Abstract. Piecewise exponential models provide a very flexible framework for modeling univariate survival data. Although in a strict sense it is a parametric model, a piecewise exponential hazard can approximate any shape of a parametric baseline hazard. Therefore, in practice, defining a prior for a piecewise exponential hazard is the same as defining prior process for the nonparametric hazard function. The hazard function for each individual may depend on a set of risk factors or explanatory variables. However, it usually not explains all such variables are known or measurable, and these variables become interesting to be considered. This unknown and unobservable risk factor of the hazard function is often termed as the individual’s heterogeneity or frailty. This paper focuses on the covariates associated with individual such as age and gender related to the survival time in a piecewise exponential model and evaluates the effect of catheter insertion on subsequent survival.

Keywords: Bayesian, frailty, Gamma Distribution, Normal Distribution, Piecewise Exponential model, Survival Analysis.

1. INTRODUCTION

There have been many increasing interests and applications in developing and implementing Bayesian statistical methods for modelling and data analysis. In medical studies, one problem intensively studied nowadays is the analysis of survival times of patients, with the main aim of modelling the distribution of failure times and the relationship with variables of interest. A survival analysis typically examines the relationship between the survival distribution and covariates.

A Cox proportional hazards (PH) model is a popular mathematical model used for modelling survival. This model was proposed by Cox and Oakes (1972) and has also been known as the Cox regression model. This model is a general class of semi-parametric hazards regression model for survival data that has been proposed by Chen and Jewell (2001). The Cox proportional hazards (PH) model can be extended to allow time dependent variables as predictors (Kleinbaum and Klein, 2005). Frailty models are extensions of the PH model which is known as the Cox model (Cox and Oakes, 1972). The aim of this model is to account for unobserved heterogeneity that is caused by unmeasured covariates (Hougaard, 1995). In statistical terms, a frailty model is a random effect model for time to event data, where the random effect (the frailty) has a multiplicative effect on the baseline hazard function (Duchateau and Janssen, 2008, Ayman and Anis, 2011).
Conditional on the frailty, the survival times are assumed to be independent on PH structure.

A useful and simple model to construct frailty model is by using a piecewise exponential model (Ibrahim et al., 2001). Piecewise exponential models and prior processes on the components provide a very flexible framework for modeling univariate survival data. Modeling the baseline hazard using prior processes is very common; see Sinha and Dey (1997) for a review. In the piecewise exponential approach a log-linear model is used to model both the effect of covariates and the underlying hazard rate function, which is approximated by a step function.

There have been many developments on the analysis of piecewise exponential model in the field of survival analysis. The advance in Bayesian paradigm (Berger, 1985, Besag et al., 1995 and Arjas & Gasbara, 1994), and Markov Chain Monte Carlo (MCMC) computational methods (Diebolt and Robert, 1994) have substantially expanded the methodology and application of piecewise exponential models. In the Bayesian context, Sinha and Dey (1997), Sahu et al. (1997) had proposed an excellent paper in this area. Paik et al. (1994) proposed a frailty model called piecewise gamma frailty for correlated survival data with random effects having a nested structure. Zhang et al. (2014) also proposed a class of frailty semi-competing risks survival models to account for the dependence between disease progression time, survival time, and treatment switching. Ismail et al. (2013) developed the new multiplicative piecewise gamma in the hazard function using OpenBugs Statistical Packages and used MCMC method in computing the Bayesian estimator.

In this paper, we consider piecewise model with baseline as normal and gamma distribution and the exponential frailty distributions. We are using Markov Chain Monte Carlo (MCMC) technique to fit these models. We apply our models to bivariate survival data set of McGilchrist and Aisbett (1991) related to Kidney infection and compare these models using Bayesian comparison techniques such as DIC model choice criteria.

The paper is organised as follows. In Section 2, we define the piecewise exponential model including frailty and Bayesian computational approach for parameter estimation. In Section 3, we apply the model using a kidney dataset. The results are discussed further in Section 4.

