Determination of Basic Reproduction Numbers using Transition Intensities Multi-state SIRD Model for COVID-19 in Indonesia

F Zuhairoh\textsuperscript{1,2}, D Rosadi\textsuperscript{1}, and A R Effendie\textsuperscript{1}

\textsuperscript{1} Department of Mathematics, Faculty of Mathematics and Natural Science, Gadjah Mada University, Indonesia
\textsuperscript{2} Mathematics Education, STKIP YPUP Makassar, Indonesia
E-mail: faihatuz.zuhairoh@mail.ugm.ac.id

Abstract. The most important quantity in infectious disease epidemiology is the basic reproduction number (\(R_0\)). \(R_0\) is the expected value of the number of infections per unit time. This paper aims to model the spread of COVID-19 in Indonesia using the multi-state SIRD model and then determine the transition intensities to construct \(R_0\). The estimation of the transition intensity uses the maximum likelihood approach with the assumption of a homogeneous time Markov chain with an exponential distribution of transition intensity and the number of transitions in a Poisson distribution. The results of the transition intensity estimation are used to construct \(R_0\) with the next generation matrix method. From the multi-state SIRD model, the largest transition is shown in the individual healing process, namely the movement from an infected to susceptible state, while the smallest transition is the transition from susceptible to dead. The \(R_0\) obtained is 1.079708 (> 1) meaning that the number of individuals infected with COVID-19 will increase until it reaches a stable point. Transition intensities is an effective way of determining \(R_0\) where the dynamics of disease transmission depends on the number of individuals transition between states and the total waiting time in a certain state. \(R_0 > 1\) states that the COVID-19 pandemic in Indonesia has not been over yet.

1. Introduction

Model COVID-19 classifies the population into susceptible individuals (Susceptibles) which consist of people who have not been infected with a virus, infected people (Infectives) which are divided into two, namely with symptoms and without symptoms, people who recover (Recovery) means someone who has been declared recovered after being infected with the coronavirus and died (Dead) means a person who died as a result of being infected with COVID-19. The plot is shown in Figure 1. In the deterministic model, the loss of an infectious disease spread is determined by the sum of the basic reproduction number (\(R_0\)). Meanwhile, the stochastic model usually inherits \(R_0\) from the deterministic model. Stochastic epidemic model using the spread of disease Markov chain model have been carried out by [1, 2] where \(R_0\) and the number of infections provide real measurements of the spread of disease since the beginning or during the process of spread.

The spread of the coronavirus in Indonesia has not yet reached the peak of the pandemic. The cumulative number of positive cases in Indonesia has now exceeded the number of sufferers in...
China. This indicates that public discipline is still lacking to comply with the security protocol COVID-19.

Interesting things in the epidemiology of the disease spread are when the spread of the disease ends, the number of cumulative cases infected, and the number of cases that can be recovered and the number of deaths that occur which will involve many states in it. Thus, the multi-state model is very useful in understanding this process by defining in each state the person’s health condition, duration of treatment, and causes of death.

The multi-state model is commonly used for continuous-time stochastic processes where if \( T = \{t|0 < t < \infty\} \) then the stochastic process has a continuous parameter and is represented by a notation where each time \( t \) occupies a finite state space. All possible rates that can occur at random intervals are called the state space. The state space is discrete when it is finite, while it is continuous when it contains intervals of real lines. Moving between states are also called transitions.

In epidemiology, modeling using a multi-state model has been carried out by [3, 4, 5], which uses Markov assumptions in modeling HIV disease involving sex and age covariates, [6] was applied to spinal cord transplantation disease, [7] in breast cancer, and [8] in Alzheimer’s disease. This Markov model is increasingly developed with various assumptions and recent research by [9] by utilizing the intensity of the transition to determine the value of \( R_0 \). Meanwhile, to calculate the value of \( R_0 \) by using the next-generation matrix has been done by [10, 11, 12, 13].

