Weibull Regression and Stratified Cox Regression in Modelling Exclusive Breastfeeding Duration

Ita Wulandari¹*, Anang Kurnia², and Kusman Sadik²

¹Polytechnic of Statistics STIS, Jl. Otto Iskandardinata No. 64C, Jakarta 13330, Indonesia
²Statistics Department, Bogor Agricultural University, Jl. Raya Dramaga Bogor, 16680, Indonesia

*ita.wulandari@stis.c.id

Abstract. Survival analysis has three approaches, parametric, non-parametric, and semi-parametric. The parametric model requires the distribution of survival time to be known. Weibull regression is one of the most popular forms of parametric approaches with advantages in flexibility and simplicity of hazard functions and survival functions. Stratified Cox regression is a semi-parametric approach method that is rightly used when the proportional hazard assumption is violated. The comparison between Weibull regression and stratified Cox regression was applied in the case of the duration of exclusive breastfeeding in infants aged 0-6 months in Indonesia. Before the model was formed, the data were tested for the Weibull distribution and the results were met properly. Whereas for testing proportional hazard assumptions the results are not fulfilled so that semi-parametric models can be used. Based on the results of the study, the Weibull regression is better than the stratified Cox regression.

1. Introduction

Survival analysis is a statistical analysis that is specifically used to analyze data or cases related to the time or length of time until a particular event occurs [1]. Here, time can be expressed in days, weeks, months, or years from the beginning of the observation on an individual until the occurrence of an event. Observed events can be in the form of negative events such as death, the occurrence of an illness, or positive events such as healing someone after surgery or when they get a job after a long time unemployed.

There are three approaches in survival analysis, parametric, non-parametric and semi-parametric. Parametric methods require that the distribution of survival times be known. Distributions that are often used in survival analysis are exponential, Weibull, and logistical. Despite the useful parametric approach, the nonparametric approach remains in practice, a most valuable, reliable, and frequently used descriptive tool. Undoubtedly, the product-limit estimator proposed by Kaplan and Meier is the most commonly employed. Furthermore, semi-parametric method, researchers do not need to look for the data distribution used, because no assumptions are underlying the distribution of survival time probabilities [2]. The assumption that it takes only proportional hazard, the function of hazard an individual against another individual hazard function [3].
If the proportional hazard assumptions are not met, which means that the linear component of the model varies depending on time, this is called non-proportional hazard. There are three options in overcoming these shortcomings, namely removing independent variables that do not meet the assumptions of the model, using the stratified Cox model, or by applying the extended Cox model.

The comparison between parametric, semi-parametric, and non-parametric models for inferencing analysis has been stated by Cox and Oakes [5]. The same study was carried out [6] by conducting three comparisons between non-parametric using the Kaplan Meier method, semi-parametric using the life table method, and parametric using the Exponential and Weibull distribution applied to health data. The semi-parametric approach is more appropriate than the parametric approach for breast cancer cases [7]. The next study concluded that concerning the Proportional Hazards models, Cox regression has better performance, while the connection with the parametric model, the Weibull distribution has the best performance among the models mentioned [8]. The comparison between stratified Cox regression and Extended Cox regression has also been carried out. The results show the stratified Cox regression is the best model.

Nutrition for infants aged 0-23 months can be divided into three conditions, exclusive breastfeeding, continued breastfeeding, and complementary feeding. In 2017, the coverage of exclusive breastfeeding in Indonesia for the first 6 months by mothers to their babies was still very low (35.7%). This number is still far from the target of exclusive breastfeeding coverage in 2019 determined by World Health Organization (WHO) and the Ministry of Health, which is 50%. We studied how parametric and semi-parametric methods differ in survival analysis of the exclusive breastfeeding duration. In this study, the parametric method uses the Weibull regression, while the semi-parametric method uses stratified Cox regression. The consideration used in comparing the two methods is that the Weibull regression has suitable in survival analysis, namely in terms of the flexibility and simplicity of the hazard function and the survival function [9]. Moreover, the model is one of the most popular forms of parametric regression models that provide estimates of the baseline hazard function [10]. Stratified Cox regression is often used and suitable when the proportional hazard is violated [11]. The differences between the two models will be studied, then the best model will be selected.

