Chapter 1

Introduction

1.1 Introduction

The statistical analysis of survival data is an important topic in many areas, including medicine, epidemiology, biology, demography, economics, engineering and other fields. A variety of techniques have been developed to analyse survival data. A common approach to the analysis of survival data is based on the assumption that the study population is homogeneous. That is, conditional on the covariates, every individual has the same risk of experiencing an event such as death or disease recurrence. Proportional hazards models and accelerated failure time models are classical models that are frequently used for the analysis of univariate (censored) survival data.

Standard survival models assume independence between the survival times. Frailty models provide a useful extension of the standard survival models by introducing a random effect (frailty) when the survival data are correlated. Frailty models can be used in the survival analysis to represent random effects or unexplained heterogeneity between individuals or groups. Multivariate or cluster failure time data are commonly encountered in the survival analysis, and finding an appropriate method to model the correlation among the ob-
servations is a very important issue. Frailty model provide an appropriate method to model the correlation among the multivariate data. Multivariate data occurs for example if lifetimes (or times of onset of a disease) of relatives (twins, parent-child) or recurrent events like infections in the same individual are considered. In such cases independence between the clustered survival times can not be assumed. A convenient choice for modeling the correlation in survival data is the frailty model, which was originally proposed by Vaupel et al. (1979). The term frailty was first introduced by Vaupel et al. (1979) in univariate survival models. Frailty models extend the Proportional Hazard model by including random effects, called frailty, to account for dependency between observations. However, it is not always reasonable to assume that all the individuals in the same cluster share exactly the same frailty. An extension of the shared frailty model is the correlated frailty model where individuals in the same cluster have different, yet correlated, frailties.

1.2 Frailty Models

The concept of frailty provides a suitable way to introduce random effects in the model to account for association and an unobserved heterogeneity. In its simplest form, a frailty is an unobserved random factor that modifies multiplicatively the hazard function of an individual or a group or cluster of individuals.

1.3 Consequences of Ignoring Frailty

It is very important to consider the effect of ignoring frailty where the existence of heterogeneity may be present. The impact of frailty in event history models differs substantially from the impact of frailty in linear regres-
sion models. In ordinary regression models, unobserved heterogeneity leads to more variability of the response compared to the case when the variables are included. In event history data, however, the increased variability implies a change in the hazard function. When the hazard rate exhibits positive duration dependence, ignoring frailty will make this duration dependence, less pronounced or even negative. When the hazard rate exhibits negative duration dependence, ignoring frailty will make this negative duration dependence stronger. Another consequence of ignoring frailty is that the effect of a covariate is biased towards zero. The consequence of ignoring frailty factors, which is very important in modeling stochastic risk factor, is that intensity estimation in credit risk models would be biased. Therefore, the exposure at the default cannot be integrated into the credit portfolio model (Grundke, 2004). This is especially a problem for market-driven instruments, such as interest rate derivatives.

It is recognised in the field of econometrics and biometrics, through empirical evidence, that if frailty is present but ignored then covariate effects will be underestimated (Lancaster, 1979; Hougaard et al.,1994 and Pickles and Crouchley, 1994,1995). Lancaster (1990) confirmed this evidence for uncensored Weibull survival data through theoretical work. Henderson and Oman (1999) showed that fitting misspecified Cox proportional hazards models to the marginal distributions (ignoring frailty) leads to regression coefficient estimates biased towards zero by an amount which depends on the variability of the frailty terms and the form of frailty distribution. He also concluded that the fitted marginal survival curves can also differ substantially from the true marginals. In the analysis of multivariate failure time data, failure to account for dependency has been shown to lead to biased parameter estimators (Klein and Moeschberger, 1988). Moreover, ignoring frailty effects with finite mean may result in a negative bias in the estimated time dependence (Lancaster and
Nickell, 1980; Yashin et al., 1985). Epidemiological relative risk measures are vulnerable to the frailty phenomenon (Aalen, 1999a 1999b), as frailty reduces the influence of known covariates on the relative risk. Consequently, this can lead to non-proportional hazards.

