J. Indones. Math. Soc.
Vol. 28, No. 03 (2022), pp. 323–344.

MIXED MODELS OF NON-PROPORTIONAL HAZARD AND APPLICATION IN THE OPEN DISTANCE EDUCATION STUDENTS RETENTION DATA

Dewi Juliah Ratnaningsih[1], Anang Kurnia[2], Asep Saefuddin[3], I Wayan Mangku[4]

[1]**Universitas Terbuka, Indonesia, djuli@ecampus.ut.ac.id**
[2]**IPB University, Indonesia, anangk@apps.ipb.ac.id**
[3]**IPB University, Indonesia, asaefuddin@gmail.com**
[4]**IPB University, Indonesia, wayanma@apps.ipb.ac.id**

Abstract The problem that arises in the Cox model is that there are more than two types of covariates and the presence of random effects is a non-proportional hazard (NPH). One example of a case that involves many factors is student retention. Low student retention can lead to dropping out of college or failure in completing studies. The purpose of this study is to overcome the problem of NPH caused by the presence of time-independent covariates, time-dependent covariates, and random effects. The research method uses simulation. Some of the modified models are the stratified Cox model, the extended Cox model, and the frailty model. The developed model is applied to distance education student retention data. The results of the study show that frailty and study programs provide considerable diversity in explaining the total diversity of the model. It can be concluded that frailty needs to be considered by UT to improve the quality of services to students. In addition, other covariates that have a significant effect on UT student learning retention modeling are age, domicile, gender, GPA, marital status, employment status, number of credits taken, and number of registered courses.

Key words and Phrases: Universitas Terbuka, stratified Cox, extended Cox, frailty, mixed effect model

1. INTRODUCTION

The Cox model assumes that one individual's hazard rate is proportional to the hazard rate for other individuals and that the population is homogeneous

2020 Mathematics Subject Classification: 62N01, 62N02
Received: 21-03-2022, accepted: 07-12-2022.

(Weinke [60], Kleinbaum and Klein [32]). This implies that one individual's hazard rate is proportional to the hazard rate of another individual with a constant (constant ratio over time). Meanwhile, a homogeneous population means that the individuals who become the samples have the same hazard factors.

In general, the covariates involved in the Cox model are time-independent covariates. The major problem in Cox model is the hazard rate of individuals with other individuals is not proportional (non-proportional hazard) caused by time-dependent covariates involvement. With time-dependent covariates, the hazard rate of an individual is not constant and can change with changes in time. The second problem is heterogeneity. This problem can occur because of the covariate is a random effect. Random effects can lead to diversity or heterogeneity of individual populations. This fact means that individuals have different hazards to survive or experience failure/death from an event. The frailty of the individual the higher the probability of dead compared to those who are not frail. (Weinke [60]). These two conditions cause non-proportional hazard problems, making the assumptions of the Cox model is not met. Several methods have been developed for handling the non-proportional hazard models, including the stratified Cox model, the extended Cox model, and the frailty model. The stratified Cox model was proposed by Abdelaal and Zakaria [1], Ata and Sozer [4], Mehrotra and Su [39]. Gellar et al. [22], Kleinbaum and Klein [32] and Saegusa et al. [50] proposed the extended Cox model. Meanwhile, the frailty model is proposed by Callegaro and Iacobelli [9], Deepapriya and Ravanan [15], Vaupel et al. [58], Wienke [60], and Yadav and Yadav [63].

In practice, the non-proportional hazard model's existence is not always satisfy, as in the case with Open and Distance Education (ODE) student retention data. Existing non-proportional hazard models are not suitable and do not provide comprehensive information regarding student retention modeling. We define the student retention is a condition in which students can survive completing their studies where students are registered continuously per semester (Arifin [3], Belawati [6]; Berger and Lyons [7], and Sembiring [52]).

Retention data related to time or analysis of time to the event. Low student retention can lead to dropping out (failure to complete studies). In the statistics model, failure in completing learning is a failure time. Thus, modeling student retention can use modeling survival time. The factors affecting student retention are complex and varied. The covariates involved in modeling student retention are numerous. These covariates are time-independent covariates, time-dependent covariates, and random-effect covariates. The survival analysis modeling using the Cox model and the existing non-proportional hazard is not suitable given these various covariates. Therefore, it is necessary to develop another non-proportional hazard modeling that can overcome data application problems, as student retention data in ODE. With modeling that is close to real conditions, the Universitas Terbuka as a provider of ODE in Indonesia, with modeling that is close to real conditions, the Universitas Terbuka as a provider of ODE in Indonesia, can increase student retention by by using appropriate and valid statistical analysis.

2. MIXED EFFECT MODEL OF NON-PROPORTIONAL HAZARD

2.1. **Previously Developed Models.**

Ratnaningsih et al. ([47] and [48]) have developed non-proportional hazard modeling by modifying several existing methods. Ratnaningsih et al. [47]) developed a model called Stratified Extended Cox (SE Cox) to overcome the existence of two covariates in the model simultaneously, namely time-dependent covariate and time-independent covariate. Meanwhile, Ratnaningsih et al. [48]) developed an advanced model and SE Cox by adding a frailty component called the Stratified Extended with Frailty (SEF) model. The SEF model was designed to overcome non-proportional hazards because of the time-dependent covariate and time-independent covariate and the observed random effect (frailty). The two models are defined as follows.

2.1.1. *Model Stratified Extended Cox.*

The SE Cox model is defined as follows:

$$\lambda_s(t,x) = \lambda_{0s}(t)\exp\left(\sum_{a=1}^{p_1}\beta_{ai}x_{ai} + \sum_{b=1}^{p_2}\alpha_{bi}x_{bi}(t_j)\right) \tag{1}$$

where:

$s =$	the order of stratum;
$=$	1, 2, ... , m (denoting the number of stratum combination)
$\lambda_{0s}(t) =$	baseline hazard function on each stratum (s = 1, 2,.., m).
$\beta_{ai} =$	fixed effect coefficient vector for covariate number a of individual number i.
$x_{ai} =$	fixed effect coefficient vector for covariate number a of individual number i.
$\alpha_{bi} =$	coefficient vector for time-dependent covariate number b of individual number i.
$x_{bi}(t_j) =$	time-dependent covariate of individual number i at time t_j.

The parameter estimation in the SE Cox model is likelihood-based by adopting the model developed by Cox [13] and Keele [31], the stratified Cox model developed by Dupuy and Leconte [19], and the extended Cox model developed by Fisher and Lin [20]. The method used is the maximum partial likelihood estimation (MPLE). The probability hazard of individual i at time t_j is defined as:

$$P_{i(t_j)} = \frac{\exp\left(\sum\limits_{a=1}^{p_1}\beta_{ai}x_{ai} + \sum\limits_{b=1}^{p_2}\alpha_{bi}x_{bi}(t_j)\right)}{\sum\limits_{j\in R(t_i)}\exp\left(\sum\limits_{a=1}^{p_1}\beta_{aj}x_{aj} + \sum\limits_{b=1}^{p_2}\alpha_{bj}x_{bj}(t_j)\right)} \tag{2}$$

$R(t_i)$ is the set of objects that have a hazard of experiencing an event until time t. The maximum likelihood function can be expressed as:

$$L_p(\phi) = \prod_{s=1}^{m} L_s = \prod_{s=1}^{m}\prod_{i=1}^{n_s} P_{i(t_j)} = \prod_{s=1}^{m}\prod_{i=1}^{n_s} \frac{\exp\left(\sum_{a=1}^{p_1}\beta_{ai}x_{ai} + \sum_{b=1}^{p_2}\alpha_{bi}x_{bi}(t_j)\right)}{\sum_{j\in R(t_i)} \exp\left(\sum_{a=1}^{p_1}\beta_{aj}x_{aj} + \sum_{b=1}^{p_2}\alpha_{bj}x_{bj}(t_j)\right)} \tag{3}$$

2.1.2. *Model Stratified Extended Cox with Frailty.*

The unobserved random effect in the survival model is called frailty (Duchateau and Janssen [18]; Lee et al. [35]; McGilchrist and Aisbett [38]; Vaupel et al. [**?**]). Frailty is an unobserved random proportionality factor that modifies the hazard function of an individual, or of related individuals. The non-proportional hazard in the SEF model is due to time-independent covariates, time-dependent covariates, and unobserved random effects (frailty). There are three-factor causes of non-proportional hazards in the SEF. They are time-dependent covariates, time-independent covariates, and unobserved random effects. The SEF model is mathematically defined as follows.

$$\lambda_s(t,x) = \lambda_{0s}(t)\exp\left(\sum_{a=1}^{p_1}\beta_{ai}x_{ai} + \sum_{b=1}^{p_2}\alpha_{bi}x_{bi}(t_j) + \delta\nu_s\right) \tag{4}$$

where:

$s =$ the order of stratum;
$\lambda_{0s}(t) =$ baseline hazard function on each stratum.
$\beta_{ai} =$ fixed effect coefficient vector for covariate number a of individual number i.
$x_{ai} =$ time-independent covariate fixed effect, number a of individual number i.
$\alpha_{bi} =$ coefficient vector for time-dependent covariate number b of individual number i.
$x_{bi}(t_j) =$ time-dependent covariate of individual number i at time t_j.
$\delta =$ frailty coefficient vector.
$\nu_s =$ frailty on stratum number s.

