Original Research Article

HIDDEN MARKOV MODEL OF DISEASE PROGRESSION AND CONTROL WITH REFERENCE TO COVID-19 SPREAD

ABSTRACT. Disease progression studies through stochastic modeling is one of the most effective approaches as different processes involved in the disease acquisition, growth, spread, and control are random. In this study, we develop a stochastic model for studying the disease spread using Markov Processes and Hidden Markov Models (HMM). This study considered two states of disease under the categories of hidden and visible. Further hidden states, as well as visible states, are classified into two groups each. This study made an attempt of relating the spread of disease in Tamil Nadu and Puducherry along with its neighboring states. Increment/Decrement in daily positive cases of Tamil Nadu and Puducherry have an influence on the Increment/ Decrement in daily positive cases of neighboring states assuming, there are regular transitions of patients from one place to another. This study develops HMM for transitions among different states (Increment/Decrement) for understanding the dynamics of positivity for a duration of consecutive two days and three days. Probability distributions of the prevalence of positivity are derived from the developed transition probability matrices. The study further derived the mathematical/ functional relations of different statistical measures through the parameters under consideration. This study will be useful for measuring the severity of the disease spread. The development of an interactive user interface for healthcare management will be the scope of this study.

Key words and phrases. Stochastic Modelling, COVID-19, Hidden Markov Models, Disease Progression, Healthcare Management.

Comment [P1]: developed

1. Introduction

In late December 2019, China has reported a new virus that affects humans and it was later named COVID-19. Further, on 31st January 2020 World Health Organization (WHO) declared this outbreak as a global pandemic. But despite the efforts, the spread of the infection worldwide was uncontrolled. In order to safeguard the people and to avoid further spread of COVID-19 government made some stringent control measures such as imposing nationwide lockdown, making use of face masks is mandatory, educating the people to maintain social distancing, and proper sanitization. The precautionary measure of nationwide lockdown played a drastic role in people's routine life. After the announcement of lockdown, the most affected population were migrant workers, who lost their livelihood and were forced to move to their native places from workplaces. Once the worker started to migrate, there was a rapid increase in the virus spread and daily reported cases surged all of a sudden. Yuan et al., (2020) used three machine learning models like the Hidden Markov chain model, the long-short-term memory model, and the Hierarchical Bayes model for the prediction of COVID-19 cases for 6 countries which includes the US, Italy, etc., for 5 days ahead [11]. Lynnette et al., (2020) used HMM to predict the people's emotion on Twitter during the COVID-19 pandemic and constructed an emotion topic, HMM to predict the user's repeated topic on Twitter [7]. Johannes et al., (2020) forecast the COVID-19 future spread between different countries by utilizing the recognized lead-lag structure and also suggested HMM can be used for future research work [8]. Abdelghofour et al., (2020) used HMM in their study in order to find out the future spread of the coronavirus from march 14, 2020 to October 5, 2020, in the Morocco context [1]. HMM has been used by Prabhu et al., (2020) to predict the future spread of the corona virus. This model was applied on Indian context. There are two types of prediction models Long-term prediction model and short-term prediction model [10]. Hongwei et al., (2021) used SEIR based short-term forecast model to predict the COVID-19 cases for two to three weeks length [6].

To predict the survival and mortality rate of the COVID-19 infected patients Aljameel et al., (2021) used some machine learning methods and analyzed the data with the help of three classification algorithms such as logistic regression, random forest, and extreme gradient boosting [2]. Ahmed Bani et al., (2020) used a stochastic model called Lotka-Volterra coupled with an extended Kalman Filter algorithm for predicting the spread of COVID-19 infections [3]. Cooper et al., (2004) predicted the spread of infection within the hospitals and also for understanding the transmission between patient to patient (or) transmission between staff to the patient by using HMM [4]. Tirupathi Rao et al (2021) used the Markov model to find the future growth of the COVID-19 disease with three different states and model behavior is studied with real-life data [9].

Comment [P2]: Source please

Comment [P3]: Depends on the choice of referencing, if APA, the author is required to cite the name of the authors in full in the first instance and use et al with the name of the first author subsequently.

2. Stochastic Model

This model is intended to derive Probability distribution functions of the number of emission states in a discrete distribution. Thus the transition states are of two kinds namely State 1: Decrement and State 2: Increment.

2.1. Notations and Terminology. Let us assume,

π_i - Initial probability for i^{th} hidden state. $\pi_i \geq 0; \forall i = 1, 2; \sum_{i=1}^{2} \pi_i = 1$

X_n - Resulting value of Hidden states at n^{th} trial

Y_n - Resulting value of Visible states with the influecne of hidden states at n^{th} trial

α_{kl} denotes the transtion probability within hidden states

$$\alpha_{kl} : P\{X_n = l | X_{n-1} = k\} \geq 0$$

$$0 \leq \alpha_{kl} \geq 1 \text{ and } \sum_{l=1}^{2} \alpha_{kl} = 1 \forall \; k = 1, 2$$

UNDER PEER REVIEW

β_{kl} denotes the emission probability in between hidden and visible states

$$\beta_{kl} : P\{Y_n = l | X_{n-1} = k\} \geq 0$$

$$0 \leq \beta_{kl} \geq 1 \text{ and } \sum_{l=1}^{2} \beta_{kl} = 1 \ \forall \ k = 1, 2$$

State 1: Decrement is $Z_{n+1} - Z_n < 0$, State 2: Increment is $Z_{n+1} - Z_n > 0$
Z_n is the number of positive cases identified on n^{th} day of study.

