International Journal of Applied Data Science
Volume1, Issue 3, 2025

Bayesian Estimation in Shared Positive Stable Frailty Models

Azar Gyasi[1], Bahareh Asadi[2]
Department of Mathematics, Faculty of mathematics, Statistics and Computer Science, Allameh Tabataba'i University, Tehran, Iran

Article Info

Received: 26-04-2025 **Revised:23-05-2025** **Accepted:20-07-2025** **Published:28-08-2025**

Abstract: The concept of frailty provides a suitable way to introduce random effects in the model to account for association and 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. In this paper, we study positive stable distribution as frailty distribution and two different baseline distributions namely Pareto and linear failure rate distribution. We estimate parameters of proposed models by introducing Bayesian estimation procedure using Markov Chain Monte Carlo (MCMC) technique. In the present study a simulation is done to compare the true values of parameters with the estimated value. We try to fit the proposed models to a real life bivariate survival data set of McGrilchrist and Aisbett (1991) related to kidney infection. Also, we present a comparison study for the same data by using model selection criterion, and suggest a better model.

Key words: Bayesian estimation, Markov chain monte carlo, Positive stable frailty, Shared frailty.
Mathematical Subject Classification: 62F15; 62N01; 62P10

Introduction:

Let T be a survival time with an absolutely continuous distribution. A non-negative random variable random variable Z is called 'frailty' if the conditional hazard function given $Z = z$ is given by

$h(t|z) = zh_0(t)$; t>0 (1.1) where $h_0(t)$ is called the baseline hazard function. Then the conditional survival function is given by,

$S(t|z) = e^{-zH(t)}$; t >0 (1.2)

where $H(t) = \int_0^t h\ (u)du$ is the cumulative baseline hazard.

And the marginal survival function

$S(t) = E[S(t|z)] = E[e^{-zH(t)}] = L_Z[H(t)]$; t >0 (1.3) where $L_Z(.\ )$ is the Laplace transform of the frailty distribution.

Hanagal (2005, 2007) proposed positive stable frailty model using bivariate Weibull distribution as a baseline distribution. Hanagal and Dabade (2013), Hanagal and Sharma (2013), Hanagal and Bhambure (2015) and Hanagal and Pandey (2015) analyzed shared gamma frailty models with different baseline distributions.

The Pareto and linear failure rate distributions serve as baseline distributions in this article, while the shared frailty model with the positive stable distribution serves as a frailty distribution. Among the suggested models, we demonstrate that the shared positive frailty with linear failure distribution as a baseline distribution is the most effective model for simulating kidney infection data. This is how the rest of the paper is structured. The characteristics of the generic shared frailty models are provided in Section 2. The shared positive stable frailty model is presented in Section 3. Section 4 provides the baseline distributions. We provide two distinct positive

stable frailty models in Section 5. The Bayesian approach for estimating the parameters of the suggested models is covered in Section 6. In section 7, we go over the various model selection criteria using the Bayesian technique. The simulation research is presented in Section 8. In Section 9, we offer a better model for the kidney infection data set by comparing the two positive stable frailty distributions above with two baseline models (without frailty). Lastly, we address the study's conclusion in Section 10.

General Shared Frailty Model:

Shared frailty models explain correlations within groups (family, litter or clinic) or for recurrent events facing the same individual. i.e. the different events within each community share a common frailty, shared by each individual within the community, each unit belongs to precisely one category. The shared gamma frailty model was suggested by Clayton (1978) for the analysis of the correlation between clustered survival times in genetic epidemiology.

Let a bivariate random variable (T_{1j}, T_{2j}) be the survival time of ith (i =1, 2) component of jth individual (j=1,2,..,n). Given the unobserved Z_j the hazard function for the (T_{1j}, T_{2j}) is given by,

$h_i(t_{ij}|z_j, \boldsymbol{X}_j) = z_j h_{0i}(t_{ij}) \exp(\boldsymbol{X}_j'\boldsymbol{\beta}), \quad i = 1,2$ (2.1)

where z_j represent frailty acting as multiplicative effect at an individual level, $h_{0i}(t_{ij})$ is the baseline hazard at time $t_{ij} > 0$ and $\boldsymbol{\beta}$ is the vector of regression coefficients with k components and $\boldsymbol{X}_j$ is the vector of observed covariates having k components.

Integrating the hazard function $h_i(t_{ij}|z_j, \boldsymbol{X}_j)$ we get, the conditional cumulative hazard function for the jth individual at the ith component survival time $t_{ij} > 0$ for the given frailty $Z_j = z_j$ is,

$H_i(t_{ij}|z_j, \boldsymbol{X}_j) = z_j H_{0i}(t_{ij})\eta_j, \quad i = 1,2$ (2.2)

where $\eta_j = \exp(\boldsymbol{X}_j'\beta)$ and $H_{0i}(t_{ij})$ is the cumulative baseline hazard function at time $t_{ij} > 0$. The conditional survival function for the jth individual at the ith component survival time $t_{ij} > 0$ for the given frailty $Z_j = z_j$ is,

$S(t_{ij}|z_j, \boldsymbol{X}_j) = e^{-Hi(t_{ij}|z_j, \boldsymbol{X}_j)}$

$= e^{-z_j H_{0i}(t_{ij})\eta_j}$ (2.3)

When T_{1j} and T_{2j} for j=1,2,..,n are independent, the bivariate conditional survival function of (T_{1j}, T_{2j}) for the given frailty $Z_j = z_j$ is the product to conditional survival function of T_{1j} and T_{2j} for the given frailty $Z_j = z_j$. Thus, we have

$S(t_{1j}, t_{2j}|z_j, \boldsymbol{X}_j) = e^{-z_j(H_{01}(t_{1j})+H_{02}(t_{2j}))\eta_j}$ (2.4) where $H_{01}(t_{1j})$ and $H_{02}(t_{2j})$ be the cumulative baseline hazard functions of first and second component at $t_{1j} > 0$ and $t_{2j} > 0$ respectively.

