Jurnal Pijar MIPA
https://jurnalfkip.unram.ac.id/index.php/JPM

ISSN 1907-1744 (Print)
ISSN 2460-1500 (Online)

Stochastic Modeling with Poisson Hidden Markov in Hepatitis B Cases

Ersya Nurul Fairuz*, Rina Widyasari, Rima Aprilia

Mathematic Study Program, Faculty of Science and Technology, North Sumatra State Islamic University, North Sumatra, Indonesia
*E-mail: ersyanurulfairuz@gmail.com

Received: October 14, 2024. Accepted: November 28, 2024. Published: November 30, 2024

Abstract: Hepatitis B is transmitted through blood or body fluids contaminated with the virus from Hepatitis B sufferers (carriers). The factors that cause a person to contract Hepatitis B are sexual intercourse, blood contact, placental contact from the mother to the baby, and saliva. The incubation period for Hepatitis B Virus (HBV) ranges from 30 - 180 days with an average of 60 - 90 days. HBV can be detected 30 - 60 days after infection and persists for a certain period. Thus, the behaviour of infectious diseases can be observed and described using mathematical modelling. Mathematical modelling is a field of mathematics that represents and explains physical systems or problems that occur in the real world and are solved in mathematical statements. The mathematical model used to overcome uncertainty in variable values is a stochastic model. These causal factors are not directly observed and form a Markov chain. The model that can be used for the uncertainty of an event is the Hidden Markov Model. The Hidden Markov Model (MHM) is a type of stochastic modelling that does not recognize the factors that trigger the problem being modelled. The Poisson Hidden Markov model is used to model Hepatitis B disease. Hepatitis B disease data is a series of observations that experience overdispersion and depend on the trigger of Hepatitis B disease, which is assumed not to be observed directly and forms a Markov chain. The aim to be achieved in this research is to model Hepatitis B disease at the Medan Haji Hospital using the Poisson Hidden Markov model and to find parameter estimates using the Expectation Maximization Algorithm. This type of research uses quantitative research methods. The conclusions that can be drawn based on the results and previous discussions are as follows: the method for determining the average number of patients in patient B can use the PHMM (Poisson Hidden Markov Model) method with the EiM (Expectation-Maximization Algorithm) algorithm, the best model for the number of Hepatitis B patients in Haji Hospital at this hospital is the model with three hidden cases with the parameter estimation value. The average number of Hepatitis B patients is 0.0324 in 1 month, and the average predicted results obtained by the hidden condition model 3 align with the original conditions in the previous months.

Keywords: Hepatitis B; Poisson Hidden Markov; Stochastic Model.

Introduction

In the world, there are many infectious diseases, including influenza, tuberculosis (TBC), dengue hemorrhagic fever (DHF), HIV/AIDS, pneumonia, measles, hepatitis, and so on. Currently, a health problem throughout the world in Indonesia is hepatitis [1]. Hepatitis is an infectious disease of liver cells caused by viral infectious microorganisms, drugs, toxins, and chemicals [2]. There are various types of hepatitis, namely Hepatitis A, B, C, D, and E. WHO (World Health Organization) said that this disease will spread to the world's population and is transmitted orally due to a lack of knowledge about clean and healthy living [3]. Indonesia is the country with the second largest endemic of Hepatitis B in the SEAR (South East Asian Region) countries after Myanmar, with 2 billion people in the world, around 240 million of whom suffer from chronic hepatitis. Currently, it is estimated that 28 million Indonesians suffer from Hepatitis B [4].

Hepatitis B is transmitted through blood or body fluids contaminated with the virus from Hepatitis B sufferers (carriers). The factors that cause a person to contract Hepatitis B are sexual intercourse, blood contact, placental contact from the mother to the baby, and saliva. The incubation period for Hepatitis B Virus (HBV) ranges from 30 - 180 days, with an average of 60 - 90 days [5]. HBV can be detected 30 - 60 days after infection and persists for a certain period [6]. Hepatitis can attack anyone, regardless of age or gender. Hepatitis B can be prevented by providing immunization. Hepatitis B immunization is given as early as possible after birth [7]. Giving Hepatitis B immunization to newborns must be based on whether the mother contains active Hepatitis B virus or not at birth. Repeat Hepatitis B immunization can be considered at 10-12 years old. If a child up to 5 years of age has not received Hepatitis B immunization, it should be given as soon as possible [8].

