ABSTRACT
COVID-19 is a new disease caused by the SARS-CoV-2 virus that first discovered at the end of 2019 in Wuhan, China. Furthermore, this virus spread to various countries and was first determined to enter Indonesia on March 2nd, 2020. Markov chain is a method that used to model variables randomly. This research uses Markov chain to model the daily number of new positive COVID-19 patients in East Java. Markov chain model can be formed for a steady state, which is a state for a long time and a stable pattern has been obtained without any significant changes. The results formed 8 states in the form of number range. The range shows the number of additions to positive COVID-19 patients in East Java by January 1st, 2021 to June 22nd, 2021. By Markov chain after we found that these states are steady, we have the probability of the number of daily positive patients was highest at range 151 to 301, with the probability of 0.446032. If it is determined as the average number of positive COVID-19 patients in steady state conditions, the median value is 226 patients. By increasing vigilance and discipline in overcoming the spread of the virus by wearing masks, diligently washing hands and keeping a distance, the number of patients due to the COVID-19 virus can be reduced, especially if the vaccine is immediately given to all people in East Java.

Keywords: COVID-19, Markov chain, Steady state.
Markov chain is a method or way of stochastically modelling the state from time to time or changes in the value of a variable. The Markov chain has various applications, one of which is to model an event that lasts long enough and is continuous. Many studies related to Markov chains and transmission have been carried out, including transmission of D. citrus in plants [1], H1N1 pandemic virus [2], and cases of Ebola virus [3]. In the study, continuous time points were modelled through the Markov chain and used as a calibration tool to model the spread of large viruses.

This research uses the Markov chain method to models the number of daily COVID-19 patient by the data of Covid-19 patient in East Java from January 1, 2021 until June 22, 2021. This research gets the probability number of COVID-19 patients in steady state is denoted by \( \{ P(X_t = j | X_{t-1} = i), \ i = 1, 2, ..., n, \ j = 1, 2, ..., n \text{ and } t = 0, 1, 2, ..., T \} \) This is known as one-step transition probability of moving from state \( i \) at time \( t - 1 \) to state \( j \) at \( t \). By definition [7], we have:

\[
\sum_j p_{ij} = 1, p_{ij} \geq 0
\]  

(1)

2.2 Ergodic State

A Markov chain can be analyzed if it has an ergodic state. State is ergodic if the state satisfies the properties of aperiodic, recurrent, and communicate [7]. If the states only one communication class, a process in a Markov chain is said to be irreducible. So the states in the process are connected so that they can transit from each state to every other state in as many steps as possible.

State \( i \) can be said to be aperiodic if and only if \( d(i) = 1 \), \( d(i) \) is defined by \( d(i) = gcd \{ n | n \geq 1, P_{ii}(n) > 0 \} \) where \( P_{ii} \) represents that the probability of state \( i \) returning to state \( i \) and \( gcd \) is greatest common divisor.

In other words, \( d(i) \) is the great common divisor (gcd) of all \( n \) possibilities that allow the process in state \( i \) to return to the same state \( i \) in \( n \) steps.

State \( i \) can be said to be recurrent if it starts in the process from state \( i \) and will return to state \( i \). can be defined as follows:

\[
f_{ii} = \sum_{n=1}^{\infty} f_{ii}^n = f_{ii}^1 + f_{ii}^2 + f_{ii}^3 ... \]  

(2)

The state \( i \) is recurrent if the probability of \( f_{ii} = 1 \)

State \( i \) is said to be positive recurrent if \( \mu_i < \infty \), where represents the average recurrence time of the state.

State \( i, j \) can be said to communicate if \( i \) can be accessed from \( j \) and \( j \) can be accessed from \( i \), \( j \rightarrow i \) and \( i \rightarrow j \) then \( i \leftrightarrow j \).

State types:

1. \( i \leftrightarrow i \), for every \( i \) (reflex)
2. \( i \leftrightarrow j \) then \( j \leftrightarrow i \) (symmetric)
3. \( i \leftrightarrow j, j \leftrightarrow k \) then \( i \leftrightarrow k \) (transitive)

2.3 Steady State Probability

The data that forming the states is taken from [8] web https://COVID19.go.id/peta-sebaran. The data used is from January 1, 2021 to June 22, 2021 with a total of 173 daily records of data on the addition of positive COVID-19 patients. Markov chain can be analyzed if it has an ergodic state. State is formed by grouping data into class intervals so that each class interval is a state in the Markov chain. The grouping of data into class intervals is used Sturge’s rule [9] as follows:

\[
P(X_t = j | X_{t-1} = i), \ i = 1, 2, ..., n, \ j = 1, 2, ..., n \text{ and } t = 0, 1, 2, ..., T
\]
Advances in Engineering Research, volume 209

\[ k = 1 + \log \log n \] (3)

\[ k = \text{the number of class interval} \]
\[ n = \text{the number of observations in the data set} \]

An ergodic state must meet the aperiodic, recurrent, and communicate. States of a Markov chain can be classified based on the transition probability \( p_{ij} \) of \( P \).

