 in Indonesia using the six probability modeling. Probability modeling will be carried out using six distributions, namely Weibull, Gamma, Log Normal, Amarendra, Rani and Akash will be used in this study. The maximum likelihood method will be used to get the estimated parameter from the models used in this study. The distribution will be selected based on some methods of Good of Fit Test namely graphical (pdf plot and cumulative distribution function (CDF) plot) and numerical (Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC), and Log Likelihood (Log-L)). In most cases, graphical methods give the same results but their AIC and BIC results are different. The most suitable result is selected as the distribution with the lowest AIC and BIC values. In general, the Weibull distribution has been chosen as the best model.*

Keywords – *Covid 19, Weibull distribution, Gamma distribution, Amarendra distribution, Log normal distribution, Rani distribution, Akash Distribution.*

1. Introduction

In December 2019, a localized outbreak of atypical pneumonia occurred in Wuhan, the capital city of Hubei province, China. It was quickly discovered to be associated with the novel coronavirus. The World Health Organization (WHO) calls it Coronavirus Disease 2019 (covid 19). Because covid 19 is highly contagious, the number of death cases in China is increasing significantly and rapidly. Therefore, on January 30, WHO declared that the event had become a Public Health Emergency of International Concern [1]. The virus is emerging rapidly in countries outside of China due to international travel and transportation, and the number of deaths is rapidly increasing worldwide. Until now it seems as if the Covid 19 virus cannot be eradicated, even deaths caused by this virus continue to occur with the number of deaths occurring randomly every day. This number of deaths certainly has a negative impact on the economic, social, health, and education fields.[2]. Therefore, the number of death data is very important to analyze, especially in analyzing the frequency of the number of deaths. Frequency analysis is related to the probability of events, therefore frequency analysis is often associated with probability density functions or probability models. Most of the research conducted using probability models is to choose the best model from the several models used. The use of several probability models to analyze the frequency of an event has been carried out by previous researchers, including, Alzaatreh and Famoye [3], Al-Babtain et al. [4], Jayakumar and Mathew [5], Cordeiro et al. [6], Oluyede et al. [7], Mansour and Mahdy [8], Maurya et al. [9], Abouammoh and Kayid [10], Chen [11], Mansour et al. [12] and Tahir et al. [13]. Modeling the probability of daily new cases and death covid 19 using statistical methods is one of the most important topics today. Researchers in various fields of study have made use of probability distributions in modeling and analyzing the frequency of daily cases covid 19 data. Below are brief literature reviews on the application of the statistical distribution, Soltani-Kermanshahi et al. [14] conducted research on probability modeling of new coronavirus data in Iran. This study compares three types of parametric distributions known as normal, log-normal, and Weibull distributions of covid 19 cases based on daily data reported. Yusuf et al. [15] Forecasting monthly data on patients with covid 19 disease that occurred in Pakistan. Zafar, J. et al [16] Compared several daily data models of covid 19 patients in Pakistan. The model used is a 2-parameter probability distribution, namely Weibull distribution (WD), Power function distribution (PFD), Log-Logistics Distribution (LLD), Log-Normal Distribution (LND), Inverse Weibull Distribution (IWD), Gumbel Distribution (GuD). , Burr III Distribution

(BIIID), Burr XII Distribution (BXIID), and Birnbaum Saunders Distribution (BSD). However, research like this has not been widely applied to data on daily deaths caused by covid 19. Several studies have started to analyze the frequency of daily death caused by covid 19 using probability models such as [17,18,19], Not many studies have been based on probability modeling on the death date of covid 19 Indonesia. This study aims to select the best model for analyzing the frequency of daily death cases covid 19 in Indonesia, six probability models consisting of two and one distribution parameters will be used in this paper, namely the distribution of Weibull, Gamma, Normal Log, Rani, Amarendra and Akash. The MLE (Maximum likelihood) method will be used in estimating Parameter. The several methods for Good of Fit test in determining the best model, such as the graphical method and the numerical method, will be carried out in this study. The graphical method will be carried out by comparing the pdf and CDF of each model while the numerical values of AIC and BIC will complete this research as a representation of the numerical method. Model selection using the graphical method often gives the same results for each model but using the numerical method will result in the selection of the best model. easy based on the lowest value generated by AIC and BIC.

