International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 • Website: www.ijfmr.com • Email: editor@ijfmr.com

A Long-Term Prediction of Covid-19 By Building a Transition Probability Matrix

Biplob Saikia

Student, Dibrugarh University, Dibrugarh, Assam, India

Abstract
The outbreak of the covid-19 pandemic caused a devastating scenario in the human lives for a long period of time. Since the outbreak of the pandemic, many researchers have published papers on the prediction of covid-19 using machine learning techniques, artificial intelligence, statistical techniques such as regression analysis, Markov chain models etc. In this paper a transition probability matrix is introduced by building a first order Markov chain. This first order Markov chain is built by generating sequences from the data. Later with the help of Chapman Kolmogorov theorem and ergodic theorem, higher order transition probabilities and limiting probabilities are obtained.

Keywords: Covid-19, transition probabilities, Markov chain, ergodic theorem.

1. Introduction

An epidemic, pandemic is a random phenomenon which cannot be predetermined. In the recent years we have faced the Covid-19 pandemic and we already seen its impact on the loss of human lives, their health, food systems. The pandemic has a huge impact on the economy in the form of losing jobs, closed enterprises, difficulty in trading among neighbouring countries, livelihoods at risk. The governments of various countries imposed lockdowns periodically, imposed social distancing, must have to wear face mask while travelling, to stop the spreading of this virus and various vaccines were made such as Pfizer, covaxin, covishield so that peoples get immune after having these vaccines. These government restrictions help a lot in not spreading the virus among the peoples and reviving the economy to a great extent. Here we considered a 3 months data in the beginning of covid-19 and wants to check the

a) Impact of the virus for a period of 3 years.
b) Impact of the virus in the long run.

2. Collection of Data

The data collected here is from the website coronavirus.data.gov.uk from date 21.03.2020 to 21.06.2020. Considering only two data sets viz. new cases by specimen date, new daily nso deaths by death date. To build the transition probability matrix there is the need of transition states. Considering new cases by specimen data, new daily nso deaths by death date as the two states i.e transition from new cases by specimen date to new cases by specimen date, transition from new cases by specimen date to new daily nso deaths by death date, transition from new daily nso deaths by death date to new cases by specimen date, transition from new daily nso deaths by death date to new daily nso deaths by death date.

3. Methodology

In this paper a transition probability matrix is developed through which we can calculate the higher order transition probabilities, and the limiting probabilities as well.

The higher order transition probabilities are obtained through the Chapman Kolmogorov theorem and the limiting probabilities are obtained through ergodic theorem. The equation for the Chapman Kolmogorov is given by

$$P^{m+n} = P^{m \cdot} P^{n}$$

A special case of the Chapman Kolmogorov for two states (say j and k) i.e to reach k from j in m+n steps is given by

$$P_{jk}{}^{m+n} = \sum_{r} p_{jr} p_{rk}$$

Where r is an intermediate state.

The equation for the ergodic theorem is given by

$$\pi = \pi' P$$

where $\pi = \begin{pmatrix} \pi_1 \\ \pi_2 \end{pmatrix}$ is a row vector and P is a transition probability matrix.

To build the transition probability matrix we generate sequences from the data as follows: -

Suppose the two data sets viz. new cases by specimen date and new daily nso deaths by death date are given in form

$N_{21,3},\ N_{22,3},\ N_{23,3},\ N_{24,3},\ N_{25,3},\ N_{26,3},\ N_{27,3},\ N_{28,3},\ N_{29,3},\ N_{30,3}, N_{31,3},\ N_{01,4}, \ldots\ldots\ldots\ldots\ldots$
$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,\ N_{20,6},\ N_{21,6}$

$D_{21,3}, D_{22,3}, D_{23,3}, D_{24,3}, D_{25,3}, D_{26,3}, D_{27,3}, D_{28,3}, D_{29,3}, D_{30,3}, D_{31,3}, D_{01,3}, \ldots\ldots\ldots\ldots\ldots\ldots$
$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,\ D_{20,6}, D_{21,6}$

Generate sequences from the data as follows

$|N_{21,3} - N_{21,3}|,\ |N_{21,3} - N_{22,3}|,\ |N_{21,3} - N_{23,3}|,\ |N_{21,3} - N_{24,3}|,\ |N_{21,3} - N_{25,3}|,\ |N_{21,3} - N_{26,3}|,\ |N_{21,3} - N_{27,3}|,\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,\ |N_{21,3} - N_{20,6}|,\ |N_{21,3} - N_{21,6}|$

$|N_{21,3} - D_{21,3}|,\ |N_{21,3} - D_{22,3}|,\ |N_{21,3} - D_{23,3}|,\ |N_{21,3} - D_{24,3}|,\ |N_{21,3} - D_{25,3}|,\ |N_{21,3} - D_{26,3}|,\ |N_{21,3} - D_{27,3}|,\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,\ |N_{21,3} - D_{20,6}|,\ |N_{21,3} - D_{21,6}|$