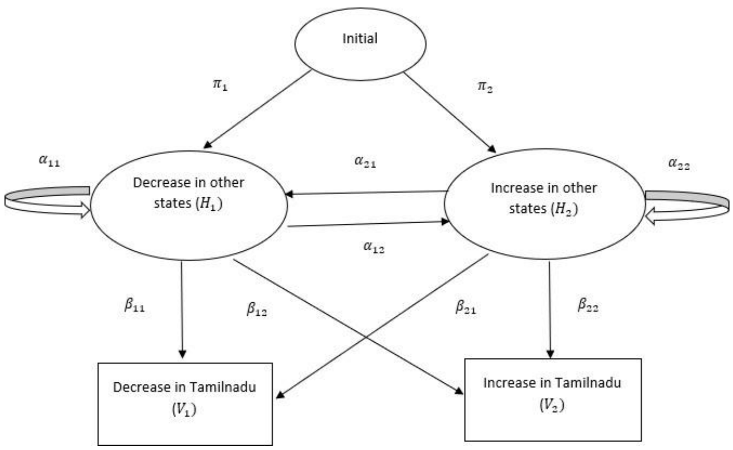


FIGURE 1. Schematic diagram for two state Hidden Markov model of COVID-19 spread

2.2. Assumptions.

(i) Hidden states have the initial probabilities $\pi_i, \forall \pi_i \geq 0; i = 1, 2.$
(ii) Transition probabilities among hidden states are of intra and inter transits in nature.
(iii) Visible/Emission states are the effects of hidden states.
(iv) The transition probabilities of visible states are initiated with hidden state.

2.2.1. *Transition probability matrix within hidden states.*

$$A = \begin{matrix} & X_n \\ X_{n-1} & \alpha_{kl} \end{matrix} \quad ; k, l = 1, 2 \tag{2.1}$$

2.2.2. *Emission/observed Probability matrix in between Hidden and Visible states.*

$$B = \begin{matrix} & Y_n \\ X_{n-1} & \beta_{kl} \end{matrix} \quad ; k, l = 1, 2 \tag{2.2}$$

2.2.3. *Initial probabilities with Hidden States:*

$$P(H_1) = \pi_1; P(H_2) = \pi_2; \sum_{i=1}^{2} n_i = n; \sum_{i=1}^{2} \pi_i = 1; \pi_i = n_i/n \qquad (2.3)$$

n_i: Number of observations in i^{th} initial state.

3. Probability Distributions for one day length of sequence

Let $X(\omega) = n$, be the random variable that denotes the occurrence of the specific state. ω represents two different states namely Decrement and Increment (i.e) ω = D (or) I. "n be the number of times the events happen in that specific state, 'n' can be $0, 1, 2, 3, ...$ Where 0 will be non-happening, 1 will be an occurrence of the event once, 2 will be an occurrence of the event twice, 3 will be an occurrence of the event thrice, and so on.

3.1. Probability Mass Function for "Decrement State" distribution.

$$P[X(D)] = \begin{cases} \sum_{i=1}^{2} \pi_i \beta_{i2}; & X(D) = 0 \\ \sum_{i=1} \pi_i \beta_{i1}; & X(D) = 1 \end{cases} \qquad (3.1)$$

3.2. Statistical Characteristics of Decrement state's Probability Distribution: Some statistical characteristics are derived in this section for the probability distribution given in the equation (3.1).

$$\text{Mean, } E[X(D)] = \sum_{i=1}^{2} \pi_i \beta_{i1} \qquad (3.2)$$

$$\text{Variance, } V[X(D)] = \sum_{i=1} \pi_i \beta_{i1} \left(1 - \sum_{i=1}^{2} \pi_i \beta_{i1}\right) \qquad (3.3)$$

Third and fourth central moments:

$$\mu_3[X(D)] = \sum_{i=1} \pi_i \beta_{i1} \left(1 - 2\sum_{i=1} \pi_i \beta_{i1}\right)\left(1 - \sum_{i=1}^{2} \pi_i \beta_{i1}\right) \qquad (3.4)$$

$$\mu_4[X(D)] = \sum_{i=1} \pi_i \beta_{i1} \left[\sum_{i=1}^{2} \pi_i \beta_{i1} - 1\right. - 3\left(\sum_{i=1}^{2} \pi_i \beta_{i1}\right)^2 + 3 \left.\sum_{i=1}^{2} \pi_i \beta_{i1} - 1\right] \qquad (3.5)$$

$$S_k[X(D)] = \left(1 - 2\sum_{i=1} \pi_i \beta_{i1}\right) \left[\sum_{i=1}^{2} \pi_i \beta_{i1} \left(1 - \sum_{i=1}^{2} \pi_i \beta_{i1}\right)\right]^{-1} \qquad (3.6)$$

Coefficient of kurtosis for Decrement state

$$\left[3\left(\sum_{i=1} \pi_i \beta_{i1}\right)^2 - 3\sum_{i=1}^{2} \pi_i \beta_{i1} + 1\right] \left[\sum_{i=1} \sum_{i=1}^{2} \pi_i \beta_{i1} \left(1 - \sum_{i=1}^{2} \pi_i \beta_{i1}\right)\right]^{-1} \qquad (3.7)$$

Characteristic Function, $\phi_x[X(D)] = 1 - \sum_{i=1}^{2} \pi_i \beta_{i1}(1 - e^{it})$ (3.8)

3.3. Probability Mass Function for "Increment State" distribution.

$$P[X(I)] = \begin{cases} \sum_{i=1}^{2} \pi_i \beta_{i1}; & for X(I) = 0 \\ \sum_{i=1}^{2} \pi_i \beta_{i2}; & for X(I) = 1 \end{cases} \quad (3.9)$$

3.4. Statistical Characteristics of Increment state's probability distribution: Some statistical characteristics are derived in this section for the probability distribution given in the equation (3.9)

Mean, $E[X(I)] = \sum_{i=1}^{2} \pi_i \beta_{i2}$ (3.10)

Variance, $V[X(I)] = \sum_{i=1}^{2} \pi_i \beta_{i2} \left(1 - \sum_{i=1}^{2} \pi_i \beta_{i2}\right)$ (3.11)

