BAREKENG: Journal of Mathematics and Its Applications
December 2023 Volume 17 Issue 4 Page 2337–2350
P-ISSN: 1978-7227 E-ISSN: 2615-3017
https://doi.org/10.30598/barekengvol17iss4pp2337-2350

THE GRADUATION OF TRANSITION INTENSITIES FROM SEMI-MARKOV PROCESSES TO PREMIUM PRICING

Faihatuz Zuhairoh[1*], **Dedi Rosadi**[2], **Adhitya Ronnie Effendie**[3]

[1]*Study Program of Mathematics Education, STKIP YPUP Makassar*
Andi Tonro Street No. 17 Makassar, 90223, Indonesia

[2,3]*Department of Mathematics, Faculty of Mathematics and Natural Sciences, Gadjah Mada University*
North Sekip Bulaksumur No 21,Yogyakarta 55281, Indonesia

*Corresponding author's e-mail: * **fzuhairoh@stkip.ypup.ac.id***

ABSTRACT

One of the important assumptions of the premium pricing of a health insurance product is the probability for someone suffers from a certain disease. In this paper, the disability income model is applied to a company using two covariates, namely age and sex. The purpose is to find out the magnitude of the probability for employees to experience disabilities due to work, a multi-state model can be used with semi-Markov assuming. There are several approaches to complete the multi-state model, one of which is the transition intensity approach. The intensity of the transition in this paper is estimated using the maximum likelihood approach, which will produce a crude estimate. Afterwards, the graduation process is performed on a crude estimate to obtain a finer shape of the transition intensity function with the Generalized Linear Model (GLM). The intensity of the transition from the graduation results is used to form transition probabilities which are eventually used as one of the assumptions in premium pricing.

Article History:

Received: 9th August 2023
Revised: 18th October 2023
Accepted: 17th November 2023

Keywords:

Generalized Linear Model;
Graduation Process;
Multi-State;
Semi-Markov;
Transition Probability.

How to cite this article:

F. Zuhairoh, D. Rosadi, and A. R. Effendie., "THE GRADUATION OF TRANSITION INTENSITIES FROM SEMI-MARKOV PROCESSES TO PREMIUM PRICING," *BAREKENG: J. Math. & App.*, vol. 17, iss. 4, pp. 2337-2350, December, 2023.

Journal homepage: *https://ojs3.unpatti.ac.id/index.php/barekeng/*
Journal e-mail: *barekeng.math@yahoo.com; barekeng.journal@mail.unpatti.ac.id*

Research Article · Open Access

1. INTRODUCTION

Multi-state models have been developed in actuarial science, one of which is in determining premium pricing. This paper explains the multi-state model of disability income using the semi-Markov assumption. Disability income insurance is divided into three states namely, healthy, disabled, and dead. Premiums are paid for the insured person and we will get benefits during disability and one-time benefits when death occurs.

In determining the premium of an insurance product, an actuary must pay attention to standard assumptions in actuarial science because the more assumptions that are violated, the greater the possibility of calculation errors. This resulted in both parties, the insured and the insurer experiencing losses. One of the essential assumptions in pricing a health insurance product is the chance of a person experiencing a specific disease or suffering from an infection. A person's health condition can be divided into several states, such as health, illness, disability, or death state.

The fundamental difference with the permanent disability model is that there is a healing period so that people who are in a state of illness will recover from the illness or even in one insurance period a person can experience pain several times until the person dies. According to our understanding of the model, this suggests that there could be multiple periods of illness preceding death, with healthy (premium-paying) intervals in between **[1]**. Health insurance pricing has also been developed based on prevalence rates **[2]**.

The multi-state model in health insurance is divided into several finite states, for example healthy, disabled and dead. There are several approaches in solving multi-state model problems, including **[3]**, namely by utilizing a deck cement table, then in subsequent studies providing an alternative solution using the Markov model assumptions, in this case, it is assumed that between each status always has a transition intensity including **[4]**. The intensity of the transition can also be used to predict future events. Several studies on forecasting with several methods were carried out by **[5]**, **[6]** using the Richards method, **[7]** using Hybrid Vector Autoregression Feedforward NN with Genetic Algorithm Model, and **[8]** using the time-temperature superposition model and SRL model.

In the Markov model, it is assumed that the intensity and probability of the transition at the time depend only on the current state **[9]**, **[10]**. Transition probability plays an important role in determining insurance premiums, both outpatient and long-term care **[11]**. In addition to pricing premiums, we developed a Markov model for modeling the spread of infectious diseases **[12]**, **[13]**. Although the current Markov model has been widely used in various fields, including health insurance, it is felt that this model still has many shortcomings. For example, in the discrete-time context, the transition time of the Markov model is not stochastic, whereas, in continuous time, the intensity of the transition is constant. Of course, this simplifies a really huge problem, so it is appropriate that the Markov model still needs to be refined. One way to do this is to add additional assumptions to the Markov model, namely, that apart from depending on the current state, the intensity and probability of the transition are also influenced by the time spent by the system in a particular state as written by **[14]**. As a result, the Markov property will be brought to a new definition, namely the semi-Markov process.

The semi-Markov process is one type of stochastic process. It can be analogized as a merger between two types of stochastic methods: Markov chains and renewal processes. The semi-Markov process does not have the limitations that the Markov model lacks, where the fact shows that the distribution function of the transition intensity in a continuous-time semi-Markov process can be of any form. In contrast, in the case of discrete-time, the transition time of a semi-Markov process is stochastic, which makes the discrete-time semi-Markov process closer to reality.

