Research Article

Statistical Analysis of the COVID-19 Mortality Rates with Probability Distributions: The Case of Pakistan and Afghanistan

Javid Gani Dar,1 Muhammad Ijaz,2 Ibrahim M. Almanjahie,3,4 Muhammad Farooq,5 and Mahmoud El-Morshedy6,7

1Department of Mathematical Sciences, IUST, Kashmir, India
2Department of Mathematics and Statistics, The University of Haripur, Haripur, KPK, Pakistan
3Department of Mathematics, College of Science, King Khalid University, Abha 62529, Saudi Arabia
4Statistical Research and Studies Support Unit, King Khalid University, Abha 62529, Saudi Arabia
5Department of Statistics, University of Peshawar, KPK, Pakistan
6Department of Mathematics, College of Science and Humanities in Al-Kharj, Prince Sattam bin Abdulaziz University, Al-Kharj 11942, Saudi Arabia
7Department of Mathematics, Faculty of Science, Mansoura University, Mansoura 35516, Egypt

Correspondence should be addressed to Muhammad Ijaz; ijaz.statistics@gmail.com

Received 16 May 2022; Accepted 5 July 2022; Published 25 July 2022

Academic Editor: Agamirza Bashirov

Copyright © 2022 Javid Gani Dar et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The COVID-19 pandemic has shocked nations due to its exponential death rates in various countries. According to the United Nations (UN), in Russia, there were 895, in Mexico 303, in Indonesia 77, in Ukraine 317, and in Romania 252, and in Pakistan, 54 new deaths were recorded on the 5th of October 2021 in the period of months. Hence, it is essential to study the future waves of this virus so that some preventive measures can be adopted. In statistics, under uncertainty, there is a possibility to use probability models that lead to defining future patterns of deaths caused by COVID-19. Based on probability models, many research studies have been conducted to model the future trend of a particular disease and explore the effect of possible treatments (as in the case of coronavirus, the effect of Pfizer, Sinopharm, CanSino, Sinovac, and Sputnik) towards a specific disease. In this paper, various probability models have been applied to model the COVID-19 death rate more effectively than the other models. Among others, exponentiated flexible exponential Weibull (EFEW) distribution is pointed out as the best fitted model. Various statistical properties have been presented in addition to real-life applications by using the total deaths of the COVID-19 outbreak (in millions) in Pakistan and Afghanistan. It has been verified that EFEW leads to a better decision rather than other existing lifetime models, including FEW, W, EW, E, AIFW, and GAPW distributions.

1. Introduction

The first case of COVID-19 infection was located in Pakistan on February 26, 2020, in Karachi—a recent returnee from Iran. From that point onward, the spread of contaminations sped up, and on March 18, 2020, it was affirmed that the infection had spread to all regions of Pakistan. More than a hundred deaths apart from more than six thousand infected people were reported in the first seven weeks of this outbreak [1]. Pakistan has the third-highest number of cases in South Asia after India and Bangladesh, while it stands 7th in Asia as of September 16, 2021, with a 26th position worldwide. The first death was reported on March 20 in Sindh province, and the community transmission was spread rapidly all over the country.

In a country like Pakistan, the graph started to follow an upward trajectory in March 2020 and peaked in June when it slowly started to decline and flattened in August and September. But again, it started to increase in October of the same year, reflecting the bathtub shape in the data. Figures 1(a) and 1(b) show the average infection rates.