The Ministry of Health of the Republic of Indonesia announced the latest situation regarding the mysterious Acute Hepatitis. On September 15, 2022, there were a total of 91 cases of acute hepatitis with unknown causes in patients under 16 years of age in 22 provinces, of which 35 were in probable or suspected status, seven were pending because they were waiting for lab test results, and 49 were discarded because they were suffering from other diseases. Of the 35 probable cases, DKI contributed the most, namely 12 cases, Yogyakarta 3 cases, West Kalimantan, Central Java, North Sulawesi, Bali, and North Sumatra 3 cases, and the others 1 case each [9].

How to Cite:
Fairuz, E. N., Widyasari, R., & Aprilia, R. (2024). Stochastic Modeling with Poisson Hidden Markov in Hepatitis B Cases. *Jurnal Pijar Mipa*, *19*(6), 1111–1117. https://doi.org/10.29303/jpm.v19i6.7510

Thus, the behavior of infectious diseases can be observed and described using mathematical modeling [10]. Modeling Mathematics is a field that represents and explains physical systems or problems that occur in the real world, which are solved in mathematical statements. Mathematical modeling is divided into four types: empirical, simulation, deterministic, and stochastic [11]. An empirical model is a model that is used based on observations without being based on theory or knowledge that generates the phenomenon. A simulation model is a written mathematical model based on rules. A deterministic model is one in which the values and variables are known with certainty. In contrast, the stochastic model is a variable value from a model that is not known with certainty in the form of a random variable [12] the four models have advantages and disadvantages and are equally needed.

The mathematical model used to overcome uncertainty in variable values is a stochastic model. The stochastic model is also called a calculated probability model of each event occurring [13]. Several factors can cause uncertainty in an event. These causal factors are not directly observed and form a Markov chain. The model that can be used for the uncertainty of an event is the Hidden Markov Model [14].

The Hidden Markov Model (MHM) is a type of stochastic modeling that does not recognize the factors that trigger the problem being modeled. The Hidden Markov model consists of a pair of stochastic processes, namely the observation process and processes that influence the observation. The cause of this observation is called state. The Poisson Hidden Markov Model (MPHM) is a Hidden Markov model with discrete time, and the observation process is assumed to be a Poisson distribution [15]. One of the characteristics of MPHM is that it is overdispersive, that is, the variety of data is greater than the average.

The Poisson Hidden Markov model is used to model Hepatitis B disease. Hepatitis B disease data is a series of observations that experience overdispersion and depend on the trigger of Hepatitis B disease, which is assumed not to be observed directly and forms a Markov chain [16].

In previous research, several studies or papers discussed epidemics using stochastic models, for example, [17] research entitled Stochastic Model Analysis of Hepatitis B Virus Transmission discusses the stochastic model of Hepatitis B virus transmission with the SIR epidemic model using two infection rates, namely the rate of acute and chronic infections follows the Wiener process. The model was then searched for an analytical solution for the Ito formula. So, the author was interested in researching the stochastic model with Poisson Hidden Markov for Hepatitis B in North Sumatra [18].

This innovation is because disease-spread models usually use SEIR or SIR and must know the causal factors. So researchers are interested in using the Poisson Hidden Markov Model method to model Hepatitis B disease at the Medan Haji Hospital [19]. The characteristic of the Poisson Hidden Markov model is its parameters. Estimating model parameters uses the Maximum Likelihood method, calculated using the Expectation Maximization Algorithm [20]. Based on research conducted by [21] researched parameter estimation and the convergence of Poisson Hidden Markov Model parameter estimates. In this research, the main case of MPHM is to estimate the parameters that maximize the Likelihood function. The Likelihood Function is calculated using the Forward-Backward algorithm. The Expectation Maximization (AEM) algorithm maximizes the Likelihood function.

The author is interested in raising the title about modeling Hepatitis B disease data at the Medan Haji Hospital using the Poisson Hidden Markov Model method and looking for parameter estimates using the Expectation Maximization Algorithm.

Research Methods

Types of research

This type of research uses quantitative research methods. Quantitative methodology is a technique that presents information or data dominated by a number structure, and the data analysis or data analysis used is statistical.

Research Variables

These research variations are everything that becomes an object of observation in research. These variations of research are used as follows:

1. State 1, data "HBsAg" Protein found inside and outside the Hepatitis B virus.
2. Statei 2, data "HBeAg" Protein, which is not present in the virus, allows the virus to circulate freely in the blood and body tissues.
3. State 3, data on "Acute Hepatitis B" for patients diagnosed with Hepatitis B in less than 6 months.
4. Statei 3, data on "Hepatitis B Chronis" patients admitted to Hepatitis B for more than 6 months.