1.4 Modeling Frailty

The generalization of the Cox proportional hazards model (Cox, 1972) is the best and widely applied model that allows for the random effect by multiplicatively adjusting the baseline hazard function. Frailty models extend Cox proportional hazards model by introducing unobserved frailties to the model. In this case, the hazard rate will not be just a function of covariates, but also a function of frailties. A frailty model is a random effects model which has a multiplicative effect on the hazard rates of all the members of the subgroups. In univariate survival models, it can be used to model the heterogeneity among individuals, which is the influence of an unobserved risk factors in a proportional hazards model. In multivariate survival models, shared frailty model is used to model the dependence between the individuals in the group. In the multivariate case, unobserved frailty is common to a group of individuals.

In a univariate frailty model, let a continuous random variable $T$ be a lifetime of an individual and the random variable $Z$ be frailty variable. The conditional hazard function for a given frailty variable, $Z = z$ at time $t > 0$ is,

$$ h(t \mid z) = z h_0(t) e^{X \beta}, \quad (1.1) $$

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

$$ S(t \mid z) = e^{- \int_0^t h(x \mid z) dx} = e^{-z h_0(t) e^{X \beta}}, \quad (1.2) $$
where $H_0(t)$ is the cumulative baseline hazard function at time $t > 0$. Inte-
grating over the range of frailty variable $Z$ having density $f(z)$, we get the marginal survival function as,

$$S(t) = \int_0^\infty S(t \mid z)f(z)dz$$

$$= \int_0^\infty e^{-zH_0(t)}e^{X\beta}f(z)dz$$

$$= L_Z(H_0(t)e^{X\beta}), \quad (1.3)$$

where $L_Z(.)$ is the Laplace transformation of the distribution of $Z$. Once we get the survival function at time $t > 0$, of life time random variable for an individual, we can obtain probability structure and make their inferences based on it.

Research on the bivariate survival models has grown rapidly several years in the past. Clayton’s (1978) random effect model of the bivariate survival was a key innovation. He introduced the notion of the shared relative risk. This model was further developed by Oakes (1982) to analyze the association between two non-negative random variables. Clayton and Cuzick (1985) added observed covariates to the bivariate survival model with the shared relative risk. Hougaard (1986) proposed the random effect models of the bivariate Weibull distributions. He also discussed several other bivariate distributions with biomedical and reliability applications. Oakes (1989) developed a shared frailty model related to the “archimedean distributions” studied by Genest and MacKay (1986). He also proposed a local time dependent association measure between bivariate life spans and discussed its use for a large class of bivariate survival functions. Vaupel (1991a, b), Vaupel et al. (1991, 1992), Nielsen et al. (1992) studied genetic and environmental influences on longevity using bivariate survival models. Hanagal (2006) discussed the gamma frailty regression model in the bivariate survival data and Hanagal (2007) also presented the gamma frailty regression models in the mixture distributions. Hanagal (2008)
proposed a bivariate Weibull regression model with heterogeneity (frailty or random effect) which is generated by the log-normal distribution. Hanagal (2010) deals with modeling heterogeneity for bivariate survival data by the compound Poisson distribution with random scale. Shared frailty models are the most commonly used frailty models in literature, where individuals in the same cluster share a common frailty.

1.5 General Shared Frailty Model

The shared frailty model is relevant to event time of the related individuals, similar organs and repeated measurements. In this model individuals from a group shares common covariates. For the shared frailty model it is assumed that survival times are conditionally independent, for given shared frailty. Shared frailty means dependence between survival times is only due to unobservable covariates or frailty. When there is no variability in the distribution of frailty variable $Z$ that implies $Z$ has a degenerate distribution and when the distribution of $Z$ is not degenerate the dependence is positive.

