International Journal of Mathematics Trends and Technology
ISSN: 2231-5373/ https://doi.org/10.14445/22315373/IJMTT-V69I1P508

Volume 69 Issue 1, 56-61, January 2023

Original Article

Mixture Distribution for Modeling Turkey Covid-19 Case

Hendri Fahrizal[1], Arisman Adnan[2], Rado Yendra[3]

[1,2]*Mathematics, University of Riau, Pekanbaru, Indonesia.*
[3]*Mathematics, UIN Sultan Syarif Kasim Riau, Pekanbaru, Indonesia.*

Received: 26 November 2022 Revised: 30 December 2022 Accepted: 09 January 2023 Published: 18 January 2023

Abstract *- The COVID-19 pandemic hit the world in early 2020 and caused many deaths in various countries, one of which was Turkey. In March 2020, the first case of COVID-19 was confirmed by the Turkish Ministry of Health. In May 2021, the rise of cases continued to occur until it reached 5.2 million cases with a death toll of 47,271 people. In this research, a model for Turkey's COVID-19 positive cases had built using a mixed distribution, namely log-normal, gamma and Weibull. Parameter estimation for each distribution was carried out using the maximum likelihood method. Then by using the Bayesian Information Criterion (BIC) and Akaike Information Criterion (AIC) tests, it was found that the 3-component gamma distribution is the best distribution to describe the pattern of spikes in positive cases of COVID-19 in Turkey.*

Keywords *- Covid-19, Gamma distribution, Log-normal distribution, Mixed distribution, Turkey, Weibull distribution.*

1. Introduction

At the beginning of 2020, Sars-Cov-2 or known as covid-19 was shocking the world. On March 11, 2020, the World Health Organization (WHO) announced covid-19 virus as pandemic. A month later, in the middle of the month April 2020, WHO reported that 311,847 death from 4,619,477 cases in the world. Until the end of May 2021, there have been 169.5 million cases with 3,530,582 deathl [16].

This virus also spread in Turkey when Turkish Ministry of Health reported the first case on March 11, 2020. Until May 2021, Turkey's Covid-19 cases have reached 5.2 million with 47,271 death cases. This situation forced Turkey to lockdown or restrictions on people's movements in April 2021.

One of the important parameters to control covid-19 is the number of positive patients. Some researchers have been modelled of the number of covid-19 positive patients using simple distributions. Qi fang [12] used a negative binomial distribution to modelling the transmission of the covid-19 virus in Shenzhen, China. But, somehow it's difficult to use that simple technique since the data may also contain more than one distribution. Therefore, we need a mixture distribution to apply in covid-19 cases.

Kremer et al [10]., have modelled the spread of secondary cases of covid-19 using the Poisson mixture distribution in Hong Kong, India and Rwanda. The study stated that mixture of Poisson and log-normal distribution was the best distribution. Then in August 2020, Polymenis [11] used an exponential mixture distribution to measure the risk level of covid-19 in Italy. In this study, we will model the first three months of positive covid-19 cases in Turkey by using the weibull, gamma log-normal mixture distribution.

2. The Data Set

We collect Turkey daily positive cases of Covid-19 data for 3 months (March to May 2020) that present in Figure 1 and Figure 2. The first case occurred on March 12, 2020 and the death cases began a week after the first positive case. On March 19, 2020, the number of positive cases reached 93 cases and fluctuated until the end of the March 2020. In April, it reached 5138 cases and forced Turkish government to implemented the large-social restrictions or known as lockdown.