In this paper, a multi-state model will be used to model the spread of COVID-19 cases in Indonesia with the Markov assumption, as it is well known that COVID-19 modeling has been carried out starting from the SIR compartment model [14, 15], SIRD [16], SEIARD [17], and the phenomenological growth model [18]. Then determine the transition intensity matrix using the Chapman-Kolmogorov differential equation. The process of estimating the transition intensities using the maximum likelihood approach is then used to construct the value of \( R_0 \).

2. Material and Method

2.1. SIRD Model

In this paper, we use the Susceptible Infectious Recovery Dead (SIRD) model, which is the Susceptible which contains all the people in the population who have not been infected with the virus, Infected contains people who have tested positive for COVID-19 after going through a series of tests that can lead to state displacement to Recovery or die. We assume population values of \( N \) are constant and various population categories are defined as \( S(t), I(t), R(t), \) and \( D(t) \).

In the SIRD model of the spread of COVID-19, several parameters are used namely \( s \) stating the greatest chance of vulnerability to COVID-19, \( i \) stating the greatest chance of infection by COVID-19, \( r \) stating the greatest chance of recovery against COVID-19 and \( d \) stating the greatest chance of death caused by COVID-19. The parameters \( s, i, r, \) and \( d \) are positive constants. In this paper, we assume that someone who has been infected with COVID-19 then dies, so that person is said to have died due to COVID-19.

To make it easier in modeling the spread of COVID-19 into a mathematical model assumptions are needed, namely closed population, born individuals not included in populations of vulnerable individuals, individuals recovering from COVID-19 can be re-infected, the incubation period is ignored, each affected individual will become infected, natural deaths and deaths caused by other diseases (besides COVID-19) are not noticed and no vaccine has been found. The stages, levels, and sequence of processes are shown in the diagram below:

The interpretation of Figure 1, explained as follows.

(i) The population of vulnerable individuals

The population of vulnerable individuals infected with COVID-19, will be the population of individuals infected with COVID-19. The population of vulnerable individuals increases
Due to individuals recovering from COVID-19, but does not have immunity to COVID-19 and decreases due to individuals infected with COVID-19. So that the following equation is obtained.

\[
\frac{dS(t)}{dt} = -iS(t)I(t) + sR(t)
\] (1)

(ii) The population of infected individuals

Increased the population infected with COVID-19, due to vulnerable individuals who come in direct contact with sufferers of COVID-19. The decreased population of infected individuals, due to individuals recovering due to natural immunity and individuals who die from COVID-19, so the following equation is obtained.

\[
\frac{dI(t)}{dt} = iS(t)I(t) - rI(t) - dI(t)
\] (2)

(iii) Individuals who recover and are not immune to the coronavirus

This population increases from infected individuals who recover and decreases because individuals who recover from COVID-19 do not have immunity to the coronavirus, so they will become vulnerable individuals again. The population of individuals recovering and not immune to the coronavirus is given by the following equation.

\[
\frac{dR(t)}{dt} = rI(t) - sR(t)
\] (3)

(iv) The population of deceased individuals

The population of individuals who died here only came from deaths caused by COVID-19, so the following equation is obtained.

\[
\frac{dD(t)}{dt} = dI(t)
\] (4)

where \( S(t), I(t), R(t) \) and \( D(t) \) are the states of the model formed, so \( S(t) + I(t) + R(t) + D(t) = N \).

In this paper, the model used is a multi-state model that follows the principles of the SIRD compartment model, where the value of the transition intensity will be sought and then used to determine the value of \( R_0 \). The data used in this paper is COVID-19 data in Indonesia from March to July, 2020 taken from the BNPB website [19]. The data used consisted of the number of Indonesians who were infected, the number who recovered, and the number who died in this time frame either due to COVID-19 or died due to other reasons. Data regarding the number of individuals who died were taken from macrotrends and worldometers [20, 21]. It also requires data on the length of time a patient is treated until healed and the length of a patient being treated until death.
2.2. Multi-state Model for COVID-19

In multi-state modeling in the case of COVID-19, to model the intensity of the transition using 4 states namely susceptible, infected, recovered and dead. The four states can be described as follows.