A state \( j \) is recurrent if the probability of being revisited from other states is 1.

A state \( j \) is aperiodic or not periodic. A periodic state with \( t > 1 \) if a return is possible only in \( t, 2t, 3t, \ldots \) steps. This means that \( p_{ij}^{(n)} = 0 \) when \( n \) is not divisible by \( t \).

A closed set of state is communicate, that is possible to go from any state to every other state in the set in one or more transitions. That means, \( p_{ij}^{(n)} > 0 \) for all \( i \neq j \), and \( n \geq 1 \).

Transition probability is a change from one state to another at the next time, is a random process or random represented by probability, \( P_{ij} \) defined as the probability of adding positive patients and recovering COVID-19 patients in state \( i \) at \( x + t \) and running up to state \( j \).

\[ S_1, S_2, \ldots, S_n \] is the state of increasing the number of patients, and the vector \( S = [S_1, S_2, S_3, \ldots, S_n] \) is called the steady state distribution for the Markov chain.

\[ S_1 + S_2 + S_3 + \ldots + S_n = 1 \] (4)

Discrete Markov chains used in predicting the value or vector \( S \) from time \( t = 1, 2, 3, \ldots, n \) where \( n \) is a large number \( (n \rightarrow \infty) \).

\[ S^{(2)} = S^{(1)} p_{ij} \]
\[ S^{(3)} = S^{(2)} p_{ij} \]
\[ \vdots \]
\[ S^{(n)} = S^{(n-1)} p_{ij} \]

The transition probability matrix

\[ P_{ij} = \begin{bmatrix}
P_{11} & P_{12} & \cdots & P_{19} \\
P_{21} & P_{22} & \cdots & P_{29} \\
\vdots & \vdots & \ddots & \vdots \\
P_{91} & P_{92} & \cdots & P_{99}
\end{bmatrix} \] (5)

Steady state is a state of equilibrium in the Markov chain. It is said to be steady state, that is, if something happens after the Markov process runs for several cycles, then the probability value of a state is fixed. In an ergodic Markov chain, steady state probabilities are defined as

\[ S_j = S_j^{(n)} \] (6)

These probabilities are independent of \( \{S_j^{(n)}\} \), can be determined from the equations:

\[ S = [S_1, S_2, S_3, \ldots, S_n] P_{ij} \]

3. Result

3.1 States in Markov Chain

The prediction of the increase in the number of positive COVID-19 patients using the Markov chain method used data on the daily increase in COVID-19 patients in East Java on January 1, 2021 until June 22, 2021. The Markov chain can be analyzed if it has an ergodic state, that is, if it meets 3 conditions, recurrent, aperiodic and communicate. State classification is based on the requirements of the ergodic Markov Chain being met or not. A state is formed through trial and error, until an ergodic state is finally found.

Absorbing chain analysis requires that a state does not need to review the state in the future. The absorbing chain model cannot be used to analyze the data. To overcome this problem, an approach method can be used where several conditions that represent an increase in positive COVID-19 patients can be formed. Data on the increase in COVID-19 patients for 6 months amounted to 173 data, determined by the number of interval classes using (3) is \( k = 1 + \log \log 173 = 8 \)

The data is arranged in 8 interval classes, and then each interval class is translated as states formed in the Markov chain. The eight states formed in the Markov chain are presented in table 1.

<table>
<thead>
<tr>
<th>Value interval</th>
<th>State</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 – 150</td>
<td>1</td>
</tr>
<tr>
<td>151 – 301</td>
<td>2</td>
</tr>
<tr>
<td>302 – 452</td>
<td>3</td>
</tr>
<tr>
<td>453 – 603</td>
<td>4</td>
</tr>
<tr>
<td>604 – 754</td>
<td>5</td>
</tr>
<tr>
<td>755 – 905</td>
<td>6</td>
</tr>
<tr>
<td>906 – 1056</td>
<td>7</td>
</tr>
<tr>
<td>≥ 1057</td>
<td>8</td>
</tr>
</tbody>
</table>

The interval class in the data above is used as a state in the Markov chain so that there are 8 cases of the number of additional COVID-19 positive patients in East Java. The group of states in Table 1 is then determined by the transition frequency of adding positive COVID-19 patients by determining the number of transitions between states with the transition scheme formed as figure-1.
Transition is a state change in the data at \( i \) with the data at \( (i + 1) \). A state change between data can be a transition to the same state. For example, from S-1 to S-1 there was 1 transition that occurred, that is happened on May 13, 2021 to May 14, 2021, the number of positive patients increased from 148 (in state 1) to 141 patients (in state 1). Likewise, for the transition from state 2 to state 3 and we write S-2 to S-3, there are 7 transitions from these existing data rows. The transition from S-2 to S-3 is marked by a change in the data at \( i \) to state \( (i + 1) \), there are 7 transitions. The complete transition between states is presented in table 2.