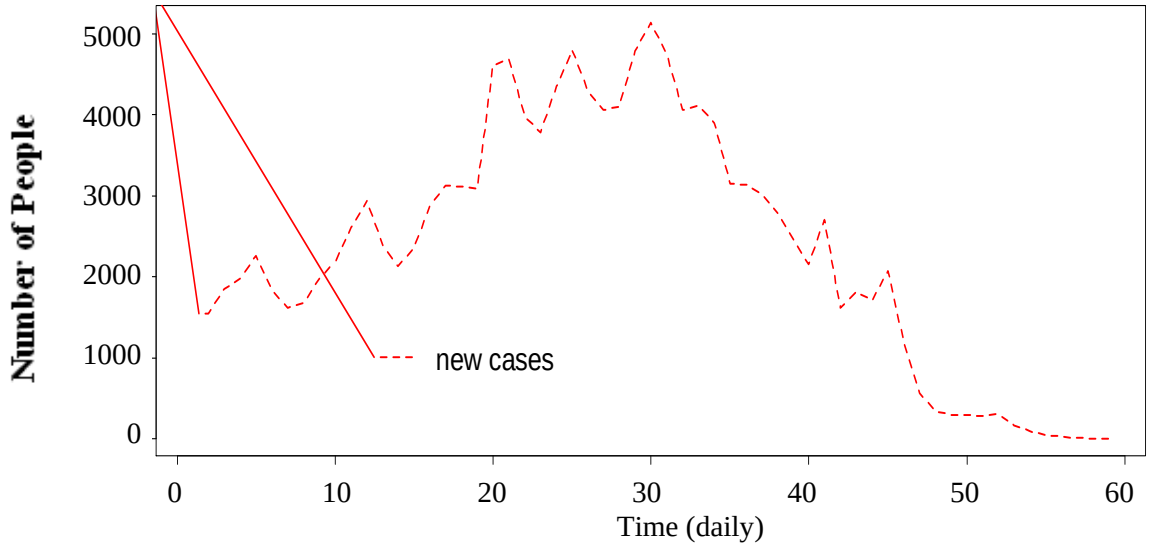

Fig. 1 Observed Turkey daily new cases covid-19

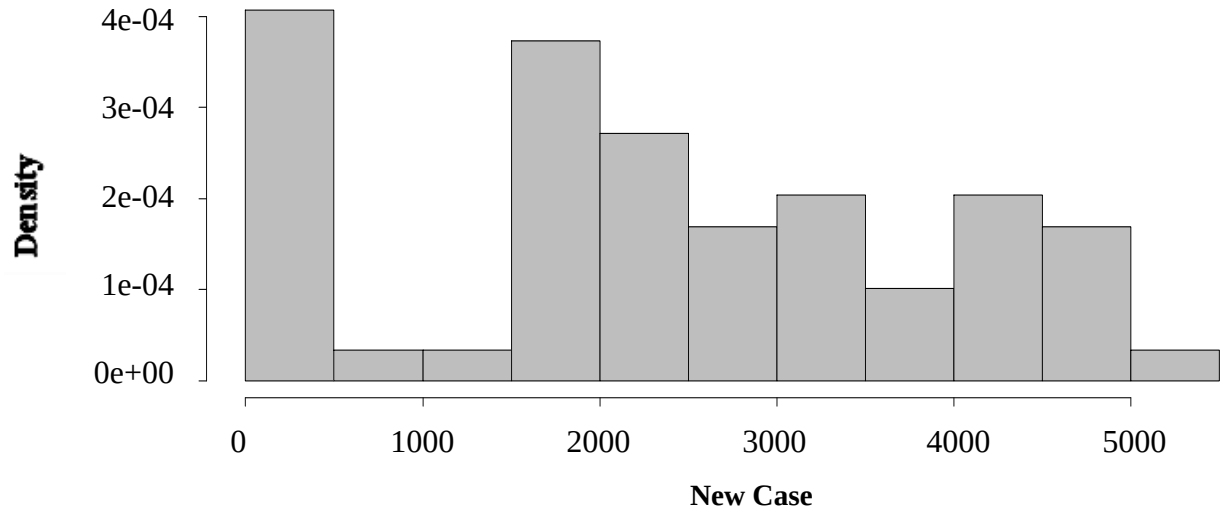

Fig. 2 The Histogram of daily new cases covid 19 of Turkey

3. Model Selection

Goodness of fit test is used to compare several different models distribution with the aim of determining the most representative model for the data. In this study, we use Akaike Information Criterion (AIC) and Bayesian test Information Criterion (BIC). The AIC defined as [1]

$$AIC = -2\,ln(L) + 2a,$$

where L is likelihood function of model and a is the number of parameter.