Research Procedures

The research procedure is based on the stages that will be carried out. There is a research procedure to make it easier for researchers or readers to learn research by seeing the steps required when carrying out research, namely:

1. Data collection
2. Counting the number of Hepatitis oil patients at RSU Haji Medan
3. Checking overdispersion to compare the value of variance and the mean value of a large amount of data. In the initial stage, the data is used to determine whether the results are natural, that is, by looking at the average value and variance of the existing data. If the variance value exceeds the average value, then the data used can be interpreted as Poisson data.
4. Stochastic modeling stage using Poisson Hidden Markov Modeling by searching for input parameter values for modeling.
   a. Determine Hidden Markov Model

This stage is modeled by intuiting the state xn = [1,2,3], and the next step is to intuit the Hidden Markov Model, which is intended to be a graph that is a transition from one state to another.

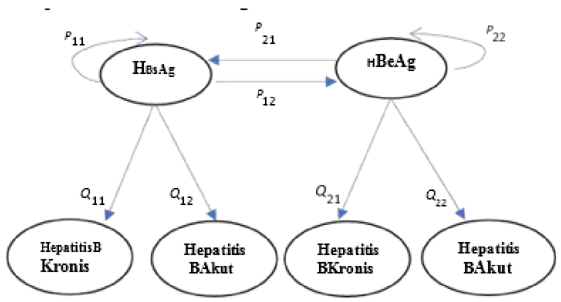

Figure 1. Graf *Hidden* Markov Model

b. Determinate Poisson Hidden Markov Model

This stage looks for the best estimation model for the number of Hepatitis B patients from 3 modes. The three modes are hidden states, namely m = (2,3,4). This study only models up to the hidden existence m = 4 because the data used does not provide coverage for the hidden existence m ≥ 5 The Poisson Hidden Markov Modeil is used, which estimates the EM algorithm.

5. Determinate Parameter

At this stage, we calculate the average number of patients per month when every λ has a Poisson distribution criterion with the initial probability of occurrence, and the transition opportunity matrix hidden existence has a size of MXM. Calculating the surface length in sample space can be done by dividing the range by the number of windows.

6. Calculate the estimates of the EM algorithm, which consists of the Ei and M phases. Then replace the missing data in the rare results E. That's it, maximized in stage M
7. Carry out an analysis of the results of the analysis that has been obtained in the best model using AIC values.

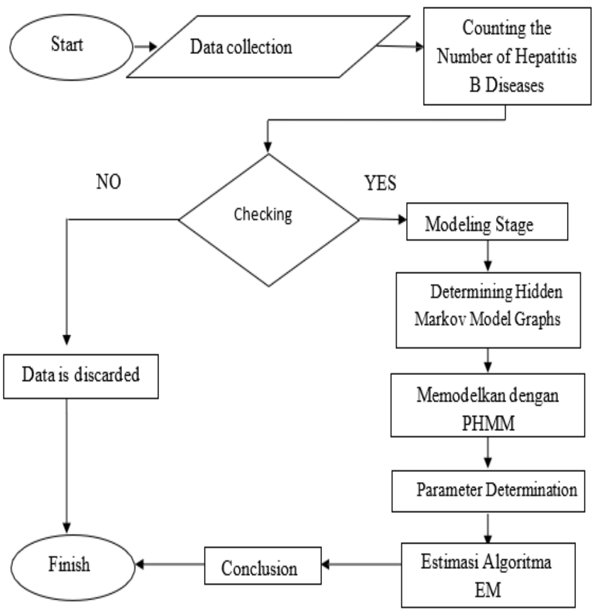

Figure 2. Research Framework

Results and Discussion

In this study, the information used is Hepatitis B patients obtained from the Haji Meidan Hospital. This information was taken over the past 3 years, from January 2020 to December 2022.

Determining overdispersion was carried out using data on the number of Hepatitis B patients over 36 months. From the data collection process on the number of Hepatitis B patients, the data characteristics follow a Poison-based data collection process. Follow the data distribution as in Figure 2.

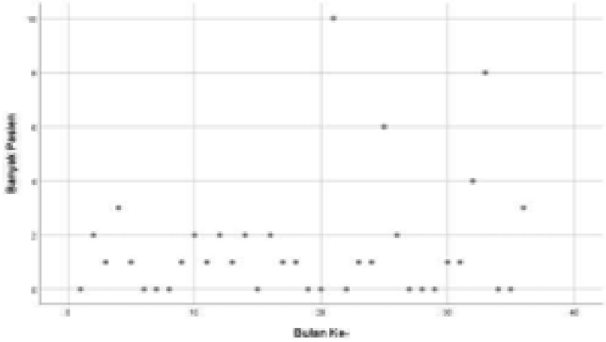

Figure 3. Distribution of Hepatitis B Patients at Haji Hospital Medan

The data used to determine whether the results are different is by looking at the average value and variance of the existing data. The data will naturally vary if the variance value exceeds the average value.