2. Data and Study Area

The daily death of covid 19 data was obtained for the period March 12 2020 to May 11, 2020, from the public report of the National Institute of Health, Indonesia. Summary of data presented in descriptive statistics as in table 1. In the table it can be concluded that the skewness value is close to zero and the kurtosis value does exceed 2, this indirectly indicates that some distributions with 2 parameters such as Weibull, Gamma, and Log normal, and 1 parameter such as Rani, Amarendra and Akash can be used in this modeling

Table 1. Desciptive Statistics of New Cases Positif Covid 19 in Indonesia

Mean	Standrard deviasion	Minimum	Maximum	Skewness	Kurtosis
17.375	12.40903	1	60	80.00035	4.67

3. Methods

3.1. Probability Density Function (PDF) and Cumulative Distribution Function (CDF)

The new death covid 19 modeling requires analysis of the number of daily death caused by covid 19 data over a number of years. The primary tools to describe the daily death covid 19 characteristics are probability density functions. Six probability density functions are divided into 3 probability density functions which have 2 parameters such as Weibull, Gamma, and Log-Normal and the others have one parameter namely Rani, Amarendra, and Akash. The pdf and CDF for each distribution that we consider are as given in Table 2, where y denotes the observed values of the random variable representing the event of interest. Several researchers have used the probability density function or distribution of Rani, Amarendra, and Akash in this study for various purposes [20, 21, 22, 23, 24, 25, 26]

Table 2. Probability Density Function (pdf) and distribution function (cdf) four distributions

	Distribution	pdf (*f(y)*) dan cdf (*F(y)*)
1	Weibull (x;η,κ)	$f(x) = \frac{\eta}{\kappa}\left(\frac{x}{\kappa}\right)^{\eta-1} e^{\left(-\left(\frac{x}{\kappa}\right)^{\eta}\right)}, x > 0,\ \ \eta, \kappa > 0$ $F(x) = 1 - e^{\left(-\left(\frac{x}{\kappa}\right)^{\eta}\right)}$
2	Gamma (x;α,β)	$f(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} x^{\alpha-1} e^{-\frac{x}{\beta}}, \quad x > 0, \quad \alpha, \beta > 0$ $F(x) = \frac{1}{\beta^{\alpha}\Gamma(\alpha)} \int_{o}^{x} x^{\alpha-1} e^{-x/\beta} dx$
3	Log Normal (x;μ,σ)	$f(x) = \frac{1}{\sqrt{2\pi\sigma x}} e^{\frac{-(\log x - \mu)}{2\sigma^2}}, x > 0, \quad \mu, \sigma > 0$ $F(x) = \Phi\left(\frac{\log x - \mu}{\sigma}\right)$ where Φ is the cumulative distribution function of the standard normal distribution (i.e., *N*(0,1)).
4	Rani (x;θ)	$f(x) = \frac{\theta^5}{\theta^5 + 24}(\theta + x^4) e^{-x\theta}, x > 0, \theta > 0$ $F(x) = 1 - \left[1 + \frac{\theta x(\theta^3 x^3 + 4\theta^2 x^2 + 12\theta x + 24)}{\theta^5 + 24}\right] e^{-\theta x}$
5	Amarendra(x;Ω)	$f(x) = \frac{\Omega^4}{\Omega^3 + \Omega^2 + 2\Omega + 6}(1 + x + x^2 + x^3) e^{-x\beta}\ , x > 0, \Omega > 0$

		$$F(x) = 1 - \left[1 + \frac{\Omega^3 x^3 + \Omega^2(\Omega+3)x^2 + \Omega(\Omega^2 + 2\Omega + 6)x}{\Omega^3 + \Omega^2 + 2\Omega + 6}\right] e^{-x\Omega}$$
6	Akash (x;λ)	$f(x) = \frac{\lambda^3}{\lambda^2+2}(1+x^2)e^{-x\lambda}, \quad x > 0, \lambda > 0$ $$F(x) = 1 - \left(1 + \frac{\lambda x(\lambda x + 2)}{\lambda^2 + 2}\right) e^{-\lambda x}$$

To select the most suitable model, parameter estimation is the first important thing to do. The parameter estimation of the distribution function is calculated using the maximum likelihood method. The maximum likelihood function for this model is implicit and complex and we will not go into details in this paper. The nonlinear equation generated by the maximum log-likelihood function (ln L) requires a numerical method, namely Newton's Rapshon, to obtain a solution to the equation. But this method has been used in iteration systems to find the solution. Several initial values have been tested for this procedure. If the initial value used causes iterations to be carried out towards a certain value or the iterations converge to a value, then that value can be considered as the selected estimation parameter. Conformity testing procedures for model selection, both numerically and graphically, are discussed.