$|D_{21,3} - N_{21,3}|,\ |D_{21,3} - N_{22,3}|,\ |D_{21,3} - N_{23,3}|,\ |D_{21,3} - N_{24,3}|,\ |D_{21,3} - N_{25,3}|,\ |D_{21,3} - N_{26,3}|,\ |D_{21,3} - N_{27,3}|,\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots,\ |D_{21,3} - N_{20,6}|,\ |D_{21,3} - N_{21,6}|$

$|D_{21,3} - D_{21,3}|,\ |D_{21,3} - D_{22,3}|,\ |D_{21,3} - D_{23,3}|,\ |D_{21,3} - D_{24,3}|,\ |D_{21,3} - D_{25,3}|,\ |D_{21,3} - D_{26,3}|,\ |D_{21,3} -$

$D_{27,3}|$, ………………………., $|D_{21,3} - D_{20,6}|$, $|D_{21,3} - D_{21,6}|$

Let

$f_{NN} = |N_{21,3} - N_{21,3}| + |N_{21,3} - N_{22,3}| + |N_{21,3} - N_{23,3}| + |N_{21,3} - N_{24,3}| + |N_{21,3} - N_{25,3}| + |N_{21,3} - N_{26,3}| + |N_{21,3} - N_{27,3}| + \ldots\ldots + |N_{21,3} - N_{20,6}|$

$f_{ND} = |N_{21,3} - D_{21,3}| + |N_{21,3} - D_{22,3}| + |N_{21,3} - D_{23,3}| + |N_{21,3} - D_{24,3}| + |N_{21,3} - D_{25,3}| + |N_{21,3} - D_{26,3}| + |N_{21,3} - D_{27,3}| + \ldots\ldots + |N_{21,3} - D_{20,6}|$

$f_{DN} = |D_{21,3} - N_{21,3}| + |D_{21,3} - N_{22,3}| + |D_{21,3} - N_{23,3}| + |D_{21,3} - N_{24,3}| + |D_{21,3} - N_{25,3}| + |D_{21,3} - N_{26,3}| + |D_{21,3} - N_{27,3}| + \ldots\ldots + |D_{21,3} - N_{20,6}|$

$f_{DD} = |D_{21,3} - D_{21,3}| + |D_{21,3} - D_{22,3}| + |D_{21,3} - D_{23,3}| + |D_{21,3} - D_{24,3}| + |D_{21,3} - D_{25,3}| + |D_{21,3} - D_{26,3}| + |D_{21,3} - D_{27,3}| + \ldots\ldots + |D_{21,3} - D_{20,6}|$

Consider

$$p_{NN} = P\left(|N_{21,3} - N_{21,6}| = i / f_{NN}\right) = \frac{f_{NN}}{f_{NN} + f_{ND}}$$

$$p_{ND} = P\left(|N_{21,3} - D_{21,6}| = j / f_{ND}\right) = \frac{f_{ND}}{f_{NN} + f_{ND}}$$

$$p_{DN} = P\left(|D_{21,3} - N_{21,6}| = k / f_{DN}\right) = \frac{f_{DN}}{f_{DN} + f_{DD}}$$

$$p_{DD} = P\left(|D_{21,3} - D_{21,6}| = l / f_{DD}\right) = \frac{f_{DD}}{f_{DN} + f_{DD}}$$

A 2×2 first order transition probability matrix is build using $p_{NN}, p_{ND}, p_{DN}, p_{DD.}$

$$P = \begin{pmatrix} p_{NN} & p_{ND} \\ p_{DN} & p_{DD.} \end{pmatrix}$$

4. Results

Higher order transition probabilities viz. new cases by specimen date to new cases by specimen date, new cases by specimen date to new daily nso deaths by death date, new daily nso deaths by death date to new cases by specimen date, new daily nso deaths by death date to new daily nso deaths by death date can be obtained using the chapman Kolmogorov theorem. Higher order transition probabilities such as new cases by specimen date to new cases by specimen date, new daily nso deaths by death date to new daily nso deaths by death date are graphically shown below.