The main contribution of this paper is how to overcome the rough estimation generated by the semi-Markov model, one of which is using graduation. Previously it had been developed only with the Markov model by **[15]**. A crude estimate of the transition intensities in this paper will be estimated with the maximum likelihood approach because it can be graduated to obtain a smoothly transition intensity. The graduation process is done with a Generalized Linear Model (GLM) **[16]**, **[17]**. The integrated transition intensity estimator is used to form transition probabilities associated with the Kolmogorov Backward and Forward differential equations **[18]**. This estimator is ultimately used as one of the assumptions in determining health insurance premiums.

This article is structured as follows. We present a theoretical framework of the multi-state model, semi-Markov process, Chapman-Kolmogorov equation, maximum likelihood method, and GLM in section 2. In Section 3, we describe the estimation procedure of the multi-state model with semi-Markov assumptions, crude transition intensity, graduation process of transition intensity using GLM, transition probability, and premium pricing. Afterward, in Section 4, we applied the multi-state model using data retrieved from www.soa.org to calculate premiums based on covariates of age and sex. Finally, in the end, namely Section 5, we give a conclusion regarding the formula to calculate the premium value of the multi-state model assuming semi-Markov with a graduation process to produce better results.

2. RESEARCH METHODS

2.1 Multi-state Model

S, T is a multi-state model with $S = 1,2,\cdots,m$ explain space state like finite set and T is a subset of the set of pairs which is $T \subseteq \{(i,j)|i \neq j;\ i,j \in S\}$ [3]. The characteristic of a multi-state process is affected by the transition probabilities between state i and j following this,

$$ {}_t p_x^{ij} = \Pr\{Y_{x+t} = j|Y_x = i\} \tag{1}$$

so that ${}_t p_x^{ij}$ is the probability that a life aged x in state i is in state j at age $x+t$, where j may be equal to i. In addition, a multi-state process can also be influenced by the intensity of the transition.

$$\mu_x^{ij} = s \lim_{h\to 0} \frac{{}_h p_x^{ij}}{h}, \qquad i \neq j \tag{2}$$

It shows the transition intensity to the direct transfer of risk state, if known in advance in state i, both p^{ij} and μ^{ij} depend on the history and processes owned.

2.2 Semi-Markov Process

A semi-Markov process is a process that makes a transition from one state to another, as in a Markov process. However, the amount of time spent in each state before transitioning to the next is any random variable that depends on the following form of a new process.

Suppose there is a stochastic process $\{S_t, H_t; t \geq 0\}$ where S_t is a random event at time t, and H_t is the sojourn time in state S_t until time t since the last transition to that state; formally [3]:

$$H_t = \max\{\tau: \tau \leq t, S_{t-h} = S_t \forall\, h \in [0,\tau]\} \tag{3}$$

Assume for the moment that $\{S_t, H_t; t \geq 0\}$ is a time-continuous. The present state, $S_t = i$, and the amount of time since the most recent transition into that state, $H_t = v$, are the most recent pieces of information that can determine the conditional probabilities for the future of the process at time t. Because of this, the aforementioned conditional probabilities are independent of any knowledge of the process's course before time t. Consequently, we employ the subsequent transition probabilities [19].

$$P^{ij}(t,u,v,w) = \Pr(S_u = j \wedge H_u \leq w \mid S_t = i \wedge H_t = v) \tag{4}$$

for $0 \leq t < u$ and $v, w \geq 0$.

2.3 Chapman-Kolmogorov Equations

The Chapman-Kolmogorov equation states that the path that starts at state i at time t goes to status j at time u through several states k continuously at any time w. The Chapman-Kolmogorov equation can be written as follows [3]

$$P^{ij}(t,u) = \sum_{k\in S} P^{ik}(t,w)P^{kj}(w,u), \qquad (t \leq w \leq u) \tag{5}$$

Proof. Using the Markov property, we have,

$$P^{ij}(t,u) \quad = \quad \Pr\{S_u = j|S_t = i\}$$

$$= \sum_{k\in S} \Pr\{S_u = j \wedge S_w = k | S_t = i\}$$
$$= \sum_{k\in S} \Pr\{S_w = k | S_t = i\} \Pr\{S_u = j | S(t) = i \wedge S_w = k\}$$
$$= \sum_{k\in S} \Pr\{S_w = k | S_t = i\} \Pr\{S_u = j | S_w = k\}$$
$$= \sum_{k\in S} P^{ik}(t,w) P^{kj}(w,u)$$

It is possible to derive two sets of differential equations for the transition probabilities. First, think about the forward differential equations for Kolmogorov:

$$\frac{d}{dt} P^{ij}(z,t) = \sum_{k=0,k\neq j}^{n} P^{ik}(z,t)\mu^{kj}(t) - P^{ij}(z,t)\mu^{j}(t) \tag{6}$$

and Kolmogorov backward differential equations:

$$\frac{d}{dz} P^{ij}(z,t) = P^{ij}(z,t)\mu^{i}(z) - \sum_{k=0,k\neq i}^{n} P^{kj}(z,t)\mu^{ik}(z) \tag{7}$$

3. RESULTS AND DISCUSSION

3.1 The Estimation Procedure Multi-States Model for Disability Income

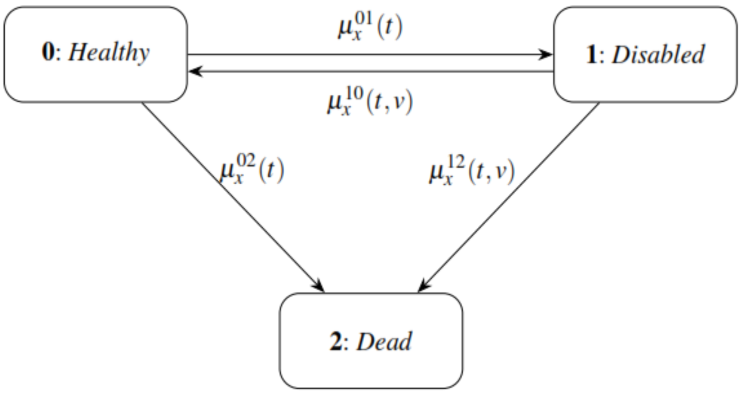


Figure 1. Multi-State Model For Disability Income

There are three states depicted in **Figure 1**: healthy, disabled, and dead. The arrows represent the transitions between the three states. Consider the intensity of the transition to be the following:

$\mu_x^{01}(t)$: Intensity of transition from healthy to disabled state;

$\mu_x^{10}(t,v)$: Intensity of transition from disabled to healthy state;

$\mu_x^{02}(t)$: Intensity of transition from health to death;

$\mu_x^{12}(t,v)$: Intensity of transition from disability to death;

It has been explained previously in **Equation (3)** that $H_t = v$ is the sojourn time in one state until the individual moves to another state. So, for the transition in **Figure 1** the sojourn time is only considered when someone is disabled. How long a person is disabled until they recover or until the person dies.