3.2. Maximum Likelihood Estimate (MLE) and Goodness of Fit Tests (GOF)

Let $(y_1, y_2, \cdots, y_n)$ be a random sample from four PDFs, The natural log likelihood (ln L) is presented in Table 3. The MLE $\hat{\theta}\ of\ \theta$ is the solution of the equation $\frac{d \ln L}{d\theta} = 0$ and thus it is the solution of the following nonlinear equation. The most appropriate distribution is identified using results found based on several goodness-of-fit tests. The GOF tests considered are based on graphical inspection probability density function (PDF) and numerical criteria Akaike's information criterion (AIC) and Bayesian information criterion (BIC) were applied to determine the goodness-of-fit criteria of the distributions. In most cases, graphical inspection gave the same result but their AIC and BIC results differed. The best fit result was chosen as the distribution with the lowest values of AIC and BIC. The formula for computing AIC and BIC are as follows:

$$\text{AIC} = -2 \ln L + 2k, \quad \text{BIC} = -2 \ln L + k \ln n,$$

where k = the number of parameters, n = the sample size

Table 3. Computed parameters of different distribution

	η	κ	α	β	μ	σ	θ	Ω	λ
Weibull	1.45	19.2							
Gamma			1.86	9.32					
Log Normal					2.56	0.87			
Rani							0.29		
Aramendra								0.23	
Akash									0.06

4. Results

In this study, the daily data histogram of new Covid 19 cases in Indonesia will be approximated by the four distributions or probability density functions used in this study. This of course can be obtained based on the estimated parameter values as presented in table 3. The parameters of the distributions are estimated using the maximum likelihood estimation method.

In this study, the histogram of the number of daily deaths due to Covid-19 in Indonesia will be approximated by the six models or distributions used in this study as depicted in Figure 1. In this figure, it can be seen that the six models used have different capabilities in approaching the histogram. Figure 1 shows that the two parameters distribution such as Weibull, Gamma, and Log Normal is able to approach the histogram or the frequency of daily death covid 19 data that occurs in Indonesia distribution is not good in approaching the frequency of daily Covid-19 case data in Indonesia, while the one parameters distribution namely Rani, Amarendra and Akash are also not very good at capturing the frequency of these data.