From the results of providing information that the number of events that occurred was 36 data $\bar{X} = 1{,}5833$ with and $s^2 = 5{,}107$. Meanwhile, the maximum value of the number of events is 10. It is assumed $s^2 = 5{,}107$ and $\bar{X} = 1{,}5833$. Suffers from overdispersion of the Poisson distribution.

Determination of the Hidden Markov Model

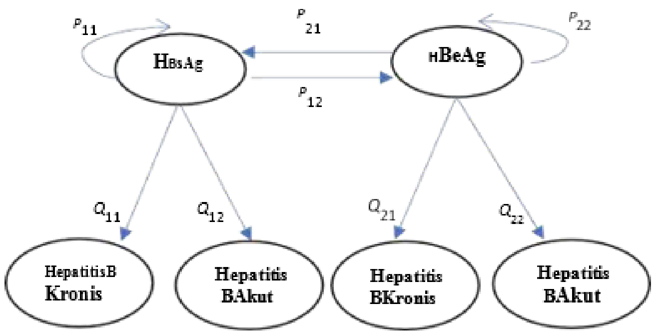

Figure 3. Graph Hidden Markov Model

Based on Figure 4.1, state transfers that may occur in the oil balance of Hepatitis B have 8 transitions, namely transfers from P state 1 to state 1, state 1 to state 2, state 2 to state 1, state 2 to state 2, etc. moving from Q state 1 to state 1, state 1 to state 2, state 2 to state 1, state 2 state 2.

Determination of the Transition Probability Matrix

In the presence of HBsAg infection, 26 patients experienced Chronic Hepatitis B, and 26 acute patients and 5 patients. Meanwhile, in the presence of HBeiAg infection, there were 18 patients suffering from Chronic Hepatitis B and 8 patients suffering from acute disease. Thus, the transition probability can be calculated. The above probability calculations are calculated using equations :

1) State 1 (HBAg) → State 1 (HBAg)
$$P_{11} = \frac{8}{9} = 0{,}89$$
2) State 1 (HBAg) → State 2 (HBeAg)
$$P_{12} = \frac{1}{9} = 0{,}11$$
3) State 1 (HBAg) → State 3 (Hepatitis B Kronis)
$$Q_{11} = \frac{26}{31} = 0{,}83$$
4) State 1 (HBAg) → State 2 (Hepatitis B Acute)
$$Q_{12} = \frac{5}{31} = 0{,}1$$
5) State 2 (HBAg) → State 1 (HBeAg)
$$P_{21} = \frac{5}{6} = 0{,}83$$
6) Statei (HBAg) → State 2 (HBAg)
$$P_{22} = \frac{1}{6} = 0{,}17$$
7) State 2 (HBAg) → State 1 (Hepatitis B Cronis)
$$Q_{21} = \frac{5}{26} = 0{,}23$$
8) State 2 (HBeiAg) → State 2 (Hepatitis B Acute)
$$Q_{21} = \frac{6}{26} = 0{,}07$$

From these results, it shows that the probability of HBsAg infection in the Hajj hospital is in state 1 (Hepatitis B) being 0.89, state 1 is in state 2 being 0.11, state 1 P with HBsAg infection being in state 1 Q Hepatitis B value 0.83, and state 1 P due to HBsAg infection leads to state 2 Q Hepatitis B Acute has a value of 0.16. And state 2 (HBeiAg) is 0.83, state 2 is 0.17 and state 2 P is 0.69, and state 2 P is 0.69, and state 2 P is 0.69, and state 2 P with HBsAg is state 2 Q Hepatitis B Acute value 0.30.

Modeling with Poisson Hidden Markov Model

This study will look for the best estimation model for the number of Hepatitis B patients from 3 models. The 3 modes are hidden states, namely m = (2,3,4). This study only models up to the hidden existence m = 4 because the data does not cover the hidden existence ≥ 5. The method used is Poisson Hidden Markov Modeil with eistimization of the EM algorithm.

