Bulletin of Environment, Pharmacology and Life Sciences
Bull. Env. Pharmacol. Life Sci., Special Issue [1]2022 : 1431-1434

Online ISSN 2277-1808
Journal's URL:http://www.bepls.com
CODEN: BEPLAD
ORIGINAL ARTICLE **OPEN ACCESS**

A Study of Statistical Distribution of Covid-19 Pandemic Cases in Second Wave with Reference to Khanapur Tehsil

Prasad Gavali[1]*, Annasaheb Suryawanshi[1], Shrikant Gimawekar[1], Pritesh Gugale[2]
[1]Assistant Professor, Department of Statistics, Balwant College, Vita
[2]Assistant Professor, Department of Zoology, Balwant College, Vita
***Corresponding Author:** Mr. P. A. Gavali; prasadgavali11886@gmail.com

ABSTRACT
The present research deals with statistical analysis of COVID 19 pandemic cases in Khanapurtahsil. For statistical analysis we consider pandemic cases in second wave of COVID 19. The data was secondary and it was taken from 1 April 2021 to 31 July 2021. The objective of this research paper is to study the statistical distribution of number of new cases as well as number of deaths per day.
Keywords: *Lognormal, Poisson, Kolmogrov-Smirnov, COVID-19, EasyFit, SPSS.*

Received 11.02.2022 Revised 19.03.2022 Accepted 20.04.2022

INTRODUCTION
It is a challenging task to properly analyze the data, statistics has become very important in the current pandemic of COVID-19. Statistician playing an important role to the fight against COVID- 19 [1]. Using statistical technique, it is possible to estimate the mortality rate, recovery rate and the potential number of new patients. Tests conducted to assess the potential preventive effect of a vaccine and its risk factors can be estimated. In Wuhan city, China the first case of COVID-19 Virus was found. After that the COVID-19 virus has spread throughout the world. It has affected the economic, social and educational systems of every country. Fever, Cough and difficulties in breathing are the main symptoms of COVID -19. The second largest populous country in the world is India. The first case of COVID-19 virus in India was reported in Kerala in January, 2020 and in Maharashtra on March 9, 2020 in Pune. The first case of COVID-19 virus was found in Sangli district at the end of March 2020 and since then the number of patients has been increasing day by day. Davit Gondauri et al. [1] studied the COVID-19 Virus spreading statistics on the cases from the different countries. They found that high correlation coefficients and determinants between the total volumes of virus spread and recovery. Palash Ghosh *et al*. [2] carried out state wise analysis and Prediction of COVID-19 in India. Abhishek Jaglan et al. studied trend analysis using Machine Learning Techniques.Most of the researchers find out the statistical model for prediction of COVID- 19 pandemic cases. In this research we try to fit the statistical distribution on number of new cases and number of deaths per day during 1st April 2021 to 31st July 2021 of COVID-19.

MATERIAL AND METHODS
Khanapur tehsil is one of the tehsils in Sangli district belongs to state of Maharashtra. According to 2011 census the total population of Khanapur tehsil is 1.7 lakhs. A total of 13874 COVID cases were found in Khanapur till 31st July 2021. For this research we consider the data from 1st April 2021 to 31st July 2021. We were collected secondary data from Sangli district government website. The totals of 10598 cases were found in Khanapur tehsil in the period of 1st April 2021 to 31st July 2021. During this period the spread of corona virus was more, hence we decided to analyse the data during this period. The analysis of data has been carried out by using statistical software's SPSS andEasyFit. In present study, the collected considered as failure time data. For this failure time data, the many reliability distributions may be fitted such as Exponential, Log-Normal, Weibull, Pareto, Gumbel Max etc. We tried to fit all these distributions for the given data but the Log-Normal distribution was good fitted for number of new cases per day. For considering the number of deaths per day, the Poisson distribution was good fitted. To test goodness of fit, the Kolmogorov-Smirnov test was used [3,4].

STATISTICAL ANALYSIS

To fit the collected data, we used two distributions for number of new cases and number of deaths per day were Log- normal and Poisson distribution respectively as follows.

