The Cox proportional Hazard model on duration of birth process

Triastuti Wuryandari,1 Sri Haryatmi Kartiko,1 Danardono1
1Department of Mathematics and Natural Sciences, Universitas Gadjah Mada Yogyakarta, Indonesia
E-mail: triastutiwuryandari@yahoo.co.id, s kartiko@yahoo.com, danardono@ugm.ac.id

Abstract. The duration of birth process, which is measured from the birth sign until baby born, is one important factor to the whole outcome of delivery process. There is a method of birth process that given relaxing and gentle treatment to the mother caled as gentlebirth. Gentlebirth is a method of birth process that combines brain science, birth science and technology to empower positive birth without pain. However the effect of method to the duration of birth process is still need empirical investigations. Therefore, the objective of this paper is to analyze the duration of birth process using the appropriate statistical methods for durational data, survival data or time to event data. Since there are many variables or factor that may affect the duration, a regression model is considerated. The flexibility of the Cox Proportional Hazard Model in the sense that there is no distributional assumption required, makes the Cox Model as the appropriate model and method to analyze the duration birth process. It is concluded that the Gentlebirth method affects on duration of birth process, with Hazard Ratio of 2.073, showing that the duration of birth process with gentlebirth method is faster than the other method.

Keyword: duration, gentlebirth, survival, Cox regression, partial likelihood, hazard ratio

1. Introduction
Survival analysis is a statistical methods for studying the occurrence and timing of events (Allisson, 2010). Survival analysis is used in the fields of health, epidemiology, demographics and actuary. Survival data is also called time-to-event data, duration data in economics, failure time data in the
industry, event history data in the social called event history data. Survival analysis is used in the biomedical to determine the probability of patients with certain diseases will persist in a certain period. In the industry to know the probability of a component will work well up to a certain period. In the economic to know the probability of a company will remain in a good rating in a certain period.

The relationship between the time to event and a set of explanatory variable is employed in survival analysis as regression models. One of regression models in survival analysis is Cox Regression Proportional Hazard Model. The Cox Regression Proportional Hazard Model is often called semiparametric model because a parametric form is assumed only for the covariate effect and the baseline hazard is treated nonparametrically (Klein, 2003). In the Cox Proportional hazard model, covariates are assumed to have a multiplicative effect on the baseline hazard. The parameter estimate in this regression is based on partial maximum likelihood. The parameters in the Cox regression can be interpreted as Hazard Ratio (HR).

This study aims was investigate whether gentlebirth methods affect the duration of birth process with adjusted for other variable using the Cox proportional hazard regression model. Gentlebirth is a convenient methods of giving birth process without pain (Aprilia, 2013). Gentlebirth combines brain science, birth science and technology to empower the positive birth through physical and mental preparation of mother. Physical preparation includes breathing exercises, massage, and healthy food consumption. Mental preparation of the mother also needs to be prepared with a routine relaxation of hypnobirthing, meditation, positive affirmations, and keep the peace of his soul. Mental preparation of the mother becomes an important that will affect the success of this gentlebirth (Aprilia, 2013). In Indonesia, genthebirth usually uses waterbirth or hipnobirthing. Waterbirth is the birth process in a tub of warm water. Hypnobirthing is one method of self-hypnosis by giving suggestions or positive intentions so that mothers are able to enjoy the smoothness of the birth process.

Duration of birth process is one of application survival analysis. Some articles on Cox regression include Application of Cox Proportional Hazard Model to the Stock Exchange Market (Ni, 2009). Dehkodi et.al (2008) compare between Cox Proportional hazard model and logistic regression on prognostic factors in gastric cancer. Hanni and Wuryandari (2013) investigate the factor that affect the survival of cervical cancer patient using the Cox Proportional Hazard. Tolosie (2014) applies the Cox proportional hazards model in case of Tuberculosis. Handayani (2017) applies Survival analysis in patient with DHF using Cox Proportional Hazard Regression.

2. Methods
2.1. Cox Regression Proportional Hazard
The Cox Regression Model is a proportional hazard regression model with baseline hazard function modeled nonparametrically and the function of the independent variable is modeled parametrically. This model is also known as Cox proportional hazards model (Cox, 1972) or Cox Semiparametric hazards
model. Survival data usually involves more than one variable. If the independent variable more than one, then to perform the analysis of survival data then used regression model. Regression model that are quite popular in survival data analysis is Cox Proportional hazard Model.