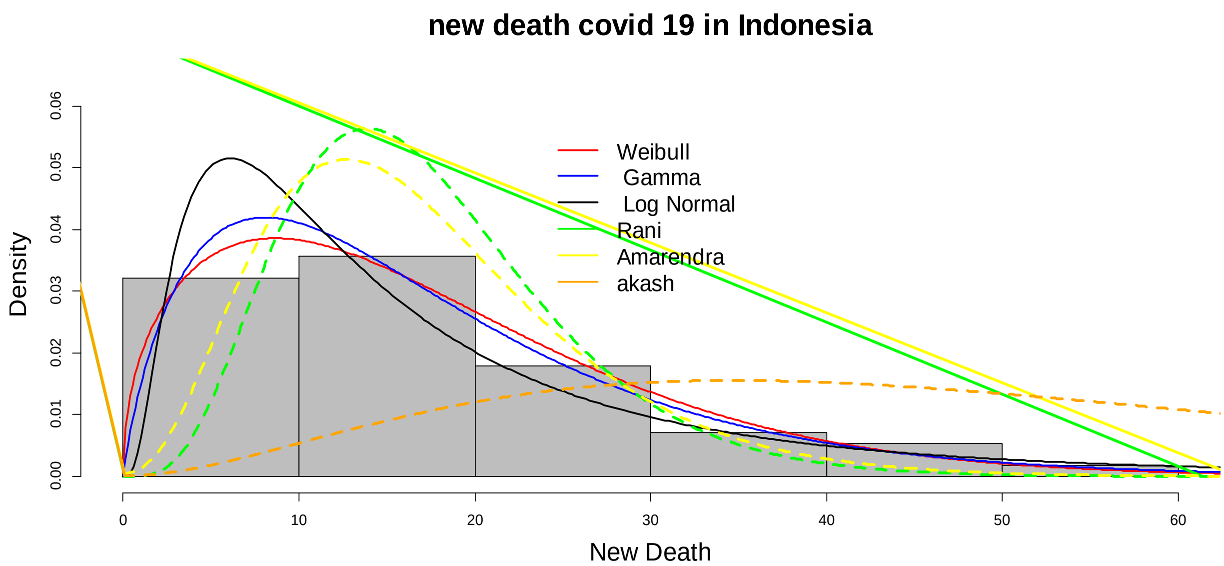

Fig. 1 The Frequency of Number Of Death Caused by Covid 19 in Indobesia and Distribution Models.

To further clarify the merits of the model used in this study, each model will be described in pairs with a data histogram as depicted in Figure 2. Figure 2 shows that the Weibull and Gamma distributions are graphically depicted in red and blue respectively. These two distributions are the best of all the models or distributions used in the study. Figure 2 also shows that the distribution that uses 1 parameter is not good at approaching the histogram data, the Akash distribution is the worst at approaching daily death data caused by Covid 19 in Indonesia.

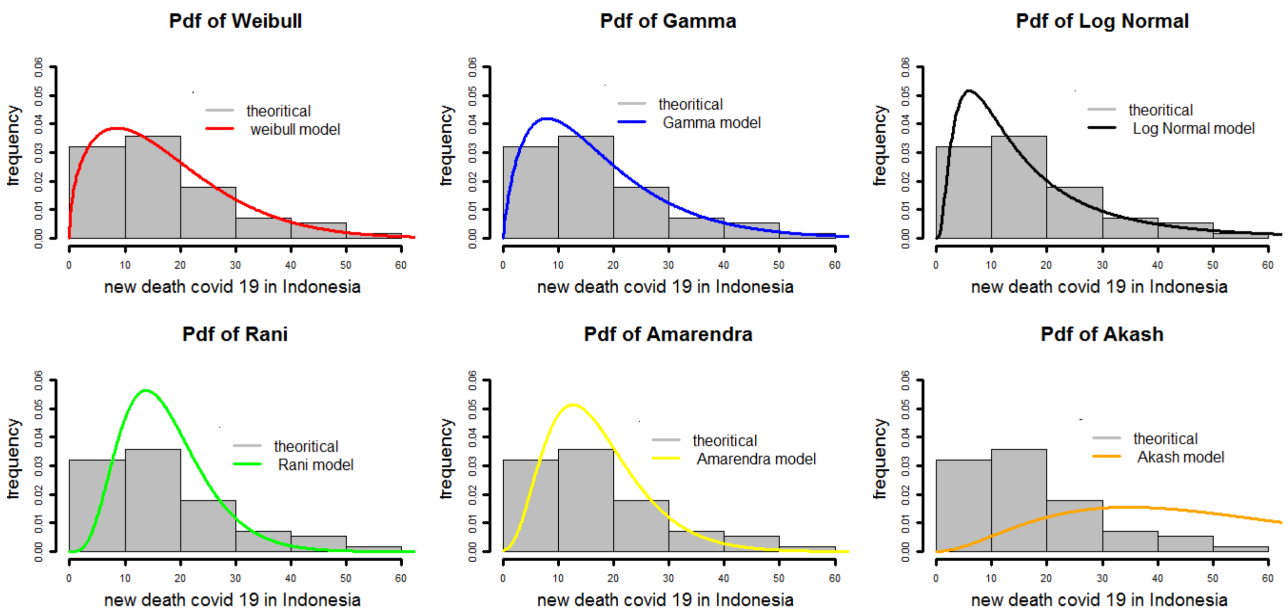

Fig. 2 Pdf plot for comparisons Predicted and observed daily death cases covid 19 in Indonesia

Furthermore, the plot of the cumulative distribution function will be used to make more convincing conclusions in choosing the best distribution in this study. Therefore Figure 3 is also presented for this purpose. From the figure it is very clear that the one-parameter distribution such us Rani, Amarendra, and Akash are not good in approaching the observed distribution function, while the opposite result is shown by a two-parameter distribution, this distribution function is very good at capturing the observation distribution function. Figure 3 also clarifies the conclusion that the Weibull and Gamma distributions are the best models produced in this study.