Many researchers have conducted various studies to investigate the COVID-19 outbreak, such as Singh et al. who explored how to predict the COVID-19 pandemic for
the top 15 countries using the ARIMA model [2]. The worldwide death rates were estimated by Chaurasia and Pal, by employing the ARIMA and regression models [3]. Chakraborty and Ghosh utilized a regression tree and ARIMA model to forecast the short time of COVID-19 cases in multiple countries and the risk of COVID-19 by finding various demographic characteristics beside some disease characteristics within these countries [4]. Yousaf et al. [5] utilized the autoregressive integrated moving average (ARIMA) model to predict infections, deaths, and recoveries. Fong et al. [6] considered small data for early forecasting, while Petropoulos and Makridakis [7] also applied the forecasting model. Chen et al. [8] designed an algorithm for predicting COVID-19 data, while Nayak et al. [9] and
Wolkewitz et al. [10] applied a probabilistic model to analyze COVID-19 data. The size of the COVID-19 epidemic has been worked out by Yue et al. [11] with the help of surveillance systems, and a similar study to estimate the final size of the COVID-19 epidemic has also been discussed by Syed and Sibgatullah [12]. Mizumoto et al. [13] estimated the asymptomatic proportion of COVID-19. Many researchers applied various statistical models to predict data analysis. For example, Sukhanova et al. [14] forecast the macroeconomic indices with the help of ARIMA, vector autoregression (VAR), and simultaneous equation system. Yu et al. [15] predict the tourism demand by utilizing the SARIMA model and neural network (NN). To examine the accuracy with which long-term scenarios can be predicted in patients with coronary artery disease, Lee et al. [16] applied Cox regression. The results showed that model-based prediction was considered better as compared to doctors’ prediction.

Many lifetime distributions are available in the literature to predict the COVID-19 data, but these distributions are unable to model the data more precisely. For example, the Weibull (W) distribution introduced by Weibull [17] and the exponential (Ex) distribution by Epstein [18] along with other lifetime distributions are unable to model the COVID-19 data or any other data related to any infections of the disease that does not follow a constant rate (monotonic data). In daily life situations, the data does not always follow a monotonic failure function; rather, it follows a nonmonotonic failure function. For example, patients with tuberculosis have a higher risk in the early stages but a lower risk later on. A similar form of nonmonotonicity occurs in infants because the hazards for infants are highest in the early stages and gradually reduce as they develop, but the danger increases again as they become older, resulting in the bathtub shape. The researchers are trying to introduce functions that are more flexible as well as and can capture the nonmonotonic hazard rate functions. For example, Cordeiro et al. [19], El-Gohary et al. [20], Ijaz et al. [21], and Farooq et al. [22] worked on introducing the new distributions. We recommend recent research studies: Ijaz et al. [21, 23] and Ijaz et al. [24].

In practice, the modeling of real phonon becomes more complex when the number of unknown’s parameters is large. There are two main significant advantages of the probability models in this paper. First, it presents a best fitted model which is more flexible with fewer unknown parameters. Secondly, it leads us to better results for various hazard rate shapes, particularly in a bathtub shape where the curves are flatted at the middle and skewed on either side. Note that the distribution in this paper may not be considered as a best fitted model for the data sets with extreme values or even when there is an outlier.
2. Material and Methodology

The current research study focuses on the best fitted probability model which has more parameters as compared to some existing models. In this paper, the best fitted model has increased a shape parameter \( d \) in the family of distributions introduced by [22]. The CDF and PDF of the proposed probability model take the following forms:

\[
F(x) = \left( \frac{e^{(1-e^{-bc})} - 1}{e^a - 1} \right)^d, \quad x > 0, \quad a, \ b, \ c, \ \text{and} \ d > 0, \tag{3}
\]

\[
f(x) = \frac{abcdx^{-d}e^{(1-e^{-bc})-bc} \left( e^{(1-e^{-bc})} - 1 \right)^{d-1}}{(e^a - 1)^d}, \tag{4}
\]

where “\( b \)” is the scale and “\( c \)” and “\( d \)” are the shape parameters.

Figure 2 defines the shapes of the CDF and PDF described in (3) and (4), respectively.