The set of simultaneous differential equations is as follows:

1. $\frac{d}{dt} P_x^{00}(z,t) = \int_z^t \left(P_x^{00}(z,r)\mu_x^{01}(r)P_x^{11}(r,t,0)\mu_x^{10}(t,t-r)\right)dr - P_x^{00}(z,t)[\mu_x^{01}(t) + \mu_x^{02}(t)]$

2. $\frac{d}{dt} P_x^{01}(z,t) = P_x^{00}\mu_x^{01}(t) - \int_z^t \left(P_x^{00}(z,r)\mu_x^{01}(r)P_x^{11}(r,t,0)\mu_x^{10}(t,t-r)\right)dr -$
$\int_z^t \left(P_x^{00}(z,r)\mu_x^{01}(r)P_x^{11}(r,t,0)\mu_x^{12}(t,t-r)\right)dr$

3. $\frac{d}{dt} P_x^{02}(z,t) = P_x^{00}\mu_x^{02}(t) + \int_z^t \left(P_x^{00}(z,r)\mu_x^{01}(r)P_x^{11}(r,t,0)\mu_x^{12}(t,t-r)\right)dr$

4. $\frac{d}{dt}P_x^{10}(z,t) = \int_z^t \left(P_x^{11}(z,r)\mu_x^{10}(r)P_x^{00}(r,t,0)\mu_x^{01}(t,t-r)\right)dr - P_x^{10}(z,t)[\mu_x^{01}(t)+\mu_x^{02}(t)]$

5. $\frac{d}{dt}P_x^{11}(z,t) = P_x^{10}\mu_x^{01}(t) - \int_z^t \left(P_x^{11}(z,r)\mu_x^{10}(r)P_x^{00}(r,t,0)\mu_x^{01}(t,t-r)\right)dr - \int_z^t \left(P_x^{11}(z,r)\mu_x^{12}(r)P_x^{00}(r,t,0)\mu_x^{02}(t,t-r)\right)dr$

6. $\frac{d}{dt}P_x^{00}(z,t) = \int_z^t \left(P_x^{11}(z,r)\mu_x^{12}(r)P_x^{00}(r,t,0)\mu_x^{02}(t,t-r)\right)dr - P_x^{10}(z,t)\mu_x^{02}(t)$

3.2 Crude Estimates of the Transition Intensities

The next step is to estimate the parameters that aim to find a function of each transition probability. The estimation method uses maximum likelihood estimation. The principle of this method is to maximize the parameter estimator so that the value will be close to the parameter.

It is assumed that during the observation, what can be observed is the time and transitions made by an individual. For age intervals, we use $(x, x+1)$ without losing generality, to assume that the transition intensity is a constant, $\mu_x^{01}, \mu_x^{02}, \mu_x^{10}, \mu_x^{12}$.

Suppose T_{im} is a continuous random variable, the time spent by individual m in state i before moving to another state (sojourn time in state i) with the hazard function $\mu(t_{im}) = \mu$. The relationship between the hazard function and the density function is obtained.

$$\begin{aligned} f(t_{im}) &= \mu(t_{im})e^{-\int_0^{t_{im}}\mu(s)ds} \\ &= \mu e^{-\int_0^{t_{im}}\mu ds} \\ &= \mu e^{-\mu\, t_{im}} \end{aligned}$$

This form is the probability density function of the Exponential distribution. So, if the transition intensity is constant, T_{im} has an Exponential distribution.

Furthermore, according to the Poisson process, the random variable N_{ijm}, namely the number of transitions from state i to state j made by individual m, has a Poisson distribution with an average $t_{im}\lambda_x^{ij}$, written.

$N_{ijm}\sim$ Poisson $(t_{im}\mu_x^{ij})$

with

$$Pr(N_{ijm} = n_{ijm}) = f_{N_{ijm}}(t_i) = \frac{e^{-\mu_x^{ij}t_{im}}(\mu_x^{ij}t_{im})^{n_{ijm}}}{n_{ijm}!}. \tag{8}$$

Then the likelihood function is

$$L(\mu_x^{ij}) = L(\mu_x^{ij}) = \prod_{i\in S}\prod_{j\in S}\left(\prod_{m=1}^{N} f(\mu_x^{ij}, t_i)\right) \tag{9}$$

for example,

n_{ij}: $\sum_{m=1}^{N} n_{ijm}$ the number of transitions from state i to state j made by all individuals in the age interval $(x, x+1)$

t_i : $\sum_{m=1}^{N} t_{im}$ the total sojourn time in state i for all individuals.