Integrate out the bivariate conditional survival function of (T_{1j}, T_{2j}) over the frailty variable Z_j having the probability function $f(z_j)$, for jth individual in order to obtain the unconditional bivariate survival function at time $t_{ij} > 0$.

$S(t_{1j}, t_{2j}|\underline{X}_j) = \int_{Z_j} e^{-z_j(H_{01}(t_{1j})+H_{02}(t_{2j}))\eta_j} f(z_j)\, dz_j$

$= L_{Z_j}((H_{01}(t_{1j}) + H_{02}(t_{2j}))\eta_j)$ (2.5) where $L_{Z_j}(.)$ is the Laplace transform of frailty the variable of Z_j for jth individual.

Similarly one can extend this result (2.5) to multivariate lifetimes with the shared frailty as

$S(t_{1j}, t_{2j}, \ldots, t_{kj}|\underline{X}_j) = \int_{Z_j} e^{-z_j(H_{01}(t_{1j})+H_{02}(t_{2j})+\cdots+H_{0k}(t_{kj}))\eta_j} f(z_j)\, dz_j$

$= L_{Z_j}((H_{01}(t_{1j}) + H_{02}(t_{2j}) + \cdots + H_{0k}(t_{kj}))\eta_j)$ (2.6)

Shared Positive Stable Frailty Model:

The shared gamma frailty model was suggested by Clayton (1978) for the analysis of the correlation between clustered survival times in genetic epidemiology. .In many situations, the gamma frailty model may not fit well, see (Shih, 1998; Glidden, 1999; Fan et al., 2000). The positive stable model (Hougaard,2000) is an useful alternative, in part because it has the attractive feature that predictive hazard ratio decreases to 1 over time (Oakes, 1989). The property is observed in familial associations of ages of onset of diseases with etiologic heterogeneity, where genetic cases occur early and long term survivors are weakly correlated. The gamma model has predictive hazard ratios which are time invariant and may not be suitable for the pattern of failures (Fine et al., 2003). Positive stable frailty is the special case of generalized gamma frailty given by Balakrishnan and Peng (2006) and is also special case of power variance function frailty given by Hougaard (2000). The aim of this paper is to analyze kidney infection data using positive stable frailty which has interesting properties as mentioned above.
The density function of a continuous random variable Z that has the positive stable distribution with two with the Laplace transform

$$L(s) = E[e^{-sZ}] = \exp\left(\frac{-us^{\sigma^2}}{\sigma^2}\right); \; z > 0, \; 0 < \sigma^2 < 1, v > 0 \qquad (3.2)$$

(See Hougaard, 2000).
In order to solve the non-identifiability problem, we take $\sigma^2 = v$. Thus, we have the following form of probability density function of the positive stable distribution and its Laplace transformation.

$$f(z) = \frac{1}{\pi} \sum_{k=1}^{\infty} \frac{\Gamma(k\sigma^2 + 1)}{k!} (-1)^{k+1} \left(\frac{1}{z}\right)^{\sigma^2 k+1} \sin(\sigma^2 k \pi); \; z > 0, 0 < \sigma^2 < 1 \qquad (3.3)$$

$$L(s) = e^{-s^{\sigma^2}} \; ; \; z > 0, \; 0 < \sigma^2 < 1$$

The unconditional bivariate survival function for the jth individual at the time t1j > 0 and t2j > 0 by replacing the above mentioned Laplace transformation in equation (2.5) we have,

$$S(t_{1j}, t_{2j}) = e^{-[\eta_j((H_{01}(t_{1j}) + H_{02}(t_{2j}))]^{\sigma^2}} \qquad (3.4)$$

where $H_{01}(t_{1j})$ and $H_{02}(t_{2j})$ are the cumulative baseline hazard functions of lifetime random variables T_{1j} and T_{2j} respectively.

To every bivariate survival function $S(t_1, t_2)$ with absolute marginal survival functions $S(t_1)$ and $S(t_2)$, corresponds a unique function $C:[0,1]\times[0,1] \to [0,1]$, is called a copula such that

$$S(t_1, t_2) = C(S(t_1), S(t_2)) \quad for \quad (t_1, t_2) \in (0, \infty) \times (0, \infty) \qquad (3.5)$$

Here $S(t_1, t_2)$, $S(t_1)$, and $S(t_2)$ are the joint and marginal survival functions respectively. Conversely, using copula, it is possible to construct a bivariate survival function having the desired marginal survivals and a chosen dependence structure (see Nelsen (2006) for details).
The joint bivariate survival function (3.4) can be expressed in terms of survival copula as

$$C(u, v) = e^{[-((-\ln(u))^{1/\sigma^2} + (-\ln(v))^{1/\sigma^2})^{\sigma^2}]} \qquad (3.6)$$

where $u = S(t_1)$ and $v = S(t_2)$.
The above copula is known as Gumbel-Hougaard copula.
Imposing different assumptions on the baseline distributions we get different shared positive stable frailty models. In this paper, we consider the two baseline distributions namely Pareto and linear failure rate distribution which yield two positive stable frailty models. The positive stable model has the advantage that it fits proportional hazards which means that if the conditional model has proportional hazards, so does the marginal distribution. This is an advantage, when considering the model as a random effect model.
When, $\sigma^2 = 1$, the frailty distribution is degenerate at $Z = 1$. The bivariate survival function in the case when the frailty variable is degenerate is given by

$$S(t_{1j}, t_{2j}) = e^{-[\eta_j((H_{01}(t_{1j})+H_{02}(t_{2j}))]} \qquad (3.7)$$