The BIC test is derived from the Bayesian framework as an estimate of the Bayes factor for two models and defined as follow [1]

$$BIC = -2\,ln(L) + s\,ln(n),$$

where n is number of sample.

4. Materials and Methods

4.1. Model Distribution

In this study, we have 9 distribution models to approximate Turkey daily covid-19 positive cases, namely: unmixed distributions (log-normal (L), gamma (G) and Weibull (W)); two-component mixture distribution models (2-component log-normal (ML2), 2-component gamma (MG2) and 2-component Weibull (MW2)) and three-components mixture distributions (3-component log-normal (ML3), 3- component gamma (MG3) and 3-component Weibull (MW3)).

Table 1. Probability density function of models

Distributions	Density Function
L	$f(x;\mu,\sigma)=\frac{1}{x\sqrt{2\pi(\sigma^2)}}e^{-\frac{(\ln x-\mu)^2}{2\sigma^2}}$
G	$f(x;\alpha,\beta)=\frac{\beta}{\Gamma(\alpha)}x^{\alpha-1}e^{-\beta x}$
W	$f(x;k,c)=\frac{k}{c}\left(\frac{x}{c}\right)^{k-1}e^{-\left(\frac{x}{c}\right)^k}$
ML2	$w_1 f(x;\mu_1,\sigma_1)+w_2 f(x;\mu_2,\sigma_2),\quad w_1+w_2=1$
MG2	$w_1 f(x;\alpha_1,\beta_1)+w_2 f(x;\alpha_2,\beta_2),\quad w_1+w_2=1$
MW2	$w_1 f(x;k_1,c_1)+w_2 f(x;k_2,c_2),\quad w_1+w_2=1$
ML3	$w_1 f(x;\mu_1,\sigma_1)+w_2 f(x;\mu_2,\sigma_2)+w_3 f(x;\mu_3,\sigma_3),\quad w_1+w_2+w_3=1$
MG3	$w_1 f(x;\alpha_1,\beta_1)+w_2 f(x;\alpha_2,\beta_2)+w_3 f(x;\alpha_3,\beta_3),\quad w_1+w_2+w_3=1$
MW3	$w_1 f(x;k_1,c_1)+w_2 f(x;k_2,c_2)+w_3 f(x;k_3,c_3),\quad w_1+w_2+w_3=1$

4.2. Maximum Likelihood

The maximum likelihood method is a method to estimate parameter of a distributor by using the maximum probability. First, make likelihood function for joint probability of density function $f(x_i|a_1,a_2,\dots,a_k)$ that defined as [9]

$$L(a|x)=\prod_{i=1}^{n} f(x_i\,|a_1,a_2,\dots,a_k).$$

Then make log-likehood function that defined as [4]

$$l(a|x)=ln\,L\,(a|x).$$

After that, using partial derivation to maximum log-likehood function

5. Results and Discussion

By using R application, we get the value of parameters for each model that present in Table 2.

Table 2. Computed parameters of each model

	L	G	W	$ML2$	$MG2$	$MW2$	$ML3$	$MG3$	$MW3$
μ	7.101	–	–	–	–	–	–	–	–
σ	1.851	–	–	–	–	–	–	–	–
α	–	0.886	–	–	–	–	–	–	–
β	–	0.00037	–	–	–	–	–	–	–
k	–	–	1.069	–	–	–	–	–	–
c	–	–	2391.22	–	–	–	–	–	–
μ_1	–	–	–	4.602	–	–	4.55	–	–
σ_1	–	–	–	2.241	–	–	2.217	–	–
μ_2	–	–	–	7.92	–	–	7.73	–	–
σ_2	–	–	–	0.379	–	–	0.281	–	–