The number of classes in the table being distributed in frequency is determined by the number of hidden circumstances included. In this process, the team takes values ranging from 0 to 11, the number of Hepatitis B patients for uniform distribution of the levels of each class. As a model example with the condition m=2, for example, if there are 2 hidden conditions with the average number of Hepatitis B patients $\lambda = (\lambda_1, \lambda_2)$, then there are 2 classes with long intervals for each class.
$$c = \frac{\text{range}}{\text{lots of classes}} = \frac{12}{2} = 6$$

Based on the value of c above calculated by the interval length for each class 6, the sample space for the number of Hepatitis B patients in the hidden state 1 is {0,1,2,3,4,5} and the remainder enters the hidden state 2.

In 3 hidden states (m= 3) with the average number of parameters event is $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ then there are 3 categories with each. The class has a long interval c=4. This means that there is a lot of sample space. The occurrence of patient Hepatitis B in hidden state 1 is {0,1,2,3}, for hidden state 2 is {4,5,6,7}, while the remainder is the sample space for hidden state 3, namely {8,9, 10,11}.

In 4 hidden states (m= 4) with the average number of parameters event is $\lambda = (\lambda_1, \lambda_2, \lambda_3, \lambda_4)$, then there are 4 categories with each.

The class has an interval length c = 3. This means that the sample space for the number of occurrences of patient B in hidden state 1 is {0,1,2}, for hidden state 2, the sample space is {3,4,5}, for hidden state 3, the sample space is {6,7,8}, while the remainder is the sample space for the hidden state 4.

The next step is to enter data in Table 4.1 for each group of hidden circumstances based on the sample space, so that determined by the parameter values $\lambda = (\lambda_1, \lambda_2)$ Parameiteir $\delta = (\delta_1, \delta_2)$, namely the initial opportunity in hidden state 1 and hidden state 2, it is calculated by calculating the number of incident frequencies in each group of hidden states and the result of the total occurrence of occurrences hide it.

Calculating the probability of hidden circumstances m = 2 hidden circumstances 1 above is calculated using :

1) State 1 (HBAg) → State 1 (HBAg)
$$P_{11} = 0{,}89$$
2) State 1 (HBsAg) → State 2 (HBAg)
$$P_{12} = 0$$
3) State 1 (HBsAg) → State 3 (Hepatitis B Cronis)
$$Q_{11} = 0{,}55$$
4) State 1 (HBsAg) → Statei 2 (Hepatitis B Acute)
$$Q_{12} = 0{,}15$$
5) State 2 (HBAg) → State 1 (HBAg)
$$P_{21} = 0$$
6) State 2 (HBAg) → State 2 (HBAg)
$$P_{12} = 0{,}17$$
7) State 2 (HBAg) → State 1 (Hepatitis B Cronis)
$$Q_{21} = 0{,}24$$
8) State 2 (HBeiAg) → State 2 (Hepatitis B Acute)
$$Q_{22} = 0{,}06$$

From these results, it shows that the probability of HBsAg infection in the Hajj hospital is at state 1 (Hepatitis B) be 0.89, state 1 is going to state 2 to stat0.11, state 1 P with HBsAg infection be at state 1 Q Hepatitis B is chronic 0.55. State 1 P deng infection HBsAg mento state 2 Q Hepatitis B Acute 0.15, and state2 (HBeiAg) 0.83, state 2 mengo state 2 is 0.17 and state 2 P infection HBeAg to state 1 Q Hepatitis B chronic 0.24, and state 2 P HBsAg infection leads to state 2 Q Hepatitis B Acute rate 0.06.

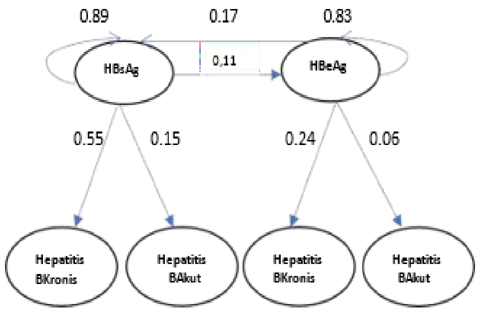

Figure 4. Hepatitis B Disease Transition State Graph

Next, calculate the probability of the hidden state m = 2 hidden circumstances 2.

