Modeling Heterogeneity for Bivariate Survival Data by Shared Gamma Frailty Regression Model

David D. Hanagal and Richa Sharma

Department of Statistics, University of Pune, Pune-411007, India.
Email: david_hanagal@yahoo.co.in; richa_stat@rediffmail.com

Abstract

In the analysis of survival data with parametric models, it is well known that the Weibull model is not suitable for modeling survival data where the hazard rate is non-monotonic. For such cases, where hazard rates are bathtub-shaped or unimodal (or hump-shaped), log-logistic, lognormal, Birnbaum-Saunders, and inverse Gaussian models are used for the computational simplicity and popularity among users. When models are inadequate and inappropriate, compound Rayleigh, arctangent, generalized Weibull, and Weibull-Pareto composite models are also used. Out of these models log-logistic (LL) model is frequently used. The log-logistic distribution (LLD) has the advantage of having simple algebraic expressions for its survivor and hazard functions and a closed form for its distribution function. In this paper, we consider gamma distribution as frailty distribution and LLD as baseline distribution for bivariate survival times. The problem of analyzing and estimating parameters of bivariate LLD with shared gamma frailty is of interest and the focus of this paper. We introduce Bayesian estimation procedure using Markov Chain Monte Carlo (MCMC) technique to estimate the parameters involved in the proposed model. We present a simulation study and two real data examples to compute Bayesian estimates of the parameters and their standard errors and then compare the true values of the parameters with the estimated values for different sample sizes. A search of the literature suggests there is currently no work has been done for bivariate log-logistic regression model with shared gamma frailty using Bayesian approach.

Key words: Bayesian Estimation, Censored sample, Gamma frailty, Gibbs Sampling, Log-logistic distribution, Markov Chain Monte Carlo (MCMC), Shared frailty.

1 Introduction

In the analysis of survival data, a useful way of expressing the experience of a group of patients under observation is by the form of the hazard rate, or mortality. This measures the instantaneous probability of dying at a given time, conditional on the patient having survived thus far. Any model which attempts to represent realistically the patients’ survival
experience may be judged for its suitability by the closeness of the theoretical hazard function
derived under the model to the “true” hazard derived empirically from the data, using life
table techniques. Models fitted to survival data may involve parametric or non-parametric
forms for the hazard function. This depends on whether this form is defined (up to a small
number of unknown parameters) as that of a known distribution, or whether it is completely
undefined. In this paper, we shall be concerned only with parametric forms which are
generally easier to fit than non-parametric forms.

The distribution used most frequently in the modeling of survival and failure time data
is the Weibull. However, its use is limited by the fact that its hazard function is monotonic,
i.e. increasing, decreasing, or constant, for whatever the values of its parameters. This
may be inappropriate where the course of the disease is such that mortality reaches a peak
after some finite period, and then slowly declines. An example of this is given by Langlands
et al. (1979) in a study of the curability of breast cancer, where peak mortality occurred
after about three years. A parametric form for such a hazard is given by the log-logistic
distribution (LLD). The LLD provides a useful alternative to the Weibull distribution for
the parametric modeling of survival data where the hazard rate is non-monotonic.

A commonly used and very general approach to the problem of modeling multivariate
survival data is to specify independence among observed data items conditional on a set of
unobserved or latent variables (random effects). A multivariate model for the observed data
is then induced by averaging over an assumed distribution for the latent variables. This
unknown or unobservable risk factor (or latent variable) is often termed as the heterogeneity
or frailty. Thus, the frailty model is a random effect model for time to event data which is
an extension of the Cox proportional hazards model. The term frailty was first introduced
by Vaupel et al. (1979) in univariate survival models and was substantially promoted by its
applications to multivariate survival data in a seminal paper by Clayton (1978).

Frailty models for multivariate survival data are derived under a conditional independ-
ence assumption by specifying latent variables that act multiplicatively on the baseline
hazard. The dependence structure in the multivariate model arises when common or depen-
dent latent variables enter into the conditional models for multiple observed data items, and
the dependence parameters may often be interpreted as variance components. This concept
provides an extension of the traditional univariate frailty model (Vaupel et al., 1979; Lan-
caster, 1979), and it allows to take the mutual dependence of life times of related individuals into account in the analysis of survival data.

There are two important approaches in this field, the shared frailty model and the correlated frailty model. In a shared frailty model, the frailty is common to the individuals in the group, and is thus responsible for creating dependence. In the following we will restrict our considerations to the bivariate case. Extensions to higher-dimensional models are straightforward in the shared frailty approach. Bivariate survival data arises when we consider life times of paired individuals or each individual experiences recurrent events. For example, if we consider paired human organs like, kidneys, eyes etc. and the recurrences of a given disease. In industrial applications, the breakdown times of dual generators in a power plant or failure times of two engines in a two-engine airplane are illustrations of bivariate survival data. So, we are interested mainly on paired data with common shared frailty. The key assumption is that the dependence between two individual lifetime variables \( T_1 \) and \( T_2 \) is caused by the frailty representing unobserved common risk factors and conditional on frailty \( T_1 \) and \( T_2 \) are independent.

Let a continuous random variable \( T \) be the lifetime of an individual and the random variable \( U \) be frailty variable. The conditional hazard function for a given frailty variable \( U = u \) at time \( t > 0 \) is,

\[
h(t \mid U, X) = uh_0(t) \exp(x'\beta)
\]

(1.1)

where \( h_0(t) \) is a baseline hazard function at time \( t > 0 \). \( X \) is a column vector of observed covariates and \( \beta \) is a column vector of corresponding regression coefficients. The conditional survival function for given frailty at time \( t > 0 \) is,

\[
S(t \mid U, X) = \exp\left[ -\int_0^t h(y \mid u, x)dy \right] = \exp\left[ -uH_0(t) \exp(x'\beta) \right]
\]

(1.2)

where \( H_0(t) \) is cumulative baseline hazard function at time \( t > 0 \). Integrating over the range
of frailty variable $U$ having density $f(u)$, we get marginal survival function as,

$$S(t) = \int_0^\infty S(t \mid u, x)f(u)du$$

$$= \int_0^\infty \exp[-uH_0(t)\exp(x'\beta)]f(u)du$$

$$= E\left[\exp[-uH_0(t)\exp(x'\beta)]\right]$$

$$= LU\left(H_0(t)\exp(x'\beta)\right)$$ (1.3)

where $LU(.)$ is a Laplace transformation of the distribution of $U$. Once we have survival function at time $t > 0$ of lifetime random variable of an individual one can obtain probability structure and can base their inference on it.

The objective of this paper is to introduce a bivariate log-logistic distribution with shared frailty which is generated by gamma distribution. It is considered as a distribution of the lifetimes of two components where each lifetime follows a log-logistic distribution. The dependence in $T_1$ and $T_2$ is induced by gamma distributed frailty variable. After integrating out frailty, $T_1$ and $T_2$ have a bivariate distribution.

Currently, in this paper Bayesian inference for the gamma frailty Cox model is mainly focused on the setting where the event times are right censored. Here we are using Markov Chain Monte Carlo (MCMC) technique to estimate parameters involved in this model. We present a simulation study along with standard errors (s.e.) to compare the true values of the parameters with the estimated values for different sample sizes. Also, we apply our model and estimation procedure to bivariate survival data set of leukemia remission times first given by Freireich et al. (1963) and to McGilchrist and Aisbett’s (1991) study related to kidney infection bivariate data set.

The remaining paper is organized as follows: In Section 2, we introduce the notion of shared frailty model with Laplace transformation followed by the Section 3 in which we give the introduction of gamma distribution with unconditional bivariate survival function evaluated at the cumulative baseline hazard. We introduce the bivariate log-logistic (BVLL) regression model as a baseline model with shared gamma frailty in section 4. The joint survival function of proposed BVLL distribution after integrating out frailty is also derived in this section. Likelihood function of the failure data given parameters is presented in
Section 5. In Section 6, the joint posterior density function of the parameters given the failure times is defined. In the same section, we also discuss how MCMC technique is used to estimate the parameters of the proposed models. In Section 7 and 8, we present final results of simulation study and the analysis of two bivariate data set, respectively. Finally, the paper ends with a discussion of our findings in Section 9.

