The comparison of proportional hazards and accelerated failure time models in analyzing the first birth interval survival data

Alfensi Faruk
Department of Mathematics, Faculty of Mathematics and Natural Science, Sriwijaya University
E-mail: alfensifaruk@unsri.ac.id

Abstract. Survival analysis is a branch of statistics, which is focussed on the analysis of time-to-event data. In multivariate survival analysis, the proportional hazards (PH) is the most popular model in order to analyze the effects of several covariates on the survival time. However, the assumption of constant hazards in PH model is not always satisfied by the data. The violation of the PH assumption leads to the misinterpretation of the estimation results and decreasing the power of the related statistical tests. On the other hand, the accelerated failure time (AFT) models do not assume the constant hazards in the survival data as in PH model. The AFT models, moreover, can be used as the alternative to PH model if the constant hazards assumption is violated. The objective of this research was to compare the performance of PH model and the AFT models in analyzing the significant factors affecting the first birth interval (FBI) data in Indonesia. In this work, the discussion was limited to three AFT models which were based on Weibull, exponential, and log-normal distribution. The analysis by using graphical approach and a statistical test showed that the non-proportional hazards exist in the FBI data set. Based on the Akaike information criterion (AIC), the log-normal AFT model was the most appropriate model among the other considered models. Results of the best fitted model (log-normal AFT model) showed that the covariates such as women’s educational level, husband’s educational level, contraceptive knowledge, access to mass media, wealth index, and employment status were among factors affecting the FBI in Indonesia.

1. Introduction
In many practical problems related to time-to-event data, the PH model is widely known as the most used multivariate survival analysis to investigate the effects of several covariates on survival time. The PH model is semi-parametric, which means that the baseline hazard is left unspecified and there is no assumption about the statistical distribution of the survival times. The main assumption of the PH model is that the hazard ratio is constant over time. The violation of this assumption may lead to several statistical problems such as misinterpretation of the estimation results and reducing the power of the related statistical tests [1].

Another model, which also uses PH assumption, is the parametric PH model [2]. This model modifies the PH model by assuming that the baseline hazard function follows one of the statistical distributions. Although it is known that parametric PH model is appropriate to many clinical problems, the parametric PH model cannot be used as an alternative model to cope with non-proportional hazards problem. On
the other hand, the accelerated failure time (AFT) model is one of parametric survival models that can be used as an alternative to PH model, especially to overcome the statistical problems due to the violation of PH assumption [3]. The AFT model accounts the effects of the covariates directly on survival times instead of the hazards rate as in the PH model. Moreover, the interpretation of the results in AFT model is easier than the results of PH model because the parameters indicate the effect of the explanatory variables on the mean survival time.

The AFT models have been widely studied by many authors. According to [4], these models fit better to the influenza data than the PH model. The AFT models are also recommended in the analysis of lifetime data, especially in medical research [5]. The comparison of the AFT models and PH model for several numbers of censored observations and the sample size was conducted through simulation study [6]. The similarity among the mentioned works is that the use of AFT models and its comparison to the PH model were conducted in relatively small sample size which is commonly found in clinical research. However, in many real life applications, it is also common that the researchers have to analyze the large data set, such as in demographic and public health studies. Therefore, the main objective of this research is to compare the performance of the Cox PH model and the AFT models in the presence of non-proportional hazards in one of demographical and public health problems, that is the FBI data in Indonesia. The FBI can be defined as the length of time spent by a married couple to get their first child after the day of marriage. The short FBI indicates the high fertility rate of a spouse. In other words, the fertility rate can be reduced by lengthening the FBI. The considered models in this work are limited to the PH model and three types of AFT models, i.e. Weibull AFT model, exponential AFT model, and log-normal AFT model.