1.1. Survival Analysis
Survival analysis is a branch of statistics to analyze the expected duration of time until one or more events happen. This analysis focuses attention on a group consisting of a collection of individuals where each individual has a survival time, the time interval until the expected event occurs. These three things must be fulfilled to determine survival time (1) there is a start time, (2) a clear time measurement scale, (3) the definition of an event must be clear [5]. In survival analysis, data from individuals who have not yet experienced the desired event are referred to as censored data. Lee and Wang divide into three types, namely type I censored, type II censored, and type III censored [12]. There are three major times of censoring: right, left and interval-censoring [4]. The survival time distribution is usually expressed in three functions, namely the probability density function, the survival function, and the hazard function particularly.

1.2. Parametric Model
The parametric model assumes that the distribution underlying survival time follows a certain distribution. To choose the parametric model in survival analysis, several tests are needed. Particularly, two approaches can be done to measure the proximity of empirical and theoretical distribution, namely the graph approach and analysis approach. The popular first approach used is the probability plot (PP and QQ plot), and stabilized probability. The next approach that is most commonly used is Kolmogorov-Smirnov, Cramer-von Mises, and Anderson Darling. Whereas to choose the best model, Akaike's Information Criterion (AIC) and Bayesian Information Criterion (BIC) can be used.
1.2.1. Weibull Distribution

Weibull distribution is a special designation used to model data with extreme values. One example that is often used is the failure time data model. The probability density function of the Weibull distribution is:

\[
f(t) = \frac{\gamma}{\lambda} \left(\frac{t}{\lambda}\right)^{\gamma-1} \exp\left(-\left(\frac{t}{\lambda}\right)^\gamma\right), \quad t > 0
\]

where \( \lambda \) is the scale parameter and \( \gamma \) is the shape parameter, also known as the Weibull slope. While the cumulative distribution function is:

\[
F(t) = 1 - \exp\left(-\left(\frac{t}{\lambda}\right)^\gamma\right), \quad t > 0
\]

The survival function is in the equation below:

\[
S(t) = 1 - P(T \leq t) = \exp\left(-\left(\frac{t}{\lambda}\right)^\gamma\right)
\]

Another function is the hazard function. This function is calculated using the survival function and the probability density function, to obtain the following equation:

\[
h(t) = \frac{f(t)}{S(t)} = \frac{\gamma}{\lambda} \left(\frac{t}{\lambda}\right)^{\gamma-1} \quad \text{with} \quad \gamma > 0, \lambda > 0
\]

The Weibull model is the most widely used survival analysis model. If \( \gamma > 1 \) the hazard increases with time. If \( \gamma = 1 \) the hazard is constant. If \( \gamma < 1 \) the hazard decreases with time. This makes the Weibull model offer greater flexibility compared to the exponential method, but the hazard function remains relatively simple\(^{(13)}\).

According to\(^{(14)}\) the flexibility of the Weibull distribution can be seen easily. When \( \gamma = 1 \), the Weibull distribution becomes an exponential distribution with the hazard rate remaining constant over time. Currently \( \gamma = 2 \) following the Rayleigh distribution. When \( 3 \leq \gamma \leq 4 \), approaching the normal distribution and when \( \gamma \) large, for example, \( \gamma \geq 10 \), it will tend to approach the distribution of the smallest extreme values. Where \( \lambda \) is the scale parameter which is explained by the variables then Weibull regression as follows:

\[
\hat{\lambda} = \exp(\beta X) = \exp(\beta_0 + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k)
\]

After obtaining the Weibull regression and estimation of the survival function, the hazard function estimation can be obtained in the Weibull regression:

\[
\hat{h}(t) = \left(\frac{\gamma}{\lambda}\right) t^{\gamma-1}
\]

The parameters of the Weibull model are estimated using the maximum likelihood estimation (MLE) method. Suppose \( Y_1, Y_2, \ldots, Y_n \) is data with \( n \) number of observations and it is assumed that the \( Y_i \) are independent random variables, the joint probability distribution or the likelihood function is:

\[
f(y_1, y_2, \ldots, y_n; \lambda, \gamma) = L(\lambda, \gamma) = \prod_{i=1}^{n} \frac{\gamma}{\lambda} \left(\frac{y_i}{\lambda}\right)^{\gamma-1} \exp\left(-\left(\frac{y_i}{\lambda}\right)^\gamma\right)
\]

The log-likelihood function is:

\[
\log L(\lambda, \gamma) = \log\left\{ \prod_{i=1}^{n} \frac{\gamma}{\lambda} \left(\frac{y_i}{\lambda}\right)^{\gamma-1} \exp\left(-\left(\frac{y_i}{\lambda}\right)^\gamma\right) \right\}
\]

The next step is to get the first and second partial derivatives. The results obtained by the method are implicit, so to get the estimated parameters can be obtained using the Newton-Raphson iteration method.