2 General Shared Frailty Model

A shared frailty model can be considered as a mixed (random effects) model in survival analysis with group variation (frailty) and individual variation described by the hazard function. In contrast, mixed models show a more symmetric handling of these two sources of variation. Because of the censored observations the Cox model and the frailty models do not belong to the class of generalized linear mixed models. The shared frailty model is a conditional independence model where the frailty is common to all individuals in a cluster and therefore responsible for creating dependence between event times. This is the reason for the concept of shared frailty. It is assumed that there is independence between the observations from different clusters. If the variation of the frailty variable is zero, this implies independence between event times in the clusters; otherwise, there is positive dependence between event times within a cluster. The shared frailty model abounds in the literature on frailty models and was extensively studied in the monographs by Hougaard (2000), Therneau and Grambsch (2000). A more detailed presentation of shared frailty models can be found in the excellent book by Duchateau and Janssen (2008).

The shared gamma frailty model was suggested by Clayton (1978) (who did not use the notion of ‘frailty’) for the analysis of the correlation between clustered survival times in genetic epidemiology. An advantage is that without covariates its mathematical properties are convenient for estimation (see Oakes, 1982, 1986). By measuring some potentially important covariates, we can examine the influence of the covariates, and we can examine whether they explain the dependence, that is, whether the frailty has no effect (or more correctly, no variation), when the covariate is included in the model. The regression model is derived conditionally on the shared frailty (U).

Suppose $n$ individuals are observed for the study and let a bivariate random variable
\((T_{i1}, T_{i2})\) be the first and second survival times of \(i^{th}\) individual \((i = 1, 2, 3, \ldots, n)\). Also suppose that there are \(p\) observed covariates collected in a vector \(X_i = (X_{i1}, \ldots, X_{ip})'\) for \(i^{th}\) individual where \(X_{ik}\) \((k = 1, 2, 3, \ldots, p)\) represents the value of \(k^{th}\) observed covariate for \(i^{th}\) individual. Here we assume that the first and second survival time \(T_1\) and \(T_2\) for each cluster share the same value of the covariates. Assuming that the frailties are acting multiplicatively on the baseline hazard function and both the survival times of individuals \(T_1\) and \(T_2\) are conditionally independent for given frailty \(U_i = u_i\). The conditional hazard model for \(i^{th}\) cluster at \(j^{th}\), \((j = 1, 2)\) survival time \(t_{ij} > 0\), for given frailty \(U_i = u_i\) has the form:

\[
h(t_{ij} | U_i, X_i) = u_i h_0(t_{ij}) \exp(x_i'\beta); \quad i = 1, 2, \ldots, n; \quad j = 1, 2 \quad (2.1)
\]

where \(U_i\) is the unobserved (random) common risk factor shared by all subjects in cluster \(i\), \(h_0(t_{ij})\) is the common baseline hazard function, \(X_i\) is a vector of observable covariates and \(\beta\) is a vector of unknown regression coefficients.

Here \(\exp(x_i'\beta)\) is the factor that gives the subject specific contribution to the hazard. Model (2.1) is called the shared frailty model because subjects in the same cluster share the same frailty factor. This model induces correlation between survival times of subjects within the same cluster. The value of the frailty \(U_i\) is common to the individuals in the group, and thus it is responsible for creating dependence. This dependence is always positive.

The conditional integrated hazard function for \(i^{th}\) individual at \(j^{th}\) survival time \(t_{ij} > 0\) for given frailty \(U_i = u_i\) is,

\[
H(t_{ij} | U_i, X_i) = u_i H_0(t_{ij}) \exp(x_i'\beta) \quad (2.2)
\]

where \(H_0(t_{ij})\) is integrated baseline hazard function at time \(t_{ij} > 0\).

The conditional survival function for \(i^{th}\) individual at \(j^{th}\) survival time \(t_{ij} > 0\) for given frailty \(U_i = u_i\) is,

\[
S(t_{ij} | U_i, X_i) = \exp\left[-H(t_{ij} | U_i, X_i)\right]
\]

\[
= \exp\left[-u_i H_0(t_{ij}) \exp(x_i'\beta)\right] \quad (2.3)
\]

Under the assumption of independence, the conditional survival function in the bivariate
case for given frailty $U_i = u_i$ at time $t_{i1} > 0$ and $t_{i2} > 0$ is,

$$S(t_{i1}, t_{i2}|U_i, X_i) = S(t_{i1}|U_i, X_i)S(t_{i2}|U_i, X_i)$$

$$= \exp[-u_i\{H_{01}(t_{i1}) + H_{02}(t_{i2})\}\exp(x'_i\beta)] \quad (2.4)$$

where $H_{0j}(t_{ij})$ is the integrated baseline hazard of $T_{ij}$, ($i = 1, 2, ..., n$; $j = 1, 2$). From this, we immediately derive the bivariate survival function by integrating out $U_i$ having the probability function $f(u_i)$, for $i^{th}$ individual.

$$S(t_{i1}, t_{i2}|X_i) = \int_{U_i} S(t_{i1}, t_{i2} | u_i, x_i) f(u_i)du_i$$

$$= \int_{U_i} \exp[-u_i\{H_{01}(t_{i1}) + H_{02}(t_{i2})\}\exp(x'_i\beta)] f(u_i)du_i$$

$$= E\left[\exp[-u_i\{H_{01}(t_{i1}) + H_{02}(t_{i2})\}\exp(x'_i\beta)]\right]$$

$$= L_{U_i}\left[(H_{01}(t_{i1}) + H_{02}(t_{i2})) \exp(x'_i\beta)\right] \quad (2.5)$$

where $L(\cdot)$ is the Laplace transform of the distribution of $U$. Thus, the bivariate survivor function is easily expressed by means of the Laplace transform of the frailty distribution, evaluated at the total integrated conditional hazard.

### 3 Gamma Frailty

One important problem in the area of frailty model is the choice of the frailty distribution. The frailty distributions most often applied are the gamma distribution (Clayton, 1978; Vaupel et al., 1979; Oakes, 1982; Hougaard, 2000; Wienke et al., 2002; Wienke et al., 2003a; Hanagal and Sharma, 2012a, 2012b), the positive stable distribution (Hougaard, 1986b), a three-parameter power variance function (PVF) distribution (Hougaard, 1986a), the compound poisson distribution (Aalen, 1988; Hanagal, 2010) and the log-normal distribution (McGilchrist and Aisbett, 1991).

We consider frailty distribution as gamma distribution because the gamma distribution fits very well to failure data from a computational and analytical point of view and it is easy to derive the closed form expression of survival, density and hazard function. This model was suggested by Clayton (1978) and Oakes (1982) and hence the model is known.
as Clayton model or Clayton-Oakes model. As the gamma variates are positive, it fits the non-negative criterion of frailties with no transformation. The gamma distribution (we use notation $\Gamma(\alpha, \kappa)$ for the two parameter distribution with shape parameter $\alpha$ and scale parameter $\kappa$) is one of the most commonly used distributions to model variables that are necessarily positive.

Gamma distributions have been used for many years to generate mixtures in exponential and Poisson models. The popularity of the model is due to the fact that the model functions are very easy to derive because of the simplicity of the derivatives of the Laplace transform. Later dependence can be estimated from the early observed values using the gamma frailty assumption.

Let a continuous random variable $U$ follows gamma distribution with parameters $\alpha$ and $\kappa$ then density function of $U$ is,

$$f(u) = \begin{cases} \frac{\kappa^\alpha u^{\alpha-1} \exp(-\kappa u)}{\Gamma(\alpha)} & u > 0, \alpha > 0, \kappa > 0 \\ 0 & otherwise \end{cases} \quad (3.1)$$

The Laplace transform of gamma-distributed random variable $U$ is of a very simple form as:

$$L(s) = E(e^{-us}) = \left(1 + \frac{s}{\kappa}\right)^{-\alpha} \quad (3.2)$$

with expectation and variance

$$E(U) = \frac{\alpha}{\kappa}, \quad Var(U) = \frac{\alpha}{\kappa^2}$$

To make the model identifiable, although we consider two parameter gamma distribution, we restrict that expectation of the frailty equals 1 and variance be finite which implies that scale parameter = shape parameter, so that only one parameter needs to be estimated.