Figure 2. Multi-state model for COVID-19

The Chapman-Kolmogorov equation states the path that starts at state $i$ at time $t$ to state $j$ at time $u$ through several states $k$ continuously at any time $w$. The Chapman-Kolmogorov equation from Figure 2 is given as follows.

$$ P_{ij}(t, u) = \sum_{k=1}^{4} P_{ik}(t, w) P_{kj}(w, u), \quad t \leq w \leq u $$

In the figure 2, there are 4 many states namely, susceptible, infected, recovered, and dead. The arrows indicate the possible transitions between the four states. If it is made in the form of a matrix, it is obtained as follows $\frac{d}{dt}P_{ij} = P_{ij}Q_{ij}$ where $P_{ij}$ is a matrix whose elements are $p_{ij}$, where $i, j = 1, 2, 3, 4$ and $Q_{ij}$ called the transition intensities matrix where the elements are $\mu_{ij}$ for $i \neq j$ and $-\sum_{j=1, j\neq i}^{4} \mu_{ij}$ for $i \neq j$.

Thus, from Figure 2 the following transition probability and transition intensity matrix are obtained.

$$ P_{ij} = \begin{bmatrix} p_{11} & p_{12} & p_{13} & p_{14} \\ p_{21} & p_{22} & p_{23} & p_{24} \\ p_{31} & p_{32} & p_{33} & p_{34} \\ 0 & 0 & 0 & 1 \end{bmatrix} $$

$$ Q_{ij} = \begin{bmatrix} -(\mu_{12} + \mu_{14}) & \mu_{12} & 0 & \mu_{14} \\ 0 & -(\mu_{23} + \mu_{24}) & \mu_{23} & \mu_{24} \\ \mu_{31} & 0 & -\mu_{31} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} $$

(6)
3. Results and Discussion

3.1. Estimates of the Transition Intensities

Assumption of constant transition intensity means that the number of transitions from one state to another per time unit does not depend on the initial time depends only on the time interval. So, the Markov chain used is a homogeneous time Markov chain. In observation, we can observe the time and type of each transition made by each individual. Assuming the intensity of the transition is constant, it results in the time spent in each state having an exponential distribution [22].

Suppose $T_{ik}$ is a continuous random variable of time spent by individual $k$ in state $i$ before moving to another state (waiting time in state $i$) with the hazard function $\mu(t_{ik}) = \mu$. By using the relationship between hazard function and density function is obtained

$$f(t_{ik}) = \mu(t_{ik})e^{-\int_0^{t_{ik}} \mu(s)ds} = \mu e^{-\int_0^{t_{ik}} \mu ds} = \mu e^{-\mu t_{ik}} \quad (7)$$

The form above is an opportunity density function (pdf) of an exponential distribution. So, if the intensity of the transition is constant, then $T_{ik}$ has an exponential distribution.

Furthermore, according to the Poisson process, the random variable $N_{ijk}$, which is the number of transitions from state $i$ to state $j$ made by individuals $k$ with a Poisson distribution with an average of $t_{ik}\mu_{ij}$, is written

$$N_{ijk} \sim \text{Poisson}(t_{ik}\mu_{ij}) \quad (8)$$

with

$$\Pr(N_{ijk} = n_{ijk}) = f_{N_{ijk}}(t_{i}) = \frac{e^{-\mu_{ij}t_{ik}}(\mu_{ij}t_{ik})^{n_{ijk}}}{n_{ijk}!} \quad (9)$$

Then the likelihood function is

$$L(\mu_{ij}) = \prod_{i \in S} \prod_{j \in S} \left( \prod_{k=1}^{N} f(\mu_{ij}t_{i}) \right)$$

For example,

$n_{ij} = \sum_{k=1}^{N} n_{ijk}$ : The number of transitions from state $i$ to $j$ made by all individuals in a given time

$t_{i} = \sum_{k=1}^{N} t_{ik}$ : total waiting time in state $i$ for all individuals