2. METHODS

2.1. Model Formulation

In this section, we define the piecewise exponential frailty model for analysing survival data. Let $T$ be a nonnegative random variable for a person’s survival time and $t$ be any specific value of interest as a realisation of the random variable $T$. Kleinbaum and Klein (2005) give some reasons for the occurrence of right censoring in survival studies, including termination of the study, drop outs, or loss to follow-up.

Initially, we denote $t_{ij}$ as the random survival time of the subjects $j$ in groups $i$ and $(\beta=(\beta_1, \beta_2, \ldots))$ the unknown parameters of the model corresponding to the data. The parameter $w_i$ represents the family random effect and $\beta=(\beta_1, \beta_2, \ldots)$ is
the vector of fixed effect coefficients. The family random effect is assumed to act on the conditional hazard \( h(t\mid \beta, w_i) \) in the following multiplicative way:

\[
h(t\mid \beta, w_i) = h_0(t\mid \beta) \exp(\mathbf{X}_i) w_i,
\]

(1)

where \( \mathbf{X}_i \) is fixed covariate vector or \( p \times 1 \) covariate vector for subjects \( i \) in groups \( j \), and time \( t \).

We need to find the distribution of the family unobserved heterogeneity \( w_i \). The classical likelihood-based method assumes that the unknown model parameters have true fixed values, which are found by optimization. The Bayesian approach, however, updates the prior of \( (\beta, w_i) \) using the data, in order to obtain the posterior distribution (which represents new beliefs after having observed the data).

We now assume that we observe possibly right-censored data for \( n \) patients and \( \delta_{ij} \) is an indicator function such that (Marin et al., 2005a) \( \delta_{ij} = 1 \), if the lifetime is uncensored, i.e., \( T_{ij} = t_{ij} \) and \( \delta_{ij} = 0 \), if the lifetime is censored, i.e., \( T_{ij} > t_{ij} \). If the survival function \( S_{ij} \) and density function \( f_{ij} \), then the contribution of subject \( j \) from group \( i \) to the likelihood function is its density function if the subject dies, and its survival function otherwise:

\[
L(t_{ij} \mid \mathbf{w}_i) = f_{ij}(t_{ij} \mid \mathbf{w}_i)^{\delta_{ij}} S_{ij}(t_{ij} \mid \mathbf{w}_i)^{1-\delta_{ij}}
\]

(2)

where \( t_{ij} = h_0(t_{ij}) \exp(\mathbf{X}_ij) \exp(\mathbf{w}_i) \), is the integrated hazard function for the fixed effects. Because of the effect of the chosen covariates on subject mortality do not have equal importance over the whole period of subjects, a piecewise exponential baseline hazard can be used.

To construct piecewise constant baseline hazard model, we first partition the time axis into \( J \) intervals with cutpoints \( 0 = v_1 < v_2 < \ldots < v_J \) define the \( k \)-th interval as \( (v_k, v_{k+1}] \). We then assume that the baseline hazard is constant within each interval, so that,

\[
h_0(t_{ij}) = k \text{ for } t \in I_k = (v_k, v_{k+1}]
\]

(3)

where \( t \) is the event time. Under that assumption, the likelihood function coincides with that of a Poisson distribution with mean \( E_{ijk} \lambda_{ijk} \) (Laird and Olivier, 1981 [11]). In that expression, \( E_{ijk} \) denotes the time lived in the interval \( \lambda_{ijk} \) from the \( j \)-th subject and the parameter \( \lambda_{ijk} \) is the corresponding hazard function.

Breslow (1972) and Breslow (1974) proposed the use of Piecewise exponential distributions in survival data analysis with successive death times to replace the base-line in (Cox, 1972). Holford (1980) and Laird and Oliver (1981) in their papers independently noted that the piecewise hazard model was equivalent to a certain Poisson regression model. Kalbfleisch and Prentice (1973) suggested that the cutpoints should be selected independently of the data and we have also assumed the same.

The observed survival times may be terminated either by failure or by censoring. It is assumed that the times of failure are independent of the times of censoring. If the individual lived beyond the end of the interval then the time lived in the interval equals the width of the interval.