Baseline Distributions:
The Pareto distribution is a skewed, heavy-tailed distribution that is sometimes used to model the distribution of incomes. This distribution is not limited to describing wealth or income, but to many situations in which an equilibrium is found in the distribution of the "small" to the "large".
The first baseline distribution is the Pareto distribution (Deshpande and Purohit, 2005). A continuous random variable T is said to follow the Pareto distribution with the scale parameter λ and the shape parameter α if its survival function is,

$$S_0(t) = \begin{cases} (\lambda t + 1)^{-\alpha} & ; \ t > 0, \lambda > 0, \alpha > 0 \\ 0 & ; \quad otherwise \end{cases} \qquad (4.1)$$

and the hazard function and the cumulative hazard function as

$$h_0(t) = \begin{cases} \frac{\alpha \lambda}{(1+\lambda t)} & ; \quad t > 0, \lambda > 0, \alpha > \\ 0 & ; \quad otherwise \end{cases} \qquad (4.2)$$

$$H(t) = \begin{cases} \alpha \log(\lambda t + 1) & ; \quad t > 0, \lambda > 0, \alpha > 0 \\ 0 & ; \quad otherwise \end{cases} \qquad (4.3)$$

Observe that $h_0(t)$ decreases with t ; $\lambda > 0, \alpha > 0$. Hence this distribution belongs to the Decreasing Failure Rate class.
The exponential and Rayleigh are the two most commonly used distributions for analyzing lifetime data. These distributions have several desirable properties and nice physical interpretations. Unfortunately the exponential distribution only has constant failure rate and the Rayleigh distribution has increasing failure rate. The linear failure rate distribution generalizes both these distributions. We consider this is the second baseline distribution.
The linear failure rate distribution of a continuous random variable T with the parameters $\alpha > 0$ and $\lambda > 0$, will be denoted by LFRD (α, λ) has the following survival function

$$S_0(t) = \begin{cases} \exp(-\alpha t - \frac{\lambda}{2} t^2) & ; \ t > 0, \alpha > 0, \lambda > 0 \\ 0 & ; \quad otherwise \end{cases} \qquad (4.4)$$

It is easily observed that the exponential distribution (ED (α)) and the Rayleigh distribution (RD (λ)) can be obtained from LFRD (a, b) by putting $\lambda = 0$ and $\alpha = 0$ respectively. Moreover, the probability density function (PDF) of the LFRD (α,λ) can be decreasing or unimodal but the failure rate function is either constant or increasing only. See for example Bain (1974), Sen and Bhattacharya (1995), Lin et al. (2006), Ghitany and Kotz (2007) and the references cited their in this connection.
The hazard function and the cumulative hazard function of linear failure rate distribution are respectively as stated below:

$$h_0(t) = \begin{cases} \alpha + \lambda t ; & t > 0, \alpha > 0, \lambda > 0 \\ 0 ; & otherwise \end{cases} \quad (4.5)$$

$$H_0(t) = \begin{cases} \alpha t + \lambda \frac{t^2}{2} ; & t > 0, \alpha > 0, \lambda > 0 \\ 0 ; & otherwise \end{cases} \quad (4.6)$$

Proposed Models:

Here we present the two positive stable frailty models say Model I and Model II by putting respectively the cumulative hazard function of the baseline distributions namely Pareto and linear failure rate distribution in the unconditional survival function of bivariate random variable (T_{1j}, T_{2j}) given in equations (4.3) and (4.6).

$$S(t_{1j}, t_{2j}) = exp\left[-\eta_j^{\sigma^2} (\alpha_1 log(\lambda_1 t_{1j} + 1) + \alpha_2 log(\lambda_2 t_{2j} + 1))\right] \quad ; t_{1j} > 0, t_{2j} > 0 \quad (5.1)$$

$$S(t_{1j}^2, t_{2j}^2) = exp\left[-\eta_j^{\sigma^2} ((\alpha_1 t_{1j} + \frac{\lambda_1 t_{1j}^2}{2}) + (\alpha_2 t_{2j} + \frac{\lambda_2 t_{2j}^2}{2}))\right] \quad ; t_{1j} > 0, t_{2j} > 0$$

(5.2)

Here onwards, equations (5.1) and (5.2) as Model I and Model II which correspond to positive frailty models with baseline Pareto and linear failure rate distributions respectively. Similarly, by putting the same cumulative hazard functions given in equations (4.3) and (4.6) of the above baselines in equation (3.5) we get, the corresponding bivariate survival functions for the degenerate frailty (without frailty) at time $t_{1j} > 0, t_{2j} > 0$ as given below. We refer these models as Model III and Model IV respectively.

$$S(t_{1j}, t_{2j}) = exp\left[-\eta_j (\alpha_1 log(\lambda_1 t_{1j} + 1) + \alpha_2 log(\lambda_2 t_{2j} + 1))\right] \quad ; t_{1j} > 0, t_{2j} > 0 \quad (5.3)$$

$$S(t_{1j}^2, t_{2j}^2) = exp\left[-\eta_j^2 ((\alpha_1 t_{1j} + \frac{\lambda_1 t_{1j}^2}{t_{1j}}) + (\alpha_2 t_{2j} + \frac{\lambda_2 t_{2j}^2}{t_{2j}}))\right] \quad ; t_{1j} > 0, t_{2j} > 0$$

(5.4)

Estimation of Parameters by Bayesian Approach:

Suppose that there are n independent pairs of components under study and the jth pair of the component have lifetimes (t_{1j}, t_{2j}) i.e. there are n individuals with a pair of components having lifetimes (t_{1j}, t_{2j}) for j=1, 2,… , n. The life time times associated with jth individual is given by,

$$(T_{1j}, T_{2j}) = \begin{cases} (t_{1j}, t_{2j}) ; & t_{1j} < c_{1j}, t_{2j} < c_{2j} \\ (t_{1j}, c_{2j}) ; & t_{1j} < c_{1j}, t_{2j} > c_{2j} \\ (c_{1j}, t_{2j}) ; & t_{1j} > c_{1j}, t_{2j} < c_{2j} \\ (c_{1j}, c_{2j}) ; & t_{1j} > c_{1j}, t_{2j} > c_{2j} \end{cases}$$

where c_{1j} and c_{2j} be the observed censoring times for jth individual (j = 1, 2, …, n) with a pair of components respectively. We assume that the lifetimes and censoring times are independently distributed.
Now the likelihood of the sample of size n is given by,

$$L(\underline{\theta}, \beta, \sigma^2) = \prod_{j=1}^{n_1} f_1(t_{1j}, t_{2j}) \prod_{j=1}^{n_2} f_2(t_{1j}, c_{2j}) \prod_{j=1}^{n_3} f_3(c_{1j}, t_{2j}) \prod_{j=1}^{n_4} f_4(c_{1j}, c_{2j})$$