1.3. Semi-parametric Model

In this model no assumptions are underlying the distribution of survival time probabilities, the assumptions needed are only proportional hazards, so called the semi-parametric model. The Cox regression model is a semi-parametric model that is widely used as a baseline hazard it need not be known, but the regression coefficients, the hazard ratio, and the curve of survival can be a good estimate.
1.3.1. Stratified Cox Regression Model

Let $T$ be a continuous variable that shows survival time and $X$ is a variable vector that is independent of time. In general, the Cox regression model is as follows:

$$ h(t, x) = h_0(t) \exp(\beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_p x_p) \tag{9} $$

where $h_0(t)$ is baseline hazard function at that time and $\beta = (\beta_1, \beta_2, \ldots, \beta_p)$ is a $1 \times p$ vector of regression parameters. Cox regression is based on several assumptions. One of them is the proportional hazard assumption which is the result of the following formula:

$$ HR = \frac{h(t, X_1)}{h(t, X_2)} = \frac{h_0(t) \exp(\beta X_1)}{h_0(t) \exp(\beta X_2)} = \exp(\beta [X_1 - X_2]) \tag{10} $$

where $X_1$ is the covariate vector of subject 1 and $X_2$ is the covariate vector of subject 2.

The proportional hazard assumption states that the hazard ratio for two subjects from different groups depends only on variable value and does not depend on time. The method used in testing proportional hazard assumptions is:

1. Graphical Approach
   Curves that are parallel between categories and do not intersect indicate the accomplishment of proportional hazard assumptions \cite{1}. One of them is by comparing the estimated curve $\ln(-\ln(S(t)))$ between categories of variables examined.

2. Test Statistics Using Goodness of Fit
   The proportional hazard assumption in a variable is considered fulfilled if the Schoenfeld residual in the variable does not depend on survival time.

   The importance of the proportional hazard assumption is stated by \cite{15} and suggests several models if the Hazard ratio is not constant over time for several variables. Stratified Cox regression model is one method that can be used if there are one or more variables that do not meet the proportional hazard assumption. This model modifies the Cox proportional hazard model by controlling variables that do not satisfy the proportional hazard assumption.

   The stratified Cox regression model modify by stratifying independent variables that do not satisfy the proportional hazard assumption. Independent variables that satisfy the proportional hazard assumption enter the model, while independent variables that do not satisfy the assumptions, which are being stratified, are not included in the model \cite{1}.

   In the model stratified Cox assumed there are as many as $p$ variables independent. Let $k$ variables satisfy the assumptions of which are non-proportional hazard denoted $X_1, X_2, \ldots, X_k$ with $k < p$. Independent variables that do not satisfy the proportional hazard assumption of $m$ are obtained from $p - k = m$ is $X_{1+k}, X_{2+k}, \ldots, X_p$ denoted $Z_1, Z_2, \ldots, Z_k$. Variables that do not meet the proportional hazard assumption $Z_i$ with $i = 1, \ldots, m$ are excluded from the Cox model to stratify the variable so that a stratifications variable $Z^*$ is obtained. Independent variables that meet the proportional hazard assumption will enter the stratified Cox model. Even so, the independent variables that are excluded from the model still have a role and by doing variable stratification it will be seen the contribution of each of the independent variables in different strata.

   The general form of the hazard function of stratified Cox regression model as follows \cite{1}:

$$ h_s(t, X) = h_{0s}(t) \exp(\beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_k X_k), s = 1,2, \ldots, m^* \tag{11} $$

where the subscript $s$ represents the strata. The strata are the different categorizations of the stratum variable. The variable $Z^*$ is not implicitly included in the model, whereas the $X$'s which are assumed to satisfy the proportional hazards assumption is included in the model. The baseline hazard function, $h_{0s}(t)$, is different for each stratum. However, the coefficients $\beta_1, \beta_2, \ldots, \beta_k$ are the same for each stratum. Since the coefficients of the $X$'s are the same for each stratum, the hazard ratios are the same for each stratum.