The main problem in this Bayesian approach is to choose a prior distribution that indicates uncertainty about parameter. The choice of the prior
distributions follows from previous studies (Ibrahim et al., 2001, Ismail et al., 2013). The following prior distributions were placed on the parameters \( w \) and \( \lambda \):

\[
    w_i \sim \text{Gamma}(\tau, \tau),
\]

where the hyper-parameter \( \tau \) is also assumed \textit{Gamma} distributed.

\[
    k_1 \sim \text{Gamma} \quad k, \quad k = 1, \ldots, g
\]

where \( \tau_0 = 1 \) and \( k_1 = E(k_1 | 1, 2, \ldots, k) \). The parameter \( k \) controls the amount of smoothness available, i.e., small \( k \) indicates less information in the smoothing of \( k \). If \( k = 0 \), then \( k \) and \( k_1 \) independent. When \( k \to \infty \), then the baseline hazard is in the same interval \( I_k \) and \( I_k \) i.e., \( k = k_1 \).

To compute the prior distribution, \( \pi(\beta, w_i) \), the matrix of fixed effects \( \beta \) is set to follow a multivariate normal distribution with zero mean and low precision: \( \beta \sim N_{m}(0, \Sigma) \), where \( \Sigma \) is a diagonal covariance matrix with large variance terms. The integer \( m \) is the total number of covariate terms in the model.

The posterior distribution of \( (\beta, w_i) \) conditioned on the data is proportional to the product of the likelihood function and the prior distribution (Gelman et al., 1995).

\[
    f(\beta, w_i | t_{ij}, t_{ij}, \mu) L(t_{ij}, \mu, \beta, w_i)
\]

Since the joint posterior distribution does not have a closed form, we use Markov Chain Monte Carlo (MCMC) methods for computation (Gilks et al., 1996). The idea is that, if a specific Markov chain is run (after a suitable initial burn-in period) for long enough time, it should reach a stationary distribution which is the same as the desired posterior distribution (Gelman et al., 2004). Following previous studies on hierarchical model (Gelman et al., 1995), the distributions of the nodes, conditional on all the parameters, are assumed independent of each other. In addition, the prior distributions of any fixed/random effects are independent.

### 2.2. Computation Method

A Gibbs sampling, as implemented in the WinBUGS (Spiegelhalter et al., 2003) software, is used to sample from the posterior density of the family random effect \( w_i \). WinBUGS saves the researcher from these complexities, and thus, allows him to concentrate on more substantive issues. An outline of the WinBUGS code written by the authors is displayed in Appendix.

Three chains with different starting values were run simultaneously. After 10000 iterations for burn-in, 100,000 iterations were performed for each chain and one out of every 100th values were used. The obtained time series alone are not sufficient criterion to conclude that the chains converge. Therefore, the Gelman-Rubin factors (Gelman and Rubin, 1992) were also examined. This factor compares the variation in the sampled parameter values within and between chains (Congdon, 2003). Thus it describes how much the increase in the number of iterations may improve the estimates.

### 3. RESULTS
The recurrence of infection at the time of catheter insertion point for kidney patients originating data will be used in this analysis. The data can be found from McGilchrist and Aisbett (1991). The number of patients is 38, in which the variables are patient, time, status, age, sex (0=male, 1=female).

In this paper, we do the analysis for two different types of the baseline hazard function in Cox regression, namely gamma prior baseline hazard function and normal prior baseline hazard function. The observation period was split into \( g = 5, g = 10 \) and \( g = 20 \).

The analysis started by choosing three parallel chains with different starting values for each model and they were carried out simultaneously. Each chain performed 100,000 iterations after 10,000 iterations for burn-in to obtain convergence to the posterior distribution. One out of every 100th values is used to reduce the autocorrelation of the chain. The convergence of the chains can be monitored via the Brooks-Gelman-Rubin (BGR) convergence-diagnostic graph.

Table 1 Summaries of parameter estimation for Cox Regression with different baseline hazard function.