The estimation method used in the SEF model is hierarchical likelihood. The SEF model estimation method adopts the hierarchical probability (H-likelihood) introduced by Ha et al. [24], Noh et al. [42], Lee et al. [34], Wang et al. [59], Ha et al. [28] and Christian et al. [12] with frailty with normal log distribution. The hierarchical likelihood function is expressed as:

$$h_s = \left(\sum_i (\delta_{si}\log\{\delta_{0s}(y_{si})) + \eta_{si}\} - \{\Lambda_{0s}(y_{si})\exp(\eta_{si})\})\right) - \frac{1}{2}\log(2\pi\theta) - \frac{1}{2\theta}v_s^2. \tag{5}$$

where:

$h_s =$	a logarithm of the joint probability function $(y_{si}, \delta_s, v_s\ )$
$\delta_{si} =$	$I(T_{si} \leq C_{si})$ where $I(.)$ is indicator function.
$\eta_{si} =$	$(\sum_{a=1}^{p1} \beta_{ai}x_{ai} + \sum_{b=1}^{p2} \alpha_{bi}x_{bi}(t_j)$
$=$	the stratified extended Cox model without baseline hazard function on each stratum.
$\delta_{0s}(y_{si} =$	the baseline hazard function for each observed data stratum and $y_{si} = min(T_{si}, C_{si})$
$\Lambda_{0s}(y_{si}) =$	the baseline cumulative hazard function for each observed data stratum.
$v_s =$	frailty on stratum number s.

2.2. **The Proposed Model.**

The mixed-effect model is a model that involves both fixed and random effects (McCulloch and Searle [37]; Dobson [17]; Stroup [53]; Goldstein [23]). Demidenco [16] states that a mixed model is used as repetitive data measurement models or hierarchical models. The mixed-effect model is used to analyze cluster or panel data, longitudinal data. The mixed-effect model is the best used in biology and medical studies where there is a heterogeneity of responses. In medicine, mixed modeling plays a significant role in modeling a case of disease or genetics (Brown and Prescott [8]). The advantage of the mixed-effect model is its ability to combine data by introducing hierarchical random effects.

The mixed influence model in survival analysis has multiple random effects (Crowther et al. [14]; Therneau and Clinic [56]; Austin [5]). The mixed-effect model developed in this study is the mixed effect model with the non-proportional hazard (MEM-NPH).

The MEM-NPH model is an extension of the SEF model (Ratnaningsih et al. [47]) by adding one observed random effect component. Thus, there are two random effects involved in the MEM-NPH model, namely frailty and random effect (observed random effect).

Mathematically, the MEM-NPH model is defined as follows.

$$\lambda_s(t, x_{ai}, x(t), z_{bi}, v_s) = \lambda_{0s}(t) \exp\left(\sum_{a=1}^{p_1} \beta_{ai}x_{ai} + \sum_{b=1}^{p_2} \gamma_{bi}z_{bi} + \sum_{c=1}^{p_3} \alpha_{ci}x_{ci}(t_j) + \upsilon\nu_s \right) \quad (6)$$

where:

$s =$	the order of stratum;
$\lambda_{0s}(t) =$	baseline hazard function in each stratum.
$\beta_{ai} =$	fixed effect coefficient vector for covariate number a of individual number i.
$x_{ai} =$	time-independent covariate fixed effect) number a of individual number i.
$\gamma_{bi} =$	random effect coefficient vector for covariate number b of individual number i.
$z_{bi} =$	observed random effect covariate number b of individual number i.
$\alpha_{ci} =$	time-dependent coefficient vector for covariate number c of individual number i.
$x_{ci}(t_j) =$	time-dependent covariate of individual number i at time t_j.
$\nu_s =$	frailty (unobserved random effect) on stratum number s.
$\upsilon =$	frailty on stratum number s.

Let $T_{si}(s = 1, 2, ..., m, i = 1, 2, ..., n_s)$ is survival time of individual number i on stratum number s and C_{si} is censored time for the individual number i on stratum number s, so that the observed data is $y_{si} = min(T_{si}, C_{si})$ and $\delta_{si} = I(T_{si} \leq C_{si})$ where $I(.)$ is the indicator function. The indicator function has two values, 0 if censored data and 1 if uncensored. Suppose that:

$$\eta^*_{si} = \sum_{a=1}^{p_1} \beta_{ai} x_{ai} + \sum_{b=1}^{p_2} \gamma_{bi} z_{bi} + \sum_{c=1}^{p_3} \alpha_{ci} x_{ci}(t_j) + \upsilon \nu_s. \tag{7}$$

Defined of ν_s is log-observed frailty. Let u_s is variable random of unobserved frailty on stratum number s, then $\nu_s = \log u_s$. Yau [64], Ha et al. [25], Ha et al. [26], Ha et al. [27], Wu [61] , and Jeon et al. [29] assume that frailty is as follows.

(1) **Assumption 1**. Given $U_i = u_i, \{(T_{si}, C_{si}), i = 1, 2, ..., n_s\}$ are conditionally independent and T_{si} and C_{si} also conditionally independent for $s = 1, 2, ..., m; i = 1, 2, ..., n_s$
(2) **Assumption 2**. Given $U_i = u_i, \{(T_{si}, C_{si}), i = 1, 2, ..., n_s\}$ is non-informative with respect to u_s
(3) **Assumption 3**. Variable ν_s are conditionally independent about other covariates.

The likelihood function approach uses hierarchical likelihood. Let $y_s = (y_{s1}, ..., y_{sn_s})^T$ and $\delta_s = (\delta_{s1}, ..., \delta_{sn_s})^T$ The hierarchical likelihood function denoted by h is sum of $h_s, s = 1, 2, ..., m$ (Lee and Nelder [33], and Lee et al. [34]). Therefore,

$$h = \sum_s h_s \tag{8}$$

where h_s is a logarithm of the joint probability function $(y_s, \delta_s, z_{bi}, \nu_s)$.

h_s can be represented in the following equation:

$$\begin{aligned} h_s(\beta, \alpha, \Lambda_s, \gamma, \upsilon^{\cdot}, y_s, \delta_s, z_{bi}, \nu_s) = &\log\{L_{1s}(\beta, \alpha, \Lambda_s, \gamma, \upsilon^{\cdot}, y_s, \delta_s | z_{bi}, u_s) \\ &\times L_{2s}(\gamma; z_{bi}) L_{3s}(\upsilon; \nu_s)\} \end{aligned} \tag{9}$$

where:

L_{1s} =	conditional probability function of (y_s, δ_s) with condition z_{bi} and u_s
L_{2s} =	probability function from z_{bi}
L_{3s} =	probability function from u_s

Because they are assumed conditionally independent and unknown parameter, for example $\omega = (\beta, \gamma, \alpha, \upsilon, \theta)^T$ then L_{1s} in equation (9) can be represented in the following equation (10).

$$\begin{aligned} L_{1s}(\omega, \Lambda_s; y_s, \delta_s | z_{bs}, u_s) &= \prod_i L_{1si}(\omega, \Lambda_s; y_s, \delta_s | z_{bs}, u_s) \\ &= \left(\sum_i (\delta_{si}\{\log(\lambda_{0s}(y_{si}) + (\eta^*_{si}))\} - \{\Lambda_{0s}(y_{si} \exp(\eta^*_{si}))\}) \right) \end{aligned} \tag{10}$$