Suppose $n$ individuals are observed for the study and let a bivariate random variable $(T_{1j}, T_{2j})$ represent the first and the second the survival times of the $j^{th}$ individual ($j = 1, 2, \ldots, n$). Also suppose that there are $k$ observed covariates collected in a vector $X_j = (X_{1j}, \ldots, X_{kj})$ for the $j^{th}$ individual where $X_{aj}$ ($a = 1, 2, \ldots, k$) represents the value of the $a^{th}$ observed covariate for the $j^{th}$ individual. Here we assume that the first and the second survival times for each individual share the same value of the covariates. Let $Z_j$ be shared frailty for the $j^{th}$ individual. Assuming that the frailties are acting multiplicatively on the baseline hazard function and both the survival times of individuals are conditionally independent for given frailty, the conditional hazard function for the $j^{th}$ individual at the $i^{th}$ ($i = 1, 2$) survival time $t_{ij} > 0$ for given frailty $Z_j$
Introduction

\[ h(t_{ij} \mid z_j, X_j) = z_j h_0(t_{ij}) e^{X_j \beta} \] (1.4)

where \( h_0(t_{ij}) \) is the baseline hazard at time \( t_{ij} > 0 \) and \( \beta \) is a vector of order \( k \), of the regression coefficients. The conditional cumulative hazard function for the \( j \)th individual at the \( i \)th survival time \( t_{ij} > 0 \) for a given frailty \( Z_j = z_j \) is,

\[ H(t_{ij} \mid z_j, X_j) = z_j H_0(t_{ij}) \eta_j \] (1.5)

where \( \eta_j = e^{X_j \beta} \) and \( H_0(t_{ij}) \) is the cumulative baseline hazard function at time \( t_{ij} > 0 \). The conditional survival function for the \( j \)th individual at the \( i \)th survival time \( t_{ij} > 0 \) for a given frailty, \( Z_j = z_j \) is,

\[ S(t_{ij} \mid z_j, X_j) = e^{-H(t_{ij} \mid z_j, X_j)} \]
\[ = e^{-z_j H_0(t_{ij}) \eta_j} \] (1.6)

Under the assumption of independence, the bivariate conditional survival function for a given frailty \( Z_j = z_j \) at time \( t_{1j} > 0 \) and \( t_{2j} > 0 \) is,

\[ S(t_{1j}, t_{2j} \mid z_j, X_j) = S(t_{1j} \mid z_j, X_j) S(t_{2j} \mid z_j, X_j) \]
\[ = e^{-z_j (H_{01}(t_{1j}) + H_{02}(t_{2j})) \eta_j} \] (1.7)

The unconditional bivariate survival function at time \( t_{1j} > 0 \) and \( t_{2j} > 0 \) can be obtained by integrating over the frailty variable \( Z_j \) having the probability function \( f_Z(z_j) \), for the \( j \)th individual.

\[ S(t_{1j}, t_{2j} \mid X_j) = \int_{Z_j} S(t_{1j}, t_{2j} \mid z_j) f_Z(z_j) dz_j \]
\[ = \int_{Z_j} e^{-z_j (H_{01}(t_{1j}) + H_{02}(t_{2j})) \eta_j} f_Z(z_j) dz_j \]
\[ = L_{Z_j} [H_{01}(t_{1j}) + H_{02}(t_{2j})) \eta_j] \] (1.8)
where $L_{Z_j}(\cdot)$ is the Laplace transform of the frailty variable of $Z_j$ for the $j^{th}$ individual. Here onwards we represent $S(t_{1j}, t_{2j} | X_j)$ as $S(t_{1j}, t_{2j})$.

Shared frailty explains correlations between subjects within clusters. However, it does have some limitations. To avoid these limitations, correlated frailty models are being developed for the analysis of multivariate failure time data, in which associated random variables are used to characterize the frailty effect for each cluster. Correlated frailty models provide not only variance parameters of the frailties as in shared frailty models, but they also contain additional parameter for modeling the correlation between frailties in each group. Frequently one is interested in construction of a bivariate extension of some univariate family distributions (e.g., gamma). For example, for the purpose of genetic analysis of frailty one might be interested in estimation of correlation of frailty. It turns out that it is possible to carry out such extension for the class of infinitely-divisible distributions (Iachine 1995a, 1995b). In this case an additional parameter representing the correlation coefficient of the bivariate frailty distribution is introduced.