1) State 1 (HBsAg) → State 1 (HBsAg)
$$P_{11} = 1$$
2) State 1 (HBsAg) → State 2 (HBeAg)
$$P_{12} = 0{,}11$$
3) State 1 (HBsAg) → State 3 (Hepatitis B Cronis)
$$Q_{11} = 0{,}33$$
4) QState 1 (HBsAg) → State 2 (Hepatitis B Acute)
$$Q_{12} = 0$$
5) State 2 (HBeAg) → State 1 (HBeAg)
$$P_{21} = 0{,}5$$
6) State 2 (HBeAg) → State 2 (HBsAg)
$$P_{22} = 0$$
7) State 2 (HBeAg) → State 1 (Hepatitis B Cronis)
$$Q_{21} = 0{,}42$$
8) State2 (HBeiAg) → State 2 (Hepatitis B Acute)
$$Q_{22} = 0{,}25$$

From these results, it shows that the probability of HBsAg infection in the Hajj hospital is at state 1 (Hepatitis B) to be 1, state 1 to state 2 to be 0.11, state 1 P with HBsAg infection to be at state 1 Q Hepatitis B is 0.33, and state 1 0.42, and state 2 P with HBsAg infection leading to state 2 Q Hepatitis B Acute is 0.25.

In the model with 4 hidden states (m = 4), it has parameters the average number of events λ in 1 month. Initial opportunities for occurrence and transition opportunity matrix:

This means if, in the current period, there is a hidden state of 1, then the chance in the next 1 month of being in a hidden state of 1 is equal to 0.82. Meanwhile, if in this period there is a hidden condition of 1, then the chance for 1 month of dreaming of being in the hidden condition of 2 is equal to 0.10, if in this period, there is a hidden situation of 1, then the chance for 1 month of dreaming of being in the hidden state of affairs hide 3 equal to 0.03. If, in the current period, it is in the hidden state of 1, then the chance for the next 1 month to be in the hidden state of 4 is approximately 0.034.

Estimating PHMM Parameters using the EM Algorithm

After determining the PHMM parameters, the next step is to calculate the estimated value of the parameters, namely λ, δ and Γ for each model using the EM Algorithm. The calculation of victimization values used is 1e-06, where the value of tolerance is the criterion.

Table 1 shows the results of estimating the PHMM parameters using the EM algorithm for each hidden situation. In the 2 hidden situations, the average and initial probability of occurrence in 1 month is the 1st hidden situation with λ = 17 and δ = 0.002 and the 2 hidden conditions with λ = 26.33 and δ = 0.001.

In the 3 hidden situations, the average and initial probability of their occurrence are 1 hidden situation with λ = 17.27 and δ = 0.0019, 2 hidden situations with λ = 1.50 and δ = 0.0006 and also the absence hidden 3 deings λ = 1.50 and δ = 0.001.

In the 4 hidden situations, the average and initial probability of their occurrence is obtained, namely hidden situation 1 with λ = 16.42 and δ = 0.0015, hidden situation 2 with λ = 24 and δ = 0, hidden situation 3 de light λ = 29 and δ = 0 as well as the hidden existence of 4 with λ = 21 and δ = 0

Table 1. PHMM Parameter Estimation Results with the EM Algorithm in Each Hidden State

Model	I	AIC	λ	δ
m=2	1	256.977	17.79	0.002
	2		26.33	0.001
m =3	1		17.27	0.0019
	2		1.50	0.0006
	3	251.980	1.50	0.001
m = 4	1		16.42	0.0015
	2		24	0
	3	315.238	29	0
	4		21	

After obtaining the results of the parameter parameters, the next step is to determine the best estimation model of the number of Hepatitis B oil events by comparing the AIC values, where the best AIC value is the best estimation model. Based on the table above, it can be seen that the real AIC value exists when 3 hidden conditions are given, namely 251,980, so it can be concluded that the model with 3 hidden conditions (m = 3) is the best model compared to those with m = 1 and m = 4. Here are the estimated results The best parameters of PHMM, namely a model with 4 hidden conditions:

λ = (17.27; 1.50; 1.50)

δ = (0.0019; 0.0006; 0.001)

$E(X_t) = \delta_1\lambda_1 + \delta_2\lambda_2 + \delta_3\lambda_3$

$E(X_t)$ = (0.0019 × 17.27) + (0.0006 × 1.50) + (0.001 × 11.50)

$E(X_t)$ = 0.03 + 0.0009 + 0.0015

$E(X_t)$ = 0.0324

So it can be concluded that of the three estimation models for the number of Hepatitis B patients in the Haji Hospital, the model with 4 hidden events is the estimation model for the number of Hepatitis B patients that occurs a lot 0.0324 = 0 events within 1 month. In this case, the average predicted result obtained by the hidden condition mode is 3 in line with the original condition in the previous months.