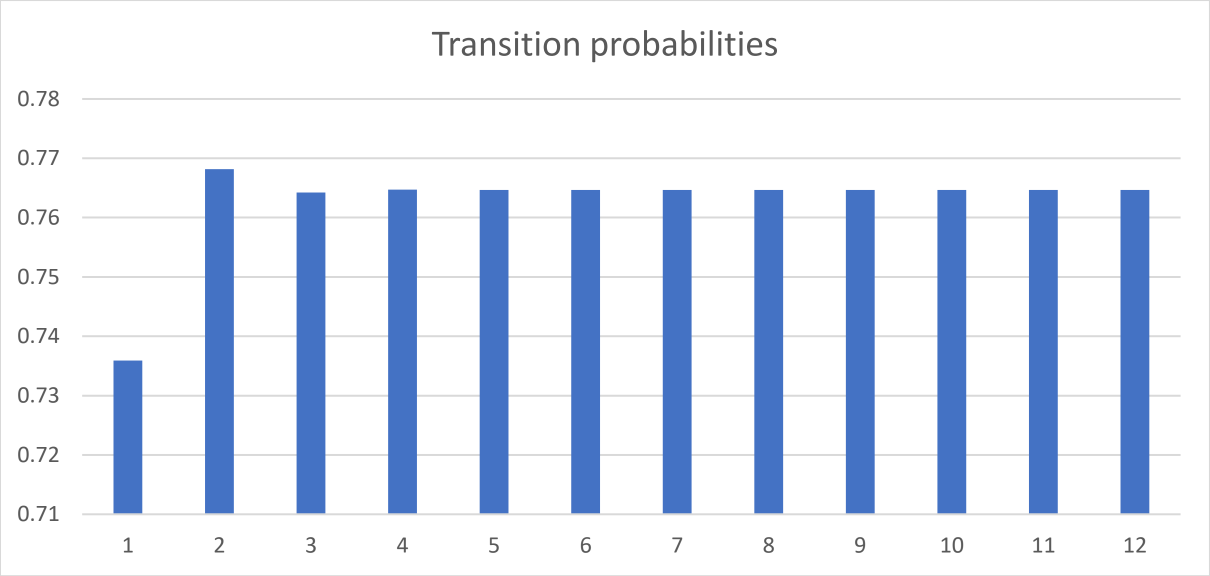

Fig 1: - transition probabilities from new cases by specimen date to new cases by specimen date is shown.

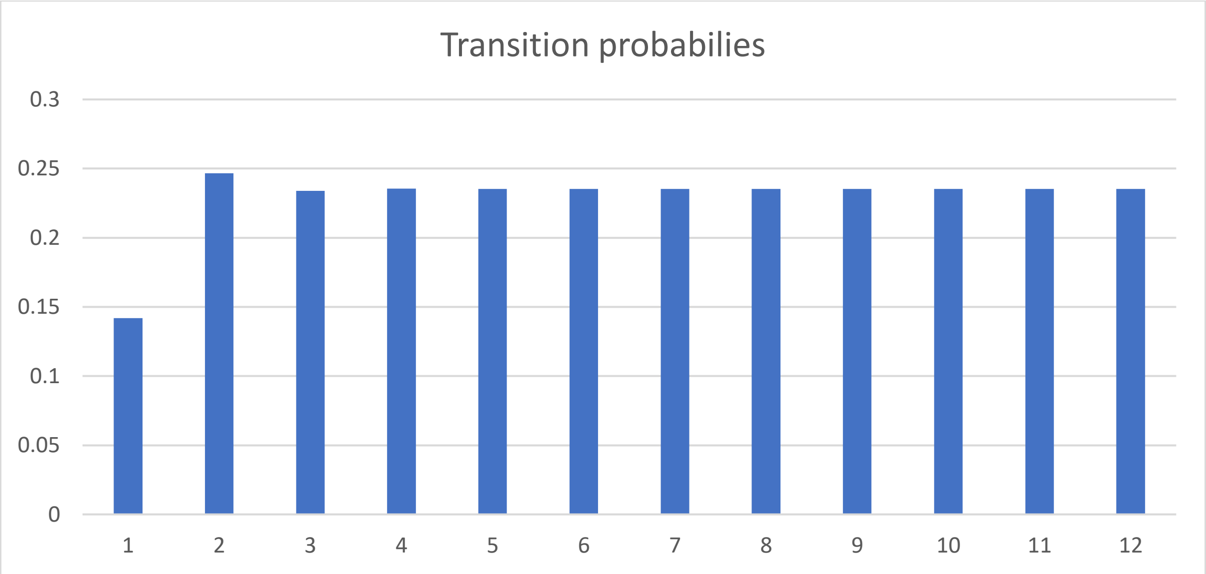

Fig 2: - transition probabilities from new daily nso deaths by death case to new daily nso deaths by death case is shown.

Fig 1. shows the transition probabilities from new cases by specimen date to new cases by specimen date within a duration of 3 months.

Similarly, fig.2 shows the transition probabilities from new daily nso deaths by death case to new daily nso deaths by death case within a duration of 3 months.

It is clear from the figures that for a period of 3 years there is low evidence in the decrement of new cases and deaths cases.

The limiting probabilities are

$$\pi_1 = 0.764666222$$
$$\pi_2 = 0.235333778$$

Thus, in the long run we can say that there is no evidence in the decrement of new cases. There is a probability of 0.764666222 being a person being exposed to covid-19. However, the probability of death being low in the long run which is quite a satisfactory result. Imposing governments restrictions will definitely help in reducing the number of new cases and death cases.

References

1. Wang. Xipeng, Ma. Yuan, Di. Junru, Murphey. L. Yi, Qiu. Shiqi, Kristinsson. Johannes, Meyer. Jason, Tseng. Finn, Feldkamp. Timothy, "Building efficient transition probability matrix using machine learning from big data for personalized route prediction", Procedia Computer Science, vol no. 53,2015.
2. Medhi. J," Stochastic Processes",4th Edition.
3. https://www.who.int, "Impact of COVID-19 on people's livelihoods, their health and our food systems."