### 1.6 General Correlated Frailty Model

The correlated frailty model is a natural extension of the shared frailty approach. In the shared frailty model frailties are the same, but for correlated frailty two individuals in a pair, frailties are not necessarily the same. Also we are assuming that the frailties are acting multiplicatively on the baseline hazard function (proportional hazards model) and that the observations in a pair are conditionally independent, given the frailties. Hence, the hazard of the individual $i$ ($i = 1, 2$) has the form,

$$h(t|X_i, Z_i) = Z_i h_0(t)e^{\beta'X_i},$$

(1.9)
where $t$ denotes age or time, $X_i$ is a vector of observed covariates, $\beta$ is a vector of regression parameters describing the effect of the covariates $X_i$, $h_0(\cdot)$ are baseline hazard functions, and $Z_i$ are frailties. Bivariate correlated frailty models are characterized by the joint distribution of a two-dimensional vector of frailties $(Z_1, Z_2)$. If the two frailties are independent, the resulting lifetimes are independent, and no clustering is present in the model. The shared frailty model can be obtained as a special case of the correlated frailty model when both the frailties are equal, (Wienke, 2011).

In order to derive a marginal likelihood function, the assumption of conditional independence of lifespans, given the frailty, is used. Let $\delta_{ij}$ be a censoring indicator for individual $i (i = 1, 2)$ in pair $j (j = 1, \ldots, n)$. Indicator $\delta_{ij}$ is 1 if the individual has experienced the event of interest, and 0 otherwise. According to 1.9, the conditional survival function of the $i$th individual in the $j$th pair is,

$$S(t|X_{ij}, Z_{ij}) = e^{Z_{ij}h_{0i}(t)e^{\beta'X_{ij}}},$$

with $H_{0i}(t)$ denoting the cumulative baseline hazard function. Here and in the following, $S$ is used as a generic symbol for a survival function. The contribution of individual $i (i = 1, 2)$ in pair $j (j = 1, \ldots, n)$ to the conditional likelihood is given by,

$$\left[ Z_{ij}h_{0i}(t)e^{\beta'X_{ij}} \right]^{\delta_{ij}} e^{Z_{ij}h_{0i}(t_{ij})e^{\beta'X_{ij}}},$$

where $t_{ij}$ stands for observation time of individual $i$ from pair $j$. Assuming the conditional independence of lifespans, given the frailty, and integrating out the frailty, we obtain the marginal likelihood function,

$$\prod_{j=1}^{n} \int \int \left[ u_{1j}h_{01}(t_{1j})e^{\beta'X_{1j}} \right]^{\delta_{1j}} e^{u_{1j}h_{01}(t_{1j})e^{\beta'X_{1j}}} f(u_{1j}, u_{2j})du_{1j}du_{2j}$$

$$\left[ u_{2j}h_{02}(t_{2j})e^{\beta'X_{2j}} \right]^{\delta_{2j}} e^{u_{2j}h_{02}(t_{2j})e^{\beta'X_{2j}}} (1.12)$$
where $f(\cdot, \cdot)$ is the probability density function of the corresponding frailty distribution. For more details see Hanagal (2011) and Wienke (2011).

1.7 Reversed Hazard Rate

Reversed hazard rate was proposed as a dual to the hazard rate by Barlow et al. (1963) as

$$m(t) = \frac{f(t)}{F(t)},$$

where $F(t)$ denotes the distribution function and $f(t)$ represents the probability density function, if it exists, of lifetime $T$. Block et al. (1998) provided a general definition of reversed hazard rate (RHR) as,

$$m(t) = \lim_{\Delta t \to 0} \frac{P(t - \Delta t < T \leq t | T \leq t)}{\Delta t}, \ t > 0. \quad (1.13)$$

The reversed hazard rate specifies the instantaneous rate of death or failure at time $t$, given that it failed before time $t$. Thus in a small interval, $m(t) \Delta t$ is the approximate probability of failure in the interval, given failure before the end of the interval $(t - \Delta t, t]$.