μ_3	–	–	–	–	–	–	8.382	–	–
σ_3	–	–	–	–	–	–	0.087	–	–
α_1	–	–	–	–	0.568	–	–	0.568	–
β_1	–	–	–	–	0.0026	–	–	0.0026	–
α_2	–	–	–	–	7.085	–	–	13.602	–
β_2	–	–	–	–	0.0023	–	–	0.0058	–
α_3	–	–	–	–	–	–	–	120	–
β_3	–	–	–	–	–	–	–	0.027	–
k_1	–	–	–	–	–	0.635	–	–	0.654
c_1	–	–	–	–	–	191.402	–	–	195.09
k_2	–	–	–	–	–	2.872	–	–	4.402
c_2	–	–	–	–	–	3318.16	–	–	2498.894
k_3	–	–	–	–	–	–	–	–	10.276
c_3	–	–	–	–	–	–	–	–	4495.856
w_1	–	–	–	0.2487	0.2235	0.2234	0.246	0.225	0.229
w_2	–	–	–	0.7513	0.77654	0.7766	0.530	0.539	0.503
w_3	–	–	–	–	–	–	0.224	0.236	0.268

The probability density function for each research model can be formed based on Table 1 and Table 2. Then, we use the graphical technique namely probability density function (pdf) plot to observe the model distributions.

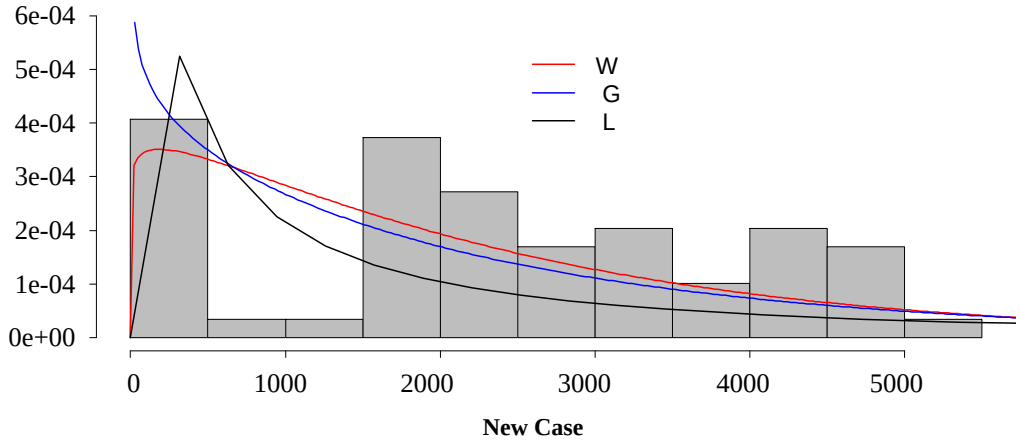

Fig. 3 Pdf plot for comparisons Predicted and observed daily new cases covid 19 in Turkey for L, G and W distribution

In Figure 3, it can be seen that the Log-normal distribution is unmixed (model L, G and W) is not fit to approximating the Turkey daily covid-19 positive cases histogram. The unmixed model only has 1 wave, while there are 3 waves in observation data. Further on, we also graph the pdf plot for two-component mixture distribution that present in Figure 4.

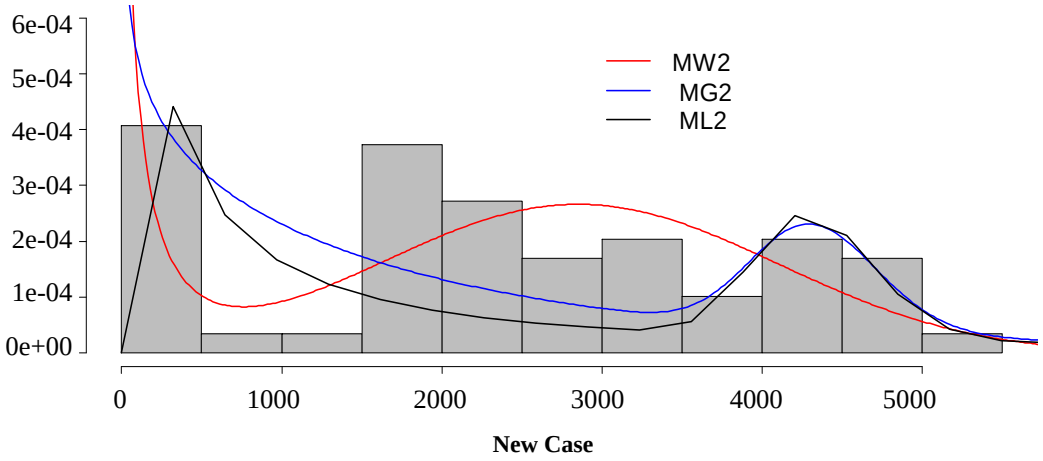

Fig. 4 Pdf plot for comparisons Predicted and observed daily new cases covid 19 in Turkey for ML2, MG2 and MW2 distribution

Figure 4 describe that the two-component mixture distribution are better than unmixed distribution since has 2 waves in pdf plot. But these models are not fit to for the data. Finally, we have 3-component mixture distribution plot in Figure 5.