Assumed that $\gamma_{bi} \sim N(0, \varphi)$ $\nu_s \sim N(0, \theta)$ then L_{2s} and L_{3s} are:

$$L_{2s}(\varphi; z_{si}) = (2\pi\varphi)^{\frac{1}{2}} \exp\left(-\frac{1}{2\varphi} z_{si}^2\right), -\infty < z_{si} < \infty \tag{11}$$

$$L_{3s}(\theta; \nu_s) = (2\pi\theta)^{\frac{1}{2}} \exp\left(-\frac{1}{2\theta} \nu_s^2\right), -\infty < \nu_s < \infty \tag{12}$$

So that equation (9) can be written as:

$$\begin{aligned} h_s = &\left(\sum_i (\delta_{si} \{\log(\lambda_{0s}(y_{si}) + (\eta^*_{si}))\} - \{\Lambda_{0s}(y_{si}) \exp(\eta^*_{si})\}) \right) \\ &- \frac{1}{2}\{\log(4\pi^2)(\varphi\theta)\} - \frac{1}{2}\left(\frac{z_{si}^2}{\varphi} + \frac{\nu_s^2}{\theta}\right) \end{aligned} \tag{13}$$

Because $h = \sum_s h_s$ then equation (9) can be written as:

$$\begin{aligned} h &= \sum_s h_s \\ &= \sum_{si} (\delta_{si} \{\log(\lambda_{0s}(y_{si}) + (\eta^*_{si}))\} - \{\Lambda_{0s}(y_{si}) \exp(\eta^*_{si})\}) - \\ &\frac{p}{2}\log(2\pi\varphi) - \frac{1}{2\varphi}\sum_s z_{si}^2 - \frac{m}{2}\log(2\pi\theta) - \frac{1}{2\theta}\sum_s \nu_s^2. \end{aligned} \tag{14}$$

2.3. **Parameter Estimation.**

In the model definition section, it has been suggested that mixed models are also called hierarchical models (Demidenco [16]). Therefore, the estimation of MEM-NPH model parameters uses a hierarchical likelihood. The hierarchical likelihood function for the model in equation (6) is as follows:

$$h = \sum_s h_s = \sum_{si} l_{1si} + \sum_s l_{2s} + \sum_s l_{3s}. \tag{15}$$

In the hierarchical probability approach procedure, to estimate parameter $\alpha, \beta, \gamma, \upsilon$ using profile hierarchical likelihood h^* by replacing λ_0 with $\tilde{\lambda}_{0s}$.

$$h^* = h|_{\lambda_0 = \tilde{\lambda}_{0s}} \tag{16}$$

where $\tilde{\lambda}_{0s}$ obtained from the results of solving the estimation equation:

$$\frac{\partial h}{\partial \hat{\lambda}_{0s}} = 0$$
$$\partial \hat{\lambda}_{0s} = \frac{d_{(i)}}{\sum\limits_{(i,j)\in R_i} \exp(\eta^*_{si})}$$

so that equation (16) can be expressed in equation (17) as follows.

$$h^* = h|_{\lambda_0 = \tilde{\lambda}_{0s}} = \sum_{sj} l_{1sj} + \sum_s l_{2s} + \sum_s l_{3s} \tag{17}$$

where: $\sum\limits_{sj} l_{1sj^*} = \sum\limits_i d_{(i)} \log \hat{\lambda}_{0i} + \sum\limits_{sj} \delta_{sj}\eta^*_{si} - \sum\limits_i d_{(i)}$. To obtain an estimate we use the Newton-Raphson numerical method.

To obtain an estimate $\beta, \gamma, \alpha, \upsilon$ what maximizes the function h^* an be done by solving the equation $\frac{\partial h^*}{\partial \beta} = 0$, $\frac{\partial h^*}{\partial \gamma} = 0$, $\frac{\partial h^*}{\partial \alpha} = 0$ and $\frac{\partial h^*}{\partial \upsilon}$ using the Newton-Raphson numerical method.

2.4. **Estimation of Model Variance Component.**

Estimation of model variance component for $\hat{\beta}$ and $\hat{\upsilon}$ is:

$$Var(\hat{\beta},\hat{\upsilon}) = -\begin{pmatrix} -\frac{\partial^2 h^*}{\partial\beta_a\partial\beta_p} & -\frac{\partial^2 h^*}{\partial\beta_a\partial\upsilon_p} \\ -\frac{\partial^2 h^*}{\partial\upsilon_a\partial\beta_q} & -\frac{\partial^2 h^*}{\partial\upsilon_a\partial\upsilon_q} \end{pmatrix}^{-1} \tag{18}$$

To estimate variance of frailty (θ) uses adjusted profile hierarchical likelihood as equation (19).

$$h_A^* = h^* - \frac{1}{2}\log\left\{\det\left(\frac{\mathbf{J}}{2\pi}\right)\right\}|_{\beta=\hat{\beta},\nu=\hat{\upsilon}} \tag{19}$$

where $J = -[H(\phi)]$.

The maximum adjusted profile hierarchical likelihood estimation for θ can be obtained by solving the equation $\frac{\partial h_A^*}{\partial\theta} = 0$ using Newton-Raphson method.

$$\frac{\partial h_A^*}{\partial\theta} = -\frac{m}{2\theta} + \frac{1}{2}\sum_s \frac{\nu_s^2}{\theta^2}^+ \left(-\frac{1}{2}tr\left(\mathbf{J}^{-1}\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix}\right)\right) \tag{20}$$

Let $\mathbf{J}^{-1}\begin{bmatrix}0 & 0\\ 0 & 1\end{bmatrix} = Q$ then second derivate of h_A^* with respect to θ is:

$$\frac{\partial h_A^*}{\partial\theta} = \frac{\partial\left(-\frac{m}{2\theta}+\frac{1}{2}\sum\limits_s\frac{\nu_s^2}{\theta^2}+\left(-\frac{1}{2}tr(Q)\right)\right)}{\partial\theta} = \frac{m}{2\theta^2} + \sum_s\frac{\nu_s^2}{\theta^3} - \frac{tr(Q)}{2\theta^2} + \frac{tr(Q)}{\theta^3}.$$

Variance of $\hat{\theta}$ is:

$$Var(\hat{\theta}) = -\left(-\frac{\partial^2 h_A^*}{\partial\theta^2}\right)^{-1}. \tag{21}$$

2.5. **Simulation Design.**

The simulation design for the non-proportional hazard model is as follows.

(1) Determine the components needed in the simulation, namely:
   (a) The stratum of the covariates are time-independent (stratum). The stratum used are educational backgrounds categorized into three categories, namely: (1) High School, (2) Diploma, and (3) Bachelor's Degree. These stratum were not entered into the model and were assumed uniform (1,3) or written, stratum $\sim Uniform(1,3)$.
   (b) Time-dependent covariate ($x1$). This covariate is analogous to the number of semester credit units taken by students which are assumed to be uniformly distributed (0, 100) or written as $x1 \sim Uniform(0,100)$.
   (c) Time-independent covariate ($x2$). In this simulation, the time-independent covariate is analogous to gender which is assumed $Binom(1,0.6)$, or written as $x2 \sim Binom(1;0.6)$.
   (d) The covariate which functions as the unobserved random effect (frailty) is given the symbol ν. Frailty in the simulation is assumed $N(0,\theta)$ with $\theta = 10,14,16,$ and 20. Written, $\nu \sim N(0,\theta)$.
   (e) The covariate which functions as the observed random effect (random effect) is given the symbol z. In this simulation, the random effect is analogous

to a study program that students are interested in. in the simulation it is assumed $N(0, 20)$ or written, $z \sim N(0, 20)$.

(f) The sample sizes tested (n) were 100, 500, and 2,000. The sample size selection is following the variation in the number of UT students representing each faculty.

(2) Generating data for stratum and covariates ($x1, x2, v$, and z). The generation of time-dependent covariates ($x1$) using the PermAlgo package developed by Sylvestre et al. [54].

(3) Determining the initial initials (the model parameter coefficient values for time-independent covariates, time-dependent covariates, frailty, and random effects) are β, α, υ, and γ, respectively. The initial initial values for $\beta = 2$ and $\alpha = \log(1, 04)$, $\upsilon = 1$, and $\gamma = 1$.

(4) Determining the value of the survival time (T_{ji}), namely the length of time until the individual gets an event for the j^{th} individual in the i-stratum. The event is analogous to the status of non-active students, namely students who do not re-register during four consecutive semesters. Thus the value of T_{ji} is set equal to 4 (T_{ji} =4).