So,

$$L(\mu_{ij}) = \prod_{i \in S} \prod_{j \in S} e^{-\mu_{ij}t_{i}}(\mu_{ij}t_{i})^{n_{ij}} \quad (10)$$

If the part that does not contain $\mu_{ij}$ is ignored then

$$L(\mu_{ij}) = \prod_{i \in S} \prod_{j \in S} e^{-\mu_{ij}t_{i}}(\mu_{ij})^{n_{ij}}$$
and log-likelihood function is

$$\log L(\mu_{ij}) = \prod_{i \in S} \prod_{j \in S} \log \left( e^{-\mu_{ij} t_i} (\mu_{ij})^{n_{ij}} \right) = \sum_{i \in S} \sum_{j \in S} -\mu_{ij} t_i + n_{ij} \log(\mu_{ij})$$  \hspace{1cm} (11)

The derivative of the above equation is

$$\frac{d \log L(\mu_{ij})}{d \mu_{ij}} = t_i + \frac{n_{ij}}{\mu_{ij}}$$

If the equation is equal to zero, it is obtained

$$\mu_{ij} = \frac{n_{ij}}{t_i}.$$  \hspace{1cm} (12)

3.2. Basic Reproduction Numbers ($R_0$)

Based on Figure 2 which consists of 4 states, namely susceptible, infected, recovered, and dead where the transition occurs from susceptible to infected, susceptible to dead, from infected to recovered, and the last from infected to dead. To calculate the basic reproduction number ($R_0$), we only use the infected state, so that the matrix $Q$ in equation 6 can be formed into

$$J = \begin{bmatrix} \mu_{12} & 0 & \mu_{14} \\ - (\mu_{23} + \mu_{24}) & \mu_{23} & \mu_{24} \\ 0 & -\mu_{31} & 0 \end{bmatrix}$$  \hspace{1cm} (13)

In calculating $R_0$ we use the Next Generation Matrix method which involves partitioning the matrix $J$ into the submatrix $F$ and $V$ where $F$ is a new non-negative infection matrix and $V$ consists of death, increased state, and other transitions.

$$J = F - V$$

$$K = FV^{-1}$$

So that the matrix $F = \begin{bmatrix} \mu_{12} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$ and $V = \begin{bmatrix} 0 & 0 & -\mu_{14} \\ \mu_{23} + \mu_{24} & -\mu_{23} & -\mu_{24} \\ 0 & \mu_{31} & 0 \end{bmatrix}$ are obtained.

Whereas the inverse of the matrix $V$ is obtained by

$$V^{-1} = \begin{bmatrix} \mu_{24} & \frac{1}{\mu_{23} + \mu_{24}} & -\frac{1}{\mu_{23}} \mu_{23} + \mu_{24} \\ \frac{1}{\mu_{14}} & 0 & \frac{1}{\mu_{31}} \mu_{23} + \mu_{24} \\ 0 & 0 & 0 \end{bmatrix}.$$  

$$K = FV^{-1} = \begin{bmatrix} \mu_{12}\mu_{24} & -\mu_{12} & \mu_{12}\mu_{23} \\ \mu_{14}(\mu_{23} + \mu_{24}) & 0 & \mu_{14}(\mu_{23} + \mu_{24}) \\ 0 & 0 & 0 \end{bmatrix}$$

By using the dominant eigenvalue it is obtained

$$R_0 = \frac{\mu_{12}\mu_{24}}{\mu_{14}(\mu_{23} + \mu_{24})}.$$  \hspace{1cm} (14)

The results of the calculations performed are as follows.
Table 1. The total transition amount and waiting time for COVID-19 patients in each state.

| $\sum n_{12}$ | $\sum n_{14}$ | $\sum n_{23}$ | $\sum n_{24}$ | $\sum n_{31}$ | $\sum t_1$ | $\sum t_2$ | $\sum t_2$ |
|--------------|--------------|--------------|--------------|--------------|------------|------------|------------|
| 108,376      | 7,250        | 65,907       | 5,131        | 65,907       | 12,424,516 | 1,726,059  | 197,721    |