2.1. The Survival \( S(x) \) and Hazard \( h(x) \) Rate Function. By definition, \( S(x) \) and \( h(x) \) functions are, respectively, defined by

\[
S(x) = 1 - F(x),
\]

\[
h(x) = \frac{f(x)}{S(x)}. \tag{5}
\]
Table 3: Data set 2: Afghanistan (total deaths per million).

| 0.026 | 0.026 | 0.026 | 0.051 | 0.077 | 0.077 | 0.103 | 0.103 | 0.103 | 0.103 |
|-------|-------|-------|-------|-------|-------|-------|-------|-------|-------|
| 0.103 | 0.103 | 0.206 | 0.257 | 0.308 | 0.385 | 0.411 | 0.411 | 0.437 | 0.462 |
| 0.462 | 0.488 | 0.565 | 0.591 | 0.745 | 0.771 | 0.771 | 0.848 | 0.925 |       |
| 0.925 | 1.028 | 1.028 | 1.105 | 1.207 | 1.336 | 1.49  | 1.516 | 1.567 | 1.644 |
| 1.747 | 1.85  | 2.183 | 2.312 | 2.44  | 2.672 | 2.723 | 2.8   | 2.954 | 3.083 |
| 3.134 | 3.262 | 3.391 | 3.494 | 3.93  | 4.316 | 4.367 | 4.444 | 4.573 | 4.829 |
| 4.984 | 5.292 | 5.574 | 5.626 | 5.651 | 5.677 | 5.857 | 6.062 | 6.345 | 6.422 |
| 6.628 | 6.833 | 7.039 | 7.655 | 7.809 | 8.04  | 8.503 | 9.273 | 9.582 | 9.967 |
| 10.506| 11.046| 11.56 | 11.688| 12.202| 12.382| 12.716| 13.05 | 14.129| 14.18 |
| 14.719| 15.028| 15.336| 15.85 | 16.389| 17.314| 17.519| 18.393| 18.701| 19.009|
| 19.318| 20.037| 20.782| 21.09 | 21.27 | 22.246| 23.119| 23.685| 24.121| 24.635|
| 24.995| 25.585| 25.996| 26.716| 27.332| 28.154| 28.694| 29.516| 29.952| 30.389|
| 30.441| 30.518| 30.62 | 31.16 | 31.519| 32.085| 32.393| 32.65  | 32.675| 32.701|
| 32.958| 32.984| 33.009| 33.035| 33.138| 33.138| 33.292| 33.42  | 33.652| 33.78 |
| 33.934| 34.14  | 34.576| 34.833| 35.064| 35.193| 35.219| 35.347| 35.398| 35.501|
| 35.553| 35.604 | 35.604| 35.604| 35.655| 35.707| 35.912| 36.015| 36.015| 36.041|
| 36.041| 36.041 | 36.143| 36.22  | 36.22 | 36.22 | 36.22 | 36.297| 36.375|       |
| 36.452| 36.503 | 36.503| 36.503| 36.503| 36.503| 36.657| 36.683| 36.94  | 36.94 |
| 36.965| 36.965 | 37.068| 37.145| 37.171| 37.197| 37.325| 37.325| 37.376| 37.376|
| 37.453| 37.505 | 37.505| 37.505| 37.505| 37.608| 37.71  | 37.736| 37.787|       |
| 37.813| 37.864 | 37.89 | 37.993| 38.044| 38.07 | 38.096| 38.096| 38.198| 38.275|
| 38.378| 38.507 | 38.558| 38.609| 38.712| 38.764| 38.866| 38.943| 39.046| 39.175|
| 39.329| 39.406 | 39.431| 39.508| 39.508| 39.663| 39.74  | 39.842| 39.997| 39.997|
| 40.048| 40.202 | 40.51 | 40.587| 40.69 | 40.947| 41.05  | 41.307| 41.615| 42    |
| 42.154| 42.334 | 42.463| 42.797| 43.105| 43.413| 43.721 | 44.055| 44.389| 44.62 |
| 44.698| 45.006 | 45.571| 46.11  | 46.804| 47.292| 47.42  | 47.42  | 47.883| 48.14 |
| 48.808| 48.962 | 49.296| 49.707| 49.964| 50.246| 50.477 | 50.58  | 51.248| 51.659|
| 52.019| 52.147 | 52.584| 53.098| 53.483| 53.843| 54.382 | 54.613 | 54.947| 55.204|
| 55.487| 55.846 | 55.975| 56.026| 56.283| 56.283| 56.283 | 56.283 | 57.465| 57.644|

Figure 4: Theoretical and empirical PDF and CDF of EFEW.
Using (3) and (4), we get
\[
S(x) = \left( e^{q} - 1 \right)^{d} - \left( e^{\theta(1-x^{1/c})} - 1 \right)^{d},
\]
\[
h(x) = \frac{abcdn^{-1} e^{\theta(1-x^{1/c}) - bx} \left( e^{\theta(1-x^{1/c})} - 1 \right)^{d-1}}{e^{q} - 1 \right)^{d} - \left( e^{\theta(1-x^{1/c})} - 1 \right)^{d}}.
\]

Figure 3 defines various shapes of the hazard rate function.