Thus, the distribution of frailty $U$ is the one parameter ($\kappa = \alpha = \theta^{-1}$) gamma distribution i.e.

$$U \sim \Gamma(\theta^{-1}, \theta^{-1}) \quad (3.3)$$

with the corresponding density function

$$f(z) = \begin{cases} \frac{z^{\theta-1} \exp(-z/\theta)}{\theta^{1/\theta} \Gamma(1/\theta)} & u > 0, \theta > 0 \\ 0 & otherwise \end{cases} \quad (3.4)$$
Under the restriction, Laplace transformation of gamma distribution result in the following simplified form,

\[
L(s) = (1 + s\theta)^{-1/\theta} \tag{3.5}
\]

Note that there is heterogeneity if \( \theta > 0 \). So the large values of \( \theta \) reflect a greater degree of heterogeneity among groups and a stronger association within groups. Replacing Laplace transformation in equation (3.5) using (2.5), we get the unconditional bivariate survival function for \( i^{th} \) individual at time \( t_{i1} > 0 \) and \( t_{i2} > 0 \). Thus, the unconditional bivariate survival function expressed as the Laplace transform of the frailty distribution, evaluated at the cumulative baseline hazard is,

\[
S_{\theta}(t_{i1}, t_{i2}) = \left[ \{H_{01}(t_{i1}) + H_{02}(t_{i2})\} \exp(x_i'\beta) \right]^{-1/\theta} \tag{3.6}
\]

where \( H_{01}(t_{i1}) \) and \( H_{02}(t_{i2}) \) are cumulative baseline hazard functions of lifetime random variables \( T_{i1} \) and \( T_{i2} \), respectively.

Once we have unconditional survival function of bivariate random variable \( (T_{i1}, T_{i2}) \) we can obtain likelihood function and estimate the parameters of the model. Since the cumulative baseline hazard \( H_0(\cdot) \) is the same for all subjects, the hazard function differences between subjects are due to either the frailty term (the cluster they belong to) or the fixed effects.

4 Bivariate Log-logistic Gamma Frailty Regression Model

In parametric proportional hazards model we assume a particular parametric function for the baseline hazard \( h_0(t) \). One of the parametric choice for \( h_0(t) \) leads to lifetimes with a log-logistic distribution function. Log-logistic distribution is often used to analyze lifetime data in survival analysis whose rate is non-monotonic.

The log-logistic distribution provides one parametric model for survival analysis. It is a commonly used lifetime distribution in lifetime data analysis since the logarithm of the lifetime variables are logistically distributed. It is used in survival analysis as a parametric
model for events whose rate increases initially and decreases later, for example mortality from cancer following diagnosis or treatment. It has a fairly flexible functional form with two parameters, denoted by \( \log L(\gamma, \lambda) \).

The distribution imposes the following functional forms on the density, survival, hazard and cumulative hazard function:

**probability density function**

\[
f_0(t) = \frac{\gamma(\gamma)}{[1 + (\frac{t}{\lambda})^\gamma]^2} \quad (\gamma > 0, \lambda > 0) \tag{4.1}
\]

**survival function**

\[
S_0(t) = \left[1 + (\frac{t}{\lambda})^\gamma\right]^{-1} \tag{4.2}
\]

**hazard function**

\[
h_0(t) = \frac{\gamma(\gamma)}{[1 + (\frac{t}{\lambda})^\gamma]^2} \tag{4.3}
\]

**cumulative hazard function**

\[
H_0(t) = \ln\left[1 + (\frac{t}{\lambda})^\gamma\right] \tag{4.4}
\]

Unlike the more commonly-used Weibull distribution, it can have a non-monotonic hazard function. It is one of the parametric survival time models in which the hazard rate may be decreasing monotonically for \( \gamma \leq 1 \), increasing, as well as hump-shaped for \( \gamma > 1 \).

This model has been criticized as a lifetime distribution because the hazard function is decreasing for large \( t \) which seems implausible in many situations. The model may fit certain cases where large values of \( t \) are not of interest. The fact that the cumulative distribution function can be written in closed form is particularly useful for analysis of survival data with censoring (Bennett, 1983).

In this paper, the two-parameter log-logistic distribution is considered. Let us assume that the independent random variables \( T_1 \) and \( T_2 \) have log-logistic distribution with parameters \( \gamma_1, \lambda_1 \) and \( \gamma_2, \lambda_2 \), respectively. In short we say \( \text{LogL}(\gamma_j, \lambda_j), (j = 1, 2) \).

Using (2.1) the resulting regression model

\[
h(t_{ij}|U_i, X_i) = \frac{u_i(\gamma_i)(b_{ij})^{\gamma_i-1}}{1 + (b_{ij} \lambda_i)^{\gamma_i}} \exp(x_i^\prime \beta), \quad (j = 1, 2; \ i = 1, \ldots, n) \tag{4.5}
\]

is indeed an extended proportional hazards model conditional on frailty and fixed factors.
The conditional survival function of the \( j \)th individual in the \( i \)th pair given as

\[
S(t_{ij}|U_i, X_i) = \exp[-u_i H_{0j}(t_{ij}) \exp(x_i' \beta)] = \exp \left[ -u_i \ln \left( 1 + \left( \frac{t_{ij}}{\lambda_j} \right)^\gamma_j \right) \exp(x_i' \beta) \right]
\]  

(4.6)

Here value of \( U \) is common to two components in a group. When there is no variability in the distribution of \( U \), that is, when \( U \) has a degenerate distribution then there is no dependency. When the distribution is not degenerate, the dependence is positive. The value of \( U \) can be considered as generated from unknown values of some explanatory variables. Conditional on \( U_i = u_i \), the bivariate survival function is

\[
S(t_{i1}, t_{i2}|U_i, X_i) = S(t_{i1}|U_i, X_i)S(t_{i2}|U_i, X_i) = \exp \left[ -u_i \left\{ \ln \left( 1 + \left( \frac{t_{i1}}{\lambda_1} \right)^\gamma_1 \right) + \ln \left( 1 + \left( \frac{t_{i2}}{\lambda_2} \right)^\gamma_2 \right) \right\} \exp(x_i' \beta) \right]
\]  

(4.7)

where \( U \) follows gamma distribution given in (3.4).

Integrating over \( U \), we get unconditional joint survival function and is given by

\[
S_\theta(t_{i1}, t_{i2}|X_i) = \left[ 1 + \theta \left\{ \ln \left( 1 + \left( \frac{t_{i1}}{\lambda_1} \right)^\gamma_1 \right) + \ln \left( 1 + \left( \frac{t_{i2}}{\lambda_2} \right)^\gamma_2 \right) \right\} \exp(x_i' \beta) \right]^{-1/\theta}
\]  

(4.8)

The log-logistic distribution is very similar in shape to the log-normal distribution, but is more suitable for use in the analysis of survival data. This is because of its greater mathematical tractability when dealing with the censored observations which occur frequently in such data. The contribution made by a right-censored observation to the likelihood is equal to the value of the survivor function at the time of censoring. This can be evaluated explicitly for the log-logistic distribution, but not for the log-normal.

Once we have unconditional survival function of bivariate random variable \((T_{i1}, T_{i2})\) we can obtain likelihood function and estimate the parameters of the model.

