C-53

by Armin Lawi

Submission date: 07-Nov-2021 12:03PM (UTC+0700)
Submission ID: 1695204369
File name: C.1 c-53 A Weibull Regression Model using Additive Frailties on Survival Data.pdf (1.41M)
Word count: 2193
Character count: 11238

A Weibull Regression Model using Additive Frailties on Survival Data

Rima Ruktiari[1], SriAstuti Thamrin[2] and **Armin Lawi[3]**
[1,2,3]*Department of Mathematics, Faculty of Mathematics and Natural SciencesHasanuddin University, Indonesia*
Email: rimaruktiarii@gmail.com

ABSTRACT

This research studies Bayesian estimation approach on proportional conditional hazard model given to frailty on multivariate survival data. The survival time model that is used in research is additive frailties containing basic hazard function with Weibull distribution.The research uses the following method in analyzing survival data; first, likelihood of the model and prior distribution should be determined thusthe posterior distribution can be formed, and the second, the Gibbs sampling algorithm is performed to generate random samples from posterior distribution. This methodologyis also applied inMcGilchristan Aisbett's kidney infection. Our research results can obtain useful information that recurrence factors affecting renal infections of patients are variable sex and the frailties.

Keywords: Bayesian estimation, frailty, proportional Hazard model, additive frailties, Weibull distribution, Gibbs sampling.

1. Introduction

Survival analysis is a method that has a relation with time. Survival analysis is a statistical analysis that models the length of time until the occurrence of an event (Kleinbaum and Klein, 2005). Often in the Survival analysis, survival time individuals who are in the same group are correlated. This is caused by some unobserved covariates. In order covariates can be observed, one way is to add covariates into a model as frailties. Of several popular existing distribution models to analyze survival, Weibull distribution is the most commonly used distribution (Evans et al., 2000) and Frailty model commonly used for univariate data. Therefore, in this study will be assessed modeling proportional hazards conditional given Frailty on survival multivariate data. Survival time model which will be used for baseline hazard function using Weibull distribution with additive frailties (Spiegelhalter et al., 1995).

The purpose of this study is to estimate the parameters of survival multivariate model based on additive frailties model and applying this model based on additive frailties model.

2. Literature Review

2.1 Basic Concepts of Survival Analysis

Survival analysis is the analysis about data obtained from the record time achieved an object until the occurrence of the events fail (failure event). The events until a failure occurs is referred to as the time. According to Lee (2003), there are three factors that must be considered in determining thetime, survivalnamely:

1. time origin / starting point of an event.
2. Failure event of the whole incident should be clear.
3. The scale of measurement as part of a time to be clear.

The difference between survival analysis with other statistical analysis is censor. If the data is said to be censored observations time survival is only partially or has expired and it was not until the failure event (Pyke & Thompson, 1986). The cause of the data censored (Lee, 2003) as follows:

1. When the object moved, died or refused to participate (lost to follow-up)
2. If treatment is stopped for any reason (drop-out)
3. When the study is end while the observed object has not reached the event failed (termination)

2.2. Cox Proportional Hazards Model

Cox proportional hazards model (Cox, 1972) is a general model is used for survival analysis. This model was first introduced by Cox. Cox proportional hazards model can be written:

$$h(y) = h_0(y)\exp(\beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_p X_p) \qquad (2.1)$$

where $h_0(y)$ is the hazard function for each object with variable estimator $X = 0$, so $h_0(y)$ is said to basic hazard function. (Collet, 1994), $\beta_1, \beta_2, \ldots, \beta_p$ are the regression parameters, and $X_1, X_2, \ldots, X_p$ are covariates.

Cox model was expanded with the addition of Frailty in the model that is known as a Frailty model (Sahu et al., 1997). The most common types of frailty model is of shared-Frailty model (Cox, 1972).

The Frailty model is a model that input survival time of individuals into subgroups (clusters subject). This model is an extension of the Cox model where the risk in each individual depends on additional unobserved random variable (Wienke, 2003).