Third and fourth central moments:

$$\mu_3[X(I)] = \sum_{i=1}^{2} \pi_i \beta_{i2} \left(1 - 2\sum_{i=1}^{2} \pi_i \beta_{i2}\right)\left(1 - \sum_{i=1}^{2} \pi_i \beta_{i2}\right) \quad (3.12)$$

$$\mu_4[X(I)] = \sum_{i=1}^{2} \pi_i \beta_{i2} \left(\sum_{i=1}^{2} \pi_i \beta_{i2} - 1\right)\left(-3\left(\sum_{i=1}^{2} \pi_i \beta_{i2}\right)^2 + 3\sum_{i=1}^{2} \pi_i \beta_{i2} - 1\right) \quad (3.13)$$

$$S_k[X(I)] = \left(1 - 2\sum_{i=1}^{2} \pi_i \beta_{i2}\right)^2 \left[\sum_{i=1}^{2} \pi_i \beta_{i2}\left(1 - \sum_{i=1}^{2} \pi_i \beta_{i2}\right)\right]^{-1} \quad (3.14)$$

Coefficient of kurtosis for Increment state

$$\left(3\left(\sum_{i=1}^{2} \pi_i \beta_{i2}\right)^2 - 3\sum_{i=1}^{2} \pi_i \beta_{i2} + 1\right)\left[\sum_{i=1}^{2} \pi_i \beta_{i2}\left(1 - \sum_{i=1}^{2} \pi_i \beta_{i2}\right)\right]^{-1} \quad (3.15)$$

Characteristic Function, $\phi_x[X(I)] = 1 - \sum_{i=1}^{2} \pi_i \beta_{i2}(1 - e^{it})$ (3.16)

4. Probability Distributions for Two days length of sequence:

4.1. Probability Mass Function for "Decrement State" distribution.

$$P[X(D)] = \begin{cases} \sum_{i,j=1}^{2} \pi_i \alpha_{ij} \beta_{j2}^2; & X(D) = 0 \\ 2\sum_{i,j=1}^{2} \pi_i \alpha_{ij} \beta_{j1} \beta_{j2}; & X(D) = 1 \\ \sum_{i,j=1}^{2} \pi_i \alpha_{ij} \beta_{j1}^2; & X(D) = 2 \end{cases} \quad (4.1)$$

4.2. Statistical Characteristics of Decrement state's Probability Distribution: In this section, some statistical characteristics are explored for the probability distribution given in the equation(4.1) by assuming, $\vartheta_1 = \sum_{i,j=1}^{2} \pi_i \alpha_{ij} \beta_{j2}^2$; $\vartheta_2 = \sum_{i,j=1}^{2} \pi_i \alpha_{ij} \beta_{j1} \beta_{j2}$; $\vartheta_3 = \sum_{i,j=1}^{2} \pi_i \alpha_{ij} \beta_{j1}^2$

Mean, $E[X(D)] = 2(\vartheta_2 + \vartheta_3)$ (4.2)

Variance, $V[X(D)] = 2(\vartheta_2 + 2\vartheta_3) - [2(\vartheta_2 + \vartheta_3)]^2$ (4.3)

Third and fourth central moments

$$\mu_3[X(D)] = 2(\vartheta_2 + 4\vartheta_3) - 4(\vartheta_2 + \vartheta_3)\left[3(\vartheta_2 + 2\vartheta_3) - 4(\vartheta_2 + \vartheta_3)^2\right] \quad (4.4)$$

$$\mu_4[X(D)] = 2(\vartheta_2 + 8\vartheta_3) - 8(\vartheta_2 + \vartheta_3)\left[2(\vartheta_2 + 4\vartheta_3) - 6(\vartheta_2 + \vartheta_3)\left[(\vartheta_2 + 2\vartheta_3) + (\vartheta_2 + \vartheta_3)^2\right]\right] \quad (4.5)$$

$$S_k[X(D)] = \frac{(\vartheta_2 + 4\vartheta_3) - 2(\vartheta_2 + \vartheta_3)\left[3(\vartheta_2 + 2\vartheta_3) - 4(\vartheta_2 + \vartheta_3)^2\right]}{2^{\frac{3}{2}-1}\left[(\vartheta_2 + 2\vartheta_3) - 2(\vartheta_2 + \vartheta_3)^2\right]^{\frac{3}{2}}} \quad (4.6)$$

Coefficient of Kurtosis for Decrement State

$$\frac{(\vartheta_2 + 8\vartheta_3) - 4(\vartheta_2 + \vartheta_3)\left[2(\vartheta_2 + 4\vartheta_3) - 6(\vartheta_2 + \vartheta_3)\left[(\vartheta_2 + 2\vartheta_3) + (\vartheta_2 + \vartheta_3)^2\right]\right]}{2^{2-1}\left[(\vartheta_2 + 2\vartheta_3) - 2(\vartheta_2 + \vartheta_3)^2\right]^{2}} \quad (4.7)$$

Characteristic Function:, $\phi[X(D)] = \vartheta_1 + e^{it}\vartheta_2 + e^{2it}\vartheta_3$ (4.8)

4.3. Probability mass function for "Increment State" distribution.

$$P[X(I)] = \begin{cases} \sum_{i,j=1}^{2} \pi_i \alpha_{ij} \beta_{j1}^2; & X(I) = 0 \\ 2\sum_{i,j=1}^{2} \pi_i \alpha_{ij} \beta_{j1} \beta_{j2}; & X(I) = 1 \\ \sum_{i,j=1}^{2} \pi_i \alpha_{ij} \beta_{j2}^2; & X(I) = 2 \end{cases} \quad (4.9)$$

4.4. Statistical Characteristics of Increment state's probability distribution: In this section, some statistical characteristics are explored for the probability distribution given in the equation(4.9) by considering, $\tau_1 = \sum_{i,j=1}^{2} \pi_i \alpha_{ij} \beta_{j1}^2$; $\tau_2 = \sum_{i,j=1}^{2} \pi_i \alpha_{ij} \beta_{j1} \beta_{j2}^2$; $\tau_3 = \sum_{i,j=1}^{2} \pi_i \alpha_{ij} \beta_{j2}^2$