Meanwhile, I Examined the implementation of the Poisson Hidden Markov Model & Expectation-Maximization Algorithm for estimating earthquake events in Indonesia [22]. The results obtained from this research are the smallest AIC value from the Poisson Hidden Markov Model estimation results using the Expectation Maximization Algorithm, which was obtained (AIC = 1101.559). Hence, the Poisson Hidden Markov model with 3 hidden states is the optimum model for parameter estimation. Of the selected models, the best model with three hidden conditions is the best model for estimating the number of earthquakes, and the best estimate is the parameter estimate that the average number of earthquakes that occur in a month is 4.08437≈4 event values.

Then, the stochastic model analysis of Hepatitis B virus transmission was examined [23]. The spread of Hepatitis B virus infection uses a deterministic SIR model, where people who recover from acute infection have temporary immunity. However, this deterministic model uses a constant virus infection rate over time. This research also aims to construct a stochastic SIR model by dividing the infection rate into 2, namely the acute and chronic infection rates following the Wiener process.

Then, researched the Poisson Hidden Markov model and its application to motor vehicle insurance claims. This research uses the Akaike Information Criteria (AIC) to obtain the best model with a two-state Poisson Hidden Markov model with the smallest value. The two states are other people's evil actions and accidents. The average claim submission due to someone else's bad action is 11,499 claims per week, while the average claim submission due to an accident is 23,935. Suppose the number of claims experiences overdispersion, and the causes of claims that are not directly observed are assumed to form a Markov chain. In that case, the number of claims can be modelled using the Poisson Hidden Markov Model [24].

Furthermore, Examined Poisson Hidden Markov modelling to predict the number of accidents on the Jakarta-Cikampek Toll Road [25-26]. The prediction results from the number of accidents can be the basis for handling accident cases at toll road locations where accidents frequently occur, such as determining the number of tow trucks, vehicles carrying vehicles that have accidents, the number of ambulances on standby, the number of patrol cars, and the number of rescue cars. Namely, a car containing equipment used in serious accidents on toll roads. For this reason, a mixed Poisson model, namely the Poisson Hidden Markov model, is used to model the number of accidents on toll roads. The results obtained from this research are the AIC values obtained by the 2 state Poisson Hidden Markov model as the best model. Using the best model, the number of accidents was predicted for the next four periods, and a MAPE value of around 34% was obtained. The MAPE values obtained indicate that the prediction results for the number of accidents include decent forecasting but are not good enough.

Conclusion

The conclusions that can be drawn based on the results and previous discussions are as follows: the method for determining the average number of patients in patient B can use the PHMM (Poisson Hidden Markov Model) method with the EiM (Expectation-Maximization Algorithm) algorithm. The best model for the number of Hepatitis B patients in the Haji Hospital at this hospital is the model with 3 hidden circumstances with the parameter estimation value. The average number of Hepatitis B patients is 0.0324 in 1 month. The average predicted results obtained by the hidden condition model 3 align with the original conditions in the previous months. In this study, researchers only analyzed to determine the best model for the number of Hepatitis B patients based on testing criteria. Therefore, the author has suggestions for other researchers interested in this material. In future research, the best parameter estimation for the number of Hepatitis B patients can be done by considering the distribution of observation locations based on the distribution of Hepatitis B patients who occur most frequently. This research can be continued again to predict Hepatitis B patients in the next period.

References

[1] Widoyono, M. P. H. (2011). Penyakit Tropis Epidemiologi, Penularan, Pencegahan & Pemberantasan. *Jakarta: Penerbit Erlangga Indonesia*..

[2] Manaqib, M., Yuliawati, A. K., & Zulkifli, D. U. (2023). MATHEMATICAL MODEL OF MEASLES DISEASE SPREAD WITH TWO-DOSE VACCINATION AND TREATMENT. *Journal of Fundamental Mathematics and Applications (JFMA)*, *6*(2), 209-223.

[3] Siswanto, S., & Octavianur, E. (2020). Epidemiologi Penyakit Hepatitis.

[4] Mustofa, S., & Kurniawaty, E. (2013). Manajemen gangguan saluran cerna panduan bagi dokter umum. *Bandar Lampung: Aura Printing & Publishing. hlm*, *272*.

[5] Idaiani, S., Yunita, I., Tjandrarini, D. H., Indrawati, L., Darmayanti, I., Kusumawardani, N., & Mubasyiroh, R. (2019). Prevalensi Psikosis di Indonesia berdasarkan Riset Kesehatan Dasar 2018. *Jurnal Penelitian Dan Pengembangan Pelayanan Kesehatan*, 9-16.