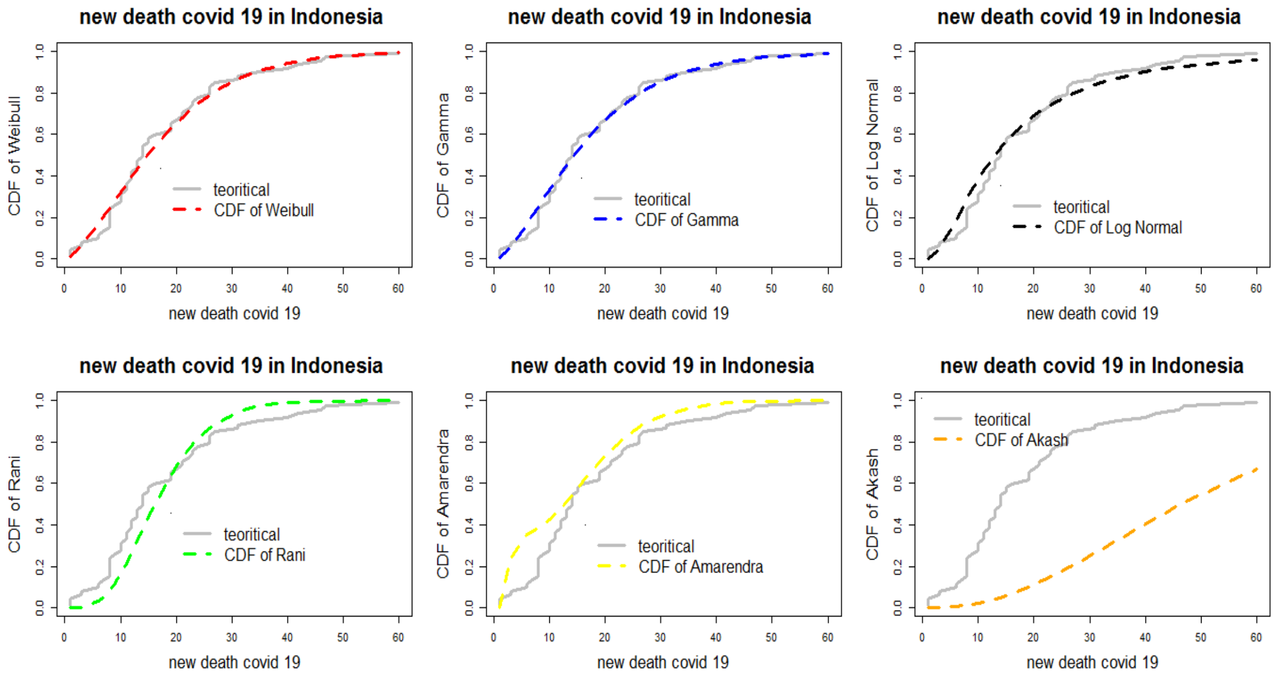

Fig. 3 Cdf plot for comparisons Predicted and observed daily death cases covid 19 in Indonesia

From the goodness of fit test of the model using this graphical method, it is very clear that the pdf and CDF plots give very clear results that the two-parameter distributions such as Weibull and Gamma are the best model in analyzing the frequency of a number of death caused by covid 19 in Indonesia. Numerical methods such as AIC and BIC values for the goodness-of-fit test were also used in this study. These two values for each distribution used will be presented in table 3. Based on the values in the table, it can be concluded that the two-parameter distribution namely the Weibull distribution is the best model because it has the smallest AIC and BIC values. Table 4 is also equipped with tests of model goodness such as log-likelihood (Log L) model, from the values presented it can also be concluded that the distribution of the two Weibull parameters is the best in this study

Table 4. The Goodness and fit test result of the daily new cases covid 19 in Indonesia