3. Statistical Properties

3.1. Quantile Function. The quantile function is defined by
\[
p(X \leq x) = q.
\]

Using (3), we get
\[
\left( e^{\theta(1-x^{1/c})} - 1 \right) = q^{1/d} (e^{q} - 1).
\]

The final result for \( X \) can be obtained as
\[
x = \left( -1 \right)^{1/c} \log \left( a - \log \left( 1 + q^{1/d} (e^{q} - 1) \right) \right),
\]
where \( q \sim U[0, 1] \).

3.2. \( r \)th Moment. The \( r \)th moment can be obtained by
\[
E(x^{r}) = \int x^{r} f(x) dx,
\]
\[
E(x^{r}) = \int_{0}^{\infty} x^{r} abcdn^{-1} e^{\theta(1-x^{1/c}) - bx} \left( e^{\theta(1-x^{1/c})} - 1 \right)^{d-1} \frac{dx}{(e^{q} - 1 \right)^{d} - \left( e^{\theta(1-x^{1/c})} - 1 \right)^{d}}.
\]

Using \( z = e^{\theta(1-x^{1/c})} \), then \( dz = abcdn^{-1} e^{\theta(1-x^{1/c}) - bx} dx \) and \( x = \left( -1/b \log \left( 1 - \log z/a \right) \right)^{1/c} \).

\[
E(x^{r}) = \frac{d(-1/b)^{r/c}}{(e^{q} - 1 \right)^{d}} \int_{1}^{e^{d}} \left( 1 - \log z/a \right)^{r/c} (z-1)^{d-1} dz.
\]

Using \( (\log (1 - \log z/a)^{r/c} = \sum_{k=1}^{\infty} \left(-1\right)^{k+1} (k^{r/c}) \right)^{d} \)

\( \left(-1\right)^{1+d+n} \binom{d-1}{n} z^{n} \),

finally, we obtained
\[
\frac{d(-1/b)^{r/c}}{(e^{q} - 1 \right)^{d}} \sum_{k=1}^{\infty} \sum_{n=0}^{\infty} \left(-1\right)^{k+1} \left(-1\right)^{1+d+n} (k^{r/c}) \right)^{d} \left( n \right)
\]

\[
\left[\left(-1\right)^{r/c} \left(-1-a\right)^{k+c} \Gamma(kr/c + 1, -(n+1) \log a) \right] \left( n + 1 \right)
\]

3.3. Order Statistics. The \( i \)th order statistic of the PDF is given by
\[
f(i, n) = \frac{n^{i}}{(i-1)! (n-i)! f(x) F(x)^{i-1} (1 - F(x))^{n-i}}.
\]

Letting Equations (3) and (4), the 1st and \( n \)th order statistics of EEW can be obtained, respectively, by using \( i = 1 \) and \( i = n \) as

\[
f(1, n) = \frac{abcdn^{-1} e^{\theta(1-x^{1/c}) - bx} \left( e^{\theta(1-x^{1/c})} - 1 \right)^{d-1} \left( e^{q} - e^{\theta(1-x^{1/c})} \right)^{(n-1)d}}{(e^{q} - 1 \right)^{d} - \left( e^{\theta(1-x^{1/c})} - 1 \right)^{d}}.
\]

\[
f(n, n) = \frac{abcdn^{-1} e^{\theta(1-x^{1/c}) - bx} \left( e^{\theta(1-x^{1/c})} - 1 \right)^{d-1} \left( e^{q} - e^{\theta(1-x^{1/c})} \right)^{(n-1)d}}{(e^{q} - 1 \right)^{d} - \left( e^{\theta(1-x^{1/c})} - 1 \right)^{d}}.
\]
Table 5: Model selection criterion for data 1.