Mean, $E[X(I)] = 2(\tau_2 + \tau_3)$ (4.10)

Variance, $V[X(I)] = 2(\tau_2 + 2\tau_3) - [2(\tau_2 + \tau_3)]^2$ (4.11)

Third and fourth central moments

$$\mu_3[X(I)] = 2(\tau_2 + 4\tau_3) - 4(\tau_2 + \tau_3)\left[3(\tau_2 + 2\tau_3) - 4(\tau_2 + \tau_3)^2\right] \quad (4.12)$$

$$\mu_4[X(I)] = 2(\tau_2 + 8\tau_3) - 8(\tau_2 + \tau_3)\left[2(\tau_2 + 4\tau_3) - 6(\tau_2 + \tau_3)(\tau_2 + 2\tau_3) + (\tau_2 + \tau_3)^2\right] \quad (4.13)$$

$$S_k[X(I)] = \frac{\left[(\tau_2 + 4\tau_3) - 2(\tau_2 + \tau_3)\left(3(\tau_2 + 2\tau_3) - 4(\tau_2 + \tau_3)^2\right)\right]^2}{\left[2(\tau_2 + 2\tau_3) - 2(\tau_2 + \tau_3)^2\right]^{3-1}} \quad (4.14)$$

Coefficient of Kurtosis for Increment State

$$\frac{(\tau_2 + 8\tau_3) - 4(\tau_2 + \tau_3)\left[2(\tau_2 + 4\tau_3) - 6(\tau_2 + \tau_3)(\tau_2 + 2\tau_3) + (\tau_2 + \tau_3)^2\right]}{2\left[(\tau_2 + 2\tau_3) - 2(\tau_2 + \tau_3)^2\right]^{2-1}} \quad (4.15)$$

Characteristic Function:, $\phi_{[X(I)]} = \tau_1 + e^{it}\tau_2 + e^{2it}\tau_3$ (4.16)

5. Probability Distributions for three days length of sequence:

5.1. Probability Mass Function for "Decrement State" distribution.

$$P[X(D)] = \begin{cases} \sum_{i,j,k=1}^{2} \pi_i \alpha_{ij} \alpha_{jk} \beta_{k2}^3; & X(D) = 0 \\ 3\sum_{i,j,k=1}^{2} \pi_i \alpha_{ij} \alpha_{jk} \beta_{k1} \beta_{k2}^2; & X(D) = 1 \\ 3\sum_{i,j,k=1}^{2} \pi_i \alpha_{ij} \alpha_{jk} \beta_{k1}^2 \beta_{k2}; & X(D) = 2 \\ \sum_{i,j,k=1}^{2} \pi_i \alpha_{ij} \alpha_{jk} \beta_{k1}^3; & X(D) = 3 \end{cases} \quad (5.1)$$

5.2. Statistical Characteristics of Decrement state's probability distribution: Statistical characteristics are explored for the probability distribution given in the equation (5.1). by considering,

$$\lambda_1 = \sum_{i,j,k=1}^{2} \pi_i \alpha_{ij} \alpha_{jk} \beta^3_{k2};\ \lambda_2 = \sum_{i,j,k=1}^{2} \pi_i \alpha_{ij} \alpha_{jk} \beta_{k1} \beta^2_{k2};$$

$$\lambda_3 = \sum_{i,j,k=1}^{2} \pi_i \alpha_{ij} \alpha_{jk} \beta^2_{k1} \beta_{k2};\ \lambda_4 = \sum_{i,j,k=1}^{2} \pi_i \alpha_{ij} \alpha_{jk} \beta^3_{k1}$$

Mean, $E[X(D)] = 3(\lambda_2 + 2\lambda_3 + \lambda_4)$ (5.2)

Variance, $V[X(D)] = 3(\lambda_2 + 4\lambda_3 + 3\lambda_4) - \left[3(\lambda_2 + 2\lambda_3 + \lambda_4)\right]^2$ (5.3)

Third and fourth central moments

$$\mu_3[X(D)] = 3(\lambda_2 + 8\lambda_3 + 9\lambda_4) - 27(\lambda_2 + 2\lambda_3 + \lambda_4)\left[(\lambda_2 + 4\lambda_3 + 3\lambda_4) - 2(\lambda_2 + 2\lambda_3 + \lambda_4)^2\right] \quad (5.4)$$

$$\mu_4[X(D)] = 3(\lambda_2 + 16\lambda_3 + 27\lambda_4) - 9(\lambda_2 + 2\lambda_3 + \lambda_4)\left[4(\lambda_2 + 8\lambda_3 + 9\lambda_4) - 9(\lambda_2 + 2\lambda_3 + \lambda_4)\left\{2(\lambda_2 + 4\lambda_3 + 3\lambda_4) + (\lambda_2 + 2\lambda_3 + \lambda_4)^2\right\}\right] \quad (5.5)$$

$$S_k[X(D)] = \left[(\lambda_2 + 8\lambda_3 + 9\lambda_4) - 9(\lambda_2 + 2\lambda_3 + \lambda_4)\left\{(\lambda_2 + 4\lambda_3 + 3\lambda_4) - 2(\lambda_2 + 2\lambda_3 + \lambda_4)^2\right\}\right]^2 \left[3\left\{(\lambda_2 + 4\lambda_3 + 3\lambda_4) - 3(\lambda_2 + 2\lambda_3 + \lambda_4)^2\right\}^3\right]^{-1} \quad (5.6)$$