So that,

$$L(\mu_x^{ij}) = \prod_{i\in S}\prod_{j\in S}\frac{e^{-\mu_x^{ij}t_i}(\mu_x^{ij}t_i)^{n_{ij}}}{n_{ij}!}$$

If the part that does not contain μ_x^{ij} is ignored, then

$$L(\mu_x^{ij}) = \prod_{i\in S}\prod_{j\in S} e^{-\lambda_x^{ij}t_i}(\lambda_x^{ij})^{n_{ij}}$$

and the log-likelihood function is

$$\log L\left(\mu_x^{ij}\right) = \prod_{j\in \boldsymbol{s}}\prod_{j\in \boldsymbol{s}} \log\left(e^{-\mu_x^{ij}t_i}\left(\mu_x^{ij}\right)^{n_{ij}}\right)$$
$$= \sum_{j\in \boldsymbol{s}}\sum_{j\in \boldsymbol{s}} -\mu_x^{ij}t_i + n_{ij}\log\left(\mu_x^{ij}\right) \quad (10)$$

The derivative of **Equation (10)** is

$$\frac{d\,log\,L\left(\mu_x^{ij}\right)}{d\mu_x^{ij}} = -t_i + \frac{n_{ij}}{\mu_x^{ij}} \quad (11)$$

If the value of **Equation (11)** is equal to zero then $\mu_x^{ij} = \frac{n_{ij}}{t_i}$. So that the estimator for μ_x^{ij} is obtained.

$$\hat{\mu}_x^{ij} = \frac{n_{ij}}{t_i} \quad (12)$$

3.3 The Graduation of Transition Intensities with Generalized Linear Model

The transition intensity estimator obtained with the maximum likelihood approach is still a crude estimate where only a certain age group is presented when using the age covariate. To smooth it out, a graduation process is carried out. Through the graduation process we can obtain estimates of all age groups without exception. Graduation guarantees that the resulting survival or multi-status model displays smoother and more representative desired traits.

The data used for the graduation process for each transition intensity consists of a set of (n_u, t_u) where n_u is the number of transitions corresponding to the central exposures t_u which is defined for each unit of u. Unit $u \equiv (x_1, x_2, x_3, \ldots)$, which is a cross classification pair of covariates. Graduation is constructed using a Generalized Linear Model (GLM) based on the response variable N_{ij}^u which is distributed with Poisson.

The variables that will be used are as follows: response variable, many transitions between states and predictor variables, age (x) at the time of disabled (in years) and gender (z). $u \equiv (x, z)$ covariate couples of age and gender. Based on the GLM framework, the function for each μ^{ij} is connected by a function g to a linear predictor.

$$g\left(\mu_{ij}^u\right) = \eta_u \quad (13)$$

where g is a bijective and differentiable function, so the inverse is there,

$$\mu_{ij}^u = g^{-1}(\eta_u) \text{ and } \eta_u = \sum_k x_{ku}\beta_k \quad (14)$$

where x_k is the covariate structure and β_k is an unknown regression parameter.

The value of β_k is estimated based on the assumption that the response variable, the number of transitions has a Poisson distribution, $N_{ij}^u \sim \text{Poi}\left(\mu_{ij}^u t_u\right)$ is independent for all $u \equiv (x, z)$ with the mean and variance as follows.

$$\lambda_u = E[N_u] = t_u\mu_{ij}^u, \qquad \text{Var}(N_u) = \lambda_u = \mu_{ij}^u t_u \quad (15)$$

If expressed in the form $\boldsymbol{\lambda} = (\lambda_u)$, the log-likehood function of the response variable N_{ij}^u is

$$l(\boldsymbol{n}.\boldsymbol{\lambda}) = \sum_u \left\{ln\frac{1}{n_u!} + n_u\,ln(\lambda_u) - \lambda_u\right\} \quad (16)$$

where $\mathbf{n} = (n_u)$. The parameter β_k enters the log-likehood by substitution through the following relationship.

$$\lambda_u = t_u\mu_{ij}^u = t_u g^{-1}(\eta_u) = t_u g^{-1}\left(\sum_k x_{ku}\beta_k\right) \quad (17)$$

and its value, $\hat{\beta}_k$ is estimated by maximizing **Equation (16)**.

Next, after the $\hat{\beta}_k$ value is obtained, the graduation process continues. The link function for the Poisson distribution is log-*link*, then

$$\eta = log(\lambda_u) = log(t_u) + log\left(\mu_{ij}^u\right) = log(t_u) + \sum_k x_{ku}\hat{\beta}_k \quad (18)$$

$log(t_u)$ is the offset, namely an additional variable with a known regression coefficient of +1. So the transition intensity, μ_{ij}^u is related to the covariate through the following relationship.

$$log(\lambda_{ij}^u) = \sum_k x_{ku}\hat{\beta}_k \Leftrightarrow \lambda_{ij}^u = \exp\left(\sum_k x_{ku}\hat{\beta}_k\right) \tag{19}$$

Therefore, the graduation formula for transition intensity is

$$\hat{\mu}_{ij}^u = \exp\left(\sum_k x_{ku}\hat{\beta}_k\right) \tag{20}$$

3.4 Transition Probability

The transition probabilities are calculated using constant intensity piece by piece because we previously assumed that the transition intensity is constant. First, for all i, j in $\mathbf{S}$, let $P_{ij}^{(m)}(z)$ denote the transition probability function associated with ll-time intervals contained in (t_{m-1}, t_m), for $m = 1,2, \ldots$ with $t_0 = 0$. Let us assume $P_{ij}(t, t+z) = P_{ij}^{(m)}(z)$ if $t_{m-1} < t < t+z \le t_m$ for any t, let m_t denote the time interval which contains t.