2. Materials and methods
The data used in this research were obtained from 2012 IDHS, which was conducted by Statistics Indonesia and United States Agency for International Development (USAID) from May 2012 until June 2012. In order to collect the data in all provinces of Indonesia, they used stratified two-stage cluster sampling technique, where the cluster blocks and the household were the first and the second stage unit sampling, respectively. The cluster blocks were based on the Indonesia population census that was held in 2010. The respondents were asked to answer the questions in the questionnaire. Not all respondents filled the questionnaire completely, therefore only the completed questionnaires were included.

To achieve the research objective, this work employed several steps. In the first step, the log cumulative hazard plots approach and Schoenfeld residual statistical test were applied to the data set in order to examine the existence of non-PH. Next, the PH model, the Weibull AFT model, exponential AFT model, and log-normal AFT model were fitted to the data set. After that, the performance of each model was evaluated by using the AIC values. Finally, as the log-normal AFT model was chosen as the best fitted model, the backward stepwise procedure was applied to select the significant covariates.

3. Basic concepts of the PH model and the AFT models
3.1. PH model
The PH model is a semiparametric regression model which can be used to measure the effects of covariates on the survival time. This model is represented by the relationship of the hazard function, the baseline hazard function, and one or more covariates in the form

\[ h(t) = h_0(t) \exp(\beta'X), \]

where \( t \) is survival time, \( h(t) \) is the hazard function, \( h_0(t) \) is the baseline hazard function which is left unspecified, \( \beta \) is a column vector of the regression coefficients, and \( X \) is a column vector of the covariates. The Cox regression model assumes that there is the proportionality of the hazard rate between any two individuals in the population. The hazard ratio is defined as the ratio of the hazard functions for two subjects with different values of covariate \( X_1 \) and \( X_2 \). The formula of the hazard ratio is given by
\[ H(t) = \frac{h_0(t) \exp(\beta_2 X_2)}{h_0(t) \exp(\beta_1 x_1)} = \frac{\exp(\beta_2 X_2)}{\exp(\beta_1 x_1)} = \exp(\beta^T (X_2 - X_1)) \].

It can be seen that the hazard ratio \( H(t) = \exp(\beta^T (X_2 - X_1)) \) is independent of time. In other words, the hazard ratio for any two individuals is constant over time. This property is also known as the PH assumption.

### 3.2. Partial likelihood (PL) method

Cox [7] introduced PL method as the procedure to estimate the regression coefficients of PH model. This method uses the partial likelihood function instead of the likelihood function as in the maximum likelihood estimation (MLE) method [8]. The first step in PL method is by arranging the observed time \( t_i, \ i = 1, 2, \ldots, k \) in ascending order so that the times are ordered, that is \( 0 \leq t_1 \leq t_2 \leq \cdots \leq t_k \). Next, let \( X_i \) is a vector of the summation of all \( p \) covariates for the individuals who experience the event of interest at time \( t_i \). If there are \( d_i \) occurrences of the event at time \( t_i \), the \( h^{th} \) element of \( X_i \) is \( X_{hi} = \sum_{r=1}^{d_i} x_{hir} \), where \( x_{hir} \) is the value of the \( h^{th} \) covariate for the \( r^{th} \) individual of \( d_i, h = 1, 2, \ldots, p \), and \( r = 1, 2, \ldots, d_i \).

The approximation of the partial likelihood function in the case that it allows more than one individual who experiences the event of interest at the observed time \( t_i \) is expressed by

\[ L_p(\beta) = \prod_{i=1}^{k} \left( \frac{\exp(\beta^T x_i)}{\sum_{j \in Y_i} \exp(\beta^T x_j)} \right)^{\delta_i} \],

where \( \delta_i \) is an indicator variable whose value equal to one when the event occurs or equal to zero when the event is censored, and \( Y_i \) is the number of individuals who experiences the event at or after time \( t_i, \ i = 1, 2, \ldots, k \). There are two reasons why equation (1) is called a partial likelihood function, i.e., it is not a full likelihood for \( \beta \) and it does not use full data but only use their ranking. The related log-likelihood function of the partial likelihood function (1) is given by

\[ \log L_p(\beta) = \sum_{i=1}^{k} \delta_i \left[ \exp(\beta^T x_i) - \sum_{j \in Y_i} \exp(\beta^T x_j) \right] \log \left( \sum_{j \in Y_i} \exp(\beta^T x_j) \right) . \]  

Then, by maximizing the log-likelihood function (2), the estimated value of \( \beta \) may be obtained by using a numerical approximation such as Newton-Raphson method [10].