**Table 2. Transition frequency of COVID-19 patient**

<table>
<thead>
<tr>
<th>S-1</th>
<th>S-2</th>
<th>S-3</th>
<th>S-4</th>
<th>S-5</th>
<th>S-6</th>
<th>S-7</th>
<th>S-8</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>66</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
<td>25</td>
<td>3</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
<td>2</td>
<td>7</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>8</td>
<td>2</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>0</td>
<td>2</td>
<td>2</td>
<td>6</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>

S-1 is state to 1

3.2 Transition Probability

The transition data between states in table 2 is then used to determine the transition probability between states in each state line. For each state line that undergoes a transition to another state, the total probability is 1. Example in state 2 there is a transition to state 1, state 2, and state 3. The probabilities that are formed are from state 2 to state 1 is \( p_{21} = \frac{66}{74} = 0.0135 \), from state 2 to state 2 is \( p_{22} = \frac{2}{74} = 0.08919 \) and from state 2 to state 3 is \( p_{23} = \frac{7}{74} = 0.0946 \). The probability value above can be seen in the second row P matrix. The probability value of each state is shown in the following probability matrix P:

\[
P = \begin{bmatrix}
0.5 & 0.5 & 0 & 0 & 0 & 0 & 0 & 0 \\
0.0135 & 0.8918 & 0.0946 & 0.0383 & 0.0279 & 0 & 0 & 0 \\
0 & 0.1429 & 0.1333 & 0.5334 & 0.1333 & 0.0667 & 0 & 0 \\
0 & 0 & 0.1429 & 0.1429 & 0.4285 & 0.2857 & 0 & 0 \\
0 & 0 & 0 & 0 & 0.3846 & 0.3846 & 0.2308 & 0 \\
0 & 0 & 0 & 0 & 0 & 0.75 & 0.25 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0.5 & 0.5 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\]

States in Markov chain is ergodic because the states are recurrent, aperiodic dan communicate.

State transition in Markov chain is recurrent because for every transition \( i \to j \) and \( i \to j \) so \( i \leftrightarrow j \).

State transition is aperiodic because \( d (1) = \gcd \{1, 2, 3, \ldots\} \) then state 1 aperiodic, \( d (2) = \gcd \{1, 2, 3, \ldots\} \) then state 2 aperiodic, \( d (3) = \gcd \{1, 2, 3, \ldots\} \) then state 3 aperiodic, \( d (4) = \gcd \{1, 2, 3, \ldots\} \) then state 4 aperiodic, \( d (5) = \gcd \{1, 2, 3, \ldots\} \) then state 5 aperiodic. \( d (6) = \gcd \{1, 2, 3, \ldots\} \) then state 6 aperiodic. \( d (7) = \gcd \{1, 2, 3, \ldots\} \) then state 7 aperiodic. \( d (8) = \gcd \{1, 2, 3, \ldots\} \) then state 8 aperiodic.

State transition is communicate because for \( 1 \leftrightarrow 2, \ P_{12} = 0.5 > 0 \) so \( 1 \to 2, \ P_{12} = 0.0135 > 0 \) so \( 2 \to 1 \) then \( 2 \leftrightarrow 3. \)

For \( 2 \leftrightarrow 3, \ P_{23} = 0.0945 > 0 \) so \( 2 \to 3, \ P_{32} = 0.1428 > 0 \) so \( 3 \to 2 \) then \( 2 \leftrightarrow 3. \) For \( 3 \leftrightarrow 4, \)

\[
P_{34} = 0.0833 > 0 \) so \( 3 \to 4, P_{43} = 0.1428 > 0 \) so \( 4 \to 3 \) then \( 3 \leftrightarrow 4. \)