Coefficient of Kurtosis for Decrement State

$$\left[(\lambda_2 + 16\lambda_3 + 27\lambda_4) - 3(\lambda_2 + 2\lambda_3 + \lambda_4)\left\{4(\lambda_2 + 8\lambda_3 + 9\lambda_4) - 9(\lambda_2 + 2\lambda_3 + \lambda_4)\left(2(\lambda_2 + 4\lambda_3 + 3\lambda_4) + (\lambda_2 + 2\lambda_3 + \lambda_4)^2\right)\right\}\right]\left[3\left\{(\lambda_2 + 4\lambda_3 + 3\lambda_4) - 3(\lambda_2 + 2\lambda_3 + \lambda_4)^2\right\}^2\right]^{-1} \quad (5.7)$$

Characteristic Function:, $\phi_{[X(D)]} = \lambda_1 + e^{it}3\lambda_2 + e^{2it}3\lambda_3 + e^{3it}\lambda_4$ (5.8)

5.3. Probability Mass Function for "Increment State" distribution.

$$P[X(I)] = \begin{cases} \sum_{i,j,k=1}^{2} \pi_i \alpha_{ij} \alpha_{jk} \beta^3_{k1}; & X(I) = 0 \\ 3\sum_{i,j,k=1}^{2} \pi_i \alpha_{ij} \alpha_{jk} \beta^2_{k1} \beta_{k2}; & X(I) = 1 \\ 3\sum_{i,j,k=1}^{2} \pi_i \alpha_{ij} \alpha_{jk} \beta_{k1} \beta^2_{k2}; & X(I) = 2 \\ \sum_{i,j,k=1}^{2} \pi_i \alpha_{ij} \alpha_{jk} \beta^3_{k2}; & X(I) = 3 \end{cases} \quad (5.9)$$

5.4. Statistical characteristics of Increment State's probability distribution. Statistical characteristics are explored for the probability distribution given in the equation (5.9). by considering,

$$\psi_1 = \sum_{i,j,k=1}^{2} \pi_i \alpha_{ij} \alpha_{jk} \beta^3_{k1};\ \psi_2 = \sum_{i,j,k=1}^{2} \pi_i \alpha_{ij} \alpha_{jk} \beta^2_{k1} \beta_{k2};$$

$$\psi_3 = \sum_{i,j,k=1}^{2} \pi_i \alpha_{ij} \alpha_{jk} \beta_{k1} \beta^2_{k2};\ \psi_4 = \sum_{i,j,k=1}^{2} \pi_i \alpha_{ij} \alpha_{jk} \beta^3_{k2}$$

Mean, $E[X(I)] = 3(\psi_2 + 2\psi_3 + \psi_4)$ (5.10)

Variance, $V[X(I)] = 3(\psi_2 + 4\psi_3 + 3\psi_4) - \left[3(\psi_2 + 2\psi_3 + \psi_4)\right]^2$ (5.11)

Third and fourth central moments

$$\mu_3[X(I)] = 3(\psi_2 + 8\psi_3 + 9\psi_4) - 27(\psi_2 + 2\psi_3 + \psi_4)(\psi_2 + 4\psi_3 + 3\psi_4) - 2(\psi_2 + 2\psi_3 + \psi_4)^2 \quad (5.12)$$

$$\mu_4[X(I)] = 3(\psi_2 + 16\psi_3 + 27\psi_4) - 9(\psi_2 + 2\psi_3 + \psi_4)\Big[4(\psi_2 + 8\psi_3 + 9\psi_4) - 9(\psi_2 + 2\psi_3 + \psi_4)\,2(\psi_2 + 4\psi_3 + 3\psi_4) + (\psi_2 + 2\psi_3 + \psi_4)^2\Big] \quad (5.13)$$

$$S_k[X(I)] = \Big[(\psi_2 + 8\psi_3 + 9\psi_4) - 9(\psi_2 + 2\psi_3 + \psi_4)(\psi_2 + 4\psi_3 + 3\psi_4) - 2(\psi_2 + 2\psi_3 + \psi_4)^2\Big]^2 \Big[3(\psi_2 + 4\psi_3 + 3\psi_4) - 3(\psi_2 + 2\psi_3 + \psi_4)^2\Big]^{3-1} \quad (5.14)$$

Coefficient of Kurtosis for Decrement State

$$(\psi_2 + 16\psi_3 + 27\psi_4) - 3(\psi_2 + 2\psi_3 + \psi_4)\Big[4(\psi_2 + 8\psi_3 + 9\psi_4) - 9(\psi_2 + 2\psi_3 + \psi_4)\,2(\psi_2 + 4\psi_3 + 3\psi_4) + (\psi_2 + 2\psi_3 + \psi_4)^2\Big]\Big[3(\psi_2 + 4\psi_3 + 3\psi_4) - 3(\psi_2 + 2\psi_3 + \psi_4)^2\Big]^{2-1} \quad (5.15)$$

Characteristic Function:, $\phi_{[X(I)]} = \psi_1 + e^{it} 3\psi_2 + e^{2it} 3\psi_3 + e^{3it} \psi_4$ (5.16)

6. Results and discussion

The data was collected with the help of internet sources (Wikipedia) by using the COVID-19 dataset (placed in annexure) the derived model behavior is studied [5]. The data consist of the number of reported positive cases as state-wise in India from 24^{th} March 2020 to 25^{th} August 2020. This model is focused on the dynamic behavior of two states. The total number of cases per day is divided into two transition states as "Decrement" and "Increment". This study primarily focused on the southern states of India and It aims to predict the COVID-19 positive cases in Tamilnadu and Puducherry that are influenced by its adjacent states namely Kerala, Karnataka, Andhra Pradesh. If the previous day cases are more when compared to the current day then that state is considered as "Decrement State". Whereas the state is mentioned as "Increment State" when the previous day cases are less compared to the current day. Here, the number of positive cases identified on each day in Kerala, Karnataka, Telangana, Andhra Pradesh are considered for the identification of Hidden states. The daily positive cases of Tamilnadu and Puducherry are considered for the identification of Visible states. Transition frequency tables and transition probability matrices are obtained for both the Hidden and Visible states.