## 5 Likelihood Specification

Some of the lifetimes are censored because it is not possible to wait until failure of all individuals in the sample. Suppose there are \( n \) independent pairs of components, for example, paired kidneys, lungs, eyes, ears in an individual under study, whose first and second observed
failure times are represented by \((t_{i1}, t_{i2})\). Let \(w_i\) be the observed censoring time for \(i^{th}\) individual \((i = 1, 2, 3, \ldots, n)\) for first and second recurrence times. We consider censoring time \((W)\) is univariate random right censoring type for both failure times \(T_1\) and \(T_2\). We use censoring scheme as given by Hanagal (1992a, 1992b). Also we assume independence between censoring scheme and life times of individuals. This assumption is sufficient for the distribution of the event times to be identifiable in inference from the censored data (Fleming and Harrington, 1991). One of the following censoring situations can happen for each data point \((t_{i1}, t_{i2})\).

\[
(T_{i1}, T_{i2}) = \begin{cases} 
(t_{i1}, t_{i2}), & \text{if } \max(t_{i1}, t_{i2}) < w_i \\
(t_{i1}, w_i), & \text{if } t_{i1} < w_i < t_{i2} \\
(w_i, t_{i2}), & \text{if } t_{i2} < w_i < t_{i1} \\
(w_i, w_i), & \text{if } w_i < \min(t_{i1}, t_{i2})
\end{cases}
\] (5.1)

Let \(I_1, I_2, I_3, I_4\), denote the following sets

\[
\begin{align*}
I_1 &= \{i : t_{i1} \leq w_i \text{ and } t_{i2} \leq w_i\}, \\
I_2 &= \{i : t_{i1} \leq w_i \text{ and } t_{i2} > w_i\}, \\
I_3 &= \{i : t_{i1} > w_i \text{ and } t_{i2} \leq w_i\}, \\
I_4 &= \{i : t_{i1} > w_i \text{ and } t_{i2} > w_i\}
\end{align*}
\] (5.2)

Discarding factors which do not contain any of the parameters, we want to estimate the parameters in the proposed model. Now the contribution of the \(j^{th}\) individual in the \(i^{th}\) pair of the conditional likelihood of data given the parameters, based on the survival function (4.8) is given by

\[
L(t_{1}, t_{2}|\zeta) = \left(\prod_{i \in I_1} f_{i1}\right)\left(\prod_{i \in I_2} f_{i2}\right)\left(\prod_{i \in I_3} f_{i3}\right)\left(\prod_{i \in I_4} F_i\right)
\] (5.3)

where \(\zeta\) is the vector of baseline parameters, frailty parameter and regression coefficients.
and

\[
f_{i1} = \frac{\partial^2 S_\theta(t_{i1}, t_{i2})}{\partial t_{i1} \partial t_{i2}}, \quad \max(t_{i1}, t_{i2}) < w_i
\]

\[
= (1 + \theta) h_{01}(t_{i1}) h_{02}(t_{i2}) S_\theta(t_{i1}, t_{i2})^{(1+2\theta)} \exp(2x'_i \beta)
\]

\[
= \frac{(1 + \theta) \gamma_1 \gamma_2 (t_{i1}^{-1} t_{i2}^{-1})}{\lambda_1^{\gamma_1} \lambda_2^{\gamma_2} [1 + (t_{i1}/\lambda_1)^{\gamma_1}] [1 + (t_{i2}/\lambda_2)^{\gamma_2}]} S_\theta(t_{i1}, t_{i2})^{(1+2\theta)} \exp(2x'_i \beta)
\]

\[
f_{i2} = -\frac{\partial S_\theta(t_{i1}, w_i)}{\partial t_{i1}}, \quad t_{i1} < w_i < t_{i2}
\]

\[
= h_{01}(t_{i1}) S_\theta(t_{i1}, w_i)^{(1+\theta)} \exp(x'_i \beta)
\]

\[
= \frac{\gamma_1 \gamma_2 (t_{i1}^{-1})}{\lambda_1^{\gamma_1} [1 + (t_{i1}/\lambda_1)^{\gamma_1}]} S_\theta(t_{i1}, w_i)^{(1+\theta)} \exp(x'_i \beta)
\]

\[
f_{i3} = -\frac{\partial S_\theta(w_i, t_{i2})}{\partial t_{i2}}, \quad t_{i2} < w_i < t_{i1}
\]

\[
= h_{02}(t_{i2}) S_\theta(w_i, t_{i2})^{(1+\theta)} \exp(x'_i \beta)
\]

\[
= \frac{\gamma_2 \gamma_3 (t_{i2}^{-1})}{\lambda_2^{\gamma_2} [1 + (t_{i2}/\lambda_2)^{\gamma_2}]} S_\theta(w_i, t_{i2})^{(1+\theta)} \exp(x'_i \beta)
\]

\[
F_i = S_\theta(w_i, w_i), \quad w_i < \min(t_{i1}, t_{i2})
\]

Substituting survival function \(S_\theta(t_{i1}, t_{i2})\) for the proposed model we get the likelihood function given by equation (5.3).

Let \(n_1, n_2, n_3\), and \(n_4\), respectively, denote the random number of observations observed to fall in the sets \(I_1, I_2, I_3\) and \(I_4\). \(f_{i1}\) is the pdf with respect to Lebesgue measure in \(R^2\) and \(f_{i2}\) and \(f_{i3}\) are the pdf with respect to Lebesgue measure in \(R^1\) in their respective regions.

### 6 Bayesian Estimation Strategies

In the Bayesian framework, the parameters of the model are viewed as random variables with some distribution known as prior distribution. To apply MCMC methods, we assume that, conditional on explanatory variables and on the entire set of parameters, observations are independent and prior distributions for all parameters are mutually independent. The distribution of a parameter can be updated by combining its prior distribution and the likelihood function, called as posterior density of a parameter. So, if \(L(Y \mid \zeta)\) is likelihood
function and \( p(\zeta) \) is prior density of a parameter \( \zeta \) then posterior density function of the same parameter \( \pi(\zeta \mid Y) \) is given by, \( \pi(\zeta \mid Y) \propto L(Y \mid \zeta)p(\zeta) \). In our case, the joint posterior density function of parameters for given failure times is given by,

\[
\pi(\lambda_1, \lambda_2, \gamma_1, \gamma_2, \theta, \beta \mid t_1, t_2) \propto L(t_1, t_2 \mid \lambda_1, \lambda_2, \gamma_1, \gamma_2, \theta, \beta) g_1(\lambda_1)g_2(\lambda_2)g_3(\gamma_1)g_4(\gamma_2)g_5(\theta) \prod_{k=0}^{p} p_k(\beta_k)
\]

where \( \beta = (\beta_0, \beta_1, \beta_2, \ldots, \beta_p)' \), \( g_i(.) \) (\( i = 1, 2, \ldots, 5 \)) indicates the prior density function which is gamma distribution with known hyper parameters of corresponding argument for baseline parameters and frailty variance; \( p_k(.) \) is prior density function for regression coefficient \( \beta_k \) which is normal with known hyper parameters and likelihood function \( L(.) \) is given by equation (5.3). Given the distribution (5.3) and the priors, all full conditional distributions of the parameters can be calculated. These full conditional distributions are used in a Gibbs sampling procedure. We assume that all the parameters are independently distributed.

Here, we use Metropolis-Hastings algorithm within Gibbs sampler to estimate the parameters of the model. Algorithm consists in successively obtaining a sample from the conditional distribution of each of the parameter given all other parameters of the model. These distributions are known as full conditional distributions. The process eventually provides samples from joint posterior distribution of the unknown parameters. In our case, full conditional distributions are not easy to integrate out. So, full conditional distributions are obtained by considering that they are proportional to the joint distribution of the parameters of the model.

Gelman-Rubin convergence statistic is based on a comparison of within and between chain variance for each variable (Brooks and Gelman, 1998). When values of this diagnostic are approximately equal to one then sample can be considered to have arisen from the stationary distribution. In this case descriptive statistics or posterior summary can be seen as valid estimates of unknown parameters. Geweke (1992) suggested a test for examining the convergence of a Markov chain in which two sub parts of Markov chain (at the end and at the beginning of the convergence period) are compared. The large standardized difference between ergodic averages at the beginning and at the end of the convergence period indicates non convergence. Sample autocorrelation plots can be used to decide autocorrelation lag.
7 Simulation Study

To evaluate the performance of the Bayesian estimation procedure we carried out a simulation study. For the simulation purpose we have considered only one covariate \( X = X_1 \) which we assume to follow normal distribution. With one covariate, the gamma frailty model given in (4.8) has six parameters. The frailty variable \( U \) is assumed to have gamma distribution with variance \( \theta \). Lifetimes \( (T_{i1}, T_{i2}) \) for \( i^{th} \) individual are conditionally independent for given frailty \( U_i = u_i \). We assume that \( T_{ij} (i, \ldots, n; j = 1, 2) \) follows the log-logistic distribution as baseline distribution. As the Bayesian methods are time consuming, we generate only fifty, seventy-five and one hundred pairs of lifetimes using inverse transform technique.