| Models | AIC          | CAIC         | BIC          | HQIC         |
|--------|--------------|--------------|--------------|--------------|
| EFEW   | 2188.224     | 2188.362     | 2202.958     | 2194.124     |
| FEW    | 2381.277     | 2381.359     | 2392.327     | 2385.702     |
| E      | 2408.907     | 2408.921     | 2412.591     | 2410.383     |
| W      | 2410.659     | 2410.701     | 2418.027     | 2413.61      |
| Ex-W   | 2415.711     | 2415.794     | 2426.762     | 2420.136     |
| AIFW   | 2462.986     | 2463.028     | 2470.354     | 2465.937     |
| GAPW   | 2386.745     | 2386.828     | 2397.796     | 2391.171     |

Figure 5: Theoretical, empirical, Q-Q plot, and P-P plot for EFEW.

Figure 6: TTT plot of the COVID-19 data.
where $\alpha$ positively skewed data, or even the data skewed to the left.

4. Special Cases

The special cases of EFEW are as follows.

### Case 1. When $d = 1$

By putting $d = 1$ in (3) and (4), we derive the CDF and PDF of the flexible exponential Weibull (FEW) distribution. The mathematical form is described as

$$F(x) = \frac{e^{\alpha (1-e^{-bx})} - 1}{e^a - 1}, \quad x > 0, \ b > 0, \ c > 0, \ a \neq 1,$$

$$f(x) = \frac{abce^{\alpha (1-e^{-bx})} - bx^c}{e^a - 1}.$$

### Case 2. When $d = 1$ and $c = 1$

Putting $d = 1$ and $c = 1$ in (3) and (4) shall refer to the CDF and PDF of the gull alpha power exponential distribution (GAPE). The mathematical form is described as

$$F(x) = \frac{e^{\alpha (1-e^{-bx})} - 1}{e^a - 1}, \quad x > 0, \ b > 0, \ a \neq 1,$$

$$f(x) = \frac{abex^{\alpha (1-e^{-bx})} - bx^c}{e^a - 1}.$$ 

### Case 3. When $d = 1$ and $c = 2$

If we replace $d = 1$ and $c = 2$ in (3) and (4), the CDF and PDF will become NF Rayleigh (NFPR) distribution. Mathematically, the CDF and PDF of NFPR are

$$F(x) = \frac{e^{\alpha (1-e^{-bx})} - 1}{e^a - 1}, \quad x > 0, \ b > 0, \ a \neq 1,$$

$$f(x) = \frac{abxe^{\alpha (1-e^{-bx})} - bx^c}{e^a - 1}.$$

5. Parameter Estimation

The log likelihood function of Equation (4) is defined by

$$\log (L) = n \log \left( \frac{abcd}{(e^a - 1)^{d}} \right) + (c - 1) \sum \log x_i$$

$$+ a \sum (1 - e^{-bx}) \sum bx^c + a(d - 1) \sum (1 - e^{-bx}).$$

The partial derivatives of (19) with respect to parameters are obtained by

$$\frac{d}{d(a)} \log (L) = (d - 1) \left( -nd \log (e^a - 1) - e^{-bx} + 1 \right) - \frac{(d - 1)ande^a}{e^a - 1} + \frac{n}{a} - e^{-bx} + 1,$$

$$\frac{d}{d(b)} \log (L) = a(d - 1) \sum x^c e^{-bx} \sum x^c + a \sum x^c e^{-bx} \sum x^c + \frac{n}{b} - \sum x^c,$$

$$\frac{d}{d(c)} \log (L) = -e^{-bx} x^c \left( (b \sum x^c \log (x))c - n) e^{bx} - abcd \sum x^c \log (x) \right),$$

$$\frac{d}{d(d)} \log (L) = a \left( -\log (e^a - 1)nd - e^{-bx} + 1 \right) - a \log (e^a - 1)n(d - 1) + \frac{n}{d}.$$
The above expressions are not in closed form, but still, the numerical solution is possible by using various mathematical techniques.