Let $\mu_{ij}(t) = \mu_{ij}^{(m)}$ if $t_{m-1} < t \le t_m$ for $m = 1,2, \ldots$, then $\mathbf{Q}^{(m)} = \left|\mu_{ij}^{(m)}\right|$ and $\mathbf{P}^{(m)}(z) = \left|P_{ij}^{(m)}(z)\right|$. To find a solution, we can determine $\mathbf{A}^{(m)}, \mathbf{D}^{(m)}, \mathbf{C}^{(m)}$, and hence the transition probability matrix $P^{(m)}(z)$ for any $z, z_{m-1} < z \le z_m$, and $m = 1,2, \ldots$

$$\mathbf{P}^{(m)}(z) = \mathbf{A}^{(m)} diag\left(e^{d_1^{(m)}z}, \ldots, e^{d_N^{(m)}z}\right)\mathbf{C}^{(m)} \tag{21}$$

then, via the Chapman-Kolmogorov equation, we can write $P_{ij}(t,u) = \sum_{h\in\mathbf{S}} P_{ih}\left(t, t_{m_u-1}\right)P_{hj}^{(m_u)}\left(u - t_{m_u-1}\right)$

$$P_{ij}(t,u) = \sum_{h=1}^{N} b_{ih}^{(m_u)} c_{hj}^{(m_u)} e^{d_h^{(m_u)(u-t_{m_u-1})}} \tag{22}$$

3.5 Premium Pricing

One of the main objectives in actuarial valuation in the insurance world is to determine the amount of the premium. Premiums can be interpreted as a sum of money paid within a certain period by the insured to the insurance party in accordance with the agreement contained in the policy. In this paper, the calculation of premiums is only based on net premiums, so other operational costs are not counted. Then in the calculation used n year term discrete life insurance valuation model.

In the outpatient health insurance application discussed in this research, a patient only needs to pay the premium once at the start of the policy. Furthermore, patients do not need to pay again if they undergo a health examination within the policy period. The outpatient health insurance model with a multi-state model developed in this research, then carried out the following things.

1. ${}_kp_x$ is replaced by the probability that an individual aged *x* at state *i* will be at state *j* at age *x* + *t*. The model with 3 states as in **Figure 1** is replaced by
   a. probability that someone aged *x*, healthy at the beginning of the period will remain healthy during *k* periods, $p_{11}(x, x+t_k) = p_x^{11}(0, t_k)$
   b. the probability that someone aged *x* suffers from a disability at the beginning of the period will remain disabled during *k* periods, $p_{22}(x, x+t_k) = p_x^{22}(0, t_k)$
2. q_{x+k} is replaced by the transition intensity of someone aged *x* from healthy status in the *k*-th period, to disability or death in the next 1 period, which is denoted by, $\mu_x^{ij}(t_{k+1})$.

Therefore, the premium formula used is:

$$APV = b_1 \sum_{t=1}^{n} v^t (p_x^{11}(0,t)\mu_{x+t}^{01}) + b_2 \sum_{t=1}^{n} v^t (p_x^{22}(0,t)\mu_{x+t}^{02} + p_x^{11}(0,t)\mu_{x+t}^{12}) \tag{23}$$

with b_1 = benefits if disabled

b_2 = benefits in case of death

$v = \frac{1}{1+i}$, i = interest rate

The data in this paper is taken from www.soa.org by making some simplifications to fit as an application of the disability income model. Data used ranging from ages 20 to 70 years who take out insurance. **Table 1** shows a summary of data on the number of changes made by insurance participants along with the total sojourn time in healthy and disabled states. The transition from **Figure 1** that is not in the data is the transition from healthy to dead.

Based on gender, the transitions that exist for the male gender are the transition from healthy to disabled, the transition from disabled to healthy, and the transition from disabled to dead, while for the female gender the transitions that exist are only the transition from healthy to disabled and transition from disability to health.

Table 1. The Total Sojourn Time and the Number of Transitions in Each State is Based on Age

Age	n_{01}	n_{02}	n_{10}	n_{12}	t_1	t_2
⋮	⋮	⋮	⋮	⋮	⋮	⋮
34	3	0	2	1	15	15
35	1	0	1	0	6	6
36	2	0	1	1	9	9
37	0	0	0	0	0	0
38	3	0	2	1	18	24
39	4	0	3	1	33	18
40	1	0	1	0	6	6
41	3	0	3	0	48	12
⋮	⋮	⋮	⋮	⋮	⋮	⋮

***Data source**: www.soa.org*

Based on the values in **Table 1**, using **Equation (12)** an estimator for the transition intensity is obtained which is given in **Table 2**. Next, a plot of the estimated transition intensity is presented in **Figure 2**. The estimated value of transition intensity obtained in **Table 2** is still a crude estimator.

Table 2. Transition Intensity Estimator

Age	μ_{01}^{x}	μ_{10}^{x}	μ_{12}^{x}
⋮	⋮	⋮	⋮
34	0.20000	0.13333	0.06667
35	0.16667	0.16667	0.00000
36	0.22222	0.11111	0.11111
37	0.00000	0.00000	0.00000
38	0.16667	0.08333	0.04167
39	0.12121	0.16667	0.05556
40	0.16667	0.16667	0.00000
41	0.06250	0.25000	0.00000
⋮	⋮	⋮	⋮

Therefore, the transition intensity μ_{ij} is graduated by the GLM method based on Poisson modeling assumptions. The variables that will be used are as follows.

1. Response variable : N_{ij}^{u}, number of transitions from state *i* to state *j*
2. Predictor variables : patient age (x) at first examination (in years) and gender (z). $u \equiv (x, z)$, pair of age and gender covariates.