1.8 Model Selection Criteria

In order to compare the proposed models we use the Bayesian Information Criteria (BIC), the Akaike Information Criteria (AIC) and the Deviance Information Criteria (DIC). The BIC was introduced by Schwarz (1978) and the BIC is defined as,

$$BIC = D(\hat{\Theta}) + p \cdot \ln(n) \quad (1.14)$$

where $p$ represents number of parameters of the model and $n$ represents number of data points. $D(\hat{\Theta})$ represents an estimate of the deviance evaluated at the
posterior mean $\hat{\Theta} = E(\Theta \mid \text{data})$. The deviance is defined by, $D(\Theta) = -2 \cdot \log L(\Theta)$, where $\Theta$ is a vector of unknown parameters of the model and $L(\Theta)$ is the likelihood function of the model. The AIC was introduced by Akaike (1973) and the AIC is defined as,

$$AIC = D(\hat{\Theta}) + 2p$$

(1.15)

DIC, a generalization of AIC was introduced by Spiegelhalter et al. (2002) and is defined as,

$$DIC = D(\hat{\Theta}) + 2 \cdot p_D$$

(1.16)

where $p_D$ is the difference between the posterior mean of the deviance and the deviance of the posterior mean of parameters of interest, that is, $p_D = \bar{D} - D(\hat{\Theta})$, where $\bar{D} = E(D(\Theta) \mid \text{data})$.

The Bayesian model examination for adequacy and model comparison can be proceeds by the predictive distribution. Let $\underline{y} = \{y_1, y_2, ..., y_n\}$ be a set of observations, where $n$ is the total number of observations and $y_{\text{obs}}$ denotes realization of $\underline{y}$. The posterior predictive density $\pi(y \mid y_{\text{obs}})$ is the predictive density of a new independent set of observations under the model, given the actual set of observations. By marginalizing $\pi(y \mid y_{\text{obs}})$, we obtain the posterior predictive density of a single observation $y_r$, $r = 1, 2, ..., n$ as follows,

$$\pi(y_r \mid y_{\text{obs}}) = \int \pi(y_r \mid \theta)\pi(\theta \mid y_{\text{obs}})d\theta.$$  

(1.17)

A simple checking for the assessment of model is the predictive interval. Suppose we generate a sample $y_{r1}, y_{r2}, ..., y_{rn}$ from the predictive density (1.17) for $r^{th}$ observation and create the $100(1 - \alpha)^0/0$ equal tailed credible interval also known as the predictive interval, then the model under consideration would be an adequate model for data if $100(1 - \alpha)^0/0$ of the $y_{r,\text{obs}}$ to fall in their respective interval.
To draw a random sample from the predictive density (1.17), suppose we have $\theta_j^*; (j = 1, 2, ..., n) n$ samples from the posterior density $\pi(\theta | y)$ possibly using one of the MCMC methods. Then a random sample $y_r^j$ drawn from $\pi(y_r | \theta_j^*)$ is a sample from the predictive density (1.17), since for given parameters $\theta$, if observations are conditionally independent then $\pi(y_r | y, \theta) = \pi(y_r | \theta)$.

Another approach for model selection is based on cross-validation predictive density. The cross-validation predictive densities are the set

\[
\{ f(y_r | Y_{(r)}); r = 1, 2, ..., n \}
\]

where $Y_{(r)}$ denotes all elements of data set $y$ except observation $y_r$ and

\[
f(y_r | Y_{(r)}) = \int f(y_r | \theta, Y_{(r)}) f(\theta | Y_{(r)}) d\theta
\]  

(1.18)

$f(y_{r, obs} | Y_{(r), obs})$ is popularly known as the conditional predictive ordinate (CPO). The smaller values of CPO does not support the model, so we prefer a model for which CPO values are higher than others. We can compare more than one models using CPO’s. We plot CPO’s versus $r$ for different models in a single graph and compare the models visually. If we plot difference of CPO values for the models A and B, i.e. $CPO_A - CPO_B$ then negative differences favour model B where positive difference favour model A. The larger the positive or negative difference better the model A or B. Sometimes CPO values are quite close to each other so that difference may be clustered at zero and plot can not be distinguishable. To overcome this situation one can use log of CPO values to plot.