<table>
<thead>
<tr>
<th>Baseline Hazard Function</th>
<th>Parameter</th>
<th>Mean</th>
<th>Standard deviation</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>beta.age</td>
<td>0.006</td>
<td>0.0129</td>
<td>(-0.018, 0.032)</td>
</tr>
<tr>
<td></td>
<td>beta.sex</td>
<td>-1.702</td>
<td>0.4882</td>
<td>(-2.695, -0.772)</td>
</tr>
<tr>
<td></td>
<td>Kappa (frailty)</td>
<td>0.5656</td>
<td>0.2899</td>
<td>(0.116, 1.245)</td>
</tr>
<tr>
<td>Gamma</td>
<td>beta.age</td>
<td>-0.003</td>
<td>0.0128</td>
<td>(-0.028, 0.021)</td>
</tr>
<tr>
<td></td>
<td>beta.sex</td>
<td>-2.009</td>
<td>0.4956</td>
<td>(-3.01, -1.055)</td>
</tr>
<tr>
<td></td>
<td>Kappa (frailty)</td>
<td>0.7233</td>
<td>0.3162</td>
<td>(0.233, 1.463)</td>
</tr>
</tbody>
</table>

Table 1 shows the posterior mean, standard deviation and 95% credible interval (CI) for \( \beta_{sex} \), \( \beta_{age} \), and Kappa for Cox regression using different types of baseline hazard functions. The parameter estimation for all types is quite similar including the log-likelihood and deviance information criterion. Figure 1(a) to (b) show the posterior trace plots for 100,000 iterations for each of three generated samples while Figure 2(a) to (b) show the density plots associated with the coefficient of the covariate.

Based on Table 1, we can see that the sex and frailty describes substantially the survival time of patients because the estimated parameters of these covariate does not contain the zero value. Meanwhile, the estimates of Kappa from different models show that there is strong posterior evidence of a high degree of heterogeneity in the population of patients. Some patients are expected to be very prone to infection compared to others with the same covariate value. This is not very surprising, as in the dataset there is a patient with infection times 8 and 16, and there is also another male patient with infection times 152 and 562. The high posterior mean of Kappa also provides evidence of a strong positive association between two infection times for the same patient. The analysis suggests that both models with different baseline hazard are very close to each other.

The convergence of the parameters has been achieved since autocorrelations decreases only after considering a lag equal to 50 and this indicates a
good convergence of the parameter space with a reasonably small number of iterations. This can be seen in Figure 3(a) to (b).

Figure 1 Different baseline hazard functions - estimated predictive trace plots associated with the coefficient of the covariate
Estimating Piecewise Exponential Frailty Model With Changing Prior For Baseline Hazards Function

**Figure 2** Different baseline hazard functions - estimated predictive history plots associated with the coefficient of the covariate

**Figure 3** Different baseline hazard functions - estimated predictive density plots associated with the coefficient of the covariate
4. CONCLUDING REMARKS

This paper has presented the piecewise exponential frailty model using Bayesian approach to fit more flexible survival models for non-informative censored data with MCMC computational methods. The case study that we considered involved kidney survival, with covariates given by age and sex. The obtained results show that the sex and frailty are substantially associated with survival in this study. In order to check the sensitivity of the results with respect to prior information, the prior distributions of the selected parameters in the Gamma and Normal distributions were changed and the analysis was re-run. Then, a comparison between estimation and standard error of these estimations was made to see any similarities or differences.

Apart from accuracy and precision criteria used for the comparison study, the Bayesian approach is coupled with MCMC and thus enable us to estimate the parameters of Weibull survival models and probabilistic inferences about the prediction of survival times. This is a significant advantage of the proposed Bayesian approach. Furthermore, flexibility of Bayesian models, ease of extension to more complicated scenarios such as a frailty model, relief of analytic calculation of likelihood function, particularly for non-tractable likelihood functions, and ease of coding with available packages should be considered as additional benefits of the proposed Bayesian approach to predict survival times.

References


Sri Astuti Thamrin: Hasanuddin University, Makassar, Indonesia.
E-mails: tuti@unhas.ac.id

Armin Lawi: Hasanuddin University, Makassar, Indonesia.
E-mails: armin@unhas.ac.id