   The parameter estimation in the stratified Cox model uses the maximum stratified partial likelihood estimation method. In the Cox-stratified model, the probability of an individual event occurring to the $i$ on the stratum to $s$ at the time of failure will be determined. The partial likelihood function is as follows:
Regression parameter estimates $\beta^T = [\beta_1, \beta_2, ..., \beta_k]$ can be obtained by multiplying together the functions performed partial likelihood of each stratum, with each function partial likelihood of each stratum is derived from the function of hazard that is appropriate.

$$L_p(\beta) = \prod_{s=1}^{m^*} L_s(\beta) = L_1(\beta) \times L_2(\beta) \times \cdots \times L_m(\beta)$$

The form of the stratification log partial likelihood function is as follows:

$$\ln L_p (\beta) = \ln L_1(\beta) \times \ln L_2(\beta) \times \cdots \times \ln L_m(\beta)$$

To get the estimation of the regression parameters $\beta^T = [\beta_1, \beta_2, ..., \beta_k]$ by maximizing the partial likelihood function, which is by completing the derivative of the likelihood function of the partial likelihood with $\beta_g$ equal to zero. The equation obtained can then be solved numerically using Newton Raphson’s method.

2. Applications to Real Data

The data used in this study are secondary data derived from the results of 2017 Indonesian Demographic and Health Survey (IDHS) data collection designed to collect fertility, family planning, and maternal and child health data. Samples used in this study were 1,725 women. The observation units used were married and aged 15-49 years, married women who were less than 15 years old and had menstruated, and married women aged more than 50 years and still menstruating. The unit of analysis in this study was infants aged 0-6 months. Variables used are child sex (SEX), childbirth assistants (BIRTH), early initiation (INITIATION), area of residence (RESIDENCE), caesarean sections (CAESAR), wealth index (WI), maternal occupation status (MOS), maternal education level (MEL).

3. Results and Discussion

3.1. Overview Characteristics of Infant Age 0-6 Months

In our data, the percentage of infants aged 0-6 months who are assisted by non-healthcare worker is a very small proportion, in other words, more than 90 percent of babies assisted by healthcare worker at birth. Moreover, infants who did not get early initiation at the beginning of their birth were also high at 49.39, which means that almost half of mothers giving birth did not do so. The majority of infants aged 0-6 months in Indonesia are female, born in families live in rural areas and born through vaginal delivery. Characteristics of families who have babies aged 0-6 months in Indonesia are live in poor families (48.93%), his mother is not working (69.86%), and his mother is educated at the secondary level (56.41%).
Table 1. Sociodemographic characteristics of infants 0-6 months in Indonesia, 2017

| Variables                  | Description | Percentage (%) | Variables                  | Description | Percentage (%) |
|----------------------------|-------------|----------------|----------------------------|-------------|----------------|
| Child sex                  | Male        | 49.62          | Wealth index               | Poor        | 48.93          |
|                            | Female      | **50.38**      | Middle                     |             | 18.55          |
|                            | Health care | **90.90**      | Rich                       |             | 32.52          |
| Childbirth assistants      | Health care worker | 9.10           | Maternal occupation status | Not working | **69.86**      |
| Early initiation           | Yes         | **50.61**      |                             |             |                |
|                            | No          | 49.39          |                             |             |                |
| Area of residence          | Urban       | 46.03          | Maternal education level    | Primary     | 23.54          |
|                            | Rural       | **53.97**      |                             | Secondary   | **56.41**      |
| Caesarean sections         | Yes         | 17.74          |                             | Higher      | 18.78          |
|                            | No          | **82.26**      |                             |             |                |

3.2. Testing Data Distribution on The Parametric Approach

Based on the previous chapter, it is stated that the parametric approach in this study will use the Weibull distribution with several considerations. Previously, we compared the Weibull distribution by several distributions, including Exponential, Logistics, and Gamma distributions. The results showed that the Weibull distribution has the smallest AIC dan BIC values. Next will be displayed the results of data testing on the Weibull distribution.