### 3.3. Accelerated failure time model

Let \( T \) is a random variable of survival times and \( X \) is a column vector of the covariates \( X_1, X_2, \ldots, X_p \), the AFT model defines the relationship of survival function for every time \( t \in T \), \( S(t|X) \), and the covariates as follows

\[ S(t|X) = S_0 \left[ t \exp(\beta^T X) \right] . \]

where \( S_0 \) is the baseline survival function and \( \beta^T = (\beta_1, \beta_2, \ldots, \beta_p) \) is a vector of regression coefficients, and \( n \in N \). The factor \( \exp(\beta^T X) \) in equation (3) is known as the accelerated factor which accelerates the survival function with covariate \( X = 0 \). The AFT model assumes that the effects of the covariate are fixed and multiplicative by the accelerated factor on the time scale of \( t \). However, it does not assume that the model holds the constant hazards assumption as in PH model.

The relationship between covariates and the survival time can be also illustrated as a linear relation between the natural logarithm of survival time and the covariate \( X \), that is

\[ Y = \ln T = \mu + \theta^T X + \sigma W , \]

where \( \mu \) is the slope, \( \sigma > 0 \) is an unknown scale parameter, \( \theta^T = (\theta_1, \theta_2, \ldots, \theta_p) \) is a vector of regression coefficients, \( \theta = -\beta, \sigma \) is a scale parameter, and \( W \) is a distribution error which is a random variable and assumed to follow a certain parametric distribution. For every distribution of \( W \), there is a related
parametric for $T$. The name for the AFT model come from the distribution of $T$ rather than the parametric distribution of $\ln T$. The commonly parametric distributions, which correspond to the AFT model, are Weibull, exponential, log-normal, gamma, and log-logistic. However, the AFT models that are considered in this research are Weibull AFT model, Exponential AFT model, and log-normal AFT model. The survival function of $T_i$, $i = 1, 2, \ldots, n$ is given by

$$S_i(t) = \Pr(T_i \geq t) = \Pr(\ln T_i \geq \ln t) = \Pr(\ln(\ln t) \geq \ln t) = \Pr(\ln t \geq \frac{\ln t - (\mu + \theta X)}{\sigma}).$$

3.4. Weibull and exponential AFT models

The Weibull distribution is one of the most popular distributions in the analysis of time-to-event data. It is a flexible model that the hazard rate can be one of the three functions namely monotone increasing, constant, or decreasing function.

The survival function of survival time $T$ for the Weibull distribution with parameters $\lambda$ and $\alpha$ can be written by $S(t) = \exp(-\lambda t^\alpha)$, and the hazard function is expressed by $h(t) = \lambda \alpha t^{\alpha - 1}$. If the $X$ in equation (4) equal to zero, then it is obtained $Y = \mu + \sigma W$, where $\mu = (-\ln(\lambda) / \alpha), \sigma = \frac{1}{\alpha} = \alpha^{-1},$ and $W$ follows the standard extreme value distribution with probability function

$$f(w) = e^{(w-e^w)}.$$  (5)

By integrating the probability density function (5) with respect to variable $u$ on the interval from $w$ and $\infty$, the survival function of $W$ is $S(w) = \int_w^\infty f(u) du = \int_w^\infty e^{(u-e^u)} du = e^{(-e^w)}$. Then, the hazard function of $W$ is $h(w) = f(w)/S(w) = e^{(w-e^w)}/e^{(-e^w)} = e^w$.