The interpretation of each symbol in the table is $\sum n_{12}$ is the number of transitions that occur from susceptible to infected means that the number of individuals infected with COVID-19 starting from March to July 2020. $\sum n_{14}$ is the number of transitions that occur from susceptible to dead states that not caused by COVID-19, this data was obtained from the number of deaths from March to July 2020 reduced by the number of deaths due to COVID-19. $\sum n_{23}$ is the number of transitions that occur from infected to recovered state means that many individuals recovered after being infected with COVID-19 starting from March to July 2020. $\sum n_{31}$ is is the number of transitions that occur from a state of recovery to being susceptible to COVID-19 and the last one is $\sum n_{24}$ which is the number of transitions that occur from infected to dead state meaning that many individuals infected with COVID-19 have died from March to July 2020.

The waiting time (duration) of a patient is in a susceptible state ($\sum t_1$) is obtained from the number of days starting from the beginning of the study until the time someone is tested positive for COVID-19 while the length of a patient is in an infected state ($\sum t_2$) is obtained from the number of days since someone tested positive for COVID-19 until the patient recovered. Because the data is not known in which patients recover or die only in numbers, it is assumed that patients who arrive early will recover or die first. The length of time the patient is in a recovered state before finally returning to susceptible state ($\sum t_3$) is obtained from the total number of patients healed multiplied by 3 because to be declared completely recovered of COVID-19, the patient must do a negative RT-PCR test plus at least 3 days no longer showing symptoms fever and respiratory problems.

So the estimation results for transition intensity are given in the following table.

Table 2. Estimated transition intensity.

| $\mu_{12}$ | $\mu_{14}$ | $\mu_{23}$ | $\mu_{24}$ | $\mu_{31}$ |
|------------|------------|------------|------------|------------|
| Estimate   | 0.008723   | 0.000584   | 0.038184   | 0.002973   | 0.333333  |

From the estimated transition intensity results in Table 2, we get the basic reproduction numbers

$R_0 = \frac{\mu_{12}\mu_{24}}{\mu_{14}(\mu_{23} + \mu_{24})} = \frac{(0.008723)(0.002973)}{0.000584(0.038184 + 0.002973)} = 1.079708$. If using initial data (only in March), the $R_0$ obtained is 1.716105, this calculation is almost the same as that obtained in research with the SIR model by [23]. The results of the calculation of reproduction numbers using several methods are presented in table 3 as a numerical comparison of the $R_0$ obtained using the transition intensity. Because the value of $R_0 > 1$ means the number of infected individuals will continue to increase until it reaches its equilibrium point.
Table 3. Reproduction number estimation results.

| Method                     | $R_0$         | $R$          | 95% CI                  |
|---------------------------|---------------|--------------|-------------------------|
| Transition Intensity      | 1.079708      |              |                         |
| Exponential Growth Rate   | 1.076883      | [1.076249, 1.077518] |                         |
| Maximum Likelihood        | 1.058728      | [1.049857, 1.067684] |                         |
| Attack Rate               | 1.058452      | [1.058097, 1.058807] |                         |

4. Conclusion
The spread of COVID-19 in Indonesia can be modeled using a multi-state model with four states, namely susceptible, infected, recovered, and dead. Then using COVID-19 data until July 2020, an estimate of the transition intensities using the maximum likelihood approach was obtained where four estimates of the intensity of the transition were obtained, namely susceptible to infected, susceptible to death, infected to recovered, infected to dead, and recovered to susceptible. The results of the transition intensity estimation are used to construct the $R_0$ value with the next generation matrix method. Because the $R_0$ obtained for the COVID-19 spread model is still more than one, it shows that the spread of COVID-19 in Indonesia has not ended.

Acknowledgments
Authors would like to thank the Lembaga Pengelola Dana Pendidikan (LPDP), Departemen Keuangan Indonesia for taking part in providing financial support.