On shared-Frailty model, we assume that Y_{ij} is survival time of $j, (j = 1,2, \ldots, m_i)$ in group $i, (i = 1,2, \ldots, n)$ and given unknown parameters of Frailty and is denoted by w_i (for group i), and the covariate vector x_{ij} so that the hazard function (multiplicative) that given frailty written as follows.

$$h(y_{ij} \mid w_i, x_{ij}) = h_0(y_{ij}) w'_i \exp(x'_{ij} \boldsymbol{\beta}), \quad (2.2)$$

where:

$\boldsymbol{\beta}$ = Parameter of regression with vector $p \times 1$ unknown regression coefficients

$h_0(y_{ij})$ = baseline hazard function common for each individual

x_{ij} = risk of failure at a certain time and vector of covariates $p \times 1$ for the jth subject in the ith group.

2.2.1 Weibull Model with Frailty Normal

Baseline hazard using Weibull distribution written as follows.

$$h_0(y_{ij}) = \mu\alpha y_{ij}^{\alpha-1}, \alpha, \mu > 0, j = 1, \ldots, m, i = 1, \ldots, n \quad (2.3)$$

Where α is the shape parameter, μ is the scale parameter with $\alpha > 0$, $\mu > 0$, and y_{ij} = Failure time, $y > 0$.

Weibull model with additive-Frailty can be formulated and assumed as follows.

$$h(y_{ij} | x_{ij}, b) = \xi_{ij} \alpha y_{ij}^{\alpha-1}, \quad (2.4)$$

where $log(\xi_{ij}) = v + \beta^T x_{ij} + b_i$. $\mu = exp\{v\}$ .We assume that a-priori $\alpha \sim Gamma(\kappa_1, \kappa_2)$ and $\mu \sim Gamma(\rho, \rho)$. b_i is the frailties in the additive model are assumed $N(0, \eta)$ and η is given $Gamma \sim (\varphi, \varphi)$. η^{-1} is the variance of Frailty.

2.3 Gibbs sampling

Gibbs sampling algorithm is contained in the MCMC method used for sampling from high-dimensional complex distribution (Gilks and Wild, 1995). The main concept of the Gibbs sampling is how to find where the univariate conditional distributions in the distribution contains all random variables with one variable to be determined value.

3. Methodology

The data used in this research is secondary data that the data time of recurrence of infection at the point of catheter insertion for renal patients using equipment portable sourced from research and Aisbett McGilchrist (1991).

The steps in this study are:

1. Determine baseline hazard function using Weibull distribution.

2. Determine frailty model.
3. Construct likelihood function based on frailty model.
4. Determine priors. Then, it can be determined posterior.
5. Estimate the parameters. In this study, to estimate the parameters will be used gibbs sampling algorithm
6. The application on data.

4. Results And Discussion

Based on the results of data exploration is obtained that the distribution pattern of of data using Weibull distribution. In this paper, the author will estimate the parameters the additive frailties model using a Bayesian approach and gibbs sampler algorithm is used to generate random samples from the posterior distribution was desired (Mukid and Sugito, 2011).

Let y_{ij} is failure event of *jth* subject in the *ith* group.in day scale. The total of patients who were observed as many as 38 peoples. Censor variables that used is the condition of the patient at the end of the study in which 0 indicates the status of the patient did not experience failure event and 1 shows the patient who experience failure event. In the sex variable, 1 for male and 2 for female.

Data processing use WINBUGS 14. Here is the parameters of model to be estimated.

$$h(y_{ij}|x_{ij}, b_i) = \xi_{ij}\,\alpha y_{ij}^{\alpha-1} = exp(v + \beta_{sex} sex_i + \beta_{age}\, age_{i.} + b_i)\,\alpha y_{ij}^{\alpha-1}$$

Estimation of the parameters model α β_{sex}, β_{age} $\sigma^2_{frailty}$ and μ where $\mu = exp\{v\}$, use a Bayesian approach that initially iterated 1000 times. Further iterations 10.000 times with different starting point to 100.000 iterations until convergent. The results of parameters estimation by this software program package are presented in Table 1.