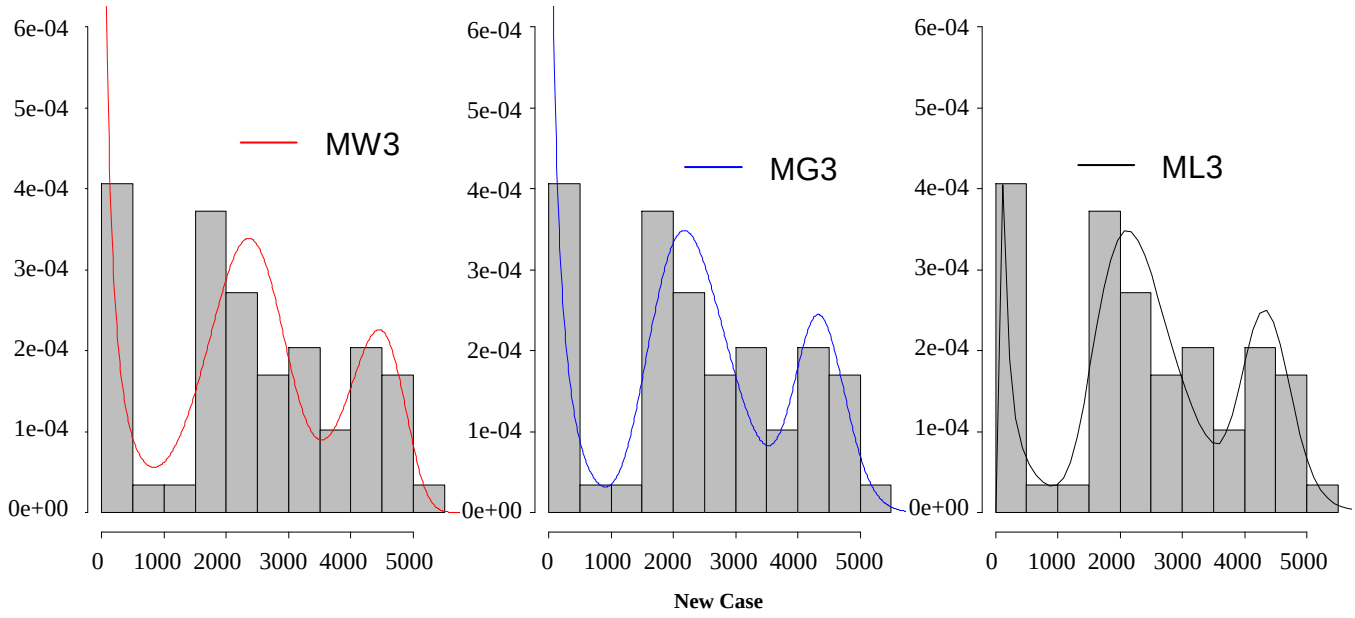

Fig. 5 Pdf plot for comparisons Predicted and observed daily new cases covid 19 in Turkey for ML3, MG3 and MW3 distribution

It is clearly seen that the 3-component mixture distributions (ML3, MG3 and MW3) approximate Turkey covid-19 cases very well. The goodness fit test was carried out by using AIC) and BIC to select the best model in Table 3.