(6.1)

where $\underline{\theta}, \beta, \sigma^2$ are the vector of parameters of the baseline distributions, the vector of regression coefficients and the frailty parameter respectively. Let the counts n_1, n_2, $n_3$and n_4 be the number of individuals for which the first and the second components failure times (t_{1j}, t_{2j}) lie in the ranges $t_{1j} < c_{1j}, t_{2j} < c_{2j}$; $t_{1j} < c_{1j}, t_{2j} > c_{2j}$; $t_{1j} > c_{1j}, t_{2j} < c_{2j}$ and $t_{1j} > c_{1j}$ $t_{2j} > c_{2j}$ respectively and

$$f_1(t_{1j}, t_{2j}) = \frac{\partial^2 S(t_{1j}, t_{2j})}{\partial t_{1j} \partial t_{2j}}$$

$$= \sigma^2 h_{01}(t_{1j}) h_{02}(t_{2j}) S(t_{1j}, t_{2j}) (1 + \sigma^2(-1 + \phi_2(t_{1j}, t_{2j}))) \frac{\phi_1(t_{1j}, t_{2j})}{\phi_2(t_{1j}, t_{2j})}$$

$$f_2(t_{1j}, c_{2j}) = -\frac{\partial S(t_{1j}, c_{2j})}{\partial t_{1j}} = \frac{\sigma^2 h_{01}(t_{1j}) S(t_{1j}, c_{2j}) \eta_j}{\phi_1(t_{1j}, c_{2j})}$$

$$f_3(c_{1j}, t_{2j}) = -\frac{\partial S(c_{1j}, t_{2j})}{\partial t_{2j}} = \frac{\sigma^2 h_{02}(t_{2j}) S(c_{1j}, t_{2j}) \eta_j}{\phi_1(c_{1j}, t_{2j})}$$

$$f_4(c_{1j}, c_{2j}) = S(c_{1j}, c_{2j})$$

(6.2)

where

$$\phi_1(t_{1j}, t_{2j}) = \eta_j (H_{01}(t_{1j}) + H_{02}(t_{2j}))^{\sigma^2} \text{and } \phi_2(t_{1j}, t_{2j}) = (H_{01}(t_{1j}) + H_{02}(t_{2j}))^2$$

Thus, we get the two likelihood functions for the two proposed positive stable frailty models namely Model I, Model II by substituting the corresponding hazard functions and cumulative hazard functions in the likelihood function given by equation (6.1) with
$S(t_{1j}, t_{2j})$ stated in equation (5.1) and (5.2).
The likelihood function for the degenerate frailty model (without frailty) is,

$$\underline{L}(\underline{\theta},\beta) = \prod_{j=1}^{} (\mathbf{G} f_1(t_{1j}, t_{2j})) \prod_{j=1}^{} (\mathbf{G} f_2(t_{1j}, c_{2j})) \prod_{j=1}^{} (\mathbf{G} f_3(c_{1j}, t_{2j})) \prod_{j=1}^{} (\mathbf{G} f_4(c_{1j}, c_{2j}))$$

(6.3)

where $\underline{\theta}$ and β are respectively the vector of baseline parameters and the vector of regression coefficients. The counts n_1, n_2, $n_3$and n_4 be the number of individuals for which the first and the second components failure times (t_{1j}, t_{2j}) lie in the ranges $t_{1j} < c_{1j}, t_{2j} < c_{2j}$; $t_{1j} < c_{1j}, t_{2j} > c_{2j}$; $t_{1j} > c_{1j}, t_{2j} < c_{2j}$ and $t_{1j} > c_{1j}$ $t_{2j} > c_{2j}$ respectively

$$f_1(t_{1j}, t_{2j}) = \frac{\partial^2 S(t_{1j}, t_{2j})}{\partial t_{1j}\, \partial t_{2j}} = \eta\, h_{01}(t_{1j})\, h_{02}(t_{2j})\, S(t_{1j}, t_{2j})$$

$$f_2(t_{1j}, c_{2j}) = -\frac{\partial S(t_{1j}, c_{2j})}{\partial t_{1j}} = \eta_j\, h_{01}(t_{1j})\, S(t_{1j}, c_{2j})$$

$$f_3(c_{1j}, t_{2j}) = -\frac{\partial S(c_{1j}, t_{2j})}{\partial t_{2j}} = \eta_j\, h_{02}(t_{2j})\, S(c_{1j}, t_{2j})$$

$$f_4(c_{1j}, c_{2j}) = S(c_{1j}, c_{2j})$$

(6.4)

Substituting the hazard functions $h_{01}(t_{1j})$, $h_{02}(t_{2j})$ and the survival function $S(t_{1j}, t_{2j})$for two baseline models (without frailty) we get the likelihood function given by equation (6.3).

The proposed models I and II consist of ten parameters computing the maximum likelihood estimators (MLEs) involves solving a ten dimensional optimization problem for these model whereas model III and IV consist of nine parameters computing the maximum likelihood estimators (MLEs) involves solving a nine dimensional optimization problem for these model. As the method of maximum likelihood fails to estimate the several parameters due to convergence problem in the iterative procedure, so we use the Bayesian approach. The traditional maximum likelihood approach to estimation is commonly used in survival analysis, but it can encounter difficulties with frailty models. Moreover, standard maximum likelihood based inference methods may not be suitable for small sample sizes or situations in which there is heavy censoring (see Kheiri et al. (2007)). Thus, in our problem a Bayesian approach, which does not suffer from these difficulties, is a natural one, even though it is relatively computationally intensive.