(5) Determining the censored time, C_{ji} is the length of time until an individual is observed (censored) for the j^{th} in the i^{th} stratum. The C_{ji} score is 20 semesters (C_{ji} =20). This value determination is based on the assumption that the average study completion rate of UT students is between 8 and 10 years.

(6) Generating a proportion of censored data (c_{ji}). $c_{ji}= I(T_{ji}|C_{ji})$ is an indicator function. The indicator function has a value of 1 if the data is uncensored (observed) and has a value of 0 if the data is censored. In this study, there are three types of proportion of censored data, namely 0%, 30%, and 50%. The censored data generation is as follows. Censored data 0%, $c_0 \sim Uniform(6, 8)$; censored data 30%, $c_{30} \sim Uniform(2, 5)$; and 50% censored data, $c_{50} \sim Uniform(0, 5)$.

(7) The number of treatments tested in the simulation is a combination of sample size and sensor type. The number of replications for each treatment is 1,000 times.

(8) The response data (y_{ji}) is $y_{ji} = \min\{T_{ji}, C_{ji}\}$ obtained from the generation of time-dependent covariate data using the syntax: $Surv(Start, Stop, Event)$ as practiced by Therneau et al. [55] and Thomas and Reyes [57]. The response data in this study is analogous to student learning retention measured in semester units.

2.6. **The Goodness of Fit Model.**

The measure of the goodness of fit model uses the bias and MSE value resulted from the model simulation (Polat and Gunay [45]).

$$Bias(\tau) = \left(\frac{1}{r} \sum_{i=1}^{r} \hat{\tau}^{(i)} \right) - \tau \tag{22}$$

$$MSE(\tau) = \frac{1}{r} \sum_{i=1}^{r} (\tau - \hat{\tau}^{(i)})^2 \tag{23}$$

where r is the number of replications,τ is the predetermined initial model parameter estimate, and $\hat{\tau}^{(i)}$ is the estimated model parameter resulting from the model simulation.

3. RESULT AND DISCUSSION

3.1. **Simulation Results.**

In the simulation, frailty is assumed to be normally distributed with $\nu \sim N(0, \theta)$, and the random effect is assumed to be normally distributed with $z \sim N(0, 20)$. In the SEF and MEM-NPH models, there were four types (θ) tested, namely: $\theta = 10$, 14, 16, and 20. The censorship and sample size used were the same as in the SE Cox model, namely: three types of censoring ($c = 0\%; c = 30\%; c = 50\%$) and three sample sizes, namely $n = 100, 500$, and 2,000. The simulation results of the three models are presented in Table 1 and Table 2.

Table 1 presents the percentage of parameter estimation bias (β, α, υ, dan γ) generated by each model in various treatment combinations. The ratio of bias in estimating the SEF model parameters is lower than that of the Cox SE model. The portion of bias in estimating the parameters of the MEM-NPH model is lower than that of the SEF model. Graphically, the parameter coefficient γ in the MEM-NPH model is relatively smaller (Figure 1). The simulation results show that the rate of bias in parameter estimation made by the MEM-NPH model is lower than the other models.

Table 1. Percentage of Parameter Bias of Models in Various Censoring and Variance

n	θ	c	Percentage of Parameter Bias									
			SE Cox Model			SEF Model			MEM-NPH Model			
			β	α	υ	β	α	υ	β	α	υ	γ
100	10	0	11.615	15.083	12.206	4.008	9.518	8.542	1.958	1.851	1.799	0.692
500	10	0	10.202	15.677	11.311	3.417	8.673	8.171	1.552	0.881	0.926	0.743
2000	10	0	9.778	14.412	11.256	3.198	8.000	8.012	1.351	0.868	0.978	0.706
100	14	0	11.533	15.608	11.203	3.893	9.518	8.777	1.948	1.873	1.499	0.683
500	14	0	10.180	14.597	12.117	3.385	9.216	7.813	1.936	0.866	0.928	0.697
2000	14	0	9.738	14.347	11.011	3.198	8.818	8.879	1.355	0.850	0.925	0.731
100	16	0	12.338	14.254	10.507	3.943	9.436	8.452	1.948	1.853	1.466	0.692
500	16	0	11.210	13.498	11.472	3.383	9.351	7.677	1.916	0.858	0.880	0.743
2000	16	0	10.696	14.401	11.358	3.194	8.348	8.375	1.455	0.857	0.839	0.706
100	20	0	12.257	14.495	11.495	4.004	8.672	6.883	1.949	1.898	1.367	0.683
500	20	0	11.174	14.992	10.712	3.403	7.676	7.913	1.935	0.893	0.899	0.784
2000	20	0	10.651	13.853	10.616	3.193	8.672	7.818	1.522	0.862	0.826	0.697
100	10	30	10.249	15.078	10.383	3.423	8.818	6.883	1.951	1.855	1.456	0.738
500	10	30	9.797	14.949	9.701	3.199	8.347	6.840	1.958	0.897	0.999	0.787
2000	10	30	11.615	13.824	11.001	4.008	7.412	8.542	1.552	0.862	0.725	0.706
100	14	30	11.621	14.986	11.613	3.981	6.677	8.661	1.934	1.861	1.359	0.680
500	14	30	10.249	14.424	10.383	3.423	7.818	6.883	1.925	0.868	0.926	0.764
2000	14	30	9.797	14.117	9.701	3.199	7.347	6.840	1.652	0.855	0.849	0.742
100	16	30	11.615	14.913	12.311	4.008	8.412	8.542	1.922	1.877	1.598	0.683
500	16	30	12.421	14.521	12.119	4.025	7.302	6.469	1.934	0.869	0.965	0.784
2000	16	30	11.263	14.256	11.211	3.409	8.397	5.270	1.453	0.850	0.839	0.697
100	20	30	10.759	14.707	10.577	3.194	8.671	5.805	1.928	1.869	1.697	0.731
500	20	30	12.361	13.982	12.465	3.925	8.934	6.962	1.922	0.886	0.925	0.692
2000	20	30	11.228	13.598	11.368	3.412	8.346	8.948	1.052	0.850	0.833	0.743
100	10	50	10.715	15.08	11.608	3.194	9.397	8.457	1.924	1.907	1.262	0.706
500	10	50	9.857	14.95	11.368	3.465	9.346	7.992	1.915	0.975	0.855	0.683
2000	10	50	9.252	14.62	11.011	3.218	8.609	7.595	0.758	0.871	0.766	0.784
100	14	50	10.877	15.49	12.465	4.163	9.609	7.303	1.921	1.897	1.265	0.697
500	14	50	9.796	14.69	11.854	3.469	8.357	7.178	1.925	0.893	0.855	0.764
2000	14	50	11.263	14.47	10.577	3.409	8.231	6.720	0.858	0.851	0.766	0.742
100	16	50	10.759	15.18	11.522	3.194	8.978	7.172	1.931	1.889	1.324	0.683
500	16	50	12.361	14.98	11.454	3.925	8.673	8.220	1.924	0.871	0.856	0.784
2000	16	50	11.228	14.64	10.761	3.412	8.171	7.195	0.951	0.869	0.765	0.697
100	20	50	10.715	14.97	11.600	3.194	9.027	7.948	1.931	1.868	0.803	0.731
500	20	50	9.757	14.51	10.596	3.440	8.909	8.110	0.953	0.873	1.360	0.692
2000	20	50	9.161	14.25	10.454	0.217	8.534	7.962	0.651	0.865	0.762	0.696

Table 2 shows the MSE values of the three models. The simulation results show that the MSE value of the MEM-NPH model is lower than the other two models. The SEF model has a lower MSE value than the SE Cox model. The MSE value of the parameter estimator in the MEM-NPH model is presented in Figure 2.