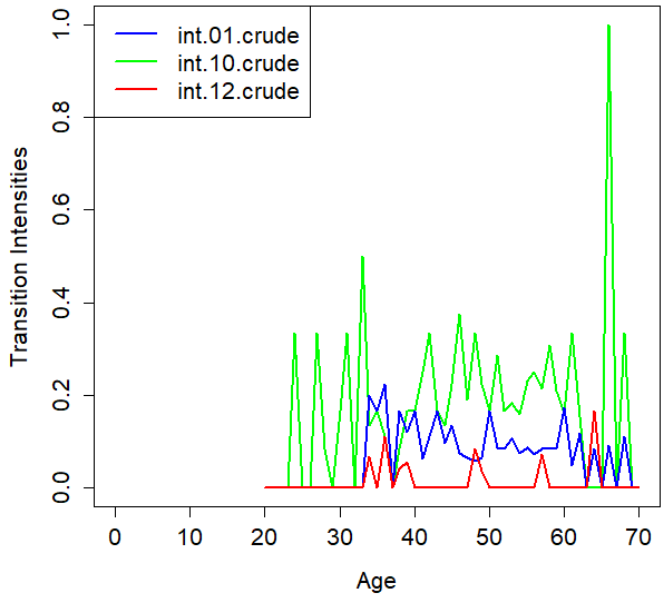


Figure 2. The Crude Transition Intensity Estimation Plot

Parameter estimation for GLM was carried out using R software with the glm package, the results in **Table 3** were obtained. **Table 3** shows parameter estimates for the transition from a healthy to disabled state, **Table 4** shows parameter estimates for the transition from a disabled to healthy state, and **Table 5** shows parameter estimates for the transition from a disabled to dead state, where for these parameter estimates there is only data for female because there is no transition from disability to death for male.

Table 3. Parameter Estimation for μ_{01}- Graduation

	Gender	Estimate	Std. Error	z-value	Pr(> |z|)	
(Intercept)	M	-1.76237	0.732826	-2.405	0.0162	*
	F	-0.362714	0.570452	0.636	0.525	
x	M	0.009364	0.015151	0.618	0.5365	
	F	-0.005703	0.012256	-0.465	0.642	
z	M	0.113014	0.018225	6.201	5.60E-10	***
	F	0.064482	0.009456	6.819	9.16e-12	***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 4. Parameter Estimation for μ_{10}- Graduation

	Gender	Estimate	Std. Error	z-value	Pr(> |z|)	
(Intercept)	M	-1.65077	0.66624	-2.478	0.0132	*
	F	-1.4417	0.670964	-2.149	0.0317	*
x	M	0.01436	0.01372	1.047	0.2952	
	F	0.006655	0.013463	0.494	0.6211	
z	M	0.10796	0.01777	6.077	1.22E-09	***
	F	0.153031	0.023359	6.551	5.70E-11	***

Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Table 5. Parameter Estimation for μ_{12}- Graduation

	Gender	Estimate	Std. Error	z-value	Pr(> |z|)	
(Intercept)	M	-	-	-	-	
	F	-2.61392	1.408836	-1.855	0.0635	.
x	M	-	-	-	-	
	F	0.001484	0.028893	0.051	0.959	
z	M	-	-	-	-	
	F	0.120776	0.05664	2.132	0.033	*

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

From the estimation results using R software, with a significance level of 5%, there are several variables that are not significant in the model, but these variables will be maintained in the model, so that based on Equation (20) the graduation formula for male for each transition intensity is obtained as following.

$$\mu_{01}^{(x,z)} = \exp(-1{,}7624 + 0{,}0094x + 0{,}1130z)$$

$$\mu_{10}^{(x,z)} = \exp(-1{,}6508 + 0{,}0143x + 0{,}1080z)$$

Meanwhile for female it is as follows.

$$\mu_{01}^{(x,z)} = \exp(-0{,}3627 - 0{,}0057x + 0{,}0645z)$$

$$\mu_{10}^{(x,z)} = \exp(-1{,}4417 + 0{,}0066x + 0{,}1530z)$$

$$\mu_{12}^{(x,z)} = \exp(-2{,}6139 + 0{,}0015x + 0{,}1208z)$$

Figure 3 shows that the graduation results provide a smoother estimate of the transition intensity. Therefore, the transition intensity from the GLM results is used to find the transition probability. The intensity of the coarse transition still shows substantial fluctuations in each age range, following the opinion of [3] that graduation ensures that the resulting multi-state model displays smoother and represents desirable properties for practical use, for example, the calculation of premiums and reserves. The obtained transition intensity estimator is calculated based on sample data from a large population, which likely contains some random fluctuations. Assuming that the intensity $\hat{\mu}_{ij}^{x}$ is the actual intensity, which is independent of each other, the crude estimates are the final estimator value. However, the correct form of intensity is that each is closely related to the other, so the next step is to carry out a graduation of crude estimates to obtain a more refined intensity. This is done systematically by revising crude estimates. Figure 3 shows a comparison plot of the crude transition intensity estimates and the graduation results.

Based on the parameter estimates above, the transition intensity values obtained from the graduation results based on gender are in Table 6 and Table 7.

Table 6. Transition Probabilities for Male

Age	P_{00}	P_{01}	P_{02}	P_{10}	P_{11}	P_{12}	P_{20}	P_{21}	P_{22}
20	0.95106	0.04894	0.0000	0.21962	0.78038	0.0000	0.0000	0.0000	1.0000
21	0.94953	0.05048	0.0000	0.22220	0.77780	0.0000	0.0000	0.0000	1.0000
22	0.94798	0.05202	0.0000	0.22480	0.77520	0.0000	0.0000	0.0000	1.0000
23	0.94643	0.05357	0.0000	0.22743	0.77257	0.0000	0.0000	0.0000	1.0000
24	0.94514	0.05486	0.0000	0.23740	0.76260	0.0000	0.0000	0.0000	1.0000
25	0.94332	0.05668	0.0000	0.23276	0.76724	0.0000	0.0000	0.0000	1.0000
26	0.94176	0.05824	0.0000	0.23546	0.76454	0.0000	0.0000	0.0000	1.0000
27	0.94048	0.05952	0.0000	0.24570	0.75430	0.0000	0.0000	0.0000	1.0000
28	0.93869	0.06131	0.0000	0.24280	0.75720	0.0000	0.0000	0.0000	1.0000
29	0.93704	0.06296	0.0000	0.24369	0.75631	0.0000	0.0000	0.0000	1.0000