Several authors have discussed Bayesian approach for the estimation of parameters of the frailty models. Some of them are, Ibrahim et al.(2001) and references theirin, Santos and Achcar (2010). Santos and Achcar (2010) considered the parametric models with Weibull and the generalized gamma distribution as the baseline distributions and gamma and log- normal as frailty distributions. Ibrahim et al. (2001) and references therein considered Weibull model and the piecewise exponential model with gamma frailty. Therefore we proposed Bayesian inferential approach in this study to estimate the parameters of the model, which is a popularly used method, because computation of the Bayesian analysis becomes feasible due to advances in computing technology.

To apply Markov chain monte carlo (MCMC) methods, we assume that, conditional on observed covariates and on the entire set of parameters, observations are independent and prior distributions for all parameters are mutually independent. We used the Metropolis- Hastings algorithm within Gibbs sampler technique which is the most basic MCMC method used in Bayesian Inference. Convergence of Markov chain to a stationary distribution is observed by the trace plots, the coupling from the past plots, the Gelman-Rubin convergence

statistic, and the Geweke test. The trace plots are used to check the behavior of the chain and the coupling from the past plots can be used to decide the burn-in period. The Gelman-Rubin convergence statistic values are approximately equal to one then sample can be considered to be come from the stationary distribution. The Geweke test examines the convergence of a Markov chain based on the sub parts of a chain at the end and at the beginning of the convergence period. The large standardized difference between ergodic averages at the beginning and at the end of the convergence period indicates non convergence. The sample autocorrelation plots can be used to decide the autocorrelation lag.

In Bayesian paradigm the parameters of the model are viewed as random variables with some distribution known as prior distribution. It enables us to combine both the prior information and the data at hand to update the information of parameter. Thus, posterior density of a parameter is the distribution of a parameter updated by combining its prior distribution and the likelihood function. 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 while applying MCMC methods.

Let $L(\theta|y)$ be the likelihood function and $p(\theta)$ be the prior density of a parameter then posterior density function of a parameter $\pi(\theta|y)$ is given by,

$\pi(\theta|y) \propto L(\theta|y)p(\theta)$ (6.5)

In our case the joint posterior density function of a parameter for given failure times (t_{1j}, t_{2j}) is $\pi(\underline{\theta} = (\alpha_1, \lambda_1, \alpha_2, \lambda_2, \sigma^2, \beta \mid (\underline{t_{1j}}, \underline{t_{2j}}))$ as

$$\pi(\underline{\theta} = (\alpha_1, \lambda_1, \alpha_2, \lambda_2, \sigma^2, \beta \mid (\underline{t_{1j}}, \underline{t_{2j}}))$$
$$\propto L(\underline{\theta} = (\alpha_1, \lambda_1, \alpha_2, \lambda_2, \sigma^2, \beta \mid (\underline{t_{1j}}, \underline{t_{2j}}))\,(g_1(\alpha_1)\, g_2(\lambda_1)\, g_3(\alpha_2)\, g_4(\lambda_2)\, g_5(\sigma^2)) \prod_{i=1}^{k} p_i(\beta_i)$$

(6.6)

where $g_i(.)$ for $i = 1, 2, ..,5$ denotes the prior density function with known hyper parameters of corresponding argument for baseline parameters and frailty variance; $p_i(\beta_i)$ is the prior density function for regression coefficients β_i for $i = 1, 2, .., k$. And the likelihood function $L(.)$ is given by equation (6.1).

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. Since the full conditional distributions are not easy to integrate out therefore full conditional distributions are obtained by considering that they are proportional to the joint distribution of the parameters of the model. The samples are then obtained from these full conditional distributions.

Model Selection Criterion by Bayesian Approach:

Bayesian model comparison is commonly performed by computing posterior model probabilities. In order to compare proposed models we use the Akaike Information Criterion (AIC), the Bayesian Information Criterion (BIC), the Deviance Information Criterion (DIC) and the Bayes factor. These are the most common methods of Bayesian model assessment.

Akaike (1973) suggested that, given a class of competing models for a data set, one choose the model that minimizes

$AIC = D(\underline{\theta}) + 2p$ (7.1)

where p represents number of parameters of the model. $D(\underline{\theta})$ represents an estimate of the deviance evaluated at the posterior mean, $\underline{\theta} = E(\underline{\theta}\,|data)$. The deviance is given by,

$D(\underline{\theta}) = -2 \log L(\underline{\theta})$, where $\underline{\theta}$ is a vector of unknown parameters of the model and $L(\underline{\theta})$ is the likelihood function of the model.

Bayesian information criterion (BIC) was suggested by Schwarz (1978). Shibata (1976) and Katz (1981) have shown that the AIC tends to overestimate the number of parameters needed, even asymptotically. The Schwarz criterion indicates that the model with the highest posterior probability is the one that minimizes

$BIC = D(\underline{\theta}) + p \log(n)$ (7.2) where n is the number of observations, or equivalently, the sample size.

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

$DIC = D(\underline{\theta}) + 2p_D$ (7.3)

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, $p_D = \bar{D} - D(\underline{\theta})$ and $\bar{D} = E(D(\underline{\theta})|data)$.

Models with smaller values of the AIC, BIC and DIC are preferred.