6.1. Initial probability matrix for Hidden states.

$$\pi = \begin{pmatrix} 0.4183 & 0.5817 \end{pmatrix}$$

6.2. Transition Probability Matrix among Hidden states: Thorough processing of identified daily positive cases in the total neighboring states of Tamil Nadu, the number of initial states, Decrement (D) and Increment (I), was carried out with the collected data. The joint states (HD, HD), (HD, HI), (HI, HD) and (HI, HI) are identified as

$$(HD, HD) : (X_{n+1} < X_n) \cap (X_n < X_{n-1})$$
$$(HD, HI) : (X_{n+1} < X_n) \cap (X_n > X_{n-1})$$
$$(HI, HD) : (X_{n+1} > X_n) \cap (X_n < X_{n-1})$$
$$(HI, HI) : (X_{n+1} > X_n) \cap (X_n > X_{n-1})$$

Transition Frequency Table:

		X_n	
		n(D)	n(I)
X_{n-1}	n(D)	16	47
	n(I)	46	43

Transition Probability Matrix:

		X_n	
		n(D)	n(I)
X_{n-1}	n(D)	0.2540	0.7460
	n(I)	0.5169	0.4831

$$A = \begin{pmatrix} 0.2540 & 0.7460 \\ 0.5169 & 0.4831 \end{pmatrix}$$

6.3. Transition Probability Matrix between Hidden and Visible states: The emission Probability matrix is obtained after processing the emission frequency matrix.The joint states (HD, VD), (HD, VI), (HI, VD) and (HI, VI) are identified as

$$(HD, VD) : (X_{n+1} < X_n) \cap (Y_{n+1} < Y_n)$$
$$(HD, VI) : (X_{n+1} < X_n) \cap (Y_{n+1} > Y_n)$$
$$(HI, VD) : (X_{n+1} > X_n) \cap (Y_{n+1} < Y_n)$$
$$(HI, VI) : (X_{n+1} > X_n) \cap (Y_{n+1} > Y_n)$$

Transition Frequency Table:

		Y_n	
		n(D)	n(I)
X_{n-1}	n(D)	28	36
	n(I)	35	54

Transition Probability Matrix:

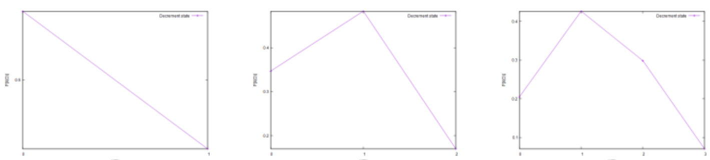
		X_n	
		n(D)	n(I)
X_{n-1}	n(D)	0.4375	0.5625
	n(I)	0.3936	0.6067

$$B = \begin{matrix} 0.4375 & 0.5625 \\ 0.3933 & 0.6067 \end{matrix}$$

Table 1. Probability distribution for Decrement State

X(D)	0	1	2	3
P[X(D)](1 day sequence)	0.5882	0.4118	-	-
P[X(D)](2 day sequence)	0.3471	0.4833	0.1697	-
P[X(D)](3 day sequence)	0.2048	0.4265	0.2985	0.0702

From table 1 it is observed that non-occurrence of Decrement state in one-day length is having more chance. The chance of happening of decrement state once in a two days sequence is more. In three-day sequence occurrence of decrement state once having more chance when compared to other. And the graph is plotted for the above probability.



(A) (B) (C)

Figure 2. Probability of Decrement State (A) 1 day sequence; (B) 2 days sequence; (c) 3 days sequence.

Table 2. Statistical results of Decrement State for different (one day, two days run and three days run) lengths of consecutive days:

Statistical measures	1 day's	2 day's	3 day's
Mean	0.4118	0.8225	1.2342
Variance	0.2422	0.4852	0.7293
3^{rd} central moment	0.0427	0.0864	0.1303
4^{th} central moment	0.0662	0.4854	1.2617
Skewness	0.1256	0.0654	0.0438
Kurtosis	1.1286	2.0618	2.3722

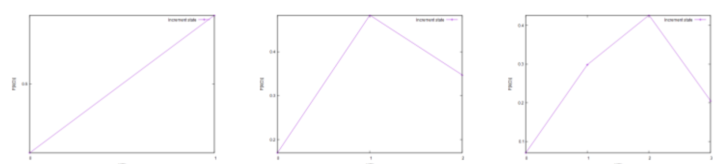
Based on the results obtained for "Decrement State" it is observed that the average happening of Decrement state is less than 1 in one day study, nearly one day in two day's study and there is a chance of occurrence of the state is more than 1 in three day's calculation. And also observed that the Decrement state is

Positively skewed. The kurtosis measure is less than 3 which means it is platy kurtic.

TablE 3. Probability distribution for Increment State

X(I)	0	1	2	3
P[X(I)](1 day sequence)	0.4118	0.5882	-	-
P[X(I)](2 day sequence)	0.1697	0.4833	0.3471	-
P[X(I)](3 day sequence)	0.0702	0.2985	0.4265	0.2048

From table 3 it is observed that happening of increment state is having more chance when compared to the non-happening of increment state. In two days sequence chance of happening of increment state once is more. The occurrence of increment state twice is having more chance in three days sequence. Graphical representation is given below.



(A) (B) (C)

FigURE 3. Probability of Increment State (A) 1 day sequence; (B) 2 days sequence; (C) 3 days sequence.

TablE 4. Statistical results for Increment state for different (one day, two days run and three days run) lengths of consecutive day's

Statistical measures	1 day's	2 day's	3 day's
Mean	0.5882	1.1774	1.7658
Variance	0.2422	0.4852	0.7293
3^{rd} central moment	-0.04274	-0.0865	-0.1302
4^{th} central moment	0.0662	0.4854	1.2617
Skewness	0.1286	0.0654	0.0438
Kurtosis	1.1286	2.0618	2.3722

From table 4, it is observed that the average happening of Increment state is less than 1 in one-day length, more than 1 in two, three day's study. It is noticed that the Increment state is positively skewed. The kurtosis measure is less than 3 which means it is platy kurtic.