The univariate probability function $Y = \ln(T)$ is

$$f(y) = \frac{1}{\sigma} \exp[(y-\mu)/\sigma - \exp[(y-\mu)/\sigma]],$$

and survival function for $Y$ is

$$S(y) = \exp\left[-\exp\left(\frac{y-\mu}{\sigma}\right)\right].$$  (6)

Equation (6) can be extended to multivariate case, that is by considering a set of $p$ covariates, $X = (X_1, X_2, \ldots, X_p)^T$ into the models. As the result, the survival function can be expressed by

$$S(y|X) = \exp\left[-\exp\left(\frac{y-(\mu + \theta^T X)}{\sigma}\right)\right].$$  (7)

The exponential distribution can be derived from Weibull distribution, that is by taking $\sigma = 1$ or $\alpha = 1$, so that the equation (6) and equation (7) become $S(y) = \exp[-\exp(y - \mu)]$ and $S(y|X) = \exp[-\exp(y - (\mu + \theta^T X))]$, where $S(y)$ and $S(y|X)$ are the survival functions for exponential AFT model in univariate and multivariate case, respectively.

3.5. Log-normal AFT model

If the random variable of the survival times $T$ is assumed to follow a log-normal distribution, the baseline survival function $S_0(t)$ and baseline hazard function $h_0(t)$ are expressed by

$$S_0(t) = 1 - \Phi\left(\frac{\ln t - \mu}{\sigma}\right),$$

and

$$h_0(t) = \frac{\phi\left(\frac{\ln t}{\sigma}\right)}{1 - \Phi\left(\frac{\ln t}{\sigma}\right)},$$

where $\Phi$ is the cumulative distribution function of the standard normal distribution.
respectively, where $\mu$ and $\sigma$ are parameters, $\varphi(x)$ is density probability function at time $x$ and $\Phi(x)$ is the cumulative distribution function at time $x$. Therefore, the survival function given a set of covariates $X = (X_1, X_2, ..., X_p)^t$ for the log-normal AFT model is $S(t) = S_0[t \exp(\beta^t X)] = 1 - \varphi((\ln t - (\mu + \theta^t X))/\sigma)$, where $t \in T$ is survival time, and $\beta^t = (\beta_1, \beta_2, ..., \beta_p)$ is a vector of coefficients.

### 3.6. Parameters estimation procedure of AFT models

Similar to other parametric models, the AFT models are also can be estimated by using MLE. The likelihood function of $n$ observed times $t_1, t_2, ..., t_n$, with unknown parameters $\beta^t = (\beta_1, \beta_2, ..., \beta_p), \mu,$ and $\sigma$, which contain $(n - r)$ right censored data, is expressed by

$$l(t; \beta, \mu, \sigma) = \prod_{i=1}^{n} f_i(t_i)^{\delta_i} * S_i(t_i)^{1-\delta_i},$$  

where $0 \leq r \leq n$, $\delta_i$ is indicator variable which is equal to one if $t_i$ is observed and equal to zero if $t_i$ is a censored observation. $f_i(t_i)$ and $S_i(t_i)$ are the density function and survival function at time $t_i$, respectively. By taking the logarithm of equation (8), it yields the log-likelihood function as follows

$$\log l(t; \beta, \mu, \sigma) = \sum_{i=1}^{n} (-\delta_i \log f_i(t_i) + \delta_i \log f_i(t_i) + (1 - \delta_i) \log S_i(W_i)),$$

where $W_i = \left(\log \frac{\mu + \beta_1 X_{1i} + \beta_2 X_{2i} + \cdots + \beta_p X_{pi}}{\sigma}\right)$. $X = \{X_{ji}\}$ is a column vector of covariates for the $j^{th}$ subject.

The estimated values of unknown parameters $\beta, \mu,$ and $\sigma$ which maximize the log-likelihood function (9) may be determined by using Newton-Rhapson method [10]. For each parametric distribution, the $f_i(t_i)$ and $S_i(t_i)$ in equation (9) are substituted with the corresponding functions. For example in log-normal AFT model, the $S_i(t_i)$ in equation (9) should be substituted with the $S(t)$ from equation (7).

