International Journal of Mathematics and
Computer Science, **17**(2022), no. 4, 1499–1506

Probability Distributions for Modeling of COVID-19 cases and deaths in Thailand

Thanittha Kowan[1], Chutima Mahalapkorkiat[2], Tiyaporn Samakrat[3], Patcharee Sumritnorrapong[4]

[1]Department of Mathematics, Statistics, and Computer
Faculty of Liberal Arts and Science
Kasetsart University, Kamphaeng Saen Campus
Nakhon Pathom, 73140, Thailand

[2,3,4]Department of Mathematics
Faculty of Science
Silpakorn University
Nakhon Pathom, 73000, Thailand

email: Thanittha_ko@ku.th, Mahalapkorkiat_c@silpakorn.edu,
Samakrat_t@silpakorn.edu, Sumritnorrapong_p@silpakorn.edu

(Received April 22, 2022, Accepted May 17, 2022)

Abstract

The COVID-19 is a pandemic and continues to mutate and spread within Thailand and throughout the world. Recently, Omicron is a new COVID-19 variant of concern because it has several mutations that may have an impact on how it behaves. It is therefore important to understand COVID-19 dynamics in order to prevent or control infections appropriately. In this study, we analyzed a model of the daily number of COVID-19 cases and deaths in Thailand using five different probability distributions. Maximum likelihood estimation (MLE) is applied to estimate parameters of the five distributions. The results indicate that the Weibull distribution and the log-normal distribution are the most suitable distributions that fit the data on daily confirmed cases and on daily confirmed deaths, respectively, by using the

Key words and phrases: Covid-19, Probability distributions, Maximum likelihood estimation.
AMS (MOS) Subject Classifications: 62E10, 62F10.
Patcharee Sumritnorrapong is the corresponding author.
ISSN 1814-0432, 2022, http://ijmcs.future-in-tech.net

Akaike information criterion (AIC) and the Bayes information criterion (BIC).

1 Introduction

The coronavirus disease 2019 or COVID-19 is a new type of coronavirus and was discovered in the middle of December 2019 in Wuhan City, Hubei Province, China. From early beginnings in Wuhan, COVID-19 has quickly spread in China and throughout the world including Thailand. Thailand was the first country to report a case outside China on January 12, 2020 [7]. After that, there were increasing confirmed cases in Thailand. A large cluster of cases was linked to Boxing matches that were held at Lumpinee Boxing stadium in Bangkok in spite of a government shutdown order issued previously. Lumpinee cases surged in many provinces in Thailand and brought the highest cumulative number of COVID cases to 4297. In December 2020, the second wave of COVID-19 in Thailand was carried by undocumented migrant workers at Central Shrimp Market in Samut Sakhon, a province southwest of Bangkok. The outbreak has affected Thailand's economy and shrimp industry because Thailand is one of the worlds 10 largest shrimp exporters. The third outbreak began severely in April 2021 with more than 500 cases reported from the Thonglor cluster at nightspots in the capital Bangkok. The total number of confirmed cases in Thailand surpassed one million on August 19, 2021. The COVID-19 virus is continuing to mutate and spread within Thailand and around the world. The fourth wave is caused by the Delta coronavirus variant and recently the fifth wave has been sweeping through Thailand by a new COVID-19 variant of concern, named Omicron. Omicron was first identified in South Africa on November 24, 2021. Thailand was the 47th nation with an Omicron case notified on December 1, 2021 [8]. Omicron has several mutations that may have an impact on how it behaves. The World Health Organization (WHO) [9] said it is not clear whether the Omicron variant is more transmissible or whether Omicron causes more severe disease compared to other variants. However, the number of people testing positive for COVID-19 impacted by Omicron has risen. This pandemic has negatively affected entire societies and the global economy. To prevent or control infections appropriately, a statistical analysis of COVID-19 is significantly used to describe and evaluate the real situation.