7. Summary and conclusion

This research study mainly focused on developing the Hidden Markov Model with the help of transition states. The main intent of the study is to examine whether the total number of positive cases registered in Tamil Nādu and Puducherry have an association with the total number of positive cases registered in other adjacent southern states like Andhra Pradesh, Telangana, Kerala, Karnataka. A transition frequency table has been obtained by considering discrete Markov processes. Transition probability matrices are also derived based on transition frequency tables. As an extended activity for understanding the model behavior, the classical formulae for probability distributions for one-day occurrence, two-day successive occurrences, and three-day successive occurrences of a single state and two states are derived. Different characteristics of the developed probability distributions are derived for the detailed study of the models. For the derived probability mass functions of the hidden Markov models and related discrete distributions, the statistical measures based on Pearson's coefficients are obtained. Model behavior is studied with the help of the COVID-19 dataset which was collected from internet sources. This data is about the occurrences of positive cases identified in TamilNadu,Puducherry,Telangana,Andhra Pradesh, Kerala, and Karnataka. Study reports on statistical measures are obtained through the numerical data sets. The intensity of the prevalence, trends of increments, or decrements fluctuations is studied here. This study will provide the indicators on disease prevalence and will be helpful to the government to take strict action plans against the migration of people from one place to other places which may be the potential reason for the outbreak of the disease.

References

1. Abdelghafour Marfak, Doha Achak, Asmaa Azizi and Chakib Nejjari„ Khalid Aboudi, , Elmadani Saad and , Abderraouf Hilali and Ibtissam Youlyouz-Marfak (2020), The Hidden Markov chain modelling of the COVID-19 spreading using Moroccan dataset, Data in brief, Vol-32,P-106067.
2. Aljameel, S. S., Khan, I. U., Aslam, N., Aljabri, M., Alsulmi, E. S. (2021). Machine learning-based model to predict the disease severity and outcome in COVID-19 patients. Scientific programming.
3. Bani Younes, A., Hasan, Z. (2020). COVID-19: modeling, prediction, and control.*Applied Sciences*, 10(11), 3666.
4. Cooper, B., Lipsitch, M. (2004). The analysis of hospital infection data using hidden Markov models. Biostatistics, Vol-5(2), P-223-237.
5. COVID-19 pandemic data/India medical cases by state and union territory, https://en.wikipedia.org/wiki/Template:COVID-19_pandemic_data/India_medical_cases_by_state_and_union_territory
6. Hongwei Zhao, Merchant, N. N., McNulty, A., Radcliff, T. A., Cote, M. J., Fischer, R. S., Ory, M. G. (2021). COVID-19: Short term prediction model using daily incidence data. PloS one, Vol-16(4).
7. Hui Xian Lynnette Ng, Roy Ka-Wei Lee and Md Rabiul Awal (2020), I miss you babe: Analyzing Emotion Dynamics During COVID-19 Pandemic, Proceedings of the Fourth Workshop on Natural Language Processing and Computational Social Science, pp 41-49
8. Johannes stubinger and Lucas Schneider (2020), Epidemiology of coronavirus covid-19: Forecasting the future incidence in different countries, Healthcare, Vol-8, no-2, pp-99

9. P.Tirupathi Rao, Kanimozhi.V, Sakkeel P.T (2021). Markov model of COVID-19 disease progression, Stochastic Modeling and Applications, vol-25(1), PP - 101-114.
10. Prabhu, S. M., Subramaniam, N. (2020). Surveillance of COVID-19 Pandemic using Hidden Markov Model. arXiv preprint arXiv:2008.07609.
11. Yuan Tian, Ishika Luthra and Xi Zhang (2020), Forecasting COVID-19 cases using Machine Learning Models, medRxiv.

8. Annexure

Date	Tamil Nadu	Puducherry	Andhra Pradesh	Telangana	Karnataka	Kerala
Mar-24	6	0	1	3	7	28
Mar-25	3	0	1	0	4	14
Mar-26	8	0	2	9	14	9
Mar-27	12	0	3	4	0	43
Mar-28	2	0	0	8	9	15
Mar-29	9	0	5	10	12	6
Mar-30	18	0	4	5	7	37
Mar-31	7	0	17	8	0	15
Apr-01	160	2	43	17	18	7
Apr-02	0	0	3	11	9	24
Apr-03	75	2	46	51	14	21
Apr-04	102	0	29	1	4	9
Apr-05	74	0	29	110	16	11
Apr-06	86	0	36	52	7	8
Apr-07	50	0	40	43	24	13
Apr-08	69	0	39	63	0	9
Apr-09	48	0	43	15	6	9
Apr-10	96	0	15	31	16	12
Apr-11	77	2	18	31	17	7
Apr-12	58	0	0	0	12	10
Apr-13	106	0	51	58	21	2
Apr-14	98	0	41	62	11	3
Apr-15	31	0	30	23	19	8
Apr-16	38	0	31	51	38	1
Apr-17	25	0	38	45	38	7
Apr-18	56	0	31	48	18	1
Apr-19	49	0	0	53	13	4
Apr-20	105	0	119	29	11	2
Apr-21	43	0	35	46	20	6
Apr-22	76	0	56	26	10	19
Apr-23	33	0	82	15	18	11
Apr-24	54	0	60	24	20	10
Apr-25	72	0	106	0	26	3
Apr-26	66	0	36	7	12	7
Apr-27	64	1	80	11	10	11
Apr-28	52	0	82	2	9	13
Apr-29	121	0	73	8	12	4
Apr-30	104	0	71	0	25	10
May-01	161	0	60	27	19	1
May-02	203	0	62	18	22	1
May-03	231	0	58	6	8	2
May-04	266	0	67	19	36	0
May-05	527	1	67	3	17	0
May-06	508	0	0	11	12	2
May-07	771	0	60	11	22	1
May-08	580	0	70	16	12	0