[6] Sarvi, F., Nadali, A., Khodadost, M., Moghaddam, M. K., & Sadeghifar, M. (2017). Application of Poisson Hidden Markov model to predict number of PM2. 5 exceedance days in Tehran during 2016-2017. *Avicenna Journal of Environmental Health Engineering*, *4*(1), 58031-58031.

[7] Widita, H., Soemohardjo, S., Muttaqin, Z., Wiguna, P. A., Rhamdiani, S. O., Wijaya, M., & Gunawan, S. (2012). Detection of HBV-DNA and its correlation with the HBeAg/Anti-HBe serological status in HBsAg-positive patients. *The Indonesian Journal of Gastroenterology, Hepatology, and Digestive Endoscopy*, *13*(2), 86-89.

[8] Sugiyarto, H. (2021). Pengantar Proses Stokastik. *Yogyakarta: Universitas Ahmad Dahlan*.

[9] Suryandaru, R. (2015). Estimasi Model Terbaik Banyaknya Gempa Bumi Menggunakan Poisson Hidden Markov Models dan Algoritma EM. *Jakarta: Universitas Islam Negeri Syarif Hidayatullah*.

[10] Lekone, P. E., & Finkenstädt, B. F. (2006). Statistical inference in a stochastic epidemic SEIR model with control intervention: Ebola as a case study. *Biometrics*, *62*(4), 1170-1177.

[11] Aprilia, R., & Panjaitan, D. J. (2022). Pemodelan Matematika.

[12] Nåsell, I. (2002). Stochastic models of some endemic infections. *Mathematical biosciences*, *179*(1), 1-19.

[13] Ceres, K. M., Schukken, Y. H., & Gröhn, Y. T. (2020). Characterizing infectious disease progression through discrete states using hidden Markov models. *Plos one*, *15*(11), e0242683.

[14] Banks, J., & Chwif, L. (2011). Warnings about simulation. *Journal of simulation*, *5*(4), 279-291.

[15] Nurhasanah, N., Setiawaty, B., & Bukhari, F. (2016). Pemodelan Poisson Hidden Markov Untuk Prediksi Banyaknya Kecelakaan Di Jalan Tol Jakarta-cikampek. *MILANG Journal of Mathematics and Its Applications*, *15*(1), 55-64.

[16] Masuku, F. N., Langi, Y. A., & Mongi, C. (2018). Analisis rantai markov untuk memprediksi perpindahan konsumen maskapai penerbangan rute manado-jakarta. *Jurnal Ilmiah Sains*, 75-79.

[17] Rabiner, L. R. (1989). A tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE*, *77*(2), 257-286..

[18] Zucchini, W., & MacDonald, I. L. (2009). *Hidden Markov models for time series: an introduction using R*. Chapman and Hall/CRC.

[19] Fernandes, A. A. R. (2016). *Pemodelan Statistika Pada Analisis Reliabilitas Dan Survival*. Universitas Brawijaya Press.

[20] Alyarahma, N. N., Nur'Adilah, S., & Alfaris, M. R. (2022). Prediksi Frekuensi Probabilitas Curah Hujan di Jambi Menggunakan Rantai Markov serta Modul Tree Forecasting Berbantuan Software Qm-V5. *Multi Proximity: Jurnal Statistika*, *1*(2), 82-92.

[21] Hasan, I. (2008). Analisis Data Penelitian Dengan Statistik. Cetakan Ketiga. *PT. Bumi Aksara. Jakarta*.

[22] Fikri, M., Setiawaty, B., & PURNABA, I. G. (2016). Pendugaan parameter dan kekonvergenan penduga parameter model poisson hidden Markov. *MILANG Journal of Mathematics and Its Applications*, *15*(1), 45-54.

[23] Irawan, A. (2021). Implementasi Poisson Hidden Markov Models & Expectation Maximization Algorithm Untuk Estimasi Kejadian Gempa Bumi Di Indonesia (Studi Kasus: Badan Meteorologi dan Geofisika Daerah Istimewa Yogyakarta).

[24] Laila, A. N. (2021). *Analisis model stokastik penularan virus hepatitis B* (Doctoral dissertation, Universitas Islam Negeri Maulana Malik Ibrahim)..

[25] Prabandari, D. (2018). Model Poisson Hidden Markov Dan Penerapannya Pada Klaim Asuransi Kendaraan Bermotor.

[26] Nurhasanah, N., Setiawaty, B., & Bukhari, F. (2016). Pemodelan Poisson Hidden Markov Untuk Prediksi Banyaknya Kecelakaan Di Jalan Tol Jakarta-cikampek. *MILANG Journal of Mathematics and Its Applications*, *15*(1), 55-64.