References
[1] Artalejo J R and Lopez-Herrero M J 2013 On the exact measure of disease spread in stochastic epidemic models *Bull. Math. Biol.* 75 1031–50
[2] Lopez-Herrero M J 2016 Epidemic transmission on SEIR stochastic models with nonlinear incidence rate *Math. Methods Appl. Sci.* 40 2532–41
[3] Commenges D 1999 Multi-state models in epidemiology *Lifetime Data Analysis* 5 315–27
[4] Longini I M Jr, Clark W S, Byers R H, Ward J W, Darrow W W, Lemp G F and Hethcote H W 1989 Statistical analysis of the stages of HIV infection using a markov model *Stat. Med.* 8 831-43
[5] Yiu S, Farewell V T and Tom B D M 2018 Clustered multistate models with observation level random effects, mover-stayer effects and dynamic covariates: modelling transition intensities and sojourn times in a study of psoriatic arthritis *J. R. Stat. Soc. Ser. C Appl. Stat.* 67 481-500
[6] Klein J P and Shu Y 2002 Multi-state models for bone marrow transplantation studies *Stat. Methods Med. Res.* 2 302-35
[7] Meier-Hirmer C and Schumacher M 2013 Multi-state model for studying an intermediate event using time-dependent covariates: application to breast cancer *BMC Med. Res. Methodol.* 13 80
[8] Zhang L, Lim C Y, Maiti T, Li Y, Choi J, Bozoki A and Zhu D C 2019 Analysis of conversion of alzheimer’s disease using a multi-state markov model *Stat. Methods Med. Res.* 28 2801-19
[9] Orowe I, Weke P, Ottieno J and Onyango N 2015 Multistate modelling vertical transmission and determination of $R_0$ using transition intensities *Appl. Math. Sci.* 9 3941-56
[10] Diekmann O, Heesterbeek J A P and Roberts M G 2010 The construction of next-generation matrices for compartmental epidemic models *J. R. Soc. Interface* 7 873-85
[11] Sato K 2019 Basic reproduction number of SEIRS model on regular lattice *Math. Biosci. Eng.* 16 6708–27
[12] Van den Driessche P 2017 Reproduction numbers of infectious disease models *Infectious Disease Modelling* 2 288-303
[13] Yang H M 2014 The basic reproduction number obtained from jacobian and next generation matrices - a case study of dengue transmission modelling *BioSystems* 126 52-75
[14] Bastos S B and Caiuceiro D O 2020 Modeling and forecasting the early evolution of the covid-19 pandemic in Brazil *Preprint* 2023.14288v3 [q-bio.PE]
[15] Tumwiine J, Mugisha J Y T and Luboobi L S 2007 A mathematical model for the dynamics of malaria in a human host and mosquito vector with temporary immunity Appl. Math. and Computation 189 1953-65
[16] Chang R, Wang H, Zhang S, Wang Z, Dong Y, Tsamlag L, Yu X, Xu C, Yu Y, Long R, Liu NN, Chu Q, Wang Y, Xu G, Shen T, Wang S, Deng X, Huang J, Zhang X, Wang H and Cai Y 2020 Phase- and epidemic region-adjusted estimation of the number of coronavirus disease 2019 cases in China Front. Med. 14 199–209
[17] de León U A-P, Pérez Á G C and Avila-Vales E 2020 A data driven analysis and forecast of an SEIARD epidemic model for covid-19 in Mexico Preprint 2004.08288v1 [q-bio.PE]
[18] Zuhairoh F and Rosadi D 2020 Real-time forecasting of the covid-19 epidemic using the richards model in south sulawesi, Indonesia Indonesian Journal of Sci. & Technology 5 456-62
[19] https://covid19.go.id/peta-sebaran, retrieved on July 31, 2020
[20] https://www.macrotrends.net/countries/IDN/indonesia/death-rate, retrieved on July 31, 2020
[21] https://www.worldometers.info/world-population/indonesia-population/, retrieved on July 31, 2020
[22] Jones B L 1994 Actuarial calculation using a markov model Transactions of the Society of Actuaries 46 227-50
[23] Fajar M 2020 Estimation of COVID-19 reproductive number case of Indonesia Preprint 10.13140/RG.2.2.32287.92328