Table 3. The Goodness and fit test result of the daily new cases covid 19 in Turkey

	L	G	W	ML2	MG2	MW2	ML3	MG3	MW3
AIC	1082.1	1037.3	1037.6	1063.8	1029.8	1003.4	999.7	**996.3**	998.8
BIC	1086.2	1041.5	1041.7	1074.2	1040.2	1013.8	1016.3	**1012.9**	1015.4
Log(L)	-539.03	-516.7	-516.8	-526.9	-509.9	-496.7	-491.8	**-490.2**	-491.4

6. Conclusion

Based on the discussion, 3-component log-normal mixture distribution, 3-component gamma mixture distribution and 3-component Weibull mixture distribution can be used for illustrates Turkey covid-19. From the goodness fit results, we obtain that the 3-component gamma distribution model is the best model.

References

[1] Henry de-Graft Acquah, "Comparison of Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) in Selection of an Asymmetric Price Relationship," *Journal of Development and Agricultural Economics*, vol. 2, no. 1, pp. 001-006, 2010.

[2] Kendall Atkinson, and Weimin Han, *Elementary Numerical Analysis*, 3rd Edition, John Wiley & Sons, New York, 2003.

[3] Lee Bain, and Max Engelhardt, *Introduction to Probability and Mathematical Statistics*, 2nd Edition, Duxbury Press, California, 1992.

[4] Sami Bourouis, Abdullah Alharbi, and Nizar Bouguila, "Bayesian Learning of Shifted-Scaled Dirichlet Mixture Models and its Application to Early COVID-19 Detection in Chest X-Ray Images," *Journal of Imaging*, vol. 7, no. 1, pp. 1-14, 2021. *Crossref*, https://doi.org/10.3390/jimaging7010007

[5] G. Casella, and R. L. Berger, *Statistical Inference*, Second Edition, Pasific Grove, California, 2002.

[6] Charles Ebeling, *An Introduction to Reliability and Maintainability Engineering*, McGraw Hill Education, New York, 1997.

[7] M. A. T. Elshahat, and A. A. M. Mahmoud, "A Study on the Mixture of Exponentiated-Weibull Distribution, Part I (The Method of Maximum Like-Lihood Estimation)," *International Journal of Mathematics and Statistics Studies*, vol. 5, pp. 9-28, 2017.

[8] Niko A. Kaciroti et al., "A Bayesian Mixture Model for Predicting the COVID-19 Related Mortality in the United States," *The American Journal of Tropical Medicine and Hygiene*, vol. 104, no. 4, pp. 1484-1492, 2021. *Crossref*, https://doi.org/10.4269/ajtmh.20-1147

[9] Robert V. Hogg, Joseph W. McKean, and Allen T. Craig, *Introduction to Mathematical Statistics*, 8th Edition, Macmillan Publishing Inc, Newyork, 2019.

[10] Cécile Kremer et al., "Quantifying Super Spreading for COVID-19 using Poisson Mixture Distributions," *Scientific Report*, Nature Portofolio, 2021. *Crossref*, https://doi.org/10.1038/s41598-021-93578-x

[11] Athanase Polymenis, "An Application of a Mixture of Exponential Distributions for Assessing Hazard Rates from COVID-19," *Journal of Population Therapeutics & Clinical Pharmacology*, vol. 27, no. SP1, pp. e58-e64, 2020. *Crossref*, https://doi.org/10.15586/jptcp.v27iSP1.721

[12] Q Qifang Bi et al., "Epidemiology and Transmission of COVID-19 in 391 Cases and 1286 of their Close Contacts in Shenzhen, China: A Retrospective Cohort Study," *Lancet Infectious Diseases*, vol. 20, no. 8, pp. 911-919, 2020. *Crossref*, https://doi.org/10.1016/S1473-3099(20)30287-5

[13] Sheldon M. Ross, *Introduction to Probability Models*, 10th Edition, Elsevier, 2010.

[14] Tabassum Naz Sindhu et al., "Estimation Method of Mixture Distribution and Modeling of COVID-19 Pandemic," *AIMS Mathematics*, vol. 7, pp. 9926-9956, 2022. *Crossref*, https://doi.org/10.3934/math.2022554

[15] World Health Organization, Coronavirus Disease COVID- Pandemic, Geneva, Switzerland: WHO, 2020. [Online]. Available: www.who.int/emergencies/diseases/novel-coronavirus-2019

[16] World Health Organization, WHO Coronavirus Disease, 2021. [Online]. Available: https://www.who.int/countries/tur/