6. Applications

In this section, the COVID-19 death data of Pakistan and Afghanistan were considered to delineate the real-life applications by means of AIC, CAIC, BIC, and HQIC.

It should be noted that the model with a fewer value of these criteria is considered as the best model among others. The data sets with the URL https://github.com/owid/covid-19-data are taken from May 2, 2020, till July 4, 2021, for Pakistan and Afghanistan. Tables 2 and 3 respectively defines the mortality rates in Pakistan and Afghanistan.

In Figure 4, both the theoretical and empirical graphs depict that the EFEW is the best fitted line as compared to other existing distributions and can be justified from Tables 4 and 5.

Figure 5 demonstrates the Q-Q and P-P plot of the COVID-19 death data. The Q-Q plot demonstrates that most of the data points, except a few points on the upper tail, follow a linear pattern on the line, while the P-P plot also indicates a reasonably good fit and indicates that the EFEW reasonably describes the empirical data distribution along with empirical and theoretical densities and their CDF.

Figure 6 depicts the pattern of the hazard rate function. The curve clearly crosses the diagonal line, and hence, the data follows a nonmonotonic hazard rate function.

Table 4 shows the Cramer-Mises (W) and Anderson-Darling (A) maximum likelihood estimates, standard errors, and log-likelihood values. Table 5 shows the best model selection criterion. The results of Tables 5 and 6 depict the smaller values for FEW among others using this goodness of fit criteria and hence show that EFEW provides a flexible fit over exponential (E), Weibull (W), Exponential-Weibull (Ex-W), Algoharai inverse flexible Weibull (AIFW), and gull alpha power Weibull (GAPW) distributions.

Figure 7 shows the theoretical and empirical PDF and CDF of EFEW distribution using the COVID-19 death data from Afghanistan. Both the theoretical and empirical graphs clearly depict that the EFEW is the best fitted line as compared to other existing distributions and can be justified from the numerical values presented in Tables 6 and 7.

Figure 8 demonstrates the Q-Q and P-P plot of the COVID-19 death data from Afghanistan. The Q-Q plot demonstrates that most of the data points, except a few points on the upper tail, follow a linear pattern on the line, while the P-P plot also indicates a reasonably good fit and indicates that the EFEW reasonably describes the empirical data distribution along with empirical and theoretical densities and their CDF.

Figure 9 follows the same pattern as Figure 6 which means that the death rate in Afghanistan also follows a non-monotonic shape.

The results of Tables 6 and 7 show that by employing these criteria, smaller values are achieved for EFEW, and hence, EFEW gives a flexible fit over FEW, E, W, Ex-W, AIFW, and GAPW.
Figure 8: Theoretical, empirical, Q-Q plot, and P-P plot for EFEW.

Figure 9: TTT plot of the COVID-19 data of Afghanistan.

Table 8: Average values of MSE and bias.

| Parameters | n   | MSE (a) | MSE (b) | MSE (c) | MSE (d) | Bias (a) | Bias (b) | Bias (c) | Bias (d) |
|------------|-----|---------|---------|---------|---------|----------|----------|----------|----------|
| a = 0.5    | 40  | 31.947  | 5.718   | 1.119   | 48.202  | 3.904    | 1.907    | 0.949    | 5.032    |
| b = 0.05   | 70  | 24.757  | 5.100   | 1.032   | 41.127  | 3.392    | 1.720    | 0.901    | 4.521    |
| c = 1.5    | 100 | 22.843  | 4.316   | 0.918   | 34.410  | 3.280    | 1.498    | 0.833    | 3.829    |
| d = 0.5    | 150 | 20.125  | 3.659   | 0.831   | 26.742  | 2.983    | 1.309    | 0.781    | 3.190    |
| a = 0.6    | 100 | 21.828  | 3.174   | 1.090   | 30.497  | 3.034    | 1.210    | 0.888    | 3.450    |
| b = 0.05   | 200 | 13.441  | 1.606   | 0.724   | 16.210  | 2.239    | 0.714    | 0.676    | 1.976    |
| c = 1.77   | 300 | 9.5134  | 1.047   | 0.577   | 11.166  | 1.768    | 0.528    | 0.581    | 1.472    |
| d = 0.6    | 400 | 7.8027  | 0.758   | 0.492   | 8.0283  | 1.558    | 0.425    | 0.513    | 1.140    |
7. Simulation Study of EFEW Distribution