Log-Normal distribution:

The Log-Normal is continuous type probability distribution and logarithm of its random variable follows Normal distribution and also Exponential form of Normal random variable has Log-Normal distribution. The Log-Normal distributions has many applications in Biology, Medicine, Social sciences and demography. It is Heavy tailed probability distribution. We consider two parameter Log Normal distribution with parameters μ and σ2. Its probability density function is given by:

$$f(x) = \begin{cases} \frac{1}{\sigma\sqrt{2\pi}x} e^{-\frac{1}{2}\left(\frac{\ln x-\mu}{\sigma}\right)^2}; & -\infty < x < \infty, -\infty < \mu < \infty, \sigma > 0 \\ 0 \quad ; & otherwise \end{cases}$$

Poisson distribution:

In Statistics, the Poisson distribution is used to show frequency of an event over a specified period. It is used to show the rare events. Here the number of deaths is counted per day is discrete type random variable. We used Poisson distribution to fit the distribution of number of deaths per day. The probability mass function of Poisson distribution with parameter ⍰ is given by:

$$P(X = x) = \begin{cases} \frac{e^{-\lambda}\lambda^x}{x!}; & x = 0,1,2,\ldots \\ 0 \quad ; & otherwise \end{cases}$$

Kolmogorov-Smirnov:

To compare a sample with reference probability distribution the Kolmogorov- Smirnov test is applicable. This test can be used instead of Chi-square goodness fit test. The null and alternative hypothesis for Kolmogorov-Smirnov test is:

i) H0: The distribution of number of new cases per day is lognormal.
H1: The distribution of number of new cases per day is not lognormal.
ii)H0: The distribution of Number of deaths per day is Poisson distribution.
H1: The distribution of number of deaths per day is not Poisson distribution.
The values of test statistic and critical value of K-S shown in Table 1

Table-1: Kolmogorov-Smirnov test Summary

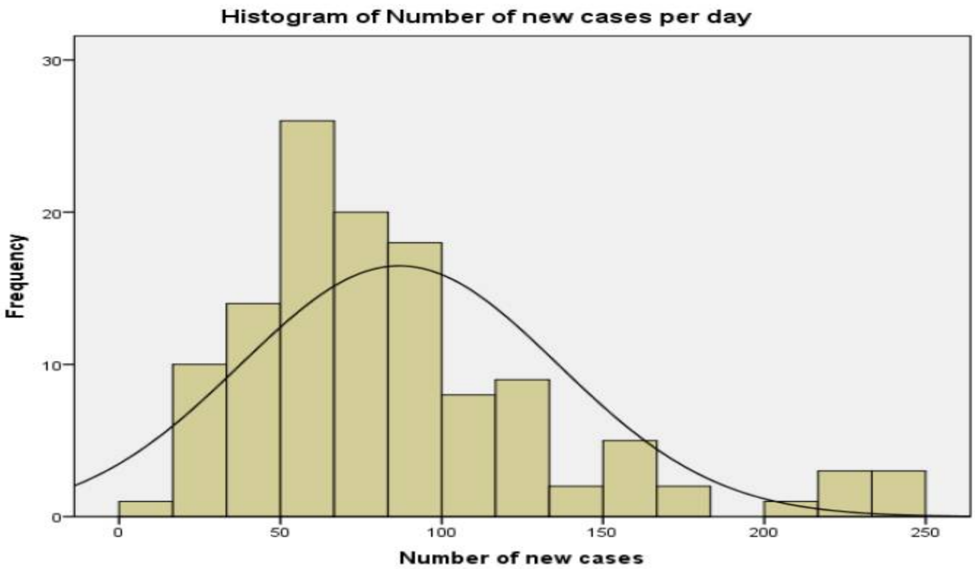
Variable	Value of K-S test statistic	Critical value**	Decision	Distribution fitted
Number of new cases per day	0.06726	0.14757	We fail to reject H_0	Lognormal
Number of deaths per day	0.14999	0.49	We fail to reject H_0	Poisson