According to the assumption, for given frailty \( U \), lifetimes of individuals are independent. So the conditional survival function for an individual for given frailty \( U = u \) and a covariate \( X \) at time \( t > 0 \) is,

\[
S(t \mid U, X) = \exp[-uH_0(t) \exp(x_i'\beta)]
\]  

Equating \( S(t \mid U, X) \) to a random number say \( r \) \((0 < r < 1)\) over \( t > 0 \) we get,

\[
t = \lambda[\exp(A) - 1]^\frac{1}{\gamma}
\]

where \( A = -\frac{\ln(r)}{u \exp(x_i'\beta)} \). Equations (7.2) is a generator to generate lifetimes for model (4.8). We have generated different random samples of size \( n = 50, 75 \) and 100 for lifetimes \( T_{i1} \) and \( T_{i2} \) using (7.2) and for frailty from gamma density function \( f(u) \).

Here, we are giving procedure for sample generation of only one sample size, say, \( n = 50 \). Samples are generated using following procedure:

1. Generate a random sample of size \( n = 50 \) from gamma distribution having density as given in equation (3.4), with \( \theta = 3.6 \) as shared frailties \( (u_i) \) for \( i^{th} \) \((i = 1, 2, \ldots, 50)\) individual.

2. Generate 50 covariate values for \( X \) from normal distribution.

3. Compute \( \exp(x_i'\beta) \) with regression coefficient \( \beta = 0.5 \).

4. Generate 50 pairs of lifetimes \( (t_{i1}, t_{i2}) \) for given covariate \( (x_i) \) using following generators,

\[
t_{i1} = \lambda_1[\exp(A_{i1}) - 1]^{\frac{1}{\gamma_1}}
\]
\[ t_{i2} = \lambda_2[\exp(A_{i2}) - 1]^{\frac{1}{\gamma_2}} \]  

(7.4)

for gamma frailty model given in (4.8), where \( A_{i1} = -\frac{\ln(r_{i1})}{u_i \exp(x_i \beta)} \) and \( A_{i2} = -\frac{\ln(r_{i2})}{u_i \exp(x_i \beta)} \); \( r_1 \) and \( r_2 \) are random variables having \( U(0, 1) \) distribution and \( \gamma_1, \lambda_1 \) are respectively shape and scale parameters of baseline distribution of first survival time and \( \gamma_2, \lambda_2 \) are that of second survival time.

5. Generate censoring time \( w_i \) from exponential distribution with failure rate 0.5.

6. Observe \( j^{th} \) survival time \( t_{ij}^* = \min(t_{ij}, w_i) \) and censoring indicator \( \delta_{ij} \) for \( i^{th} \) individual \( (i = 1, 2, \ldots, 50 \text{ and } j = 1, 2) \), where

\[
\delta_{ij} = \begin{cases} 
1, & t_{ij} \leq w_i \\
0, & t_{ij} > w_i 
\end{cases}
\]

Thus we have data consists of 50 pairs of survival times \((t_{i1}^*, t_{i2}^*)\) and censoring indicators \( \delta_{ij} \).

We run two parallel chains for the proposed model with the different starting points using Metropolis-Hastings algorithm within Gibbs sampler based on normal transition kernels. We iterate both the chains for 95,000 times. Prior distributions that we have assumed for the parameters are, respectively, \( \Gamma(0.0001, 0.0001) \) for baseline parameters \( \lambda_1, \lambda_2, \gamma_1 \) and \( \gamma_2 \); \( \Gamma(0.0001, 0.0001) \) for frailty parameter \( \theta \) and \( N(0, 1000) \) for regression parameter \( \beta \). Here \( \Gamma(a, b) \) is gamma distribution with shape parameter \( a \) and scale parameter \( b \) and \( N(\mu, \sigma^2) \) represents normal distribution with mean \( \mu \) and variance \( \sigma^2 \). For both the chains the results were somewhat similar so we present here the analysis for only one chain (i.e. chain I) for the resulting model.

For the proposed model, Table 1 gives Gelman-Rubin convergence statistic values and in Table 2 Geweke test values with corresponding p-values are given. From the Tables 1 and 2 we can observe that, Gelman-Rubin convergence statistic values are nearly equal to one, also Geweke test values are quite small and corresponding p-values are large enough. So, we can say that the chain attains stationary distribution. Simulated values of parameters have autocorrelation of lag \( k \) (values given in Table 3), so every \( k^{th} \) iteration is selected as a sample from posterior distribution. The posterior mean and standard error with 95% credible intervals are reported in Table 3. From the Table 3, it can be observed that estimated values
of parameters reach quite close to true values of the parameters with decreasing standard errors as the sample size goes on increasing. We have used R statistical software to perform this simulation study.

8 Applications

We illustrate the proposed model with two well-known examples. The first examines the effect of a clinical trial of a drug 6-mercaptopurine (6-MP) versus a placebo in 42 children with acute leukemia. The second deals with the survival of 38 kidney catheters patients.

Example 1.: Acute Leukemia Data (Freireich et al., 1963).

Here we demonstrate the method using the well-known leukemia data, consisted of 21 pairs matched of leukemia patients, analyzed by Cox (1972), Hougaard, (2000) (subsection 1.5.4), Ibrahim et al., (2001) (example 3.4), Spiegelhalter et al., (2004), among others. The leukemia remission times data, first given by Freireich et al. (1963), have been cited many times, without reference to their origin from a paired study.

The random variable of interest consists of remission times (in weeks) of the patients assigned to treatment with a 6-MP drug or a placebo during remission maintenance therapy. After having been judged to be in a state of partial or complete remission for the primary treatment with prednisone, a patient was paired with a second patient in the same state. One randomly chosen patient in each pair received the maintenance treatment 6-MP and the other a placebo. It was assumed that deaths at a given time always preceded censoring at the same time, and other ties were broken by randomization. Success (failure) was defined to occur in the $i^{th}$ pair if the time from remission to relapse or censoring for the patient on 6-MP (placebo) exceeded the time to relapse for the patient on placebo (6-MP). The trial was stopped once the number of successes or failures had reached significance. These data have been used in many articles, but in most of them neglecting the pairing. Out of 21 patients in treatment group, 9 failed during the study period and 12 were censored. In contrast, none of the data are censored in placebo group; that is, all 21 patients in the placebo group went out of remission during the study period. The data set contains a single covariate $x$ with value 0 or 1 indicating remission status (0=partial, 1=complete).

First we check goodness-of-fit of the data for baseline distribution and then apply the
Bayesian estimation procedure. In the analysis of this study, we have used the R package. To check goodness-of-fit of data set, firstly we consider graphical procedure to check appropriateness of the model. To assess model graphically, we plot the graph of \( \ln\left(\frac{1 - S(t)}{S(t)}\right) \) versus \( \ln(t) \). If the resulted plot is roughly a straight line then we can say that the underline log-logistic model is appropriate. The graph of \( \ln\left(\frac{1 - S(t)}{S(t)}\right) \) versus \( \ln(t) \) for log-logistic distribution for placebo and 6-MP treatment using R-program are shown in the Figure 1(a) and 1(b), respectively. From the Figures, we can observe that many of the points are nearly on the straight line.

We have considered two statistical tests also viz. Kolmogorov-Smirnov (K-S) test and Hollander & Proschan’s (HP) test (proposed by Hollander and Proschan (1979)) for testing composite goodness-of-fit hypotheses to the log-logistic distribution. Both tests are based on the empirical distribution function. Test procedure is developed for testing goodness-of-fit with data subjected to random right censoring. The p-values of K-S and HP test for baseline distribution are presented in Table 4 which are very large approximately close to 1. Although the p-value of HP test for T6 (time to relapse for 6-MP patients) is not enough large as compared to the same for TP (time to relapse for placebo patients) but that may be due to censoring because there is no censored observation in placebo group. Thus, from different goodness-of-fit graphs and p-values of K-S and HP test (Table 4) we can say that there is no statistical evidence to reject the hypothesis that data are from log-logistic distribution.