Our data, based on a sample of size n, consists of the triple \((T_j, \delta_j, X_j(t))\); \(j = 1, \ldots, n\) where \(T_j\) is the time on study for the \(j\)th patient, \(\delta_j\) is the event indicator for the \(j\)th patient (\(\delta_j = 1\) if the event has occurred and \(\delta_j = 0\) if the lifetime is censored) and \(X_j(t) = (X_{j1}(t), \ldots, X_{jp}(t))'\) is the vector of covariates or risk factors for the \(j\)th individual at time \(t\) which may affect the survival distribution of \(X\). Here the \(X_{jk}(t)\); \(k = 1, \ldots, p\) may be time-dependent covariates whose value changes over time or may be constant values known at \(t = 0\). The Cox Model is as follows:

\[
h(t \mid X) = h_0(t) \exp \left( \sum_{k=1}^{p} \beta_k X_k \right)
\]

Cox proportional hazard used to predict the distribution of the time to some event from a set of explanatory variable (Klein, 2003). The Cox model is often called a proportional hazards model because, if we look at two individuals with covariate values \(X\) and \(X'\), the ratio of their hazard rates is constants. Hazard Ratio (HR) is define as

\[
HR = \frac{h(t \mid X)}{h(t \mid X')} = \frac{h_0(t) \exp \left( \sum_{k=1}^{p} \beta_k X_k \right)}{h_0(t) \exp \left( \sum_{k=1}^{p} \beta_k X_k \right)} = \exp \left( \sum_{k=1}^{p} \beta_k \left( X_k - X'_k \right) \right)
\]

Where \(h_0(t)\) is an arbitrary baseline hazard, and \(\beta = (\beta_1, \beta_2, \ldots, \beta_p)\) is parameter vector. The parameter estimation is based on partial maximum likelihood. Likelihood function is

\[
L(\beta) = \prod_{i \in D} \frac{\exp \left( \sum_{k=1}^{p} \beta_k X_{(i)k} \right)}{\sum_{j \in R(t_i)} \exp \left( \sum_{k=1}^{p} \beta_k X_{(j)k} \right)}
\]

Where \(D\) is the set of index \(j\) of all the time occurrences. \(R(t_i)\) is the set of individuals at risk. The set of risks in the partial likelihood is the set of all individuals who are likely to get events just before \(t\).

The log likelihood function from (3) is as follows

\[
\lambda(\beta) = \sum_{i \in D} \sum_{k=1}^{p} \beta_k X_{(i)k} - \sum_{i \in D} \ln \left( \sum_{j \in R(t_i)} \exp \left( \sum_{k=1}^{p} \beta_k X_{(j)k} \right) \right)
\]
The first derivation of (4) is called the score function as

\[
U_h(\beta) = \sum_{i=1}^{p} \sum_{k, j \in \mathcal{R}(i)} X_{(i)k} \exp \left( \sum_{l=1}^{p} \beta_l X_{jk} \right) \sum_{j \in \mathcal{R}(i)} \exp \left( \sum_{l=1}^{p} \beta_l X_{jk} \right)
\]

(5)

The partial maximum likelihood estimates are found by solving the set of \( p \) nonlinear equations \( U_h(\beta) = 0 \). This can be done numerically, using a Newton–Raphson technique from (4) and (5). The information matrix is the negative of the matrix of second derivatives of (4) and is given by

\[
I(\beta) = I_{gh}(\beta) \quad \text{with the (g,h)th element given by}
\]

\[
I_{gh}(\beta) = \sum_{j \in \mathcal{R}(i)} X_{jk} X_{jk} \exp \left( \sum_{l=1}^{p} \beta_l X_{jk} \right) \sum_{j \in \mathcal{R}(i)} \exp \left( \sum_{l=1}^{p} \beta_l X_{jk} \right) - \sum_{j \in \mathcal{R}(i)} X_{jk} \exp \left( \sum_{l=1}^{p} \beta_l X_{jk} \right) \left[ \sum_{j \in \mathcal{R}(i)} X_{jk} \exp \left( \sum_{l=1}^{p} \beta_l X_{jk} \right) \right]^{-1} \left( \sum_{j \in \mathcal{R}(i)} X_{jk} \exp \left( \sum_{l=1}^{p} \beta_l X_{jk} \right) \right)
\]