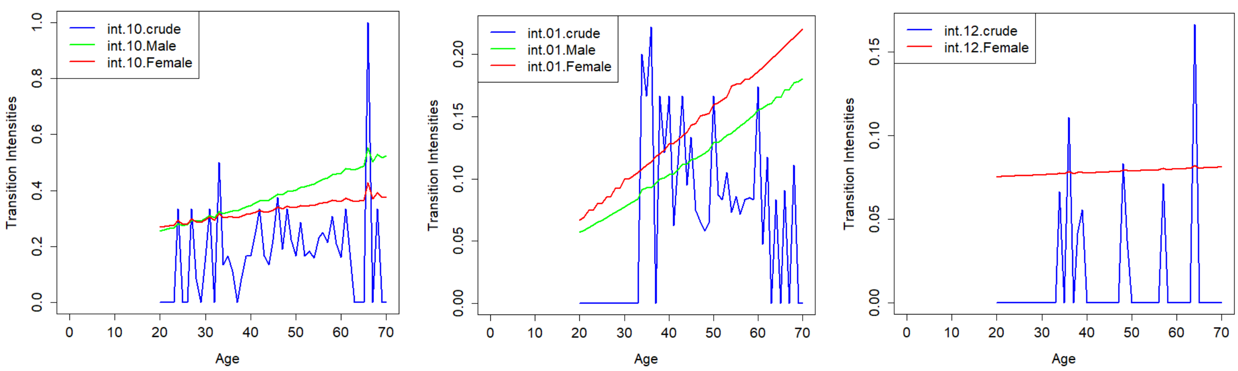
Age	P_{00}	P_{01}	P_{02}	P_{10}	P_{11}	P_{12}	P_{20}	P_{21}	P_{22}
30	0.93562	0.06438	0.0000	0.25033	0.74967	0.0000	0.0000	0.0000	1.0000
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 6 shows the probability of transition between states for the male gender, where it is known that there is no transition from the healthy to the dead state or from people with disabilities to the dead state. The resulting transition probability value increases with age in the transition from a healthy to a disabled state and from a disabled to a healthy state. In contrast, the transition probability value decreases with increasing age in self-transitions in a healthy and disabled state. The transition probability of remaining in the death state is equal to 1 because it is an absorption state, meaning that anyone who has entered a specific state will not be able to leave that state.

Table 7. Transition Probabilities for Female

Age	P_{00}	P_{01}	P_{02}	P_{10}	P_{11}	P_{12}	P_{20}	P_{21}	P_{22}
20	0.94313	0.05467	0.00220	0.22123	0.71479	0.06398	0.0000	0.0000	1.0000
21	0.94098	0.05673	0.00229	0.22222	0.71375	0.06403	0.0000	0.0000	1.0000
22	0.93661	0.06092	0.00247	0.22293	0.71229	0.06408	0.0000	0.0000	1.0000
23	0.93610	0.06141	0.00249	0.23414	0.71174	0.06412	0.0000	0.0000	1.0000
24	0.93274	0.06463	0.00264	0.23511	0.70114	0.06475	0.0000	0.0000	1.0000
25	0.93228	0.06506	0.00265	0.23621	0.60958	0.06481	0.0000	0.0000	1.0000
26	0.93808	0.06910	0.00283	0.23694	0.60880	0.06496	0.0000	0.0000	1.0000
27	0.92850	0.06868	0.00286	0.23853	0.60760	0.06587	0.0000	0.0000	1.0000
28	0.92299	0.07397	0.00304	0.23938	0.60437	0.06525	0.0000	0.0000	1.0000
29	0.92246	0.07448	0.00307	0.24008	0.60353	0.06639	0.0000	0.0000	1.0000
30	0.91739	0.07933	0.00328	0.24580	0.60297	0.06523	0.0000	0.0000	1.0000
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮

The same thing can also be explained in Table 7, where the transition between states for the female gender increases with age in different states.



(a) (b) (c)

Figure 3. (a) Comparison of estimated crude transition intensity and graduation from disabled to healthy; (b) Comparison of estimated crude transition intensity and graduation from healthy to disabled; and (c) Comparison of estimated crude transition intensity and graduation from disabled to dead

After knowing the transition probability values in Table 6 and Table 7, the premium for each age will be calculated using Equation (23) with an SBI interest rate of 6%, then the premium table in Table 8 is obtained. In this example case, the benefit value when disabled is used. 5 million rupiah and benefits upon death of 7 million rupiah.

Table 8. Average Premium Prices for Age Groups

Age	Premium Price	
	Male	Female
≤ 30	315,846	758,150
31-40	426,118	888,349
41-50	526,528	1,008,465
51-60	631,739	1,127,991
> 60	740,034	1,238,489

Table 8 shows the average premium price for each age group, where the amount of premium that must be paid will increase with increasing age for both males and females. The premium paid for females is more significant than for males because, according to the data, the number of disabled patients is dominated by females.

4. CONCLUSIONS

The intensity of the transition is the rate of change in a person's state from one state to another in a unit of time, the value or form of its function is unknown, so it needs to be estimated using statistical data. One method for estimating transition intensity is the maximum likelihood approach that varies by age. The general form of transition intensity estimator is $\hat{\mu}_x^{ij} = \frac{n_{ij}}{t_i}$. The estimator obtained is still a crude estimate value, then graduated to obtain the transition intensity function with the GLM. The graduation formula for transition intensity is $\hat{\mu}_{ij}^{u} = \exp\left(\sum_k x_{ku}\hat{\beta}_k\right)$ which is different for each age and gender. Graduation ensures that the resulting multi-state model displays smoother and represents desirable properties for practical use, for example, the calculation of premiums and reserves. The obtained transition intensity estimator is calculated based on sample data from a large population, which likely contains some random fluctuations.