![Cullen and Frey graph](image)

Tests using a graphical approach sometimes lead to subjectivity (figure 1 and figure 2), for this reason, statistical testing is done. Data distribution testing was performed using the Kolmogorov-Smirnov test. According to the results of this study, the \( p \)-value is greater than \( \alpha = 0.05 \) that which means the data used in this study follows the Weibull distribution.
Figure 2. Empirical and theoretical cdf’s for Weibull distribution

Table 2. Results goodness of fit statistics and the p-value for Weibull distribution

| Goodness-of-fit statistics   | Weibull |
|------------------------------|---------|
| Kolmogorov-Smirnov           | 0.19156 |
| p-value                      | 0.92016 |

3.3. Duration of Exclusive Breastfeeding Model Using Weibull Distribution

The Weibull regression model can be applied to the real data, based on the results of the graphical approach and statistical tests. Looking at table 3, significant variables are early initiation, maternal working status, and maternal educational level. The hazard ratio value in the early initiation was 0.5759 which means that the baby who didn’t initiation had 1.7364 times the possibility to stop being given exclusive breastfeeding compared to infants who received early initiation at the beginning of birth. The variable of the maternal occupation status has a hazard ratio of 1.2375 which means infants with non-working mothers have 1.2375 times the possibility to stop being given exclusive breastfeeding compared to infants with working mother status. When viewed from the maternal education level, the higher the level of mother’s education the tendency of mothers to stop giving exclusive breastfeeding to their babies the higher.

This regression model that is obtained based on a Weibull distribution as follows:

\[
\hat{\lambda} = \exp(0.0534 - 0.0408\text{SEX}(2) - 0.1839\text{BIRTH}(2) - 0.5517\text{INITIATION}(2) \\
+ 0.1017\text{RESIDENCE}(2) - 0.0348\text{CAESAR}(2) + 0.0499\text{WI}(3) + 0.0151\text{WI}(2) \\
+ 0.2131\text{MOS}(2) - 0.7101\text{MEL}(4) + 0.5408\text{MEL}(2) + 0.4427\text{MEL}(3)
\]

and probability density function:

\[
\hat{f}(t) = \frac{1.6416}{\hat{\lambda}^{2.5901}} t^{1.6416 - 1} \exp \left( - \left( \frac{t}{\hat{\lambda}} \right)^{1.6416} \right)
\]

Shape parameter estimates (\(\gamma\)) which is worth 1.6416 substituted on Weibull Regression Hazard function as follows:

\[
\hat{h}(t) = \left( \frac{1.6416}{\hat{\lambda}^{1.6416}} \right) t^{1.6416 - 1}
\]
Table 3. Results of the Weibull regression model

| Variables       | Coefficient | Standard error | Hazard ratio | Lower bounds | Upper bounds | p-value |
|-----------------|-------------|----------------|--------------|--------------|--------------|---------|
| Scale ($\lambda$) | 0.0534      | 0.0208         | -            | -            | -            | <2e-14  |
| Shape ($\gamma$) | 1.6416      | 0.0511         | -            | -            | -            | <2e-14  |
| SEX(2)          | -0.0408     | 0.0727         | 0.9601       | 0.8326       | 1.1070       | 0.575   |
| BIRTH(2)        | -0.1839     | 0.1282         | 0.8320       | 0.6471       | 1.0697       | 0.151   |
| INITIATION(2)   | -0.5517     | 0.0758         | 0.5759       | 0.4964       | 0.6682       | 9.6e-13 |
| RESIDENCE (2)   | 0.1017      | 0.0820         | 1.1071       | 0.9427       | 1.3000       | 0.215   |
| CAESAR(2)       | -0.0348     | 0.0940         | 0.9658       | 0.8033       | 1.1611       | 0.711   |
| WI(3)           | 0.0499      | 0.0987         | 1.0511       | 0.8662       | 1.2756       | 0.614   |
| WI(2)           | 0.0151      | 0.1036         | 1.0152       | 0.8287       | 1.2437       | 0.884   |
| MOS(2)          | 0.2131      | 0.0872         | 1.2375       | 1.0432       | 1.4680       | 0.015   |
| MEL(4)          | 0.7101      | 0.3743         | 2.0342       | 0.9767       | 4.2368       | 0.058   |
| MEL(2)          | 0.5408      | 0.3651         | 1.7174       | 0.8396       | 3.5130       | 0.139   |
| MEL(3)          | 0.4427      | 0.3644         | 1.5569       | 0.7622       | 3.1801       | 0.224   |