Where \( g,h = 1,2, ..., p \)

There are three tests for hypotheses about regression parameters. These are Likelihood Ratio, Wald Test and Score Test. Test statistic for each method is

A. Likelihood Ratio : \( \chi^2_{LR} = 2\left( \ell(\hat{\beta}) - \ell(\beta_0) \right) \)

B. Wald Test : \( \chi^2_{W} = (\hat{\beta} - \beta_0)^T I(\hat{\beta})(\hat{\beta} - \beta_0) \)

C. Score Test : \( \chi^2_{SC} = U(\beta_0)^T I^{-1}(\beta_0) U(\beta_0) \)

There are three methods to test the proportional hazard assumption. One of them is Goodness of Fit (GOF). The GOF approach uses Z test statistics for large samples or calculated Chi-Square for each model. P-value is calculated for each variable. If \( p\_value > \alpha \) then proportional hazard assumption is true (Kleinbaum, 2005).

2.2. Data

The research uses data from two clinics, these are “bidan kita” clinic which located in Klaten, Central Java with gentlebirth method and a clinic with the other method which located in Sukoharjo, Central Java. The number of samples are used 95 births, 54 mothers with gentlebirth and 41 mothers the other method. Dependent variable is the duration of the birth process. Independent variables are mother’s Age
(X1), Gravida (X2), Paritas (X3), Abortus (X4), Pregnancy age (X5), Baby’s Weight (X6), Baby Sex (X7) and Method in the birth process (X8). Gravida is the number of pregnancies, Paritas is the number of birth and Abotus is the number of abortion. X1, X2, X3, X4, X5, X6 are continue, while X7, X8 are categoric. Sex variable consists of two categories, these are boy and girl. Method variable consists of two categories, these are Gentlebirth (GB) and the other method.

3. Result

The parameter estimation, Hazard Ratio, inference for parameter and test of proportional hazard assumption in Table 1 and Table 2 that its data analysis using software R.

### Table 1: Parameter Estimation

| Variable      | Coef | Exp(coef) | Se(coef) | Z     | Pr(>|z|) |
|---------------|------|-----------|----------|-------|----------|
| Age           | 0.0083 | 1.0083 | 0.3021 | 0.274 | 0.78401 |
| Gravida       | -0.2694 | 0.7638 | 0.9450 | -0.285 | 0.77557 |
| Paritas       | 0.1559 | 1.1687 | 0.9314 | 0.167 | 0.86709 |
| Abortus       | 0.7047 | 2.0233 | 0.9972 | 0.707 | 0.47977 |
| Pregnancy age | -0.0946 | 0.9097 | 0.1097 | -0.862 | 0.38852 |
| Weight        | 0.0006 | 1.0006 | 0.0003 | 1.774 | 0.07613 |
| Sex P         | 0.2663 | 1.3052 | 0.2283 | 1.167 | 0.24334 |
| Method GB     | 0.7287 | 2.0725 | 0.2388 | 3.052 | 0.00227 |

The value Likelihood Ratio, Wald Test and Score Test is

A. Likelihood Ratio = 17.46 on 8 df, \( p=0.02568 \)
B. Wald Test = 18.19 on 8 df, \( p=0.01987 \)
C. Score(logrank) Test = 19.12 on 8 df, \( p=0.01421 \)

According \( p \) value from LR test, Wald test and Score test, then the model is significant.

According to Table 1, gentlebirth method has an effect on the duration of birth process because \( \left( P(>|z|) < 0 \right. \).

Estimation for Methods is \( \hat{\beta}_h = 0.7287 \).

Cox regression model can be written as follows

\[
\hat{h}(t|X) = h_0(t) \exp(0.7287 \text{method})
\]

Hazard Ratio (HR) is \( HR = \exp(\hat{\beta}_h) = \exp(0.7287) = 2.0724 \), its means gentlebirth method has faster duration of birth process than the other method.