A simulation study has been performed to check the consistency of the parameters of the EFEW distribution. We consider two set of parameter values, i.e., $a = 0.5$, $b = 0.05$, $c = 1.5$, and $d = 0.5$ and $a = 0.6$, $b = 0.05$, $c = 1.77$, and $d = 0.6$. A simulation is performed with 1000 replications. A sample of sizes $n = 40, 70, 100, 150$ and $n = 100, 200, 300, 400$ are drawn, respectively, and the bias and mean square error (MSE) are estimated. The mathematical forms are described as

$$\text{MSE} = \frac{1}{W} \sum_{i=1}^{W} (\hat{\alpha}_i - \alpha)^2,$$

$$\text{Bias} = \frac{1}{W} \sum_{i=1}^{W} (\hat{\alpha}_i - \alpha).$$

(21)

Table 8 defines the average mean square errors and biases of each parameter using small and large sample sizes taken from EFEW. It is quantified that when we increase the sample of size $n$, the average values of mean square errors and bias decrease with different values of parameters.

8. Conclusion

In this article, the best fitted model (EFEW) is pointed out for modeling the death rates of coronavirus. Various statistical properties of the proposed model have been discussed. The significance of EFEW has been evaluated using the death data of COVID-19 in Pakistan and Afghanistan. It has been verified that the EFEW model is capable of modeling both the monotonic and nonmonotonic failure data better than the existing models. Moreover, the findings consistently lead to better results and increase the model flexibility compared to the existing probability distributions. Hence, the inclusion of the parameter ($d$) to the existing model plays an important role and hence is a better choice in making predictions of deaths among infected patients of coronavirus than the other models.

It is expected that the present class of expressions, along with its special forms, will attract the researchers towards its learning, common errors and how to avoid them, and various others. Moreover, the present study can be extended to neutrosophic statistics. A future research study may also be conducted on the Bayesian analysis of the model parameters under various loss functions.

Data Availability

The simulated data used to support the findings of this study are included within the article.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this article.

Acknowledgments

The authors thank and extend their appreciation to the Deanship of Scientific Research at King Khalid University for funding this work through the Research Groups Program under grant number R.G.P. 2/132/43.

References

[1] N. Ilyas, R. E. Azuine, and A. Tamiz, “COVID-19 pandemic in Pakistan,” International Journal of Translational Medical Research and Public Health, vol. 4, no. 1, pp. 37–49, 2020.

[2] R. K. Singh, M. Rani, A. S. Bhagavathula et al., “Prediction of the COVID-19 pandemic for the top 15 affected countries: advanced autoregressive integrated moving average (ARIMA) model,” JMIR Public Health and Surveillance, vol. 6, no. 2, article e19115, 2020.

[3] V. Chaurasia and S. Pal, “COVID-19 pandemic: ARIMA and regression model-based worldwide death cases predictions,” SN Computer Science, vol. 1, no. 5, pp. 1–12, 2020.

[4] T. Chakraborty and I. Ghosh, “Real-time forecasts and risk assessment of novel coronavirus (COVID-19) cases: a data-driven analysis,” Chaos, Solitons & Fractals, vol. 135, article 109850, 2020.

[5] M. Yousaf, S. Zahir, M. Riaz, S. M. Hussain, and K. Shah, “Statistical analysis of forecasting COVID-19 for upcoming month in Pakistan,” Chaos, Solitons & Fractals, vol. 138, article 109926, 2020.

[6] S. J. Fong, G. Li, N. Dey, R. G. Crespo, and E. Herrera-Viedma, “Finding an accurate early forecasting model from small dataset: a case of 2019-ncov novel coronavirus outbreak,” 2020, https://arxiv.org/abs/2003.10776.