### 4. Results

#### 4.1. Description of the data set

In this study, a total of 27488 ever-married women aged 15-49 at the time of interview from 33 provinces in Indonesia were selected for further analysis. Of the total of respondents, there were 25564 (93%) women who gave birth and 1924 (7%) women who did not give birth until the end of the interview. The FBI of ever-married women aged 15-49 years old was defined as the response variable and measured in months.

The research subjects were limited to the women who were not pregnant and did not have a child at the time of their first marriage. It means that the less than nine months and negative FBI were not considered as the right censored data. Eight covariates which consisted of one continuous covariate and seven categorical covariates were used in this work. These covariates were socioeconomic and demographic characteristics of the respondents (table 1).

#### 4.2. Checking PH assumption by log cumulative hazard plots

In order to identify the presence of non-proportional hazards, it is common to use log cumulative hazard plots. This procedure can be implemented by plotting $\log(t)$ versus $\log(-\log(S(t)))$ for all groups in every categorical covariate, where $t$ and $S(t)$ are survival time and survival function at time $t$, respectively. If the plot does not yield the parallel curves separated by $\beta$ over the log(time), then the PH assumption is not correct. In other words, the PH assumption is violated in the model.

The difficulty of this approach will arise if a covariate has too many categories or infinite categories as in continuous covariate. This problem can be solved by dividing the subjects into a few categories. However, this procedure is not applied in this work because the presence of non-proportional hazards can be shown by just checking the categorical covariates. The plots of all categorical covariates based on the data set are depicted in figure 1. These plots were built by using package ‘survival’ with plot() and survfit() functions in RStudio version 1.1.383.
Table 1. Respondents’ characteristics based on 2012 IDHS

| Covariate                      | Categories                                | Measurement Scale |
|--------------------------------|-------------------------------------------|-------------------|
| Age at the first birth         | Continuous                                |                   |
| Place of residence             | Rural                                     | Nominal           |
|                                | Urban                                     |                   |
| Women's educational level      | No education                              | Ordinal           |
|                                | Primary                                   |                   |
|                                | Secondary & above                         |                   |
| Husband's educational level    | No Education                              | Ordinal           |
|                                | Primary                                   |                   |
|                                | Secondary & above                         |                   |
| Contraceptive knowledge        | No                                        | Nominal           |
|                                | Yes                                       |                   |
| Access to mass media           | No                                        | Nominal           |
|                                | Yes                                       |                   |
| Wealth index                   | Poor                                      | Ordinal           |
|                                | Middle & above                            |                   |
| Employment status              | No                                        | Nominal           |
|                                | Yes                                       |                   |

Figure 1. The log cumulative hazard plots of categorical covariates. The curves in (a), (b), (c), (e), (f), and (g) are not parallel. The parallel curves are shown only in (d).
Based on figure 1, the plot that yields two or more parallel curves is just for contraceptive knowledge. Meanwhile, the plots for the other categorical covariates are clearly not parallel because there are at least two curves which intercept each other. It means that by using log cumulative hazard plot, the PH assumption is violated for modelling the FBI data in Indonesia.

4.3. Checking PH assumption by the goodness of fit (GOF) testing approach

The GOF testing approach is very attractive because it provides the statistical tests in order to check the presence of the non-PH. Consequently, the decision about PH assumption is more objective than the log cumulative hazard plots approach which is already discussed in the previous subsection.

One of the most popular GOF testing methods is a GOF test based on the Schoenfeld residual [9].