	Weibull	Gamma	Log Normal	Rani	Amarendra	Akash
AIC	424.75	424.99	434.77	465.56	439.11	571.93
BIC	428.80	429.05	438.82	441.14	441.14	573.95
Log (L)	-210.37	-210.49	-215.38	-231.78	-218.55	-284.96

5. Conclusion

This study is focused to analyze the frequency of the number of daily death data caused by covid 19 in Indonesia, to identify the appropriate six models or distributions that can be used to describe the distribution of the daily number of death data covid 19. It is concluded that the Weibull two parameters distribution returned better results when compared with other well-known distributions. This conclusion is based on widely used goodness of fit test models such as AIC and BIC. The graphical technique namely pdf and CDF plots were also observed comparing the empirical distributions with the adjusted by the Weibull two parameters distribution. In addition, through the best model in this study, we can use the distribution of the quantile function to simulate the number of daily death data of covid 19 for the future. These results are interesting for estimating the number of hospital beds during resource allocation planning or social isolation policies

References

[1] WHO, Statement on the Second Meeting of the International Health Regulations, Emergency Committee Regarding the Outbreak of Novel Coronavirus (2019-nCoV), 2005. [Online]. Available: https://www.who.int/news-room/detail/

[2] Jeffrey Chu, "A Statistical Analysis of the Novel Coronavirus (COVID-19) in Italy and Spain," *PLoS ONE*, vol. 16, no. 3, p. e0249037, 2021. *Crossref*, https://doi.org/10.1371/journal.pone.02490ss37

[3] Ayman Alzaatreh, Carl Lee, and Felix Famoye, "A New Method for Generating Families of Continuous Distributions," *Metron*, vol. 71, pp. 63–79, 2013. *Crossref*, https://doi.org/10.1007/s40300-013-0007-y

[4] Abdulhakim A. Al-Babtain et al., "A New Modified Kies Family: Properties, Estimation Under Complete and Type-II Censored Samples and Engineering Applications," *Mathematics*, vol. 8, no. 8, pp. 1345, 2020. *Crossref*, https://doi.org/10.3390/math8081345

[5] K. Jayakumar, and Thomas Mathew, "On a Generalization to Marshall–Olkin Scheme and Its Application to Burr Type XII Distribution," *Statistical Papers,* vol. 49, pp. 421–439, 2008. *Crossref,* https://doi.org/10.1007/s00362-006-0024-5

[6] Gauss M. Cordeiro et al., "A New Wider Family of Continuous Models: The Extended Cordeiro and de Castro Family," *Hacettepe Journal of Mathematics and Statistics*, vol. 47, no. 151, p. 25, 2016. *Crossref,* https://doi.org/10.15672/HJMS.201615122074

[7] Broderick O. Oluyede, Gayan Warahena-Liyanage, and Mavis Pararai, "A New Compound Class of Log-Logisticweibull–Poisson Distribution: Model, Properties and Applications," *Journal of Statistical Computation and Simulation,* vol. 86, no. 7, pp. 1363–1391, 2015. *Crossref,* https://doi.org/10.1080/00949655.2015.1064409

[8] Mahmoud M. Mansour, and Salah M. Mohamed, "A New Generalized of Transmuted Lindley Distribution," *Applied Mathematical Sciences*, vol. 9, no. 55, pp. 2729–2748, 2015. *Crossref,* http://dx.doi.org/10.12988/ams.2015.52158

[9] Sandeep K. Maurya et al., "A New Method of Proposing Distribution and Its Application to Real Data," *Imperial Journal of Interdisciplinary Research*, vol. 2, no. 6, pp. 1331–1338, 2016.