Table 1. Results of the estimated parameters α, β_{sex}, β_{age}, and $\sigma^2_{frailty}$, μ

Parameters	Mean	Deviation Standart	(95% CI)
α	1.202	0.1706	(0.9361,1.5)
β_{sex}	-1.655	0.5449	(-2.596,-0.8106)
β_{age}	0.00602	0.01401	(-0.01608,0.02908)
$\sigma^2_{frailty}$	0.7853	0.5349	(0.1129,1.782)
μ	0.0132	0.0129	(0.001621,0.03703)

Table 1 given informations that β_{sex} affect the timing of infection. It can be seen from the value of the confidence interval for each parameter. The parameter confidence interval contains the value zero does not affect the timing of the infection. Conversely, if the value of the confidence interval does not contain zero value then affects the timing of the infection of a patient. Comparison of survival time between male and female patients served by the following graph.

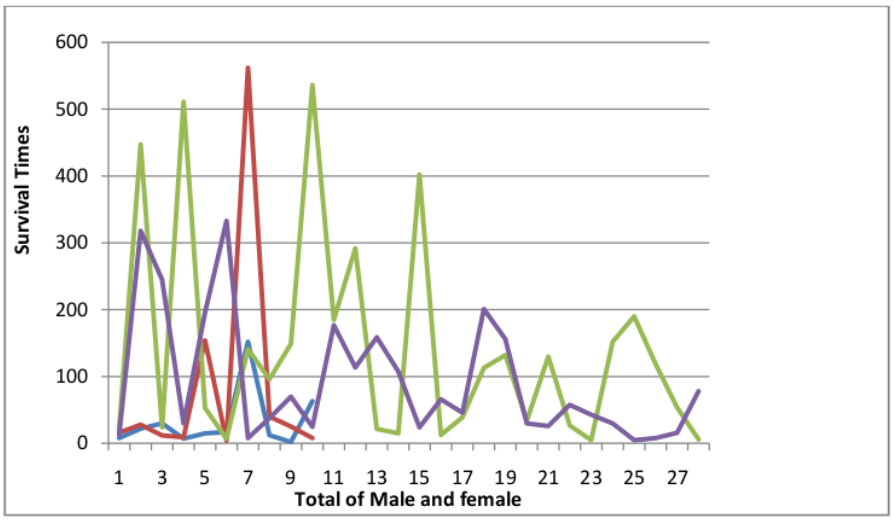

Figure 4.1 The comparisons survival time graph between male and female patients (blue and red = male, Green and purple=female)

From **Figure 4.1** it can be seen that the mean of survival time for men on first and second study is lower than female patients of the first and second studies. This means that female patients have a lower risk of infection than male patients.

$\sigma^2 > 0$ is within the interval (0.1129, 1.782) which means that there is heterogeneity in the data that affect the survival time of patients. Convergence for posterior for each parameter is shown in **Figure 4.2, 4.3 and 4.4** in the form density, trace plots, and history.