For \( 4 \leftrightarrow 5, P_{45} = 0.3751 > 0 \) so \( 4 \to 5, P_{54} = 0.1333 > 0 \) so \( 5 \to 4 \) then \( 4 \to 5. \) For \( 5 \leftrightarrow 6, P_{56} = 0.1333 > 0 \) so \( 5 \to 6, P_{65} = 0.1428 > 0 \) so \( 6 \to 5 \) then \( 5 \leftrightarrow 6. \) For \( 6 \leftrightarrow 7, P_{67} = 0.2857 > 0 \) so \( 6 \to 7, P_{76} = 0.3846 > 0 \) so \( 7 \to 6 \) then \( 6 \leftrightarrow 7. \) For \( 7 \leftrightarrow 8, P_{78} = 0.2307 > 0 \) so \( 8 \to 7 \) then \( 7 \leftrightarrow 8. \)

According to transitive properties \( i \leftrightarrow j, j \leftrightarrow k \) then \( i \leftrightarrow k \) so we can conclude that \( 1 \leftrightarrow 8. \)

Analyzing the steady state on the Markov chain because it is ergodic. Based on the probability matrix of positive COVID-19 patients, from equation (7) we get the following system of equations.

\[
S_1 = 0.5 \cdot S_1 + 0.0135 \cdot S_2 \\
S_2 = 0.5 \cdot S_1 + 0.8918 \cdot S_2 + 0.1944 \cdot S_3 \\
S_3 = 0.0945 \cdot S_2 + 0.6499 \cdot S_3 + 0.1428 \cdot S_4 + 0.1333 \cdot S_5 \\
S_4 = 0.0833 \cdot S_3 + 0.5 \cdot S_4 + 0.1333 \cdot S_5 + 0.1428 \cdot S_6 \\
S_5 = 0.0277 \cdot S_3 + 0.3571 \cdot S_4 + 0.5333 \cdot S_5 + 0.1428 \cdot S_6 \\
S_6 = 0.1333 \cdot S_4 + 0.4285 \cdot S_5 + 0.3846 \cdot S_7 \\
S_7 = 0.0666 \cdot S_5 + 0.2857 \cdot S_6 + 0.3846 \cdot S_7 + 0.75 \cdot S_8 \\
S_8 = 0.2307 \cdot S_7 + 0.25 \cdot S_8 \\
S_1 + S_2 + S_3 + S_4 + S_5 + S_6 + S_7 + S_8 = 1
\]

We get the solution of the above system of equations are:
\[ S_1 = 0.012401 \]
\[ S_2 = 0.446032 \]
\[ S_3 = 0.216359 \]
\[ S_4 = 0.079953 \]
\[ S_5 = 0.094164 \]
\[ S_6 = 0.065839 \]
\[ S_7 = 0.065198 \]
\[ S_8 = 0.020055 \]

The system solution can be defined as the probability of each state as shown in Table 3.

**Table 3.** Probability of daily number of COVID-19 patient

<table>
<thead>
<tr>
<th>Numbers of COVID-19 patient</th>
<th>Probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>0 – 150</td>
<td>0.012401</td>
</tr>
<tr>
<td>151 – 301</td>
<td>0.446032</td>
</tr>
<tr>
<td>302 – 452</td>
<td>0.216359</td>
</tr>
<tr>
<td>453 – 603</td>
<td>0.079953</td>
</tr>
<tr>
<td>604 – 754</td>
<td>0.094164</td>
</tr>
<tr>
<td>755 – 905</td>
<td>0.065839</td>
</tr>
<tr>
<td>906 – 1056</td>
<td>0.065198</td>
</tr>
<tr>
<td>( \geq 1057 )</td>
<td>0.020055</td>
</tr>
</tbody>
</table>

The probability of daily new cases of COVID-19 patients per day which is potentially quite high is in state 2 with the addition of positive COVID-19 patients as many as 151 to 301 patients with a probability value of 0.446032. While the next highest probability is in state 3 with the addition of positive COVID-19 patients as many as 302 to 452 patients, the probability value is 0.216359. For other states the probability is quite low, which is below 0.1. It can be concluded that in steady state conditions, namely conditions without any new disturbances or significant improvement efforts, the probability of adding new COVID-19 patients in East Java is at the highest probability, namely in the 151-301 value interval. We can take the median value as an estimate of the daily increase in COVID-19 patients in East Java, which is 226 patients.

**4. Conclusion**

The highest probability of daily positive COVID-19 patients is the addition of 151 to 301 patients with a probability of 0.446032. If the average is calculated at the interval through the median value, the number of positive COVID-19 patients will reach a daily average of 226 patients.

**AUTHORS’ CONTRIBUTIONS**

Yuliani: conceptualization, method drafting manuscript, finishing article and Gita Dwi Safitri: data curation, data visualization and editing.

**ACKNOWLEDGMENTS**

The authors wish to acknowledge to the governor of East Java for trying to accelerate the delivery of vaccines to all residents of East Java.

**REFERENCES**