[10] Abdulrahman Abouammoh, and Mohamed Kayid, "A New Family of Extended Lindley Models: Properties, Estimation and Applications," *Mathematics*, vol. 8, no. 12, p. 2146, 2020. *Crossref,* https://doi.org/10.3390/math8122146

[11] Zhenmin Chen, "A New Two-Parameter Lifetime Distribution with Bathtub Shape or Increasing Failure Rate Function," *Statistics & Probability Letters,* vol. 49, pp. 155–161, 2000. *Crossref,* https://doi.org/10.1016/S0167-7152(00)00044-4

[12] Mahmoud M. Mansour, "A New Log-Logistic Lifetime Model with Mathematical Properties, Copula, Modified Goodness of-Fit test for Validation and Real Data Modeling," *Mathematics*, vol. 8, no. 9, p. 1508, 2020. *Crossref,* https://doi.org/10.3390/math8091508

[13] M. H. Tahir, M. Adnan Hussain, and Gauss M. Cordeiro, "A New Flexible Generalized Family for Constructing Many Families of Distributions," *Journal of Applied Statistics,* vol. 49, no. 7, pp. 1615-1635, 2021. *Crossref,* https://doi.org/10.1080/02664763.2021.1874891

[14] Elham Gholami, Kamyar Mansori, and Mojtaba Soltani-Kermanshahi, "Statistical Distribution of Novel Coronavirus in Iran," *International Journal of One Health*, vol. 6, no. 2, pp. 143-146, 2020. *Crossref,* https://doi.org/10.14202/IJOH.2020.143-146

[15] Muhammad Yousaf et al., "Statistical Analysis of Forecasting Covid-19 for Upcoming Month in Pakistan," *Chaos Solitons Fractals,* vol. 138, p. 109926, 2020. *Crossref,* https://doi.org/10.1016/j.chaos.2020.109926

[16] Muhammad Ahsan-ul-Haq et al., "Modeling of COVID-19 Cases in Pakistan Using Lifetime Probability Distributions," *Annals of Data Science,* vol. 9, pp. 141–152, 2022. *Crossref,* https://doi.org/10.1007/s40745-021-00338-9

[17] Sami Khedhiri, "Statistical Modeling of COVID-19 Deaths with Excess Zero Counts," *Epidemiol. Methods,* vol. 10, no. s1, pp. 1-13, 2021. *Crossref,* https://doi.org/10.1515/em-2021-0007

[18] Mahmoud M. Mansour et al., "Modeling the COVID-19 Pandemic Dynamics in Egypt and Saudi Arabia," *Mathematics,* vol. 9, no. 8, p. 827, 2021. *Crossref,* https://doi.org/10.3390/math9080827

[19] Ogunwale Olukunle Daniel, Ayeni Taiwo Michael, and Odukoya Elijah Ayooluwa, "Modeling Covid-19 Deaths in Nigeria Using Exponential-Gamma Distribution," *International Journal of Multidisciplinary Research and Analysis,* vol. 5, no. 6, pp. 1456-1460, 2022. *Crossref,* https://doi.org/10.47191/ijmra/v5-i6-32

[20] Rama Shanker, "Rani Distribution and its Application," *Biometrics & Biostatistics International Journal*, vol. 6, no. 1, pp. 256-265, 2016. *Crossref,* https://doi.org/10.15406/bbij.2017.06.00155

[21] Rama Shanker, "Akash Distribution and Its Applications," *International Journal of Probability and Statistics,* vol. 4, no. 3, pp. 65-75, 2015. *Crossref,* https://doi.org/10.5923/j.ijps.20150403.01

[22] Rama Shanker, "Shanker Distribution and its Applications," *International Journal of Statistics and Applications,* vol. 5, no. 6, pp. 338-348, 2015. *Crossref,* https://doi.org/10.5923/j.statistics.20150506.08

[23] Rama Shanker, "Aradhana Distribution and its Applications," *International Journal of Statistics and Applications,* vol. 6, no. 1, pp. 23-34, 2016. *Crossref,* https://doi.org/10.5923/j.statistics.20160601.04

[24] Rama Shanker, "Sujatha Distribution and its Applications," *Statistics in Transition-New Series*, vol. 17, no. 3, pp. 391-410, 2016. *Crossref,* https://doi.org/10.21307/stattrans-2016-029

[25] Rama Shanker, "Amarendra Distribution and its Applications," *American Journal of Mathematics and Statistics,* vol. 6, no. 1, pp. 44-56, 2016. *Crossref,* https://doi.org/10.5923/j.ajms.20160601.05

[26] R. Shanker. "Devya Distribution and its Applications," *International Journal of Statistics and Applications,* vol. 6, no. 4, pp. 189-202, 2016. *Crossref,* https://doi.org/10.5923/j.statistics.20160604.01