RESULTS AND DISCUSSION

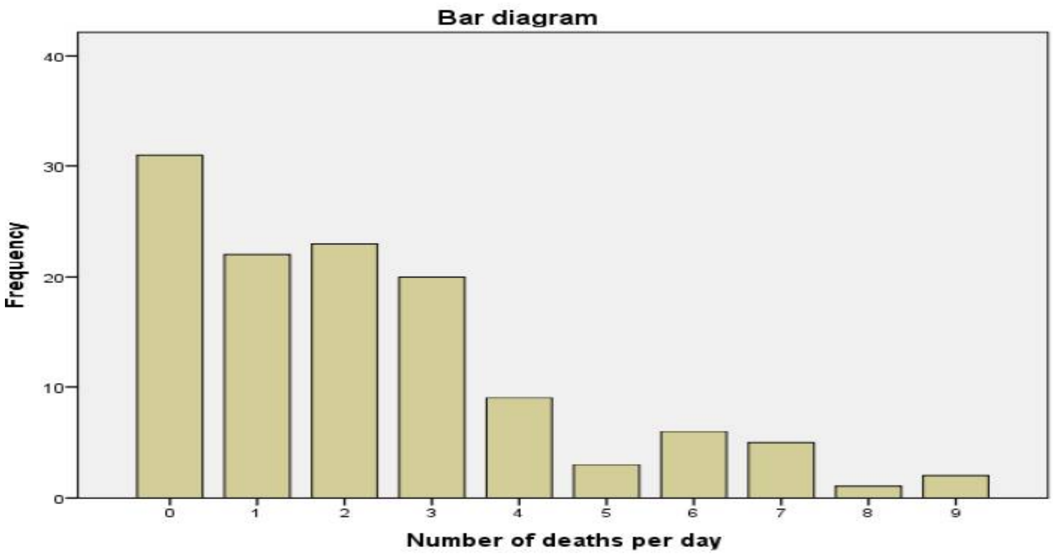
1. The value of test statistic of Kolmogorov-Smirnov test for testing the null hypothesis that the number of new cases per day follows lognormal distribution is 0.06726 which is less than critical value of Kolmogorov-Smirnov test. Hence, we fail to reject the null hypothesis at 1% level of significance. Thus, there is insufficient evidence to conclude that the number of new cases per day does not follow lognormal (i.e. The distribution of number of new cases per day is lognormal).
2. The value of test statistic of Kolmogorov-Smirnov test for testing the null hypothesis that the number of deaths per day follows Poisson distribution is 0.14999 which is less than critical value of Kolmogorov-Smirnov test. Hence, we fail to reject the null hypothesis at 1% level of significance. Thus there is insufficient evidence to conclude that the number of deaths per day does not follow Poisson (i.e.The distribution of Number of deaths per day is Poisson distribution).

GRAPHICAL REPRESENTATION

For Graphical presentation, we used histogram and bar diagram. Histogram and Bar diagram drawn with the help of SPSS software. For showing frequency distribution of number of new cases per day, we used histogram (graph- 1) and the frequency distribution of number of deaths per day shown graphically using bar diagram (graph-2).



Graph-1: Distribution of Number of new cases per day



Graph-2: Distribution of Number of deaths per day

CONCLUSION

The fitted distribution for number of new cases per day follows Log-Normal distribution. Log normal distribution is the right tailed distribution. It shows that fewer cases were found in more days. Similarly the number of deaths per day follows Poisson distribution. Poisson distribution also positively skewed and the given data analysis shows that the maximum number of days reported no deaths. On the basis of fitted distribution we can predict the possibility of occurrence of new cases as well as number of deaths.

REFERENCES

1. Davit Gondauri, EkaterineMikautadze, Mikheil Batiashvili (2020). "Research on based on the Examples of the Cases from Different Countries" Electronic Journal of General Medicine 2020, 17(4).

2. Palash Ghosh, Rik Ghosh and Bibhas Chakraborty (2020). “COVID-19 in India: Statewise Analysis and Prediction” JMIR Public Health and Surveillance 2020, Jul- Sep; 6(3).
3. Abhishek Jaglan, Daksh Trehan and PriyanshSinghal (2020).“COVID-19 Trend Analysis using Machine Learning Techniques” International Journal of Scientific & Engineering Research; Volume 11, Issue 12.
4. https://sangli.nic.in

CITATION OF THIS ARTICLE

P Gavali, A Suryawanshi, S Gimawekar, P Gugale. A Study of Statistical Distribution of Covid-19 Pandemic Cases in Second Wave with Reference to Khanapur Tehsil. Bull. Env.Pharmacol. Life Sci., Spl Issue [1] 2022 : 1431-1434