Kadane and Lazar (2004) review model selection from Bayesian and frequent perspectives. The Bayes factor B_{jk} for a model M_j against M_k for given data $D = (t_{1j};\ t_{2j});\ j = 1,2,\dots,n$ is

$$B_{jk} = \frac{P(D|M_j)}{P(D|M_k)} \quad (7.4)$$

where $P(D|M_k) = \int_S P(D|M_k, \theta_k)\pi(\theta_k|M_k)\, d\theta_k$

θ_k is the parameter vector under model M_k and $\pi(\theta_k|M_k)$ is prior density and S is the support of the parameter θ_k. Raftery (1994), following Jeffreys (1961), propose 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,

$$I_k = \left\{\frac{\sum_{i=1}^{N} P(D|\theta^{(i)})^{-1}}{N}\right\}^{-1} \quad (7.5)$$

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

Simulation Study:

A simulation study is done to evaluate the performance of the Bayesian estimation procedure. For the simulation purpose we have considered only one covariate X1 and we assume that it follows normal distribution. The frailty variable Z is assumed to have positive

stable distribution with variance 3.6. Life times (T1j , T2j) for the jth individual are conditionally independent for the given frailty Zj = zj . We considered that Tij (i = 1, 2; j = 1, 2, … , n) follows one of the baseline distributions, namely, Pareto (Model-I) or linear failure rate (Model-II) distribution respectively. As the Bayesian methods are time consuming, we generate only twenty, fourty and sixty pairs of lifetimes using inverse transform technique. Here we have generated different random samples of size n = 20, 40 and 60 for lifetimes T1j and T2j. But here we are giving procedure for sample generation of only one sample size, say, n = 20. Samples are generated using the following procedure:

Generate a random sample of size 20 from the positive stable distribution with the same shape and the scale parameter =1/3.6 as shared frailties (zj) for jth (j = 1, 2, …, 20) individual.

Generate 20 covariate values for X1 from the normal distribution.

Compute $\eta = e^{-x_1\beta}$ with the regression coefficient $\beta = 0.5$.

Generate 20 pairs of lifetimes (t1j , t2j) for the given frailty (zj) using the following generators, for Model-I and Model-II respectively,

$$t_{1j} = \left(\frac{A_{1j}}{1}\right) - 1), \quad t_{2j} = \left(\frac{A_{2j}}{1}\right) - 1)$$

$$\frac{1j}{\lambda_1}\left(exp\left(\frac{-2j}{\alpha_1}\right)\right) \quad \frac{}{\lambda_2}\left(exp\left(\frac{}{\alpha_2}\right)\right) \tag{8.1}$$

$$t_{1j} = \frac{1}{\lambda_1}\left(-\alpha_1 - \alpha_1 - 4\sqrt{\left(\frac{\lambda}{2}\right)A}\right)\frac{1}{1j} \quad t_{2j} = \frac{1}{\lambda_1}\left(-\alpha_1 - \alpha\sqrt{-24\left(\frac{\lambda}{1\ 2}\right)\underline{A}}\right)_{1j} \tag{8.2}$$

where $A_{1j} = \frac{-\log r_1}{z_j\eta_j}$ and $A_{2j} = \frac{-\log r_2}{z_j\eta_j}$ and r_{ij} i=1,2 are random sample from U(0,1).

Generate the censoring times (c_{1j} and c_{2j}) from the exponential distribution.

Observe the ith survival time t_{ij} = min (t_{ij} ,c_{ij}) and the censoring indicator $\delta_{ij}$for the jth individual (i = 1; 2 and j = 1, 2, …., 20), where

$$\delta_{ij} = \begin{cases} 1 & ; \quad t_{ij} < c_{ij} \\ 0 & ; \quad otherwise \end{cases}$$

Thus we have data consists of 20 pairs of the survival times (t_{1j}; t_{2j}) and the censoring indicators δ_{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. In our study we use same non-informative prior for the frailty parameter $\sigma^2$and the regression coefficient β_1. Since we do not have any prior information about baseline parameters, $\alpha_1, \lambda_1, \alpha_2$ and λ_2 , prior distributions are assumed to be flat. A widely used prior for frailty parameter σ^2 is the gamma distribution

with mean one and large variance, G (ϕ, ϕ) say with a small choice of ϕ and for the regression coefficients β_i i=1, 2, .., k, we use the normal prior with mean zero and large variance ϵ^2.

For both the chains the results were somewhat similar so we present here the analysis for only one chain (i.e. chain 1) for the resulting model. Also due to lack of space we are not providing graphs. Simulated values of the parameters have the autocorrelation of lag k, so every kth iteration is selected as a sample from posterior distribution. The posterior mean and standard error with credible intervals for different sample sizes are reported in Table 1 and 2

for Model-I, Model-II respectively. From these Tables, it can be observed that the estimated values of the parameters are close to the true values of the parameters and the standard errors decrease as the sample size increases.

Analysis of Kidney Infection Data:

To study the Bayesian estimation procedure we use kidney infection data of McGilchrist and Aisbett (1991). The data is regarding recurrence times to infection at point of insertion of the catheter for 38 kidney patients using portable dialysis equipment. For each patient, first and second recurrence times (in days) of infection from the time of insertion of the catheter until it has to be removed owing to infection is recorded. The catheter may have to be removed for reasons other than kidney infection and this regard as censoring. So survival time for a patient 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 the common frailty component. The data consists of three risk variables age, sex and disease type GN, AN and PKD where GN, AN and PKD are short forms of Glomerulo Neptiritis, Acute Neptiritis and Polycyatic Kidney Disease. Let T1 and T2 be represents first and second recurrence time to infection. Five covariates age, sex and presence or absence of disease type GN, AN and PKD are denoted by $X_1, X_2, X_3, X_4,$ and X_5.