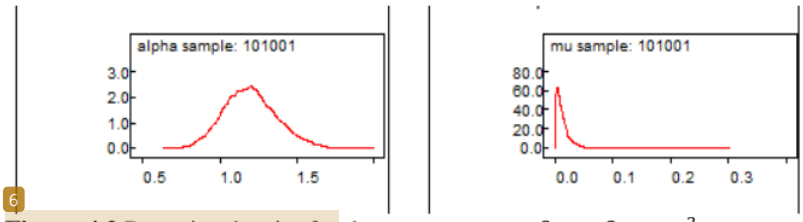

Figure 4.2 Posterior density for the parameters α, β_{sex} , β_{age} , σ^2, μ

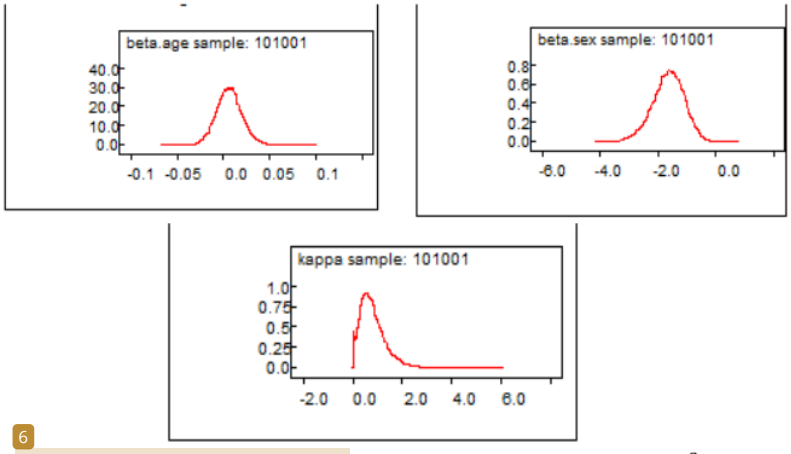

Figure 4.2 Posterior density for the parameters α, β_{sex} , β_{age} , σ^2, μ

Estimate for the parameters α, β_{sex}, β_{age}, σ^2 and μ given good results because the plot of the density tends to smooth shape. Based on estimates on Table 1 show that $\alpha > 1$. α In accordance with the shape of the density, because the value of $\alpha > 1$, Weibull distribution follows the shape of the normal distribution. $\alpha > 1$ indicate that the failure rate increases with time.

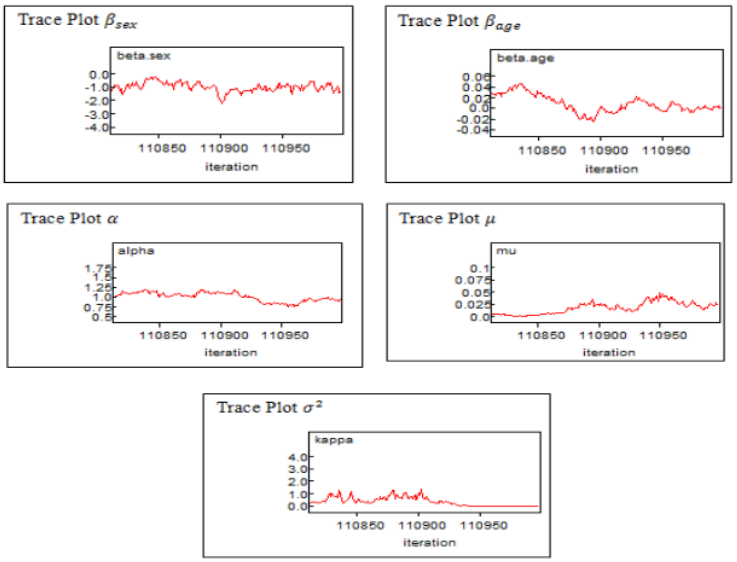

Figure 4.3 Trace Plots posterior

From **Figure 4.3** it can be seen that there is no trend in the Trace plot of each parameter so that it can be said that the model has converged.