Date	Tamil Nadu	Puducherry	Andhra Pradesh	Telangana	Karnataka	Kerala
May-10	526	0	43	30	41	2
May-11	669	0	50	33	54	7
May-12	798	3	38	79	14	7
May-13	716	1	72	51	63	5
May-14	509	0	47	41	34	10
May-15	447	0	68	47	28	26
May-16	434	0	102	40	69	16
May-17	477	0	48	55	36	11
May-18	639	0	52	42	55	14
May-19	536	5	67	46	99	29
May-20	688	0	58	37	151	12
May-21	743	0	70	27	65	24
May-22	776	2	45	38	143	24
May-23	786	6	62	62	138	42
May-24	759	0	48	52	216	63
May-25	765	15	66	41	130	52
May-26	805	0	287	66	93	49
May-27	646	5	61	71	101	67
May-28	817	0	0	107	135	41
May-29	827	5	80	158	115	84
May-30	874	0	185	169	248	62
May-31	938	0	133	74	141	58
Jun-01	1149	19	110	199	299	61
Jun-02	1162	4	104	94	187	57
Jun-03	1091	8	115	99	388	86
Jun-04	1286	0	182	129	267	82
Jun-05	1384	0	143	127	257	94
Jun-06	1438	17	80	143	515	111
Jun-07	1458	0	207	206	378	108
Jun-08	1515	0	198	154	239	107
Jun-09	1562	28	143	92	308	91
Jun-10	1685	0	219	178	161	91
Jun-11	1927	0	199	191	120	64
Jun-12	1875	30	160	209	204	83
Jun-13	1982	0	251	164	271	78
Jun-14	1989	19	285	253	308	85
Jun-15	1974	18	198	237	176	54
Jun-16	1843	8	293	219	213	82
Jun-17	1515	14	385	213	317	79
Jun-18	2174	29	230	269	204	75
Jun-19	2141	26	447	352	210	97
Jun-20	2115	15	443	499	337	118
Jun-21	2396	0	491	546	416	127
Jun-22	2532	80	547	730	453	133

UNDER PEER REVIEW

Date	Tamil Nadu	Puducherry	Andhra Pradesh	Telangana	Karnataka	Kerala
Jun-23	2710	17	373	872	249	138
Jun-24	2516	19	630	879	322	141
Jun-25	2865	59	329	891	397	152
Jun-26	3509	41	553	920	442	123
Jun-27	3645	0	605	985	445	150
Jun-28	3713	117	796	1087	918	195
Jun-29	3940	0	956	983	1267	118
Jun-30	3949	0	650	975	1105	0
Jul-01	3943	95	704	945	947	253
Jul-02	3882	0	657	1018	1272	151
Jul-03	4343	88	845	1213	1502	160
Jul-04	4329	0	837	1892	1694	211
Jul-05	4280	0	765	1850	1839	240
Jul-06	4150	0	998	1590	1925	225
Jul-07	3827	0	1322	1831	1843	193
Jul-08	3616	128	1178	1879	1498	272
Jul-09	3756	78	1062	1924	2062	301
Jul-10	4231	143	1555	1410	2228	339
Jul-11	3680	121	1608	1278	2313	416
Jul-12	3965	65	1813	1178	2798	488
Jul-13	4244	81	1933	1269	2627	435
Jul-14	4328	50	1935	1550	2738	449
Jul-15	4526	63	1916	1524	2496	608
Jul-16	4496	65	2432	1597	3176	623
Jul-17	4549	147	2593	1676	4169	722
Jul-18	4538	89	2602	1478	3693	791
Jul-19	4807	62	3963	1284	4537	593
Jul-20	4979	105	5041	1296	4120	821
Jul-21	4985	93	4074	1198	3648	794
Jul-22	4965	87	4944	1431	3649	720
Jul-23	5849	121	6045	1554	4764	1038
Jul-24	6472	120	7998	1567	5030	1078
Jul-25	6785	95	8147	1640	5007	885
Jul-26	6988	139	7813	0	5072	1103
Jul-27	6986	132	7627	1593	5199	927
Jul-28	6993	86	6051	3083	5324	702
Jul-29	6972	139	7948	0	5536	1167
Jul-30	6426	166	10093	1764	5503	903
Jul-31	5864	121	10167	1811	6128	506
Aug-01	5881	174	10376	1986	5483	1310
Aug-02	5879	134	9276	2083	5172	1129
Aug-03	5875	200	8555	1891	5532	1169
Aug-04	5609	176	7822	2269	4752	962
Aug-05	5063	165	9747	2012	6259	1083
Aug-06	5175	286	10128	2092	5619	1195
Aug-07	5684	188	10328	2207	6805	1298

Date	Tamil Nadu	Puducherry	Andhra Pradesh	Telangana	Karnataka	Kerala
Aug-08	5880	241	10171	2256	6670	1251
Aug-09	5883	261	10080	1982	7178	1420
Aug-10	5994	259	10820	1256	5985	1211
Aug-11	5914	242	7665	1896	4267	1184
Aug-12	5834	276	9024	1897	6257	1417
Aug-13	5871	481	9597	1931	7883	1212
Aug-14	5835	299	9996	1921	6706	1564
Aug-15	5890	315	8943	1863	7908	1569
Aug-16	5860	359	8732	1102	8818	1608
Aug-17	5950	378	8012	894	7040	1530
Aug-18	5890	297	6780	1682	6317	1725
Aug-19	5709	367	9652	1763	7665	1758
Aug-20	5795	354	9742	1724	8642	2333
Aug-21	5986	542	9393	1967	7385	1968
Aug-22	5995	302	9544	2474	7571	1983
Aug-23	5980	518	10276	2384	7330	2172
Aug-24	5975	410	7895	1842	5938	1908
Aug-25	5967	337	8601	2579	5851	1242