Figure 3(a) and 3(b), which plots the Kaplan-Meier estimates and hypothesized theoretical distribution \( S_0(t) \), demonstrates a close agreement between the two. The sample mean of the relapse time for the 6-MP drug treated patients is 17.095 weeks while the same for the placebo treated patient is 8.667 weeks. This may indicate a possible 6-MP drug effect. It clearly conveys the effectiveness of the drug in maintaining remission.

In the analysis, we have used the R program. Given the model assumptions, this program performs the Gibbs sampler by simulating from the full conditional distributions. The Bayesian estimators were obtained through the implementation of the Metropolis-Hastings algorithm within Gibbs sampling scheme described in the earlier section. We implemented 95,000 iterations of the algorithm. To generate the Gibbs posterior samples, we choose to use two parallel chains. The chains should start from over-dispersed initial values to ensure
good converge of parameter space. To deminish the effect of the starting distribution, we generally discard the early iterations of each sequence and focus attention on the remaining. We described the first 2,000 iterations as a burn-in period.

Monitoring convergence of the chains has been done via the Brooks and Gelman (1998) convergence-diagnostic. Hence, once convergence has been achieved, 93,000 observations are taken from each chain after the burn-in period. Gelman-Rubin convergence statistic values and Geweke test statistic values with corresponding p-values for all the parameters are given in Table 5 and Table 6, respectively. On inspection of the Brooks and Gelmans diagnostic, we find the BGR (Brooks and Gelman Ratio) convergent to one, this show that the convergence for the coefficient of regression $\beta_1$, the variance of frailty $\theta$ and other parameters has obtained. Also, the Geweke test statistic values are quite small and corresponding p-values are large enough to say the chains attains stationary distribution. Thus, our diagnostic suggests that the MCMC chains are mixing very well.

We have taken the independent prior as $\lambda_i \sim \Gamma(1, 10^{-4})$, $\gamma_i \sim \Gamma(1, 10^{-4})$, $(i = 1, 2)$; $\beta_1 \sim N(0, 10^5)$ and $\theta \sim \Gamma(10^{-4}, 10^{-4})$. As in the simulation here also results for chain I chain II are similar so we present result for only one chain (i.e. chain I). Figure 5 shows the trace plot, coupling from past plot and sample autocorrelation plot for the parameter $\beta_1$ for chain I. For other parameters graphs have similar pattern so due to lack of space we are not presenting graphs for other parameters. Trace plots for all the parameters show zigzag pattern which indicates that parameters move and mix more freely. Thus, it seems that the Markov chain has reached the stationary state. However, a sequence of draws after burn-in period may have autocorrelation. Because of autocorrelation consecutive draws may not be random, but values at widely separated time points are approximately independent. So, a pseudo random sample from the posterior distribution can be found by taking values from a single run of the Markov chain at widely spaced time points (autocorrelation lag) after burn-in period. Simulated values of parameters have autocorrelation of lag $k$ (see Table 7), so every $k^{th}$ iteration is selected as sample to thin the chain and discarding the rest. The autocorrelation of parameters become almost negligible after the defined lag, given in Table. Table 7 shows the autocorrelation lag, posterior mean, standard deviation, and 95% credible intervals for all baseline parameters, frailty variance and regression coefficient. The posterior estimate of $\theta = 0.4941678$ shows that there exists significant heterogeneity in population of
patients even though each patient share the same value of the covariate. The posterior mean of $\beta_1$ is $-1.766374$ with 95% credible interval $(-2.698818, -0.8937898)$. Thus, patients who had a complete remission of their leukemia have significant effect of induced treatment. In other words, we can say that there is lower risk of acute leukemia relapse for complete remission patients.

**Example 2.: Kidney Infection Data (McGilchrist & Aisbett, 1991)**

The following data set is presented in McGilchrist and Aisbett (1991). The data set consists of times to the first and second recurrences of infection at point of insertion of the catheter in 38 kidney patients using a portable dialysis machine. Infections can occur at the location of insertion of the catheter. The catheter is later removed if infection occurs and can be removed for other reasons, which we regard as censoring. So, survival times for patients given may be first or second infection time or censoring time.

After the occurrence or censoring of the first infection sufficient (ten weeks interval) time was allowed for the infection to be cured before the second time the catheter was inserted. So, the first and second recurrence times are taken to be independent apart from there common frailty component. The survival times from the same patient are likely to be related because of frailty describing the patient’s effect.

The data set consists of three risk variables age, sex (0= male and 1= female) and disease type GN, AN and PKD where GN, AN and PKD are short forms of Glomerulo Neptiritis, Acute Neptiritis and Polycyatic Kidney Disease. The infection times from each patient share the same value of the covariates. Let $T_1$ and $T_2$ be represents first and second recurrence time to infection or censoring. Five covariates age, sex and presence or absence of disease type GN, AN and PKD are represented by $X_1, X_2, X_3, X_4,$ and $X_5$, respectively.

method. Santos et al. (2010) used MCMC method to estimate the parameters of parametric regression model with Weibull and generalized gamma distribution as baseline and gamma and log-normal as frailty distributions. Boneg (2001) considered Cox proportional hazards model and also parametric frailty models. In parametric frailty models he considered Weibull distribution as the baseline and log-normal, Weibull as frailty distributions. He applied MHL and RMHL methods to estimate the parameters of the models.

First, we check goodness-of-fit of the data for the baseline distribution and then apply the Bayesian estimation procedure. Firstly, we consider graphical procedure to check appropriateness of the models. To assess model graphically, we plot the graph of $\ln[(1 - S(t))/S(t)]$ versus $\ln(t)$. If the resulted plot is roughly a straight line then we can say that the underline log-logistic model is appropriate. Figure 2(a) and Figure 2(b) represents goodness-of-fit plots for log-logistic distribution for first and second recurrence times, respectively. From the figures, we can observe that many of the points are nearly on the straight line.

Now, to check goodness-of-fit test of kidney data set we consider two statistical tests viz. Kolmogorov-Smirnov (KS) Test and Hollander & Proschan’s (HP) Test. Also, p-values of KS and HP test statistic for the baseline distribution for first and second recurrence times (see Table 8) are quite large. So there is no statistical evidence to reject the hypothesis that data is from the log-logistic distribution. Also, Figure 4(a) and 4(b), which plots the Kaplan-Meier estimates and hypothesized theoretical distribution $S_0(t)$, demonstrates a close agreement between the two.

In the analysis, we have used the R program. Given the model assumptions, this program performs the Gibbs sampler by simulating from the full conditional distributions. A widely used prior for frailty parameter $\theta$ is gamma distribution with mean one and large variance, $\Gamma(\phi, \phi)$, say with a small choice of $\phi$ and the prior for regression parameter $\beta$ is normal with mean zero and large variance say $\sigma^2$. Similar types of prior distributions are used in Ibrahim et al. (2001), Sahu et al. (1997) and Santos et al. (2010). So, in our study also we use same prior for $\theta$ and $\beta$’s. Since, we do not have any prior information about baseline parameters, $\lambda_1, \gamma_1, \lambda_2,$ and $\gamma_2$ prior distributions are assumed to be informative. For baseline parameters $\lambda_1, \lambda_2, \gamma_1$ and $\gamma_2$, the prior is $\Gamma(a, b)$, for the frailty parameter $\theta$, the prior is $\Gamma(\phi, \phi)$, and for regression coefficients $\beta_i$, ($i = 1, \cdots, 5$), the prior is $N(0, \sigma^2)$. Here $\Gamma(a, b)$ is gamma distribution with shape parameter $a$ and scale parameter $b$. All the hyper-parameters $\phi, a,$
$b$, and $\sigma^2$ are known. We set hyper-parameters as $\phi = 0.0001$, $a = 1$; $b = 0.0001$, and $\sigma^2 = 10^5$.