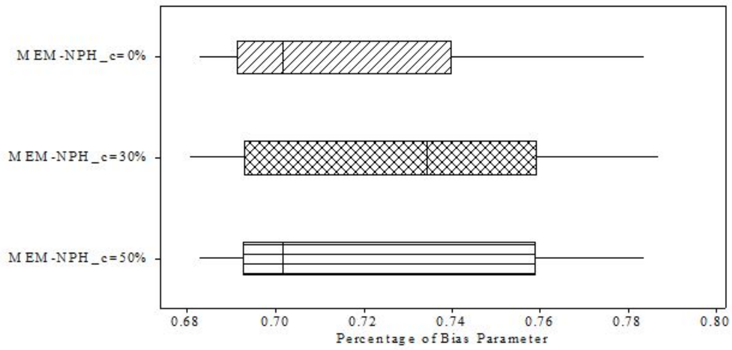

Figure 1. The Percentage of Bias Parameter for γ in Various Censoring

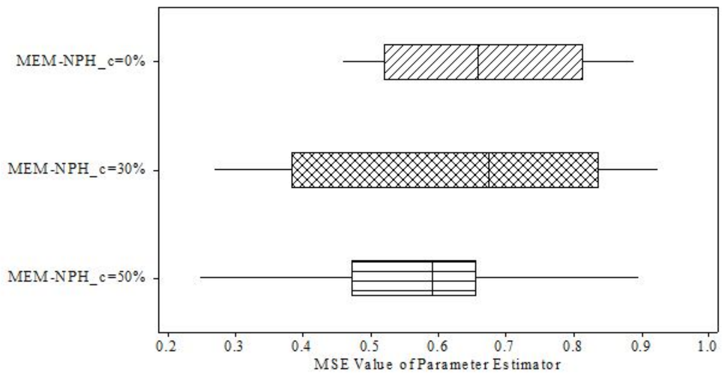

Figure 2. The Percentage of MSE Value for γ in Various Censoring

Based on two model criteria, bias and MSE, it shows that the MEM-NPH model provides a percentage of bias and smaller parameter estimator MSE values than the SEF model and the SE Cox model. In modeling involving two types of covariates and random effects, the MEM-NPH model provides better model merit. Thus, the MEM-NPH model can be used to model the unmatched risk caused by the presence of time-independent covariates and time-dependent covariates, and random effects.

Table 2. Percentage of MSE Value of Models in Type of Censoring and Variance

n	θ	c	MSE Value of Parameter of Models									
			SE Cox Model			SEF Model			MEM-NPH Model			
			β	α	υ	β	α	υ	β	α	υ	γ
100	10	0	0.529	0.636	0.558	0.374	0.488	0.451	0.209	0.321	0.325	0.554
500	10	0	0.578	0.621	0.548	0.329	0.469	0.404	0.189	0.320	0.311	0.689
2000	10	0	0.595	0.608	0.552	0.337	0.482	0.402	0.123	0.294	0.304	0.814
100	14	0	0.537	0.651	0.550	0.386	0.493	0.448	0.152	0.312	0.298	0.658
500	14	0	0.578	0.642	0.512	0.361	0.431	0.425	0.131	0.291	0.265	0.466
2000	14	0	0.595	0.570	0.563	0.324	0.428	0.410	0.113	0.261	0.292	0.659
100	16	0	0.536	0.591	0.547	0.374	0.474	0.468	0.133	0.310	0.297	0.592
500	16	0	0.580	0.556	0.517	0.290	0.470	0.457	0.193	0.313	0.242	0.889
2000	16	0	0.595	0.565	0.528	0.321	0.410	0.416	0.193	0.290	0.256	0.509
100	20	0	0.517	0.578	0.569	0.362	0.451	0.457	0.193	0.259	0.286	0.866
500	20	0	0.578	0.574	0.568	0.363	0.385	0.437	0.235	0.272	0.272	0.459
2000	20	0	0.595	0.543	0.526	0.314	0.375	0.428	0.204	0.251	0.258	0.806
100	10	30	0.537	0.596	0.495	0.399	0.466	0.394	0.285	0.314	0.289	0.666
500	10	30	0.577	0.593	0.481	0.386	0.448	0.373	0.288	0.335	0.274	0.797
2000	10	30	0.595	0.556	0.474	0.377	0.471	0.336	0.262	0.310	0.271	0.441
100	14	30	0.533	0.597	0.468	0.393	0.475	0.373	0.288	0.305	0.260	0.696
500	14	30	0.577	0.591	0.435	0.374	0.452	0.347	0.271	0.282	0.283	0.897
2000	14	30	0.595	0.585	0.462	0.369	0.435	0.324	0.254	0.295	0.193	0.294
100	16	30	0.530	0.594	0.467	0.333	0.485	0.380	0.285	0.256	0.231	0.924
500	16	30	0.577	0.574	0.412	0.337	0.469	0.329	0.271	0.293	0.243	0.641
2000	16	30	0.595	0.589	0.396	0.393	0.454	0.320	0.263	0.258	0.196	0.364
100	20	30	0.539	0.549	0.456	0.373	0.474	0.374	0.287	0.265	0.258	0.850
500	20	30	0.580	0.579	0.442	0.326	0.464	0.320	0.294	0.249	0.197	0.686
2000	20	30	0.595	0.566	0.428	0.333	0.432	0.290	0.239	0.237	0.184	0.270
100	10	50	0.521	0.637	0.389	0.415	0.384	0.367	0.233	0.292	0.217	0.646
500	10	50	0.573	0.598	0.374	0.371	0.371	0.357	0.182	0.257	0.207	0.509
2000	10	50	0.593	0.580	0.371	0.425	0.349	0.317	0.145	0.261	0.167	0.460
100	14	50	0.518	0.680	0.360	0.367	0.376	0.318	0.115	0.215	0.168	0.642
500	14	50	0.573	0.571	0.383	0.358	0.293	0.323	0.173	0.202	0.173	0.539
2000	14	50	0.593	0.599	0.373	0.422	0.358	0.310	0.122	0.190	0.150	0.304
100	16	50	0.525	0.620	0.341	0.311	0.366	0.324	0.125	0.215	0.164	0.537
500	16	50	0.576	0.585	0.343	0.365	0.352	0.316	0.183	0.199	0.166	0.758
2000	16	50	0.593	0.564	0.364	0.432	0.270	0.301	0.115	0.214	0.151	0.246
100	20	50	0.525	0.599	0.328	0.365	0.368	0.303	0.151	0.236	0.153	0.657
500	20	50	0.575	0.581	0.347	0.363	0.356	0.283	0.124	0.194	0.153	0.897
2000	20	50	0.543	0.552	0.324	0.431	0.358	0.292	0.196	0.190	0.152	0.654

3.2. **An Application on Real Data.**

Universitas Terbuka (UT) is a public university in Indonesia that implements a ODE system. As a higher education institution, UT can expand the reach of learning assistance services and address gaps in access to education constrained by distance, space, time, geographical conditions, and heterogeneous communities. One indicator of institutional accountability in implementing educational programs is the student retention. Low student retention indicates the inability of the institution to improve the quality of educational services.

UT can view student retention problems as problems related to timing or analysis of time events (time to event). Low student retention can lead to dropping out (failure to complete studies). The students who fail to complete the course can be viewed as failure times. Therefore, UT can use the MEM-NPH model for modeling student retention. The random effect observed in the student retention data was the study program that students took. The treatment given by study programs to students varies. Therefore, the student graduation rate in the course or student retention in each study program also varies.

The following describes the application of the MEM-NPH model to UT student retention data. The data used in this study is the retention data of UT students who made their first registration in 2008 semester 1 (2008.1) to 2015 semester 2 (2015.2). The amount of data observed was 4,483 students from 10 study programs. The study programs used as examples in this research are 10 of the 27 undergraduate study programs offered. The sampling of these 10 study programs is based on considering that survival analysis will produce good predictions if the percentage of censorship is at least 50%.In this study, the sample study program is a study program that has a censoring of 65%. Of the ten selected study programs, the number of students observed was 4,483 students consisting of 1,574 people (35.11%) who were censored and 2,909 (64.89%) uncensored. Students are said to be censored if they are still actively studying or have graduated or moved study programs. Meanwhile, students who are uncensored are non-active students.

Meanwhile, the unobserved random effect (frailty), ν is assumed that the random effect follows the normal distribution with an average of 0 and a range of 20 or written as $\nu \sim N(0, 20)$. Frailty in UT student retention can be considered motivation, independent learning culture, management of study time, availability of learning facilities, or historical participation in tutorials face-to-face and online), student social interaction, student learning styles, institutional understanding of needs. Students and the culture or organizational culture in the institution affects UT student retention.

The response variable in this study was student retention which is measured in semester units. UT student retention data is observed from the time the student registers the first until the event occurs. In this study, an event is defined as a change in students' academic status from active students to non-active students. The censoring level is censored (given a 0) score if the student is active or graduated, while uncensored (provided a score of 1) if the student is non-active.