To analyze kidney data set, various models have been applied by different researchers. Some of them are, McGilchrist and Aisbett (1991), McGilchrist (1993), Sahu et al. (1997), Boneg (2001), Yu (2006) and Santos and Achcar (2010). McGilchrist and Aisbett (1991) considered semi-parametric Cox proportional hazards model with log-normal frailty distribution and applied Newton-Raphson iterative procedure to estimate the parameters of the model. McGilchrist (1993) and Yu (2006) both considered the same model as in McGilchrist and Aisbett(1991) but McGilchrist (1993) estimated the parameters of the model using BLUP, ML and REML methods and Yu (2006) proposed modified EM algorithm and penalized partial likelihood method. Sahu et al. (1997) considered four parametric models first two are piecewise exponential model with constant baseline hazard say λ_k within each

interval having gamma prior for λ_k for Model-I and for Model-II normal prior to $\log \lambda_k$. Both

the models have frailty distribution as gamma. Other two models are multiplicative gamma frailty as Model-III and additive frailty as Model-IV with Weibull as baseline distribution. Santos and Achcar (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 the goodness of fit of the data for positive stable frailty distributions with two baseline distributions and then apply Bayesian estimation procedure. To check the goodness of fit of kidney infection data set, we consider Kolmogrove-Smirnov (K-S) test for two frailty models. Table 3 gives K-S statistics and p-values of K-S statistics for Models I and II. Thus from the p-values of K-S test we can say that there is no substantial evidence to

reject the hypothesis that data are from Model I and Model II in the univariate case and we assume that they also fit for the bivariate case. Figure 1 shows the parametric plot versus non- parametric plot of the survival function and these two graphs are close to each other which is also the evidence that the proposed models fit very well.
We run two parallel chains for both models using two sets of prior distributions with the different starting points using Metropolis-Hastings algorithm and Gibbs sampler based on normal transition kernels. On the similar line of simulation, here also we assume same set of prior distributions. We iterate both the chains for 95000 times. We present here the analysis for only one chain with G (a1; a2) as prior for baseline parameters, for both the proposed models. Gelman-Rubin convergence 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 attain stationary distribution. Simulated values of parameters have autocorrelation of lag k, so every kth iteration is selected as sample.
The trace plots for all the parameters show zigzag pattern which indicates that the parameters move and mix more freely. Thus it seems that the Markov chain has reached the stationary state. Burn-in period is decided by using coupling from the past plot. However, a sequence of draws after burn-in period may have autocorrelation. Because of the 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. The autocorrelation of the parameters become almost negligible after certain lag. We present trace plots, the coupling from the past plots and autocorrelations plots for the parameters of Model II only due to lack of space in the Figures 2, 3 and 4 respectively.
The posterior mean and standard error with 95% credible intervals for baseline parameters, frailty parameter and regression coefficients for the Model I and II are presented in Table 4. Table 5 contain the posterior mean and standard error with 95% credible intervals for baseline parameters, and regression coefficients for the Model III and IV.
The AIC, BIC and DIC values for both the models are given in Table 6. The Bayes factors for the proposed models are also computed as given in Table 7.
From the Tables 4 and 5, we can observe that for all other models except Model II , the value zero is not a credible value for the credible interval of the regression coefficient X_2 and
X_5 so X_2 that is sex and disease type PKD variable seems significant. Negative value of β_2 indicates that the female patients have a slightly lower risk of infection and Negative value of β_5 indicates that the patients with the absence of the disease type PKD have lower risk of infection. Whereas, these two covariates X_2 and X_5 are observed insignificant for the Model
II. The remaining covariates X_1 , X_3 and X_4 are insignificant for all the four models.
The estimate of σ^2 from two models Model I and II shows that there is a strong evidence of high degree of heterogeneity in the population of patients. Some patients are expected to be very prone to infection compared to others with same covariate value. This is not surprising, as seen in the data set there is a male patient with infection time 8 and 16, and there is also male patient with infection time 152 and 562.
The comparison between two proposed models is made using AIC, BIC and DIC values given in Table 6. If we order AIC values from smallest to largest then Model II (the linear failure rate distribution) gets first rank then Model I (the Pareto distribution). Same order hold for BIC and DIC. The comparison among the four models is done using AIC, BIC and DIC

Bayes factor is calculated for all pairs of model and some of them given in Table 7. From Table 7, we can observe that, between Model II and Model I, there is strong positive evidence against Model I, so Model II is better model for modeling kidney infection data. The estimate of the frailty parameter σ^2 for Model I is 0.9905874 and for Model II is 0.09947308 which is not equal to zero. This shows that there is heterogeneity between the first and second recurrence times of infection for each patient. Similarly the Bayes factor for Model II against Model IV is 4.509948 which shows that Model I and Model II are better fit for kidney infection data set as compared to Model III and Model IV, respectively. This is also a Bayesian test for testing $\sigma^2 = 0$ against $\sigma^2 > 0$ and which supports the alternative hypothesis i.e., positive stable frailty models are better than positive stable models without frailty models. Also AIC, BIC and DIC values for two baseline models namely the Pareto (without frailty) and the linear failure rate (without frailty) are larger than that of positive stable frailty models for the two baseline distributions.

Now, we are in a position to say that, we have suggested a new shared positive stable frailty model with linear failure rate distribution as the baseline distribution is best among the four proposed models for modeling of kidney infection data. For simulation study and to analysis the kidney infection data we used R software.

Conclusions

The findings for the four suggested models of inverse Gaussian frailty—Pareto, linear failure rate distribution, and baseline distributions—are discussed in this article. The main aim of our study is to check which distribution (with the positive stable frailty or without frailty) fits better. We compare the four models and provide a mathematical setup. Models III and IV have no frailty, whereas Models I and II have positive frailty. We choose the Bayesian technique since the maximum likelihood method is unable to estimate the parameters and likelihood equations do not converge. We evaluate kidney infection data and conduct simulation studies using the Bayesian technique. The best model, Model-II, has an estimate of $\sigma 2$ of 0.09947308, indicating a high degree of heterogeneity in the patient group. With the exception of positive stable frailty with a baseline distribution that is the linear failure rate distribution, the covariate sex and the illness category PKD are significant for all models. Female patients have a somewhat decreased risk of infection when the regression coefficient ($\beta 2$) of covariate sex is negative. Negative value of the regression coefficient ($\beta 5$) of covariate, the disease type PKD indicates that the patient with the absence of this diseases have a slightly lower the risk of infection. On the basis of AIC, BIC, DIC and Bayes factor, the Model II, the shared positive stable frailty with linear failure rate distribution is best model for modeling kidney infection data among the proposed models. We can now declare that the models suggested for modeling kidney infection data are inferior than the new shared positive stable frailty model with the linear failure rate distribution as baseline distribution. The techniques discussed in this paper may be extended into other shared frailty models and correlated frailty models with different baseline distributions, using the Bayesian approach, provided the models fit to data. The suggested model may easily be expanded to include shared frailty models with many components.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. Second International Symposium on Information Theory, eds. B. N. Petrox and F. Caski, Budapest: Akademiai Kiado, 267-281.