For the cross-validation predictive density, in general we can write,

\[
f(y_r | Y_{(r)}) = \frac{f(y)}{f(Y_{(r)})}
\]

\[
= \frac{1}{\int \frac{f(y_r | \theta, Y_{(r)})}{f(y_r | \theta, y)} f(\theta | y) d\theta}
\]

\[
= \frac{1}{\int f(y_r | y_{(r)}, Y_{(r)}) f(\theta | y_r, Y_{(r)}) d\theta}.
\]
Introduction

So an immediate Monte Carlo estimate of CPO, is given by,

\[ \sum_{j=1}^{n} \frac{1}{f(y_{r,obs} | \theta_{j}^*)} \]

which is the harmonic mean of the conditional density function of \( y_r \) evaluated at the posterior sample values. See Gelfand(1996) for more details.

The Bayes factor \( B_{jk} \) for a model \( M_j \) against \( M_k \) for a given data \( D = (t_{1j}, t_{2j}); (j = 1, 2, \ldots, n) \) is,

\[ B_{jk} = \frac{P(D|M_j)}{P(D|M_k)} \]

where \( P(D|M_k) = \int_{S} P(D|M_k) \pi(\theta_k|M_k) d\theta_k; (k = 1, 2, \ldots, m) \) where \( \theta_k \) is the vector of unknown parameters of model \( M_k \), \( \pi(\theta_k|M_k) \) is the prior density and \( S \) is the support of the parameter \( \theta_k \). Here \( m \) represents the number of models.

Raftey (1994), following Jeffreys (1961), proposes the rules of thumb for interpreting twice the logarithm of the Bayes factor. For two models of substantive interest, \( M_j \) and \( M_k \), twice the log of the Bayes factor is approximately equal to the difference in their BIC approximations.

To compute Bayes factor we need to obtain \( I_k = P(D|M_k) \). By considering one of the approaches given in Kass and Raftery (1995), we obtain the following MCMC estimate of \( I_k \) which is given by,

\[ \hat{I}_k = \left\{ \frac{\sum_{i=1}^{N} P(D|\theta^{(i)})^{-1}}{N} \right\}^{-1} \]

which is harmonic mean of the likelihood values. Here \( N \) represents the posterior sample size and \( \{\theta^{(i)}, i = 1, 2, \ldots, N\} \) is the sample from the prior distribution.

In present study, we introduced some new shared frailty models and correlated frailty models in hazard rate and reverse hazard rate set-up. A comparison between all the introduced models is done and the best model is suggested.
For our work we restricted ourself to bivariate survival data only. Similar work can be extended to higher dimensional cases.

1.9 Chapterwise Summary

The frailty factor is random and therefore a frailty distribution needs to be specified in the frailty model. So in this study we have proposed different shared frailty models and correlate frailty models based on hazard and reversed hazard rate with different baseline distributions. We have considered simulation study to check the performance of the models. We applied models to two real life data sets and the best model is suggested for the data.

In chapter two we have proposed shared gamma frailty models with generalized Weibull distribution and generalized log-logistic distribution as baseline under random censoring in usual hazard rate set-up. We generated posterior sample using Metropolis and Hastings algorithm and Gibbs sampler and obtained posterior summary. We used different techniques to check the convergence of stationary distribution. We considered two set of priors to see the effect of prior assumption. In simulation study, we generated posterior sample of different sample sizes. We have applied these models to a real life data set of kidney infection by McGilchrist and Aisbett (1991) and the best model is suggested. Also we have considered testing of hypothesis for the significance of frailty and the regression coefficients. Part of this work has been published as “Gamma Frailty Models for Bivariate Survival Data” in the Journal of Statistical Computation and Simulation.