MEM-NPH mathematical model used for modeling UT student retention data is as follows.

$$\lambda_s(t, x(t), z_{bi}, \nu_s) = \lambda_{0s}(t) \exp\left(\sum_{a=1}^{4} \beta_{ai} x_{ai} + \sum_{b=1}^{10} \gamma_{bi} z_{bi} + \sum_{c=1}^{4} \alpha_{ci} x_{ci}(t_j) + \upsilon \nu_s\right), \quad (24)$$

The program used in the MEM-NPH model data analysis is the R program with the frailtyHL program package. Testing of hypothesis of the equation (24) with a significance level of 10% for each parameter is stated as follows.

(1) At the time-independent covariate coefficient, the hypothesis test is:
$H_0 : \beta_a = 0$
$H_1 : \beta_a \neq 0$
This hypothesis test is intended to determine whether time-independent covariates such as gender, domicile, employment status, and marital status affect UT student retention.

(2) At the time-dependent covariate coefficient, the hypothesis test is:
$H_0 : \gamma_a = 0$
$H_1 : \gamma_a \neq 0$
Hypothesis testing is intended to determine whether time-dependent covariates such as age, credits taken per semester, number of courses registered per semester, and GPA affect UT student retention.

(3) At the observed random effect coefficient, the hypothesis test is:
$H_0 : \alpha_c = 0$
$H_1 : \alpha_c \neq 0$
This hypothesis test is intended to determine whether covariates that have an observed random effect, such as the study program of interest, affect UT student retention.

(4) At the unobserved random effect (frailty) coefficient, the hypothesis test is:
$H_0 : \upsilon = 0$
$H_1 : \upsilon \neq 0$
Hypothesis testing is intended to determine whether covariates that have an unobserved random effect. Such as motivation, independent learning culture, management of learning time, availability of learning facilities, or historical participation in tutorials (both face-to-face and online, student social interaction, style student learning, institutional understanding of student needs, and organizational culture or culture in the institution affect UT student retention.

The coefficient or estimator value generated from modeling is then compared with the p-value. If the p-value $<$ significance level, the observed covariate is significant. This value means that the covariate effects UT student retention.

Using MEM-NPH in the model is adequately applied to UT student learning retention data. It can be seen from the value of the correlation coefficient in the classroom (intraclass correlation coefficient, ICC). The ICC value can indicate whether or not a mixed-effect model is required in modeling. McCulloch and Searle [37] stated that the ICC value shows the proportion of random influence on the model's total variability. The greater the ICC value, the units in the same group/cluster have the same characteristics. This fact shows that the grouping in the model is exact.

Table 3 shows the ICC values. From Table 3, the ICC value for the study program is 0.5690 and for frailty is 0.4307. The total random effect ICC value is 0.9997. This shows that the mixed effect model in this case is very adequate and the grouping in this model is very precise.

Table 3. Variance component of MEM-NPH model

Model Effect	Std. Error	Variance	ICC
Fixed effect		70.12501	
Random effect			
- Study Program	22.7112	515.7989	0.5690
- Frailty	19.7585	390.3976	0.4307
Residual		0.2861	0.0003
Total		906.4826	

The estimation of UT student retention data parameters using the MEM-NPH model is presented in Table 4. Table 4 shows the results of parameter estimation using the MEM-NPH model. The most significant factors on UT students' retention are study programs, frailty, work status, marital status, age, cumulative grade point average (GPA), number of course credits (credit hours, CH), and the number of courses registered per semester.

Table 4. The result of parameter estimation uses the MEM-NPH model

Parameter	Estimate Value	Hazard Ratio	Std.Error	Z-Value	Pr $(>\|Z\|)$
Fixed effect					
(Intercept)	0.5472	1.7284	0.0409	13.373	$< 2e-16$***
Age 35-45 years old	0.0032	1.0032	0.0287	-0.113	0.9103
Age > 45 years old	-0.1082	0.8975	0.0354	3.054	0.0023**
Home area (city)	0.0038	1.0038	0.0185	0.206	0.8367
Gender (male)	-0.0264	0.9739	0.0150	-1.762	0.0780*
$1,00 < GPA \leq 2,00$	-0.5811	0.5593	0.0221	26.256	$< 2e-16$***
$2,00 < GPA \leq 3,00$	0.4487	1.5663	0.0284	15.782	$< 2e-16$***
$GPA > 3,00$	0.1907	1.2100	0.0798	2.390	0.0168*
Employed	-0.0102	0.9898	0.0286	0.357	0.0721*
Married	-0.0187	0.9814	0.0176	1.064	0.0287*
$75 \leq CH \leq 120$	0.8460	2.3304	0.0254	33.355	$< 2e-16$***
$CH > 120$	1.2392	3.4528	0.0207	59.972	$< 2e-16$***
$5 \leq$ Courses ≤ 8	-0.0097	0.9903	0.0205	-0.474	0.6355
Courses > 8	-0.0275	0.9729	0.0323	-6.228	$4.73e-10$***
Random effect					
(Intercept)	-9.4295	0.0001	3.4605	-3.881	0.0001***
Study Program	0.8014	2.2287	0.3392	2.735	0.0062**
Frailty	1.5816	4.8627	0.2951	-0.521	0.0603*

The estimated value of the study program is 0.8014 ($e^{0.8014}$). This value means that management in the study program provides high learning retention for UT students or the risk of student learning retention by 2.23 times. This value indicates that students who choose a study program according to their interests or educational background will survive to continue their studies 2.23 times than students who prefer a study program not according to their interests and educational experience. This fact is following Olliveira [43] course at the Open University of Brazil, which states that students' study program significantly affects the sustainability of student studies.

The frailty component has a significant influence on the learning retention of UT students. The estimated parameter value for frailty is 1.5816. That is, frailty has a risk of 4.86 times compared to other UT students' learning retention. This condition is consistent with the statements of Merriam and Caffarella [40], Cercone [11], and Kara et al. [30] that learners in ODE are unique and follow adult learning styles. Adults learn according to their needs and responsibilities. Furthermore, Merriam and Caffarella [40] state that adult learning needs to consider various aspects, such as the learning environment, student characteristics, and student learning styles. This fact supports the statement of Schuemer [51], who argues that in the ODE system, the student learning process is more complicated because, in general, ODE students are elderly, work, and have families.

The age of UT students over 45 years influences student retention. The estimated parameter value of -0.1082 indicates that students who is 45 years old or above at risk of having low learning retention of 0.8975 times compared to other generations. This condition is following the study of Xenos et al. [62] and Pierrakeas et al. [44] in Greece as well as Schuemer [51] and Rovai [49]. Rovai [49] states that one of the reasons for dropping out of college at ODE is old age.

Gender is quite influential if the alpha level used is 10%. From Table 3, it can be seen that the gender parameter estimator coefficient is -0.0264. This value shows that male students tend to have low learning retention or 0.97 times compared to women. This condition is possible because male students generally work. They may find it difficult to divide their time between work and study.

The Grade Point Average (GPA) has a significant influence on student retention. Student with GPA between 1.00 and 2.00 tend to have low retention of learning. His learning retention risk was 0.5593 compared to other GPAs. However, students who have a GPA of more than 2.00 tend to have a high retention of learning. From Table 4, it can be seen that the coefficient is positive. This coefficient means that student retention of learning in the GPA range above 2.00 ($2.00 < \text{GPA} \leq 3.00$) has a risk of surviving to learn by 1.57 times, and for students who have a GPA > 3.00, the risk of surviving to learn is equal to 1.21 times. Student work has a significant effect on learning retention. Students who work tend to have lower retention than students who do not work. The value risk of learning is 0.99 times that of students who do not work. This reality is very reasonable because, in general, UT students work. According to Schuemer [36] and Rovai [49], the ODE system allows a more involved student learning process because, in general, ODE students are elderly, work, and have families. The uniqueness of ODE students' condition can cause problems because they are required to be able to coordinate various aspects, such as family, work, and free time with study time.

Marriage also has a significant contribution to student retention. Students who are married have low learning retention ($e^{0.0187}$) or have 0.98 times lower risk of learning retention than unmarried students. The number of credits taken has a significant contribution to student learning retention. Students who have taken 75 to 120 credits have a high retention of learning ($e^{0.8460}$). This means that the risk of having high learning retention is 2.33 times. Meanwhile, students who had taken more than 120 credits had a higher tendency to survive. The risk of students who have taken more than 120 credits has a high retention of learning of 3.45 times compared to students who have just taken less than 75 credits.