The basic idea of the test is that if the PH assumption for a covariate is not violated, there is no correlation between the Schoenfeld residuals and the survival time. In other words, it will test the hypothesis \( H_0: \rho = 0 \) (there is no correlation between the covariate and time). If the hypothesis \( H_0 \) is rejected, it can be concluded that the covariate does not hold the PH assumption. In this research, the results of GOF testing with four different forms of survival time for all covariates in the FBI data are obtained by using package ‘survival’ in RStudio version 1.1.383 (table 2).

| Covariate                          | P-value | Kaplan-Meier | Log(time) | Time ranking | Time |
|------------------------------------|---------|--------------|-----------|--------------|------|
| Age at the first birth             | 0.00^a  | 0.00^a       | 0.00^a    | 0.00^a       |
| Place of residence                 | 0.068   | 0.039^a      | 0.068     | 0.086        |
| Women's educational level          | 1.46e-09| 8.80e-08^a   | 1.35e-09^a| 0.001^a      |
| Husband's educational level        | 0.003^a | 0.002^a      | 0.003^a   | 0.01^a       |
| Contraceptive knowledge            | 0.373   | 0.221        | 0.380     | 0.263        |
| Access to mass media               | 0.481   | 0.190        | 0.497     | 0.041^a      |
| Wealth index                       | 0.456   | 0.448        | 0.455     | 0.425        |
| Employment status                  | 0.683   | 0.553        | 0.690     | 0.337        |
| (Global test)                      | 0.00^a  | 0.00^a       | 0.00^a    | 0.00^a       |

^a: significant at 5% level

In this study, the four different forms of survival time are Kaplan-Meier (KM) transform, log(time) transform, time ranking (or ordering the time), and time (or no transformation). In general, the tests are twofold namely per variable test and global test. From table 2, it can be seen that the per variable tests for all transformation forms consist of at least one significant covariate at 5% level. It means that there is no correlation between the covariate and a specific transformation form. In other words, such covariate does not hold the PH assumption. Moreover, the global tests for all considered survival time transformation forms show the strong evidence of non-PH (p-value = 0.00).

4.4. Fitting the PH model and the AFT models

It has already shown in previous subsections that the PH assumption is violated in the FBI data. Therefore, the AFT models, which has no assumption about PH, are used as the alternative to PH model.

In this research, partial likelihood and MLE approach are employed to estimate the regression coefficients of the PH model and AFT models, respectively. The computation of the estimation of both models for FBI data in Indonesia is performed by RStudio version 1.1.383 with package ‘survival’ and package ‘flexsurv’ [10]. The parameters estimation and its corresponding p-values of PH model, Weibull AFT model, exponential AFT model, and log-normal AFT model are given in table 3.
insignificant at 5% level for Weibull AFT model and exponential AFT model. The place of residence (urban) and access to mass media (yes) are insignificant at 5% level for log-normal AFT model.

Table 3. Results of fitting the PH model and the AFT models

| Categories                  | PH model | Weibull AFT model | Exponential AFT model | Log-normal AFT model |
|-----------------------------|----------|-------------------|-----------------------|----------------------|
| (Intercept)                 | -        | -                 | -                     | -                    |
| Age at the first birth      | -        | 4.493             | 0.00 a                | 4.449                | 0.00 a                |
| Place of residence          | Rural b  | 1.004             | 0.005 a               | -0.002               | 6.28e-02              |
|                             | Urban b  | 0.993             | 0.604                 | 0.039                | 2.23e-03 a            |
| Women's educational level   | Primary b| 1.079             | 0.032 a               | -0.122               | 1.30e-04 a            |
|                             | Secondary & Above | 1.267         | 0.97e-10 a            | -0.295               | 1.06e-18 a            |
| Husband's educational level | Primary b| 1.151             | 0.001 a               | -0.209               | 3.87e-08 a            |
|                             | Secondary & Above | 1.224         | 3.14e-06 a            | -0.277               | 1.11e-12 a            |
| Contraceptive knowledge     | Yes b    | 1.706             | < 2e-16 a             | -0.652               | 1.03e-34 a            |
| Access to mass media        | No b     | 1.017             | 0.463                 | -0.020               | 3.29e-01              |
| Wealth index                | No b     | 1.045             | 0.002 a               | -0.027               | 3.91e-02 a            |
| Employment status           | Yes b    | 0.938             | 5.53e-07 a            | 0.095                | 3.96e-16 a            |
|                             | No b     | -                 | -                     | -                    | -                    |

^a Significant at 5% level  
^b Reference category

In many cases, it is also important to compare the efficiency of different models. AIC is one of the common criteria to select the best model. AIC can be computed by

\[
AIC = -2 \log(\text{likelihood}) + kp, \quad (10)
\]

where \( \log(\text{likelihood}) \) is the loglikelihood function, \( p \) is the number of parameters, and \( k \) is the predetermined constant (we shall take as 2). The AIC values based on equation (10) of the PH model and the three AFT models for the data set are computed by using RStudio version 1.1.383 with package ‘survival’ (table 4).