To estimate the parameters we run two parallel Markov chains with the different starting points using Metropolis-Hastings algorithm within Gibbs sampling. We iterate both the chains for 95,000 times and discard 2,500 observations as burn-in. As in the simulation here also results for chain I chain II are similar so we present result for only one chain (i.e. chain I).

Figure 6 shows the trace plot, coupling from past plot and sample autocorrelation plot for the parameter $\beta_2$ for chain I. For other parameters graphs have similar pattern so due to lack of space we are not presenting graphs for other parameters. Trace plots for all the parameters shows zigzag pattern which indicates that parameters move more freely. Table 9 and Table 10 show that Gelman Rubin test statistic values are nearly equal to one and Geweke test statistic values are quite small and corresponding p-values are large enough to say the chains attains stationary distribution. Simulated values of parameters have autocorrelation of lag k (values given in Table 11), so every kth iteration is selected as sample. The posterior mean and standard error with 95% credible intervals for baseline parameters, frailty parameter and regression coefficients are presented in Table 11.

In particular, the sex effect $\beta_2 = -0.9148345$ indicates that in the frailty models the female patients have a significantly lower infection rate than male patients. It is also clear from Table 11 that the effect of sex and age covariates are significant only but the effect of other covariates are not. Also, estimate of $\theta = 0.1640389$ shows that there exists heterogeneity in population of patients. As per result, we can say that some patients are expected to be more prone to infection as compared to others with the same covariate value. Posterior estimate of $\theta$ also provide evidence of positive dependence between two infection times for the same patient.

9 Conclusions

The present study focuses on parametric models, which implies parametric specification of the baseline hazard and the distribution of the frailty. This paper presents some results for the shared frailty models. In this paper, we have considered two failure times by allowing
for potential dependence in the random quantities corresponding to each failure time which is induced by frailty. Here we have considered the log-logistic distribution for modeling the lifetime of two components and frailties are assumed to follow a gamma distribution. The log-logistic distribution provides a useful alternative to the Weibull distribution for the parametric modeling of survival data where the hazard rate is non-monotonic. The log-logistic distribution has a non-monotonic hazard function which makes it suitable for modeling some sets of cancer survival data. The model may be fitted easily to right and left-censored data.

We have discussed the Bayesian estimation procedure including Gibbs sampling for computing the estimation of the unknown parameters by simulating samples of different sizes $n = 50, 75, \text{ and } 100$. We have clearly written the steps involved in the iteration procedure. As expected, the estimations for the larger sample size are far more accurate. As Bayesian methods proved to be very time-consuming, we have not generated large sample sizes, say, more than 100 for the simulation study. We have provided 95,000 iterations to perform simulation study.

We have run two parallel chains from different starting points and considered the burn-in interval for each chain. The quality of convergence was checked by Gelman-Rubin statistics (see Brooks and Gelman, 1998). The values of the Gelman-Rubin statistics in this case are quite close to one. Thus, the sample can be considered to have arisen from stationary distribution. The simulation results and real data analysis indicate that the performance of Bayesian estimation method is quite satisfactory.

ACKNOWLEDGMENTS

We thank both the referees for their valuable suggestions and comments.

References


Appendix A: Figures and Summary Tables

Figure 1: Graphs for goodness-of-fit for Acute Leukemia data

(a) Placebo treatment
(b) 6-MP drug treatment

Figure 2: Graphs for goodness-of-fit for Kidney Infection data

(a) Recurrence Time1
(b) Recurrence Time2
Figure 3: Comparison of Non Parametric Survival Function with Log-logistic Survival Function for Acute Leukemia data

Figure 4: Comparison of Non Parametric Survival Function with Log-logistic Survival Function for Kidney Infection Data
Figure 5: (a) Trace plot (b) Coupling from past plot and (c) Sample autocorrelation plot for the parameter $\beta_1$ for Acute Leukemia Data
Figure 6: (a) Trace plot (b) Coupling from past plot and (c) Sample autocorrelation plot for the parameter $\beta_2$ for Kidney Infection Data Set.
Table 1: Gelman-Rubin Convergence Statistic Values for Simulation

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\lambda_1$</th>
<th>$\lambda_2$</th>
<th>$\gamma_1$</th>
<th>$\gamma_2$</th>
<th>$\theta$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>True values</td>
<td>3.0</td>
<td>3.0</td>
<td>1.5</td>
<td>1.5</td>
<td>3.6</td>
<td>0.5</td>
</tr>
<tr>
<td>n = 50</td>
<td>1.006485</td>
<td>1.000876</td>
<td>1.001707</td>
<td>1.000411</td>
<td>0.9999929</td>
<td>1.000884</td>
</tr>
<tr>
<td>n = 75</td>
<td>1.000467</td>
<td>1.006047</td>
<td>1.002898</td>
<td>1.000835</td>
<td>1.000178</td>
<td>1.006903</td>
</tr>
<tr>
<td>n = 100</td>
<td>1.011923</td>
<td>1.002351</td>
<td>1.006485</td>
<td>1.01742</td>
<td>1.00376</td>
<td>1.032138</td>
</tr>
</tbody>
</table>

Table 2: Geweke Test Values and Corresponding p-values for Simulation

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\lambda_1$</th>
<th>$\lambda_2$</th>
<th>$\gamma_1$</th>
<th>$\gamma_2$</th>
<th>$\theta$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>True values</td>
<td>3.0</td>
<td>3.0</td>
<td>1.5</td>
<td>1.5</td>
<td>3.6</td>
<td>0.5</td>
</tr>
<tr>
<td>n = 50</td>
<td>0.001568422</td>
<td>-0.00146471</td>
<td>-0.004875977</td>
<td>-0.00305441</td>
<td>-0.004216671</td>
<td>-0.004235707</td>
</tr>
<tr>
<td>p-value</td>
<td>0.5006257</td>
<td>0.4994157</td>
<td>0.4980548</td>
<td>0.4987815</td>
<td>0.4983178</td>
<td>0.4983102</td>
</tr>
<tr>
<td>n = 75</td>
<td>-0.003089062</td>
<td>0.004712401</td>
<td>0.008740293</td>
<td>-0.00215305</td>
<td>0.005469704</td>
<td>0.002981995</td>
</tr>
<tr>
<td>p-value</td>
<td>0.4987676</td>
<td>0.50188</td>
<td>0.5034868</td>
<td>0.4991411</td>
<td>0.5021821</td>
<td>0.5011896</td>
</tr>
<tr>
<td>n = 100</td>
<td>-0.02232292</td>
<td>-0.009077437</td>
<td>0.02401819</td>
<td>0.008079622</td>
<td>0.0158848</td>
<td>0.001896258</td>
</tr>
<tr>
<td>p-value</td>
<td>0.4910952</td>
<td>0.4963787</td>
<td>0.509581</td>
<td>0.5032233</td>
<td>0.5063369</td>
<td>0.5007565</td>
</tr>
</tbody>
</table>
Table 3: Parameters Estimates for a Shared Gamma Frailty Model with Log-logistic Baseline Hazards for Simulation

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\lambda_1$</th>
<th>$\lambda_2$</th>
<th>$\gamma_1$</th>
<th>$\gamma_2$</th>
<th>$\theta$</th>
<th>$\beta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>True values</td>
<td>3.0</td>
<td>3.0</td>
<td>1.5</td>
<td>1.5</td>
<td>3.6</td>
<td>0.5</td>
</tr>
</tbody>
</table>

$n = 50$

| Estimates | 2.709626 | 3.206486 | 1.316447 | 1.268816 | 3.35878  | 0.4036909 |
| S.E. | 0.5129031 | 0.5479131 | 0.1924458 | 0.1769939 | 0.3020328 | 0.1117505 |
| Lower Lt. | 2.03228  | 2.012305  | 0.9554537 | 0.9505256 | 3.03909  | 0.2864242 |
| Upper Lt. | 3.898679 | 3.9336    | 1.997687  | 1.947903  | 4.127484 | 0.7552391 |

$n = 75$

| Estimates | 3.10579 | 2.833898 | 1.665907 | 1.585118 | 3.503197 | 0.5868682 |
| S.E. | 0.4913551 | 0.5024563 | 0.1508131 | 0.1579397 | 0.2954702 | 0.08595581 |
| Lower Lt. | 2.12463 | 2.056955 | 1.362487 | 1.273275 | 3.057737 | 0.3342988 |
| Upper Lt. | 3.840332 | 3.875176 | 1.966371 | 1.803529 | 4.110113 | 0.7518857 |

$n = 100$

| Estimates | 2.968941 | 2.888897 | 1.592575 | 1.440733 | 3.576461 | 0.5787032 |
| S.E. | 0.4686344 | 0.421927 | 0.1445806 | 0.1300566 | 0.2842008 | 0.082332 |
| Lower Lt. | 2.212934 | 2.075327 | 1.391031 | 1.275807 | 3.131777 | 0.410188 |
| Upper Lt. | 3.741437 | 3.82639  | 1.95811  | 1.705194 | 4.101382 | 0.7307128 |
Table 4: p-values for Kolmogorov-Smirnov Test Statistic and Hollander & Proschan’s Test Statistic for Acute Leukemia data set.