Due to the outbreak of COVID-19, there were several research studies using both local and global statistical COVID-19 models. One of the most extensive studies was to evaluate probability distributions to describe infec-

tions and deaths from COVID-19 and also compare probability distributions to determine which distribution best fits the data. In 2020, Gholami, Mansori, and Soltani-Kermanshahi [4] studied three types of distributions, known as normal, log-normal, and Weibull distributions, of COVID-19 cases in Iran and they showed that the Weibull distribution was the best fit to the daily reported data. Likewise, AhsanulHaq et al. [1] analyzed data on daily confirmed COVID-19 cases and deaths in Pakistan. They considered nine lifetime probability distributions and concluded that the Weibull distribution fit the data best. In 2021, Mansour et al. [5] developed a new generalized family of distributions to describe and compare the COVID-19 pandemic dynamics between Egypt and Saudi Arabia. Hincal et al. [3] predicted the death rates of ten countries which are Argentina, Austria, Brazil, France, Iran, Italy, Sweden, Turkey, United Kingdom, and United States of America by using the gamma, binomial, and posterior distributions as extensions of Bayes theorem.

Thailand, as the first country outside of China to record a case of COVID-19 at the beginning of the outbreak, has been directly affected. Since COVID-19 dynamics diverge between different regions, we are interested in studying probability distributions of the numbers of COVID-19 infections and deaths in Thailand. In this research, we study the COVID-19 pandemic dynamics in Thailand by using five lifetime probability distributions which are normal, log-normal, Weibull, log-logistic distribution, and inverse Weibull distributions. We collect the data of the fifth wave of COVID-19 with highly contagious Omicron in Thailand from January 1, 2022 until April 7, 2022 and we then find values of unknown model parameters using the maximum likelihood estimation (MLE). We also determine the best-fit distribution to the data on daily confirmed COVID-19 cases and deaths in Thailand by using the Akaike information criterion (AIC) and the Bayesian information criterion (BIC).

2 Preliminaries

A life distribution [6] is an important probability distribution to model lifetime data such as analytical data on life annuity in economics and data of selected Fluorophores in biomedical sciences. Life distributions are applied in many fields including actuarial science, environmental sciences, and public health. For analyzing life data, there are several life distributions such as exponential, log-normal, normal, and Weibull distributions. In this study, we consider two-parameter probability distributions as follows:

1. The **Normal distribution** has two parameters: the location parameter μ which is the mean or expectation of the distribution, and the scale parameter σ which is the standard deviation of the distribution. The normal distribution is the most popular and important distribution because it fits many phenomena. For example, the normal distribution is used to model asset's expected returns in finance and to predict the level of solar radiation in Physics.

2. The **Log-normal distribution** is described by the location parameter μ which is the expectation of the variable's natural logarithm and the scale parameter σ which is the variance of the variable's natural logarithm. The log-normal distribution is perfect for modeling data that cannot take negative values and is widely used in many fields. For instance, the quantity of rainfall, the extent of periods to which any infectious disease, and the size of tissue are analyzed by using log-normal distribution.

3. The **Weibull distribution** is characterized by two parameters: the shape parameter k which is known as the Weibull slope and the scale parameter λ which is defined as the value at the 63.2th percentile. The Weibull distribution is used to model failure times and reliability such as analysis of infant mortality rate, failure analysis in engineering, and analysis of wind speed data.

4. The **Log-logistic distribution** is a 2-parameter distribution with positive shape parameter α and positive scale parameter β which is the mean of the distribution. The log-logistic distribution is known to be useful to model an event whose rate increases initially and decreases later such as the mortality rate from cancer after diagnosis and flood frequency used in hydrology.

5. The **Inverse Weibull distribution** is a 2-parameter distribution with positive shape parameter α and positive scale parameter β. The Inverse Weibull distribution can be used to model a variety of failure rates such as infant mortality and degradation of mechanical components.

There are many methods for estimating parameters of a probability distribution such as method of moment, maximum likelihood estimation, and method of least square. In this study, we use the maximum likelihood estimation (MLE) because the MLE is a widely used method using some observed

data. The goal of this method is to find estimators by maximizing a likelihood function.

To evaluate how well our models fit the data on daily confirmed COVID-19 cases and deaths in Thailand, we use the Akaike information criterion (AIC) and the Bayes information criterion (BIC). Distributions with the lowest AIC and BIC values are considered the best. The formulas of the Akaike information criterion and the Bayes information criterion are given as follows:

$$\text{AIC} = 2k - 2\log L_n(\hat{\theta}) \qquad \text{and} \qquad \text{BIC} = k\log n - 2\log L_n(\hat{\theta}),$$

where k is the number of parameters, $L_n(\theta)$ is the likelihood function, n is the number of observations, and $\hat{\theta}$ is the MLE of parameter θ.