Table 4. Comparison of the PH model and the AFT models using AIC criteria

| Model                  | AIC Value |
|------------------------|-----------|
| PH model               | 477950.2  |
| Weibull AFT model      | 221728.4  |
| Exponential AFT model  | 222249.5  |
| Log-normal AFT model   | 207266.9  |

The AIC values in table 4 are obtained in the case of the full model which means that all covariates are included in the model. According to table 4, the AIC value of PH model is 477950.2 and it doubles the average AIC values of all proposed AFT models. It proves that the AFT models are much better than the PH model to fit the data set. In addition, it can be seen also in table 4 that the log-normal AFT model has the lowest AIC value. In other words, the log-normal AFT model is the best fitted model among the other considered models.
4.5. Variable selection using backward stepwise

In this subsection, the backward stepwise procedure is applied to select the best combination variables for log-normal AFT model. This procedure is implemented by using package ‘survival’ with step() function in RStudio version 1.1.383 (table 5). In addition, the computation also yields that the AIC value of the log-normal AFT model in table 5 is 207265, which is less than the full log-normal AFT model (AIC= 207266.9) as shown in table 4. It means that the log-normal AFT model in table 5 is better than the full model in table 4.

Table 5. The best log-normal AFT model based on the backward stepwise procedure

| The Estimate of Coefficients | Women's educational level (primary) | Women's educational level (secondary & above) | Husband's educational level (primary) | Husband's educational level (secondary & above) | Contraceptive knowledge (yes) | Access to mass media (yes) | Wealth index (middle & above) | Employment status (yes) |
|------------------------------|-------------------------------------|-----------------------------------------------|--------------------------------------|-----------------------------------------------|-----------------------------|---------------------------|---------------------------|--------------------------|
| (Intercept)                  | 3.903                               | -0.083                                        | -0.203                               | -0.133                                        | -0.445                      | -0.027                     | 0.046                     | -0.009                   |

5. Discussion

Published studies in various practical problems tend to use the PH model rather than the AFT models. However, there are only a few of the studies which investigated the required PH assumption [11]. The violation of the PH assumption can cause the model to be unreliable and biased [12]. Therefore, the AFT models can become one of the solutions to deal with the presence of non-PH in the data set. Moreover, several works have proved that the AFT models better than the PH model to fit the data set with non-PH [3], [4], [6].

In order to check the violation of PH assumption in the FBI data set, this work applied two methods namely log cumulative hazard plots and the GOF test based Schoenfeld residual. The results of both methods showed that the non-PH existed in the data set (figure 1 and table 2). This finding is similar to [13], which also showed the presence of non-PH in the data set.

In this study, the results of the PH model and the AFT models were compared to analyze the FBI data in Indonesia. To compare these models, this work used the AIC value. The estimation of the full models (table 3) showed that the PH model and the AFT models had no much differences in testing the significance of the covariates. However, the AIC value of PH model (AIC= 477950.2) approximately doubles the AIC values of each considered AFT models, i.e. Weibull AFT model (AIC= 221728.4), exponential AFT model (AIC=22249.5), and log-normal AFT model (AIC=207266.9). It means that, in this circumstance, the performance of the AFT models is much better than the PH model. Based on the AIC value, it can be seen also that the log-normal is the best fitted model among the other considered models. These results are consistent with the results of the other works [5], [14], which showed that the log-normal AFT model was the best fitted model to their data set.