Each estimated transition intensity value will be utilized to calculate transition probabilities. The general form of transition probability estimator with the backward and forward Kolmogorov differential equations. These transition probabilities are used to calculate premium rates for disability income models. The premium is obtained by following the flow of term life insurance premiums and multiple decrements using $NSP = b_1 \sum_{t=1}^{n} v^t (p_x^{11}\mu_{x+t}^{01}) + b_2 \sum_{t=1}^{n} v^t (p_x^{22}\mu_{x+t}^{02} + p_x^{11}\mu_{x+t}^{12})$.

REFERENCES

[1] D. C. M. Dickson, M. R. Hardy, and H. R. Waters, *Actuarial Mathematics for Life Contingent Risks*. United States of America: Cambridge University Press, 2009.

[2] F. Baione and S. Levantesi, "A health insurance pricing model based on prevalence rates: Application to critical illness insurance," *Insur. Math. Econ.*, vol. 58, no. 1, pp. 174–184, 2014, doi: 10.1016/j.insmatheco.2014.07.005.

[3] S. Haberman and E. Pitacco, *Actuarial Models for Disability Insurance*. London, New York: CRC Press LLC, 1999.

[4] A. S. Anggraeni, A. Listiani, K. Alim, and A. R. Effendie, "Morbidity-mortality table construction for eleven chronical diseases (ECD) using constant force assumption," *J. Phys. Conf. Ser.*, vol. 1341, no. 062030, pp. 1–9, 2019, doi: 10.1088/1742-6596/1341/6/062030.

[5] F. Zuhairoh and D. Rosadi, "Real-time forecasting of the COVID-19 epidemic using the Richards model in South Sulawesi, Indonesia," *Indones. J. Sci. Technol.*, vol. 5, no. 3, pp. 456–462, 2020, doi: 10.17509/ijost.v5i3.26139.

[6] F. Zuhairoh and D. Rosadi, "Real-time prediction for multi-wave COVID-19 outbreaks," *Commun. Stat. Appl. Methods*, vol. 29, no. 5, pp. 499–512, 2022, doi: 10.29220/CSAM.2022.29.5.499.

[7] R. E. Caraka, R. C. Chen, H. Yasin, Suhartono, Y. Lee, and B. Pardamean, "Hybrid vector autoregression feedforward neural network with genetic algorithm model for forecasting space-time pollution data," *Indones. J. Sci. Technol.*, vol. 6, no. 1, pp. 243–268, 2021, doi: 10.17509/ijost.v6i1.32732.

[8] B. A. Budiman *et al.*, "Prediction of the remaining service lifetime of inflatable rubber dam with deep hole damage," *Indones. J. Sci. Technol.*, vol. 5, no. 3, pp. 366–381, 2020, doi: 10.17509/ijost.v5i3.24936.

[9] N. L. Bowers, H. U. Gerber, J. C. Hickman, D. A. Jones, and C. J. Nesbitt, *Actuarial Mathematics*. Schaumburg, Illinois, United States of America: The Society of Actuaries, 1997. doi: 10.1080/00029890.1986.11971867.

[10] B. L. Jones, "Actuarial calculations using a Markov model," *Trans. Soc. Actuar.*, vol. 46, pp. 227–250, 1994.

[11] Q. Guibert and F. Planchet, "Non-Parametric Inference of Transition Probabilities Based on Aalen-Johansen Integral Estimators for Semi-Competing Risks Data - Application to LTC Insurance," *Insur. Math. Econ.*, vol. 82, pp. 21–36, 2018, doi: https://doi.org/10.1016/j.insmatheco.2018.05.004.

[12] F. Zuhairoh, D. Rosadi, and A. R. Effendie, "Determination of Basic Reproduction Numbers using Transition Intensities

Multi-state SIRD Model for COVID-19 in Indonesia," *J. Phys. Conf. Ser.*, vol. 1821, no. 1, p. 012050, 2021, doi: 10.1088/1742-6596/1821/1/012050.

[13] F. Zuhairoh, D. Rosadi, and A. R. Effendie, "Multi-state Discrete-time Markov Chain SVIRS Model on the Spread of COVID-19," *Eng. Lett.*, vol. 30, no. 2, pp. 598–608, 2022.

[14] Y. Foucher, E. Mathieu, P. Saint-Pierre, J. F. Durand, and J. P. Daurès, "A semi-markov model based on generalized Weibull distribution with an illustration for HIV disease," *Biometrical J.*, vol. 47, no. 6, pp. 1–9, 2005, doi: 10.1002/bimj.200410170.

[15] Istiqomah, D. Rosadi, and A. R. Ronnie, "The Graduation of Transition Intensities using Generalized Linear Models in Multiple State Model and it's Application to Premium Pricing of Health Insurance," Universitas Gadjah Mada, 2011. Accessed: Jul. 14, 2019. [Online]. Available: http://etd.repository.ugm.ac.id/penelitian/detail/52394

[16] S. Haberman and A. E. Renshaw, "Generalized linear models and actuarial science," *J. R. Stat. Soc. Ser. D Stat.*, vol. 45, no. 4, pp. 407–436, 1996, doi: 10.2307/2988543.

[17] J. A. Rice, *Mathematical Statistics and Data Analysis, Third Edition*. Belmont, USA: Thomson, 2007.

[18] K. Alim, A. Listiani, A. S. Anggraeni, and A. R. Effendie, "Critical illness insurance pricing with stochastic interest rates model," *J. Phys. Conf. Ser.*, vol. 1341, no. 062026, 2019, doi: 10.1088/1742-6596/1341/6/062026.

[19] A. Asanjarani, B. Liquet, and Y. Nazarathy, "Estimation of semi-Markov multi-state models: A comparison of the sojourn times and transition intensities approaches," *Int. J. Biostat.*, vol. 18, no. 1, pp. 243–262, 2022, doi: https://doi.org/10.1515/ijb-2020-0083.