3 Main results

First, we collect data on daily COVID-19 cases confirmed with RT-PCR or ATK and data on daily deaths directly from Thailands Department of Disease Control website [2]. The daily COVID-19 confirmed case and mortality data are gathered from January 1, 2022 until April 7, 2022. Next, we analyze our data by using the five probability distributions: normal, log-normal, Weibull, log-logistic distribution, and inverse Weibull distributions. Parameters of these five distributions are estimated by MLE. To determine the best distribution that fits our data, we apply two criteria: AIC and BIC.

In the following two tables we present the parameter estimates and the goodness of fit of five distributions of daily confirmed COVID-19 cases and of five distributions of daily COVID-19 deaths, respectively.

Distributions	Parameter estimates		AIC	BIC
Normal	$\hat{\mu} = 28761.690$	$\hat{\sigma} = 17931.160$	2179.367	2184.517
Log-Normal	$\hat{\mu} = 10.006$	$\hat{\sigma} = 0.784$	2173.274	2178.424
Weibull	$\hat{k} = 1.619$	$\hat{\lambda} = 32268.794$	2162.521	2167.670
Inverse Weibull	$\hat{\alpha} = 1.260$	$\hat{\beta} = 14774.407$	2193.547	2198.697
Log-Logistic	$\hat{\alpha} = 2.09003$	$\hat{\beta} = 23318.062$	2182.038	2187.188

Table 1: The parameter estimates and the goodness of fit of distributions for daily confirmed cases

Distributions	Parameter estimates		AIC	BIC
Normal	$\hat{\mu} = 42.165$	$\hat{\sigma} = 28.798$	931.175	936.324
Log-Normal	$\hat{\mu} = 3.484$	$\hat{\sigma} = 0.737$	895.949	901.098
Weibull	$\hat{k} = 1.524$	$\hat{\lambda} = 47.067$	900.886	906.036
Inverse Weibull	$\hat{\alpha} = 1.545$	$\hat{\beta} = 22.634$	898.603	903.753
Log-Logistic	$\hat{\alpha} = 2.189$	$\hat{\beta} = 32.272$	906.654	911.804

Table 2: The parameter estimates and the goodness of fit of distributions for daily deaths

According to Table 1, the Weibull distribution is the best fit data on daily confirmed cases because of its lowest AIC and BIC. Similarly, Table 2 shows that the log-normal distribution is the best fit data on daily confirmed deaths.

To visualize and compare the five distributions on data in vision, we next presents the histograms of daily confirmed cases and deaths in Figure 1 and Figure 2, respectively.