Chapter three we have studied inverse Gaussian shared frailty models with the same baseline distributions namely, generalized Weibull distribution and generalized log-logistic distribution under random censoring in usual hazard rate set-up. Here also we used Bayesian approach to estimate the parameters of
these models. To study effect of prior assumption, we considered same two sets of prior distribution as in chapter one. For simulation study, we generated posterior sample of sample sizes. We applied these models to the kidney infection data of McGilchrist and Aisbett (1991) and also tested of the significance for frailty and the regression coefficients. Part of this work has been accepted for publication as “Inverse Gaussian shared frailty for modeling kidney infection data” in the Advances in Reliability.

In Chapter four we have proposed correlated gamma frailty models with generalized Weibull distribution and generalized log-logistic distribution as baseline under random censoring in usual hazard rate set-up. We generated posterior sample using Metropolis and Hastings algorithm and Gibbs sampler and obtained posterior summary. We used different techniques to check the convergence of Markov chain. We considered two sets of priors to see the effect of prior assumption. In simulation study, we generated posterior sample of different sample sizes. We have applied these model to the kidney infection by McGilchrist and Aisbett (1991) and the best model is suggested. We have considered testing of hypothesis for significance of frailty and regression coefficients. Some part of this work has been send for publication to a reputed journal.

In Chapter five we have introduced gamma shared frailty models with reversed hazard rate set-up. We use three different baseline namely modified inverse Weibull distribution, and generalized log-logistic distribution type I and generalized log-logistic distribution type II. In this chapter we derive shared gamma frailty models based on reversed hazard rate. Estimation of parameters for parametric models is done using MCMC techniques. The convergence of Markov chain to stationary distribution is considered using different techniques. We have considered simulation study to see the performance of models. We generated posterior sample of sample sizes. We applied these models to the
Australian twin data given in Duffy et al. (1990) and testing of significance for frailty and regression coefficients is considered. A part of this work has been published as “Gamma Shared Frailty Model Based on Reversed Hazard Rate for Bivariate Survival Data” in the Statistics and Probability Letters (88, 190-196). Another part of this work has been accepted for publication as “Shared frailty models based on reversed hazard rate for modified inverse Weibull distribution as a baseline distribution” in the Communications in Statistics, Theory and Methods.

In Chapter six we have introduced inverse Gaussian shared frailty models with reversed hazard rate set-up. We use modified inverse Weibull distribution and generalized log-logistic distribution type I and generalized log-logistic distribution type II as baseline distributions. In this chapter we derive shared gamma frailty models based on reversed hazard rate. Estimation of parameters for parametric models is done using MCMC techniques and the convergence to stationary distribution is also considered using different techniques. In this chapter we have considered simulation study to see the performance of models. We generated posterior sample of sample sizes. We applied these models to Australian twin data given in Duffy et al. (1990) and also testing of significance for frailty and regression coefficients is considered. A part of this work has been accepted for publication as “Shared frailty models based on reversed hazard rate for modified inverse Weibull distribution as a baseline distribution” in the Communications in Statistics, Theory and Methods.

In Chapter seven we have introduced correlated gamma frailty models with reversed hazard rate set-up. We use three different baseline distributions namely modified inverse Weibull distribution and generalized log-logistic distribution type I and generalized log-logistic distribution type II. We generated posterior sample using Metropolis and Hastings algorithm and Gibbs sampler and obtained posterior summary. We used different techniques to check the
convergence of Markov chain. We considered two set of priors to see the effect of prior assumption. In simulation study, we generated posterior sample of different sample sizes. We applied these models to the Australian twin data given in Duffy et al. (1990) and the best model is suggested. Also we have considered testing of hypothesis for the significance of frailty and the regression coefficients. A part of this work has been send for publication to a reputed journal.

In the last Chapter eight a comparison between all the models is considered and the best model amongst all is suggested for the two data sets. Also the conclusion is mentioned.