Description: Bold type is significant at: *: 0.05; **: 0.10; ***: 0.15

The hazard ratio value is obtained by calculating the exponential value of the estimated parameter value. The result of the study can be stated that babies of female sex (SEX(2)) are likely to experience a decreased rate of getting exclusive breastfeeding which is 0.9601 times faster than male babies. Another condition is that infants residing in intersection (RESIDENCE (2)) have the possibility of experiencing an increased rate of getting exclusive breastfeeding by 1.1071 times faster than infants residing in urban areas. For babies who have mothers with a higher education level (MEL(4)), the possibility of experiencing an increase in the rate of getting exclusive breastfeeding is 2.0342 times faster than babies who have mothers who do not attend school. Other conditions can be interpreted according to the hazard ratio obtained.

3.4. Testing Assumptions Proportional Hazard on Semi-parametric Approach

Procedures to check the assumption of proportional hazards is used to determine whether the data is suitable in a semi-parametric approach. In this study, a graphical approach and goodness of fit test are used to determine whether the variable is constant or changes depending on time.
Figure 3. Plot of the Kaplan Meier survival function curve for the data

The bottom line on each Kaplan Meier survival curve shows the time of survival of a female baby, birth attendant with medical staff, getting early initiation, living in a rural area, being born by means other than a cesarean, being in a rich family, working mother status, not have higher education compared to other categories.

The plot of the Kaplan Meier survival function curve (figure.3) shows that all variables used in each category coincide at certain points. To ensure there is a violation of the proportional hazard assumption an approach will be used the goodness of fit test.

The graphical approach sometimes produces different decisions for each researcher. For that, we need a test that can produce more precise decisions with the same standard. One approach that can be used is the Schoenfeld residual. The results of the tests are shown in Table 4.

Table 4. Test results using goodness of fit

| Variables      | p-value |
|----------------|---------|
| SEX(2)         | 0.7642  |
| BIRTH(2)       | 0.1336  ***|
| INITIATION(2)  | 0.0045  *|
| RECIDENCE(2)   | 0.1002  **|
| CAESAR(2)      | 0.7926  |
| WI(3)          | 0.3063  |
| WI(2)          | 0.4550  |
| MOS(2)         | 0.4533  |
| MEL<6(2)       | 0.2283  |
| MEL<2(2)       | 0.6109  |
| MEL<3(2)       | 0.04969 |
| GLOBAL         | 0.0094  *|

Description : Bold type is significant at : *: 0.05; **: 0.10; *** : 0.15

The simultaneous test shows that the p-value of 0.0094 is smaller than $\alpha = 0.05$, which means that there is at least one variable that does not meet the proportional hazard assumption. The partial test shows that variables that do not meet the proportional Hazards assumption are childbirth assistants, early initiation, and residential areas. The Schoenfeld residual graph can be used to see whether the variables used in the study meet the proportional hazard assumption or not. Looking at figure 4, each
variable tends to show a pattern, it can be concluded that the proportional hazard assumption is not met.

![Schoenfeld residual graph](image)

**Figure 4.** Schoenfeld residual graph

### 3.5. Exclusive Breastfeeding Duration Models with Semi-parametric Approach

There are three variables that do not meet the assumptions, childbirth assistants, early initiation, and area of residence. The three variables were stratified. The first step that must be used in determining the model was to test whether there are interactions between stratified variables, childbirth assistants, early initiation, and area of residence. The results show that there is no interaction between the three variables. The best model is the stratified Cox regression without interaction model with area of residence as strata.

Table 5. Results of stratified Cox regression without interactions with strata of the area of residence