The number of courses registered per semester also has a significant influence on learning retention. Students who take more than eight courses per semester tend to have low retention of learning. In other words, the risk of having low retention of knowledge

($e^{0.0275}$) or 0.97 times compared to students who take less than eight courses per semester. This fact is in line with the study conducted by Cambruzzi et al. [10] in Brazil, which states that many students drop out of college because the credit load does not match the students' abilities. For example, the institution recommends that the credits taken per semester are only 12 credits. However, many students take up to 20 credits because they consider learning with the ODE system easy and can speed up their studies. Allen et al. [2] state that many students take courses, pay tuition fees, and then drop out.

3.3. Discussion.

The application of the MEM-NPH model to UT students' learning retention data indicates that the mixed effect model is perfect for describing the real data retention phenomenon of UT students. Through MEM-NPH modeling, it is appropriate. This fact is evident from the ICC value of 0.4307 for frailty and 0.5690 for the study program. With this grouping, most of the members in the group have similar characteristics.

By using the MEM-NPH model, it can be shown that frailty and study programs play a huge role in modeling UT student learning retention. The contribution of frailty and study programs provides considerable diversity in explaining the total variety of models. Thus, through this modeling, UT needs to consider random effects as the ODE organizing institution to increase UT students' retention. UT should pay more attention to unexpected influences in developing programs to improve services for students.

The application of the MEM-NPH model to UT student retention data shows that the random effect, both unobserved (frailty) and unobserved, contributes quite a lot to the total model's diversity. The existence of frailty and study programs needs to be considered and get special attention from the ODE organizing institutions to provide services to students. Based on the MEM-NPH model, other covariates that affect UT student learning retention modeling at the alpha level of 10% are age, domicile, gender, GPA, marital status, employment status, number of credits taken, and number of registered courses. Greece, Pierrakeas et al. [44] said that younger students (<30 years) tend to drop out of school. This fact is understandable because they may not have had independent study experience, and they tend to underestimate the effort and workload required for university studies. Meanwhile, Gaytan [21], McCormic and Lucas [36], and Ratnaningsih [46] argued that the GPA greatly influences student retention and is a determining factor for the sustainability of university studies. In addition, concerning marital status Rovai [49], states that in general, the factors that cause dropouts experienced by ODE students include advanced age, lack of study time, difficulties in accessing the internet, lack of feedback from tutors, work, family, external stimuli, and personal financial problems

In this study, the frailty used is assumed to be a regular spread. In the future, it is necessary to conduct research where the frailty model uses other distributions closer to the data case, for example, the Gamma or Weibull distribution. Besides, there needs to be a unique study involving a spatial component. This condition is possible because UT is located in 39 cities of the Distance Learning Program Unit and one of the Distance Learning Program Unit in Abroad, which manages UT students in 39 countries. The spatial component needs to be considered in the development of the next model.

4. CONCLUSION

Student retention problems can be viewed as problems related to timing or analysis of the timing of events. Retention is related to the success or failure of completing a study. Success and failure in completing the survey can be viewed as modeling survival time, which can be analyzed using survival analysis.

To model retention data, ODE students can use an unequaled risk model, namely the mixed effect model on non-proportional hazard, abbreviated as MEM-NPH. This model is an extension of the existing Cox model and peer-to-peer risk models. The MEM-NPH model is a combination of the stratified Cox model and the extended Cox model by adding two random effect components, namely the observed random effect and the unobserved random effect.

The MEM-NPH model can determine the variability of random effects on the total model variability. Based on the MEM-NPH model analysis, it can be shown that frailty and study programs provide a reasonably large diversity in explaining the complete variety of the model. From the results of this analysis, it can be concluded that UT's random effect needs to be considered to improve students' quality of service. Besides, other covariates that significantly affect modeling UT student learning retention at the alpha level of 10% are age, domicile, gender, GPA, marital status, employment status, number of credits taken, and number of registered courses.

REFERENCES

[1] Abdelaal, M.M. and Zakaria, S.H., Modeling survival data by using cox regression model". *American Journal of Theoretical and Applied Statistics.* **4(6)** (2015), 504-512, doi: 10.11648/j.ajtas.20150406.21.

[2] Allen, I.E., Seaman, J., Poulin, R. and Straut, T.T., *Online report card: Tracking online education in the United States*, Babson Park, MA: Babson Survey Research Group Retrieved from http://onlinelearningsurvey.com/re-ports/onlinereport-card. (2016).

[3] Arifin, H., The role of student support services in enhancing student persistence in the open university context: lesson from Indonesia Open University", *Turkish Online Journal of Distance Education-TOJDE*, **19(3)**(2018), 156-168.

[4] Ata, N. and Sozer, M.T., Cox regression models with nonproportional hazards applied to lung cancer survival data", *Hacettepe Journal of Mathematics and Statistics*, **36(2)** (2007), 157-167.

[5] Austin, P.C., A tutorial on multilevel survival analysis: methods, models and applications", *International Statistical Review.* **85(2)** (2017), 185-203. doi:10.11-11/insr.12214.

[6] Belawati, T., Increasing student persistence in Indonesian postecondary distance education", *Distance Education*, **19(1)**(1998), 81-108, doi: 10.1080/-0158791980190107.

[7] Berger, J. and Lyon, S., *Past to present: A historical look at retention.* In A. Seidman (Ed.), College student retention, 130. Westport, CT: Praeger, 2005.

[8] Brown, H. and Prescott, R., *Applied Mixed Models in Medicine*, Second Edition. New York: John Wiley & Sons, 2006.

[9] Callegaro, A. and Iacobelli, S., The Cox shared frailty model with log-skew-normal frailties", *Statistical Modelling*, **12(4)** (2012), 399-41, doi:10.1177/14-71082X-12460146.

[10] Cambruzzi, W., Rigo, S.J. and Barbosa, J.L.V., Dropout prediction and reduction in distance education courses with the learning analytics multitrail approach", *Journal of Universal Computer Science*, **21(1)** (2015), 23-47, doi: 10.3217/jucs-021-01-0023.

[11] Cercone, K., Characteristics of adult learners with implication for online learning design", *AACE Journal*, **16(2)** (2008), 137-159.
[12] Christian, N.J., Ha, I.D. and Jeong, J.H., Hierarchical likelihood inference on clustered competing risk data", *Statistics in Medicine*, **35** (2016), 251-267.
[13] Cox, D.R., Regression models and life-tables", *Journal of the Royal Statistical Society Series B (Methodological)*, **34(2)** (1972), 187-220.
[14] Crowther, M.J., Look, M.P and Riley, R.D., Multilevel mixed effects parametric survival models using adaptive Gauss-Hermite quadrature with application to recurrent events and individual participant data meta-analysis", *Statistics in Medicine*, **33** (2014), 38443858.
[15] Deepapriya, S. and Ravanan, R., Gamma frailty model to evaluate heterogeneity for a survival data", *International Journal of Scientific Research*, **5(4)** (2015), 2-5.
[16] Demidenco, E., *Mixed Models Theory and Application With R*, Second Edition, New Jersey: John Wiley & Sons, 2013.
[17] Dobson A.J., *An Introduction to Generalized Linear Models*, Second Edition, New York: Chapman & Hall/CRC, 2002.
[18] Duchateau, L. and Janssen, P., *The Frailty Model*, New York, NY: Springer, 2008.
[19] Dupuy, J.F. and Leconte, E., *Estimation in a Partially Observed Stratified Cox Model*, accessed 2016 December 05 on https://scholar.google.com/cita-tions?user=cQLxz-X8AAAAJ&hl=en, 2006.
[20] Fisher, L.D. and Lin, D.Y., Time-dependent covariates in the Cox proportional-hazard regression model", *Ann. Rev. Public Health*, **20**(1999), 145-157.
[21] Gaytan J., Comparing faculty and student perceptions regarding factors that affect student retention in online education", *American Journal of Distance Education*,**29(1)** (2015), 56-66.
[22] Gellar, J.E., Colantuoni, E., Needhan, D.M. and Crainiceanu, C.M., Cox regression models with functional covariates for survival data". *Statistical Modelling*. **15(3)** (2015) , 123, doi:10.1177/1471082X14565526.
[23] Goldstein, H., *Multilevel Statistical Models*, United Kingdom: John Wiley & Sons, 2011.
[24] Ha, I.D., Lee, Y. and Song, J.K., Hierarchical likelihood approach for frailty models", *Biometrics*, **88(1)** (2001), 233-243.
[25] Ha, I.D., Lee, Y. and Song, J., Hierarchical-likelihood approach for mixed linear models with censored data", *Lifetime Data Analysis*, **8** (2002), 63-176.
[26] Ha, I.D., Lee. Y. and MacKenzie, G., Model selection for multi-component frailty models", *Statistics in Medicine*, **26(26)** (2007), 4790-807.
[27] Ha I.D., Legrand C., MacKenzie G., Sylverter R., Frailty modelling for survival data from multi-centre clinical trials", *Statistics in Medicine*, **30(17)** (2011), 2144-59.
[28] Ha, I.D., Lee, Y. and Noh, M., FrailtyHL: Frailty Models using H-Likelihood", *R package version*, **1.1** (2012), 28-30.
[29] Jeon J., Hsu L., Gorfine M., Bias correction in the hierarchical likelihood approach to the analysis of multivariate survival data", *Biostatistics*, **13(3)** (2012), 384-397.
[30] Kara, M., Erdogdu, F., Kokoc, M. and Cagiltay, K., Challenges faced by adult leaners in online distance education: a literature review", *Open Praxis International Council for Open and Distance Education*, **11(1)** (2019), 5-22, doi: https://doi.org-/10.5944/openpraxis.11.1.929.
[31] Keele, L., Proportionally difficult: testing for nonproportional hazards in cox models", *Political Analysis*, **18** (2010), 189205, doi:10.1093/pan/mpp044.
[32] Kleinbaum, D.G. and Klein, M., *Survival Analysis: A Self-Learning Text*, Third Edition, New York (US): Springer, doi: 10.1007/978-1-4419-6646-9, 2012
[33] Lee, Y. and Nelder, J.A., Hierarchical generalised linear models: A synthesis of generalised linear models, random-effect models and structured dispersions", *Biometrika*, **88(4)** (2001), 987-1006.
[34] Lee, Y., Nelder, J.A. and Pawitan, Y., *Generalized Linear Models with Random Effects: Unified Analysis via Hlikelihood*, London: Chapman & Hall, (2006).