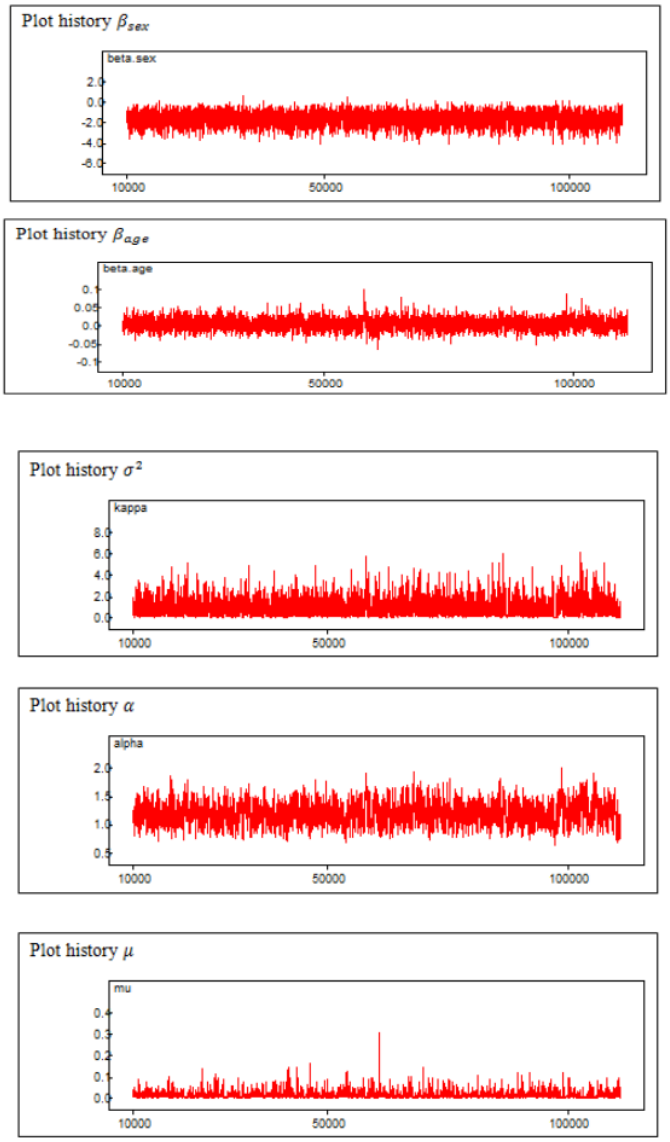

Figure 4.4 history plot posterior

The model can also be seen converging from the history plot. The model is said to converge if history seems tight and able to respond to all of the parameters.

5. CONCLUSION

Based on the writing of the above, it can be concluded that:

1. Parameter estimation of additive frailties model with Bayesian approach. Likelihood function construction is formed by additive frailties model. Prios is using non-informatif prior. Posterior distribution is proporsional likelihood times prior.
2. Based on data processing, only sex and frailty influence the survival time of the patient. From the observation between men and women, women have a lower risk than man to infection. This is cause by women have the mean of survival time longer than men.

6. REFERENCES

Arjas, E. and Gasbarra, D. 1994. Nonparametric Bayesian Inference for right-censored survival data using the Gibbssampler.Statistica Sinica, vol. **4.**505-524.

Bain, LJ and Engelhardt, M. 1992. Probability and Mathematical Lntroduction toStatistics.Belmont Duxburry Press.

Berger, JO 1985. Statistical Decision Theory and BayesianAnalysis.Springer-Verlag. New York.

Clayton, D. 1978. A models for association in bivariate life tables and its application in epidemiological studies in chronic tendency offamiliar diseaseincidence.Biometrics, vol.**65.**141-151.

Clayton, D and Cuzick,J.1985.Multivariate generalizations of the proportional hazards model (withdiscussion).J.Roy. Statist. Soc., Vol.**148.**82-117.

Congdon, Peter. 2003. Applied BayesianModelling.England: Wiley.

Cox, DR (1972). Regression Models and Life Table. J Roy Stat SocB,34, 187-202.

Gill, Jeff.2002. Bayesian Methods: A Social and Behavioral Sciences Approach. A CRC Press Company, Washington DC

Kleinbaum, DG and Klein, M. 2005. Survival Analysis: A SelfText-learning,2nd Edition, Springer, New York

Lawless, JF 2003. Statistical Models and Methods for Lifetime Data, 2nd Edition, John Wiley and Sons, Inc. New York.

Lee, ET 2003. Statistical Methods for Survival Data Analysis 3rdEdition.John Wiley & Sons, Inc. Canada.

Nurmalitasari. 2011. Estimation of Model Parameters Inar (1) Using Bayes Method. Surakarta.

Oakes, D. 1982. A models for association in bivariate survivaldata.J. Roy. Statist. Soc., B, vol. **44.**414-422.

Sahu, SK, Dey, DK and Aslanidou, H., Sinha, D. 1997. A Weibull Regression Model with Gamma Frailty Models for Multivariate SurvivalData.Life Time Data Analysis, vol. **3.**123-137.

Spiegelhalter, D. J, Abrams, KR and Myles, JP,2004.Bayesian Approaches to Clinical Trials and Health CareEvaluation.John Wiley and Sons.

Wienke, A. 2011. Frailty Models in SurvivalAnalysis.A CRC Press Company, Washington DC

C-53

ORIGINALITY REPORT

7%	5%	4%	2%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	repository.unhas.ac.id Internet Source	2%
2	David D. Hanagal. "Modeling Survival Data Using Frailty Models", Springer Science and Business Media LLC, 2019 Publication	1%
3	Submitted to Liverpool John Moores University Student Paper	1%
4	hedibert.org Internet Source	1%
5	www.sthda.com Internet Source	1%
6	Broemeling, . "Agreement and Correlated Observations", Chapman & Hall/CRC Biostatistics Series, 2009. Publication	1%
7	Submitted to University of Strathclyde Student Paper	1%

Exclude quotes	On	Exclude matches	< 1%
Exclude bibliography	On		