| Variables | Coefficient | Standard error | Hazard ratio | Lower bounds | Upper bounds | p-value   |
|-----------|-------------|----------------|--------------|--------------|--------------|-----------|
| SEX(2)    | -0.0326     | 0.0727         | 0.9679       | 0.8387       | 1.1154       | 0.6540    |
| BIRTH(2)  | -0.2432     | 0.1286         | 0.7842       | 0.6145       | 1.0182       | 0.0587**  |
| INITIATION(2) | -0.6140    | 0.0757         | 0.5412       | 0.4663       | 0.6273       | 0.0000*   |
| CAESAR(2) | -0.0364     | 0.0936         | 0.9643       | 0.7985       | 1.1525       | 0.6974    |
| WI(3)     | 0.0144      | 0.0917         | 1.0146       | 0.8770       | 1.2946       | 0.8748    |
| WI(2)     | 0.0043      | 0.1016         | 1.0043       | 0.8417       | 1.2647       | 0.9660    |
| MOS(2)    | 0.2581      | 0.0874         | 1.2945       | 1.0884       | 1.5335       | 0.0031*   |
| MEL(4)    | 0.7688      | 0.3746         | 2.1572       | 1.0391       | 4.5165       | 0.0402*   |
| MEL(2)    | 0.6234      | 0.3652         | 1.8652       | 0.9124       | 3.8223       | 0.0879**  |
| MEL(3)    | 0.4905      | 0.3645         | 1.6331       | 0.8115       | 3.3939       | 0.1748    |

Description: Bold type is significant at: *, 0.05; **, 0.10; ***, 0.15
The results of the model without the interaction of the area of residence are shown in Table 5. Variables that have a significant influence on the model are delivery assistant, early initiation, mother's employment status, and mother's education level. The hazard ratio value in the birth attendant variable is 0.7842, which means infants with birth attendants who are not health workers have 1.2752 times the chance to stop being exclusively breastfed. The early initiation variable has a hazard ratio of 0.54, in other words, infants who do not receive early initiation at the onset of childbirth have a 1.8477 times chance of stopping exclusive breastfeeding compared to infants who receive early initiation. The variable of maternal employment status shows that babies with mothers who do not work are 1.2945 times more likely to stop being exclusively breastfed compared to mothers who are working. For the level of maternal education, the conclusion is the same as the previous model, namely the higher the level of education of the mother, the higher the tendency for mothers not to exclusively breastfeed their children.

The samples formed for Cox stratified regression with area of residence strata are:

**Urban model**

\[ h_1(t) = \hat{h}_0(t) \exp(-0.0326SEX_{(2)} - 0.2432BIRTH_{(2)} - 0.6140INITIATION_{(2)} - 0.0364CAESAR_{(2)} + 0.0144WI_{(3)} + 0.0043WI_{(2)} + 0.2581MOS_{(2)} + 0.7688MEL_{(4)} + 0.6234MEL_{(2)} + 0.4905MEL_{(3)}) \]

**Rural model**

\[ h_2(t) = \hat{h}_0(t) \exp(-0.0326SEX_{(2)} - 0.2432BIRTH_{(2)} - 0.6140INITIATION_{(2)} - 0.0364CAESAR_{(2)} + 0.0144WI_{(3)} + 0.0043WI_{(2)} + 0.2581MOS_{(2)} + 0.7688MEL_{(4)} + 0.6234MEL_{(2)} + 0.4905MEL_{(3)}) \]

In this study, the best model is determined from the smallest AIC and -2Loglikelihood values. Weibull regression is the best model with an AIC value of 4,058,625 and 2Loglikelihood of 4,032.6.

### Table 6. Results of AIC and -2Loglikelihood models of Weibull regression and stratified Cox regression

| Model | AIC         | -2*Loglikelihood |
|-------|-------------|------------------|
| **Parametric** |             |                  |
| Regresi Weibull | **4,058.625** | **4,032.6** |
| **Semi-Parametric** |         |                  |
| Stratified Cox Regression (Area of residence) | 11,007.170 | 10,977.6 |

Description: Bold type is the best model

**4. Conclusion**

In this study, the results of the graphical and the analysis approach to the Weibull models were fulfilled, whereas in the semiparametric approach the proportional hazard assumption is violated so that the Weibull regression and the stratified Cox regression model can also be applied. A comparison of the two models in the case of the duration of breastfeeding in infants aged 0-6 months in Indonesia in 2017 shows that the Weibull regression model is better than the stratified Cox regression model. The results in this study are not feasible to conclude in a general manner that the Weibull model is better than the stratified Cox model, so in subsequent studies simulation data can be used. To produce a better analysis of stratified Cox regression alternative approaches can be used by making separate estimates in the Hazard ratio log for each stratum, then specific estimates for stratum are combined for the entire inferencing using a sample size (strata weight).
References

[1] Kleinbaum D G and Klein M 2012 Survival Analysis: A Self-Learning (London: Springer Third Edition)