[35] Lee M, Ha I.D., Lee Y., Frailty modeling for clustered competing risks data with missing cause of failure", *Statistical Methods in Medical Research*, **0(0):118** (2014), doi: 10.1177/0962280214545639.

[36] McCormick N.J., Lucas M.S., Student retention and success: Faculty initiatives at Middle Tennessee State University", *Journal of Student Success and Retension*, **1(1)** (2014), 1-12.

[37] McCulloch, C.E. and Searle, S.R., *Generalized, Linear, and Mixed Models*, New York: John Wiley & Sons, Inc. 2001.

[38] McGilchrist, C.A. and Aisbett, C.W., Regression with frailty in survival analysis", *Biometrics*. **47**(1991), 461466.

[39] Mehrotra, D.V. and Su, S.C., An efficient alternative to the stratified cox model analysis", *Statistics Medicine*, **31(17)** (2012), 1849-56, doi: 10.1002/sim.5327.

[40] Merriam, S.B. and Caffarella, R.S., *Learning in Adulthood (2nd ed.)*, San Francisco: Jossey-Bass, 1999.

[41] Muljana, P.S. and Luo, T., Factors contributing to student retention in online learning and recommended strategies for improvement: A systematic literature review", *Journal of Information Technology Education*, **18** (2019), 19-57.

[42] Noh. M., Ha, I.D. and Lee, Y., Dispersion frailty models and HGLMs", *Statistics in Medicine*, **25(8)** (2006), 13411354, doi:10.1002/sim.2284.

[43] Olliveira, P.R., Oesterreich, S.A. and Almeida, V.L., School dropout in gradutae distance education: evidence from a study in the interior of Brazil", *Educ. Pesqui., So Paulo*, **4(e165786)** (2018), 1-20, doi: http://dx.doi.org/10.1590/S1678463420-1708165786.

[44] Pierrakeas, C., Xenos, M., Panagiotakopoulos, C. and Vergidis, D., A comparative study of dropout rates and causes for two different distance education courses", *International Review of Research in Open and Distance Learning*, **5(2)** (2004), 1-15.

[45] Polat, E. and Gunay, S., A new approach to robust of partial least squares regression analysis", *International Journal of Mathematics Trends and Technology*, **9(3)** (2014), 132-141.

[46] Ratnaningsih D.J., Saefuddin A., Wijayanto H., An analysis of drop-out students survival in distance higher education", *Journal of Open Distance Education*, **9(2)** (2008), 01110.

[47] Ratnaningsih, D.J., Saefuddin, A., Kurnia, A. and Mangku, I.W., Stratified-extended Cox model in survival modeling of non proportional hazard", *IOP Conference Series: Earth and Environmental Science*, **299(1)** (2019), 012023, doi:10.1088/1755-1315/299/1/012023.

[48] Ratnaningsih, D.J., Saefuddin, A., Kurnia, A. and Mangku, I.W., Stratified-Extended Cox with Frailty Model for Non Proportional Hazard: A Statistical Approach to Student Retention Data from Universitas Terbuka in Indonesia", *Thailand Statistician Journal*, (2020), Article in Press (ID: TS 192490).

[49] Rovai, A.P., In search of higher persistence rates in distance education online program", *The Internet and Higher Education*, **6(1)** (2008), 1-16, doi:10.1016/S1096-7516(02)00158-6.

[50] Saegusa, T., Di, C. and Chen, Y.Q., Hypothesis testing for an extended cox model with time-varying coefficients". *Biometrics*. **70** (2014), 619628, doi:10.1111/bi-om.12185

[51] Schuemer, R., Some Psychological Aspects of Distance Education", *Hagen, Germany: Institute for Research into Distance Education*. **ED 357 266** (1993).

[52] Sembiring, M.G., Modeling the determinants of student retention in distance education institutions", *International Journal of Continuing Education and Lifelong Learning*, **6(2)** (2014), 15-28.

[53] Stroup, W.W., *Generalized Linear Mixed Models Modern Concepts, Methods and Applications*, New York: Taylor & Francis Group, 2013.

[54] Sylvestre, M.P., Edens, T., MacKenzie, T. and Abrahamowicz, M., *PermAlgo: Permutational Algoritm to Simulate Survival Data*, Package R can accessed in https://CRAN.R-project.org/package=PermAlgo, 2015.

[55] Therneau, T.M., Grambsch, P.M. and Pankratz, V.S., Penalized survival models and frailty", *Journal of Computational and Graphical Statistics*, **12(1)** (2003) , 156-175.

[56] Therneau, T. and Clinic, M., *Mixed Effect Cox Models*, [Accessed 10 Februari 2017]. On: https://cran.rproject.org/web/packages/coxme/vig-nettes/cox-me.pdf. 2015.

[57] Thomas, L. and Reyes, E.M., Tutorial: Survival estimation for Cox regression models with time-varying coefficients using SAS and R", *Journal of St0atistical Software*, **61** (2014), 1-2, doi: 10.18637/jss.v061.c01.

[58] Vaupel, J.W., Manton, K.G. and Stallard, E., The impact of heterogeneity in individual frailty on the dynamics Panof mortality", *Demography*, **16**(1979), 439-454.

[59] Wang, N., Xu, S. and Fang, J., Hierarchical likelihood approach for the Weibull frailty model", *Journal of Statistical Computation and Simulation*, **81(3)** (2011), 343-356.

[60] Wienke, A., *Frailty Models in Survival Data*, First Edition, New York: Chapman and Hall, 2011.

[61] Wu, L., *Mixed Effects Models for Complex Data (Monographs on Statistics and Applied Probability; 113)*, Boca raton: CRC Press, Taylor & Francis Group, 2010.

[62] Xenos, M., Pierrakeas, C. and Pintelas, P., A survey on student dropout rates and dropout causes concerning the students in the Course of Informatics of the Hellenic Open University", *Computers & Education*, **39(4)** (2002), 361 377.

[63] Yadav, A.K. and Yadav, R.J., A study on childhood mortality using shared frailty modeling approach", *Journal of Applied Sciences*, **16(1)**(2016), 11-17, doi: 10.3923/jas.2016.11.17.

[64] Yau, K.W., Multilevel models for survival analysis with random effects", *Biometrics*, **57(1)** (2001), 96-102.