### Table 2: Test of proportional hazard assumption

| Variable | Chisq | p_value |
|----------|-------|---------|
| Age      | 1.1745 | 0.278   |
| Variable          | p-value 1 | p-value 2 |
|-------------------|-----------|-----------|
| Gravida           | 0.0129    | 0.910     |
| Paritas           | 0.0381    | 0.845     |
| Abortus           | 0.3980    | 0.528     |
| Pregnancy age     | 0.2219    | 0.638     |
| weight            | 0.5492    | 0.459     |
| sex P             | 0.4680    | 0.494     |
| Method GB         | 0.9831    | 0.321     |
| GLOBAL            | 4.8204    | 0.777     |

According \( p \_\text{value} \) from Table 2 to each variable, then proportional hazard assumption is true because the \( p \_\text{value} \) of all variable is more than 0.05.

4. Discussion

In this paper, Cox’s regression is used to investigate factors that may affect the duration of birth process and gentlebirth is interesting. The research on gentlebirth has not been done specifically in terms of statistics. Some author apply the Cox proportional hazard on some various field [5], [6], [7], [11] and [12]. Gentlebirth methods is known had the duration of birth is faster than the nongentlebirth method and the other than painless birth.

5. Conclusion

In Cox regression, the assumption to be fulfilled is its proportional hazard. According \( p \_\text{value} \) from Table 2 to each variable, then proportional hazard assumption is true because the \( p \_\text{value} \) of all variable is more than 0.05. In this study, the Cox model applies to the duration of the birth process in a clinic that applying the genthebirth methods. Based on the previous discussion, then the factors that affect the duration of the birth process is gentlebirth method. The duration of birth process with gentlebirth method is faster than the non gentlebirth method.

6. Recommendation

Further research can be done to determine whether there is a difference the duration of birth in the way gentle birth method used such as waterbirth or hipnobirthing

Reference

[1] Allison, P.D., *Survival Analysis Using SAS A Practical Guide*. Second Edition. 2010. SAS Institute Inc

[2] Aprilia, Y dan Ritchmond, B.L., 2013, *Gentlebirth Melahirkan Nyaman Tanpa Rasa Sakit*, Penerbit Gramedia Widiasarana Indonesia, Jakarta.
[3] Collett, D., 2003, *Modelling Survival Data in Medical Research*, Second Edition, Chapman and Hall.

[4] Cox, D.R., 1972, *Regression Model and Life Table*, Journal Royal Statistical Society, London.

[5] Dehkordi, M., Safee, A., Tabei, SZ., 2009, *A Comparison between Cox Proportional Hazard Model and Logistic Regression on Prognostic Factor in Gastric Cancer*, East Africa Journal Public Health.

[6] Handayani, L., Fatekurohman, M., Anggraeni, D., 2017, *Survival analysis in patient with DHF using Cox Proportional Hazard Regression.*, International Journal of Advanced Engineering Research and Science (IJAERS)

[7] Hanni, T and Wuryandari, T., 2013, *Model Regresi Cox Proporsional Hazard pada Data Ketahanan Hidup.*, Media Statistika Vol 6, No 1.

[8] Kleinbaum, D.G., and Klein, M., 2005, *Survival Analysis A Self-Learning Text*, New York USA.

[9] Klein, J.P. and Moeschberger, M.L., 2003, *Analysis Techniques for Censored and Truncated Data*. Second Edition, New York.

[10] Lin, D.Y. and Wei, L.J., 1989, *The robust inference for the Cox proportional hazard model*, J.Amer. Stat. Assoc. Vol 84.

[11] Martinussen, T. and Scheike, T.H., 2006, *Dynamic Regression Models for Survival Data*. New York.

[12] Ni, J., 2009, *Application of Cox Proportional Hazard to the Stock Exchange*. Undergraduate Mathematics Exchange, Vol 6 No 1.

[13] Tolosie, K. and Sharma, M.K., 2014, *Aplication of Cox Proportional Hazards Model in Case of Tuberculosis Patients in Selected Addis Ababa Health Centres, Ethiopia*. Hindawi Publishing