The further analysis of the best fitted model (the log-normal AFT model) showed that the variables of women’s educational level (primary), women’s educational level (secondary & above), husband’s educational level (primary), husband’s educational level (secondary & above), contraceptive knowledge (yes), access to mass media (yes), wealth index (middle & above), and employment status (yes) (table 5). This result is different with the result by [15], which concluded that age at the first birth is significant at 10% level. However, the significance of employment status, contraceptive knowledge, and women education are consistent with [16].

6. Conclusions

In this paper, the PH model and three AFT models were applied to the FBI data in Indonesia. Based on the log cumulative hazard plots and GOF test, it was clear that the constant hazards assumption for the FBI data in Indonesia was violated. Although the procedure of the log cumulative hazard plots was relatively simple, it was difficult to make the decision about the PH assumption for any covariate which consists of many categories. This drawback could be overcome by applying GOF test which provides more powerful results using a statistical test. The results showed that the log-normal AFT model had the
lowest AIC value comparing to the PH model, Weibull AFT model, and exponential AFT model. It means that the log-normal AFT is the best fitted model for the FBI data in Indonesia among the other considered models.

References
[1] Schenper M 1992 Cox analysis of survival data with non-proportional hazard functions J. R. Stat. Soc. Series D The Stat., 41 455-465
[2] Lee E T and Wang J W 2003 Statistical methods for survival data analysis (New Jersey- John Wiley & Sons, Inc.)
[3] Wei L J 1992 The accelerated failure time model: a useful alternative to the Cox regression model in survival analysis Stat. Med. 11 1871-79
[4] Kay R and Kinnersley 2002 On the use of the accelerated failure time model as an alternative to the proportional hazards model in the treatment of time to event data : a case study in influenza Drug Inf. J. 36 571-79
[5] Khanal S P, Sreenivas V and Acharya S K 2014 Accelerated failure time models : an application in the survival of acute liver failure patients in India Int. J. Sci. Res. 3 161–66
[6] Orbe J, Ferreira E and Núñez-Antón V 2002 Comparing proportional hazards and accelerated failure time models for survival analysis Stat. Med. 21 3493–3510
[7] Cox D R 1972 Regression models and life-tables J. R. Stat. Soc. Ser. B 34 187–220
[8] Faruk A 2014 Estimasi parameter data tersensor tipe I berdistribusi log-logistik menggunakan maximum likelihood estimate dan iterasi Newton-Rhapson Proc. Seminar Nasional MIPA 2014 Palembang pp 21-25
[9] Schoenfeld D 1982 Partial residuals for the proportional hazards regression model Biometrika 69 239–41
[10] Jackson C H 2016 Flexsurv: a platform for parametric survival modeling in R J. Stat. Softw. 70 1–33
[11] Altman D G, De Stavola B L, Love S B and Stepniewska K A 1995 Review of survival analyses published in cancer journals Br. J. Cancer 72 511–18
[12] Zare A, Hosseini M, Mahmoodi M, Mohammad K, Zeraati H and Naieni K H 2015 A comparison between accelerated failure-time and cox proportional hazard models in analyzing the survival of gastric cancer patients Iran. J. Public Health 44 1095–1102
[13] Rahman M A 2015 Fitting time to first birth using extended Cox regression model in presence of non-proportional hazard Dhaka Univ. J. Sci. 63 25-30
[14] Adelian R, Jamali J, Zare N, Ayatollahi S M T, Pooladfar G R and Roustei N 2017 Comparison of Cox’s regression model and parametric models in evaluating the prognostic factors for survival after liver transplantation in Shiraz during 2000–2012 Int. J. Org. Transplant Med. 6 119-125
[15] Faruk A 2017 Aplikasi regresi Cox pada selang kelahiran anak pertama di provinsi Sumatera Selatan J. Mat. 7 19-29
[16] Gebeyehu A 2015 Survival analysis of time-to-first birth after marriage among women in Ethiopia : application of parametric shared frailty model (Jimma- Jimma University)