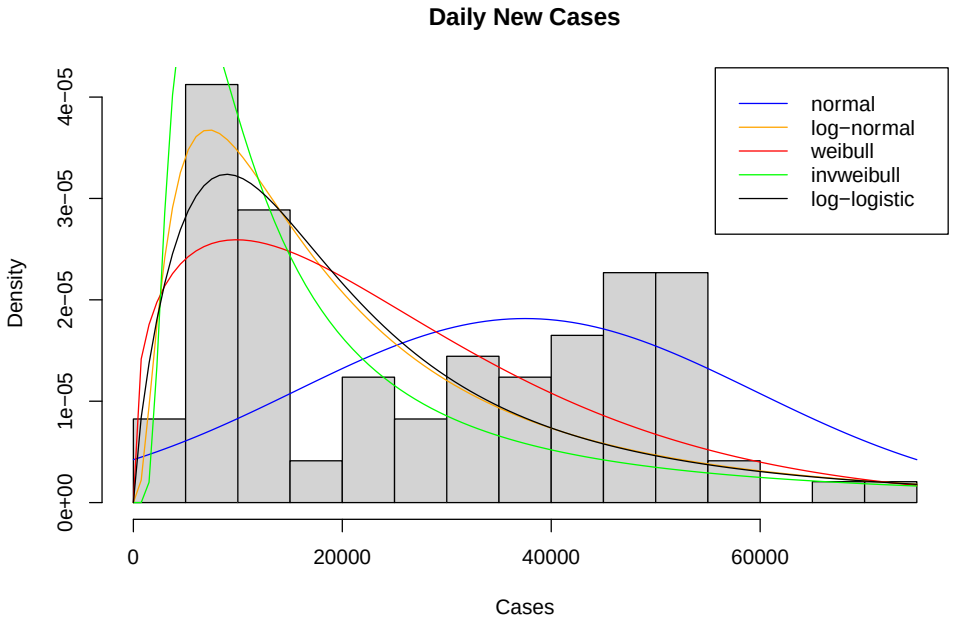

Figure 1: Histogram Covid-19 daily cases and fitted distributions

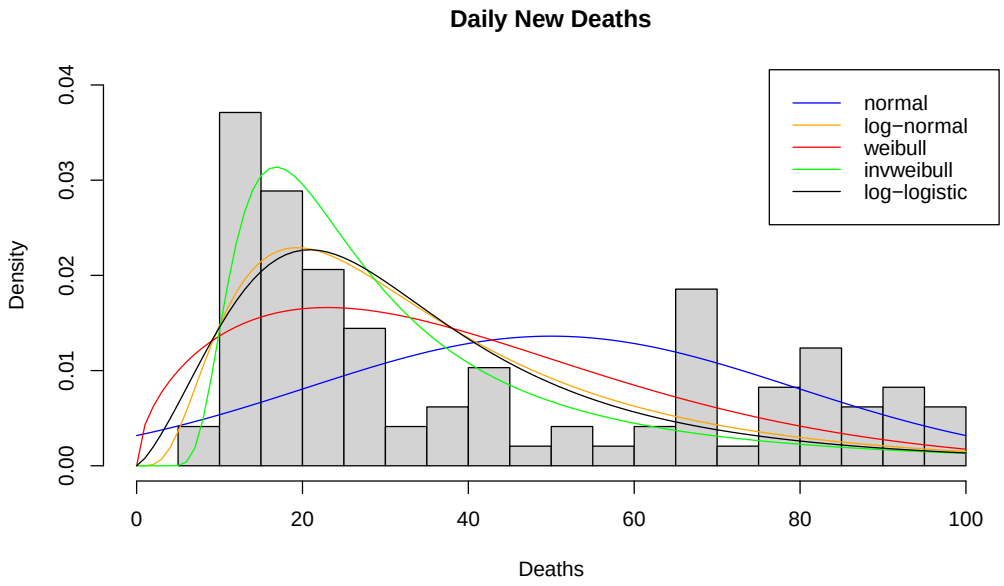

Figure 2: Histogram Covid-19 daily deaths and fitted distributions

4 Conclusions

In this research, the daily COVID-19 confirmed case data are accumulated and the mortality data in Thailand are gathered during the fifth wave from January 1, 2022 until April 7, 2022. We analyzed a model of the daily number of COVID-19 cases and deaths using five different probability distributions: normal, log-normal, Weibull, log-logistic distribution, and inverse Weibull distributions. The parameter estimators were obtained by using the MLE. The results indicate that the Weibull and the log-normal distributions were the most suitable distributions that fit the data on daily confirmed cases and on daily confirmed deaths, respectively, by using the AIC and the BIC.

References

[1] M. Ahsan-ul-Haq, M. Ahmed, J. Zafar, P. L. Ramos, Modeling of covid-19 cases in Pakistan using lifetime probability distributions, Ann Data Sci., (2021), 1–12.

[2] Department of Disease Control, Corona Virus Disease (COVID-19), https://ddc.moph.go.th/viralp-neumonia/index.php.

[3] E. Hincal, B. Kaymakamzade, F. N. Suren, N. Gokbulut, *Estimating Covid-19 eaths by using binomial model*, In AIP Conference Proceedings, 2325, AIP Publishing LLC, (2021), 020050-1–020050-4.

[4] K. Mansori, E. Gholami, M. Soltani-Kermanshahi, Statistical distribution of novel coronavirus in Iran, *International Journal of One Health*, **6**, no. 2, (2020), 143–146.

[5] M. M. Mansour, M. A. Farsi, S. M. Mohamed, E. M. Abd Elrazik, Modeling the Covid-19 pandemic dynamics in Egypt and Saudi Arabia, *Mathematics*, (2021), **9**, 827, 13 pages.

[6] A. Marshall, I. Olkin, *Life Distributions, Structure of Nonparametric, Semiparametric, and Parametric Families*, Springer, 2007.

[7] A. Ratnarathon. Coronavirus infectious disease-2019 (COVID-19): a case report, the first patient in Thailand and outside China, *Journal of Bamrasnaradura Infectious Diseases Institute*, **14**, no. 2, (2020), 116–123.

[8] Thai Residents, (2021), First Omicron case in Thailand–Timeline, https://thairesidents.com/health/first-omicron-case-in-thailand-timeline

[9] World Health Organization, (2021), Update on omicron. World Health Organization. https://www.who.int/news/item/28-11-2021-update-on-omicron.