<table>
<thead>
<tr>
<th>Test</th>
<th>TP</th>
<th>T6</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-S</td>
<td>0.9625</td>
<td>0.9634</td>
</tr>
<tr>
<td>HP</td>
<td>1.0000</td>
<td>0.1662</td>
</tr>
</tbody>
</table>

Table 5: Gelman-Rubin Convergence Statistic Values for Acute Leukemia data set.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>(\lambda_1)</th>
<th>(\lambda_2)</th>
<th>(\gamma_1)</th>
<th>(\gamma_2)</th>
<th>(\theta)</th>
<th>(\beta_1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>stat. value</td>
<td>1.009167</td>
<td>0.999928</td>
<td>1.009546</td>
<td>1.008413</td>
<td>1.015725</td>
<td>1.009753</td>
</tr>
</tbody>
</table>

Table 6: Geweke Test Values and Corresponding p-values for Acute Leukemia data set.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>(\lambda_1)</th>
<th>(\lambda_2)</th>
<th>(\gamma_1)</th>
<th>(\gamma_2)</th>
<th>(\theta)</th>
<th>(\beta_1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>test value</td>
<td>0.0007616</td>
<td>0.0102423</td>
<td>0.0028076</td>
<td>0.0013007</td>
<td>-0.0000832</td>
<td>-0.003081</td>
</tr>
<tr>
<td>p – value</td>
<td>0.5003038</td>
<td>0.5040861</td>
<td>0.5011201</td>
<td>0.5005189</td>
<td>0.4999668</td>
<td>0.498771</td>
</tr>
</tbody>
</table>
### Table 7: Parameters Estimates for Acute Leukemia data set.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Autocorrelation Lag</th>
<th>Estimates</th>
<th>S.E.</th>
<th>LowerLt.</th>
<th>UpperLt.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_1$</td>
<td>35</td>
<td>0.0597621</td>
<td>0.0090631</td>
<td>0.0358654</td>
<td>0.0696451</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>20</td>
<td>0.0507481</td>
<td>0.0142916</td>
<td>0.0182246</td>
<td>0.0693433</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>130</td>
<td>1.190208</td>
<td>0.4950038</td>
<td>0.5393428</td>
<td>2.453389</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>120</td>
<td>0.5274395</td>
<td>0.2643718</td>
<td>0.2070135</td>
<td>1.203408</td>
</tr>
<tr>
<td>$\theta$</td>
<td>180</td>
<td>0.4941678</td>
<td>0.2878033</td>
<td>0.0598422</td>
<td>1.165824</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>130</td>
<td>-1.766374</td>
<td>0.4719272</td>
<td>-2.698818</td>
<td>-0.8937898</td>
</tr>
</tbody>
</table>

### Table 8: p-values for Kolmogorov-Smirnov Test Statistic and Hollander & Proschan’s Test Statistic for Kidney Infection data set.

<table>
<thead>
<tr>
<th>Test</th>
<th>RecurrenceTime1</th>
<th>RecurrenceTime2</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-S</td>
<td>0.9720</td>
<td>0.9662</td>
</tr>
<tr>
<td>HP</td>
<td>0.9709</td>
<td>0.9022</td>
</tr>
</tbody>
</table>

### Table 9: Gelman-Rubin Convergence Statistic Values for Kidney Infection data set.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>$\lambda_1$</th>
<th>$\lambda_2$</th>
<th>$\gamma_1$</th>
<th>$\gamma_2$</th>
<th>$\theta$</th>
</tr>
</thead>
<tbody>
<tr>
<td>stat. value</td>
<td>1.004909</td>
<td>1.000077</td>
<td>1.032916</td>
<td>1.018915</td>
<td>1.002538</td>
</tr>
<tr>
<td>Parameters</td>
<td>$\beta_1$</td>
<td>$\beta_2$</td>
<td>$\beta_3$</td>
<td>$\beta_4$</td>
<td>$\beta_5$</td>
</tr>
<tr>
<td>stat. value</td>
<td>1.000426</td>
<td>1.005524</td>
<td>1.016267</td>
<td>1.000706</td>
<td>1.004309</td>
</tr>
</tbody>
</table>
Table 10: Geweke Test Values and Corresponding p-values for Kidney Infection data set.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>test value</th>
<th>p - value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_1$</td>
<td>-0.004168567</td>
<td>0.498337</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>0.01184535</td>
<td>0.5047255</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>-0.008293802</td>
<td>0.4966913</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>-0.007371464</td>
<td>0.4970592</td>
</tr>
<tr>
<td>$\theta$</td>
<td>-0.00295372</td>
<td>0.4988216</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>0.003982676</td>
<td>0.5015889</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>0.004766793</td>
<td>0.5019017</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>-0.005057773</td>
<td>0.4979822</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>-0.004597094</td>
<td>0.498166</td>
</tr>
<tr>
<td>$\beta_5$</td>
<td>-0.0002992744</td>
<td>0.4998806</td>
</tr>
</tbody>
</table>

Table 11: Parameters Estimates for Kidney Infection data set.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Autocorrelation Lag</th>
<th>Estimates</th>
<th>S.E.</th>
<th>LowerLt.</th>
<th>UpperLt.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\lambda_1$</td>
<td>45</td>
<td>0.06009937</td>
<td>0.008760602</td>
<td>0.03679422</td>
<td>0.06970225</td>
</tr>
<tr>
<td>$\lambda_2$</td>
<td>45</td>
<td>0.05789328</td>
<td>0.00989742</td>
<td>0.03365606</td>
<td>0.06961944</td>
</tr>
<tr>
<td>$\gamma_1$</td>
<td>180</td>
<td>0.8740553</td>
<td>0.2885247</td>
<td>0.4469254</td>
<td>1.507527</td>
</tr>
<tr>
<td>$\gamma_2$</td>
<td>180</td>
<td>0.7458508</td>
<td>0.2603728</td>
<td>0.3869899</td>
<td>1.377523</td>
</tr>
<tr>
<td>$\theta$</td>
<td>100</td>
<td>0.1640389</td>
<td>0.1214242</td>
<td>0.05290306</td>
<td>0.4979691</td>
</tr>
<tr>
<td>$\beta_1$</td>
<td>190</td>
<td>-0.02924777</td>
<td>0.009013905</td>
<td>-0.04772919</td>
<td>-0.01289544</td>
</tr>
<tr>
<td>$\beta_2$</td>
<td>70</td>
<td>-0.9148345</td>
<td>0.315259</td>
<td>-1.525267</td>
<td>-0.3017047</td>
</tr>
<tr>
<td>$\beta_3$</td>
<td>100</td>
<td>0.05695645</td>
<td>0.4394808</td>
<td>-0.8129304</td>
<td>0.9164792</td>
</tr>
<tr>
<td>$\beta_4$</td>
<td>120</td>
<td>0.2982645</td>
<td>0.4603029</td>
<td>-0.6184608</td>
<td>1.1687</td>
</tr>
<tr>
<td>$\beta_5$</td>
<td>80</td>
<td>-0.2886985</td>
<td>0.6299306</td>
<td>-1.560267</td>
<td>0.8128706</td>
</tr>
</tbody>
</table>