Bain, L. J. (1974). Analysis for the linear failure-rate life-testing distribution. Technometrics, 16(4), 551 - 559.

Balakrishnan, N. and Peng, Y. (2006). Generalized gamma frailty model. Statistics in Medicine, 25, 2797-2816.

Boneg, Y. T. (2001). Weibull frailty for modeling heterogeneity. Unpublished Ph.D. thesis, University of Gueiph.

Clayton, D.G. (1978). A model for association in bivariate life tables and its applications to epidemiological studies of familial tendency in chronic disease incidence. Biometrika, 65, 141-51.

Deshpande, J. V. and Purohit, S. G. (2005). Life Time Data: Statistical Models and Methods. World Scientific, New Jersey.

Fan, J.J., Hsu, L., and Prentice, R.L. (2000). Dependence estimation over a finite bivariate failure time region.

Lifetime Data Anal., 6, 343-355.
Fine J.P., Glidden DV, Lee K.E. (2003). A simple estimator for a shared frailty regression model. J. R. Stat. Soc., B, 65, 317-329.

Ghitany, M. E. and Kotz, S. (2007). Reliability properties of extended linear failure-rate distributions, Probability in the Engineering and Information Sciences, 21, 441 - 450.

Glidden, D.V. (1999). Checking the adequacy of the gamma frailty model for multivariate failure times. Biometrika, 86, 381-394.
Hanagal, D. D. (2005). A positive stable frailty regression model in bivariate survival data. Journal of Indian Society for Probability and Statistics, 9, 35-44.
Hanagal, D. D. (2007). Positive stable frailty regression models in mixture distributions. Proceedings of 3rd International Conference on Reliability and Safety Engineering, December 17-19, 2007. Udaipur, India. Pp. 350-356.

Hanagal, D. D. and Bhambure, S. M. (2015). Comparison of shared frailty models using Bayesian approach. Model Assisted Statistics & Applications, 10, 25-41.

Hanagal, D. D. and Dabade, A. D. (2013). Bayesian estimation of parameters and comparison of shared frailty models. Communications in Statistics, Theory & Methods, 42(4), 910-931.

Hanagal, D. D. and Pandey, A. (2013). Gamma frailty models for bivariate survival data.
Journal of Statistical Computation and Simulation, 85(15), 3172-3189.

Hanagal, D. D. and Sharma, R. (2013). Modeling heterogeneity for bivariate survival data by shared gamma frailty regression model. Model Assisted Statistics & Application, 8, 85-102.

Hougaard, P. (2000). Analysis of Multivariate Survival Data. Springer-Verlag NewYork I Inc.
Ibrahim, J.G, Ming-Hui Chen and Sinha, D. (2001). Bayesian Survival Analysis.
Springer, Verlag.
Jeffreys, H. (1961). Theory of Probability. 3rd edn., Oxford: Oxford University Press. [20]Kadane, J. and Lazar, N. (2004). Methods and criteria for model selection. Journal of the
American Statistical Association, 99(465), 279-90.

Katz, R. W. (1981). On some criteria for estimating the order of a Markov chain.
Technometrics, 23, 243-249.
Kheiri, S., Kimber, A. and Meshkani, M. R. (2007). Bayesian analysis of an inverse Gaussian correlated frailty model. Computational Statistics and Data Analysis, 51, 5317-5326.

Lin, C. T. , Wu, S. J. S. and Balakrishnan, N. (2006). Monte carlo methods for Bayesian inference on the linear hazard rate distribution, Communications in Statistics - Theory and Methods, 35, 575 - 590.

McGilchrist, C. A. and Aisbett, C. W. (1991). Regression with frailty in survival analysis.
Biometrics, 47, 461-466.

McGilchrist, C. A. (1993). REML estimation for survival models with frailty. Biometrics, 49,221-225.
Nelsen, R. B. (2006). An introduction to Copulas, 2nd edition. Springer.

Oakes, D. (1989). Bivariate survival models induced by frailties. J. Amer. Stat. Assoc., 84, 487-493.
Raftery, A. E. (1994). Approximate Bayes factors and accounting for model uncertainty in generalized linear models. Biometrika, 83(2), 251-266.

Sahu, S. K., Dey, D. K. Aslanidou, H. and Sinha, D. (1997). A Weibull regression model with gamma frailties for multivariate survival data. Life Time Data Analysis, 3, 123- 137.

Santos, C. A. and Achcar, J. A. (2010). A Bayesian analysis for multivariate survival data in the presence of covariates. Journal of Statistical Theory and Applications, 9,233- 253.
Schwarz, G. E. (1978). Estimating the dimension of a model. Annals of Statistics, 6(2), 461-464.

Sen, A. and Bhattacharya, G. K. (1995). Inference procedure for the linear failure rate model, Journal of Statistical Planning and Inference, 46, 59 – 76.
Shibata, R. (1976). Selection of the order of an autoregressive model by Akaikes information criterion. Biometrika, 63, 117-126.
Shih, J.H. (1998). A goodness-of- fit test for association in a bivariate survival model.
Biometrika, 85, 189-200.
Spiegelhalter, D., Best, N., Carlin, B. and Linde, A. V. D. (2002). Bayesian measures of model complexity and fit. J. Roy. Statist. Soc., B, 64, 583-639.