[2] Berglund A P 2011 An Overview of Survival Analysis using Complex Sample Data (Statistics and Data Analysis)

[3] Guo S 2010 Survival Analysis-Pocket Guides to Social Work Research Methods (New York:Oxford University Press Inc)

[4] Collet D 1994 Modelling Survival Data in Medical Research (London: Chapman and Hall)

[5] Cox D R and Oakes D 1984 Analysis of Data Survival Data (London: Chapman and Hall)

[6] Guvercin A C Y Tekindal M A Kaymaz O G and Guvercin C H 2017 Metodological Comparison of Survival Analysis Methods in Censored Medical Data (Biomedical Research vol. 28 Issue 10)

[7] Bouncher K M, Asselain B Tsodikov A D and Yakovlev A Y 2004 Semiparametric Versus Parametric Regression Analysis Based on the Bounded Cumulative Hazard Model; An Application to Breast Cancer Recurrence (New York: Springer Scince and Business Media)

[8] Georgousopoulou E N Pitsavos C Yannakoulia C M and Panagiotakos D B 2015 Comparisons between Survival Models in Predicting Cardiovascular Disease Events: application in the ATTICA Study (2002-2012) (Journal of Statistics Applications and Probability 4, No.2, pp 203-210)

[9] Crumer A M 2011 Comparison Between Weibull and Cox Proportional Hazards Models (Mahattan, Kansas: Kansas State University)

[10] Zhang Z 2016 Parametric Regression Model for Survival Data: Weibull Regression Model as an example (Big-Data Clinical Trial Column)

[11] Mehrotra D V Su S S and Li X 2012 An Efficient Alternative to The Stratified Cox Model Analysis (Statistics in Medicine John Wiley & Sons Ltd)

[12] Lee E T and Wang J W 2003 Statistical Methods for Survival Data Analysis (New Jersey: Third Edition, John Wiley and Sons Inc)

[13] Dobson A J 2001 An Introduction to Generalized Linear Models (London: Second Edition, Chapman and Hall/CRC)

[14] Nelson W 1982 Applied Life Data Analysis (New York: John Wiley & Son Inc)

[15] Hosmer D W and Lemeshow S 1999 Applied Survival Analysis-Regression Modelling of Time to Event Data. (New York: John Wiley & Sons Inc)

[16] Mehrotra D V and Raikar R A 2012 Minimum risk weights for comparing treatments in stratified binomial trials (Statistics in Medicine vol. 19(6)) pp 811-825

[17] Adejumo A and Ahmadu A O 2016 A Study of The Slope of Cox Proportional Hazard and Weibull Models: Simulated and Real Life Data Approach (Science World Journal vol 11 No. 3)

[18] Ata N and Sozer M T 2007 Cox Regression Models With Nonproportional Hazards Applied To Lung Cancer Survival Data (Journal of Mathematics and Statistics vol. 36 (2)) pp 157-167

[19] Lai Chin and Diew 2006 Weibull Distribution and Their Applications (Springer Handbook)

[20] McCullagh P and Nelder J A 1989 Generalized Linear Models (London: Second edition, Chapman and Hall)

[21] Muller M L D and Dutang C 2015) Fitdistplus: An R Package for Fitting Distribution (Journal of Statistical Software vol.64 Issue 4)

[22] Nielsen M A 2011 Parameter Estimation for the Two-Parameter Weibull Distribution (Provo: Brigham Young University)

[23] Nikulin M S Balakrishnan N Mesbah M and Limnios N 2004 Parametric and Semiparametryic Models with applications to Reliability, Survival Analysis and Quality of Time (New York: Statistics for Industry and Technology Springer Scince and Business Media)

[24] Richards J 2010 A Diagnostic Function to examine Candidate Distributions to Model Univariate
Data (Manhattan Kansas: Kansas State University)

[25] Schoenfeld D 1982 Partial Residuals for The Proportional Hazards Regression Model (Biometrika vol. 69, No. 1) pp 239-241

[26] Stanley C C Mukaka M and Molyneux 2016 Comparison of Performance of Exponential, Cox Proportional Hazards, Weibull and Frailty Survival Models for Analysis of Small Sample Size (Data Journal of Medical Statistics and Informatics vol. 2 Article 2) pp 2053-7662