[7] F. Petropoulos and S. Makridakis, “Forecasting the novel coronavirus COVID-19,” PLoS One, vol. 15, no. 3, article e0231236, 2020.

[8] Y. Chen, J. Cheng, X. Jiang, and X. Xu, “The reconstruction and prediction algorithm of the fractional TDD for the local outbreak of COVID-19,” 2020, https://arxiv.org/abs/2002.10302.

[9] S. R. Nayak, V. Arora, U. Sinha, and R. C. Poonia, “A statistical analysis of COVID-19 using Gaussian and probabilistic model,” Journal of Interdisciplinary Mathematics, pp. 1–14, 2020.

[10] M. Wolke, M. Lombart, M. von Cube et al., “Statistical analysis of clinical covid-19 data: a concise overview of lessons learned, common errors and how to avoid them,” Clinical Epidemiology, vol. 12, pp. 925–928, 2020.

[11] M. Yue, H. E. Clapham, and A. R. Cook, “Estimating the size of a COVID-19 epidemic from surveillance systems,” Epidemiology (Cambridge, Mass.), vol. 31, no. 4, pp. 567–569, 2020.

[12] F. Syed and S. Sibgatullah, “Estimation of the final size of the COVID-19 epidemic in Pakistan,” MedRxiv., 2020.

[13] K. Mizumoto, K. Kagaya, A. Zarebski, and G. Chowell, “Estimating the asymptomatic proportion of coronavirus disease 2019 (COVID-19) cases on board the Diamond Princess cruise ship, Yokohama, Japan, 2020,” Eurosurveillance, vol. 25, no. 10, p. 2000180, 2020.

[14] E. I. Sukhanova, S. Y. Shirnaeva, and A. G. Mokronosov, “Econometric models for forecasting of macroeconomic indices,” International Journal of Environmental and Science Education, vol. 11, no. 16, pp. 9191–9205, 2016.
[15] Y. Yu, Y. Wang, S. Gao, and Z. Tang, 2017, vol. 2017, Statistical modeling and prediction for tourism economy using dendritic neural network, Computational intelligence and neuroscience, 2017.

[16] K. L. Lee, D. B. Pryor, F. E. Harrell Jr. et al., “Predicting outcome in coronary disease statistical models versus expert clinicians,” The American Journal of Medicine, vol. 80, no. 4, pp. 553–560, 1986.

[17] W. Weibull, “A statistical distribution function of wide applicability,” Wide applicability. Journal of applied mechanics, vol. 18, no. 3, pp. 293–297, 1951.

[18] B. Epstein, The Exponential Distribution and Its Role in Life Testing. Wayne State Univ Detroit MI, 1958.

[19] G. M. Cordeiro, E. M. Ortega, and A. J. Lemonte, “The exponential–Weibull lifetime distribution,” Journal of Statistical Computation and Simulation, vol. 84, no. 12, pp. 2592–2606, 2014.

[20] A. El-Gohary, A. H. El-Bassiouny, and M. El-Morshey, “Inverse flexible Weibull extension distribution,” International Journal of Computer Applications, vol. 115, no. 2, pp. 46–51, 2015.

[21] M. Ijaz, S. M. Asim, M. Farooq, S. A. Khan, and S. Manzoor, “A gull alpha power Weibull distribution with applications to real and simulated data,” PLoS One, vol. 15, no. 6, article e0233080, 2020.

[22] M. Farooq, M. Ijaz, S. F. Shah, and M. Kilai, “A novel approach to increase the goodness of fits with an application to real and simulated data sets,” Mathematical Problems in Engineering, vol. 2021, 11 pages, 2021.

[23] M. Ijaz, M. Asim, and A. Khalil, Flexible Lomax distribution, 2019.

[24] M. Ijaz, W. K. Mashwani, and S. B. Belhaouari, “A novel family of lifetime distribution with applications to real and simulated data,” PLoS One, vol. 15, no. 10, article e0238746, 2020.