Journal of Reliability and Statistical Studies; ISSN (Print): 0974-8024, (Online):2229-5666
Vol. 8, Issue 1 (2015): 159-180

BAYESIAN ESTIMATION IN SHARED COMPOUND POISSON FRAILTY MODELS

***David D. Hanagal and Asmita T. Kamble**
Department of Statistics, Savitribai Phule Pune University, Pune, India
Modern College of Arts, Science and Commerce, Pune, India
E Mail: *david@stats.unipune.ac.in; asmita.stats@gmail.com

Received March 16, 2015
Modified June 13, 2015
Accepted June 25, 2015

Abstract

In this paper, we study the compound Poisson distribution as the shared frailty distribution and two different baseline distributions namely Pareto and linear failure rate distributions for modeling survival data. We are using the Markov Chain Monte Carlo (MCMC) technique to estimate parameters of the proposed models by introducing the Bayesian estimation procedure. In the present study, a simulation is done to compare the true values of parameters with the estimated values. 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 frailty model out of two proposed frailty models.

Key Words: Bayesian Estimation, Compound Poisson Frailty, Markov Chain Monte Carlo, Shared Frailty.

1. Introduction

Frailty is a random component designed to account for heterogeneity caused by unobserved individual-level factors that are otherwise neglected by the other predictors in the model. Vaupel et al. (1979) suggested frailty models to account for the variations due to unobserved covariates. Several distributions like gamma, inverse Gaussian, positive stable distribution, power variance function, Weibull, compound Poisson are used as frailty models for heterogeneity in the populations.

A class of random effect models which proved useful in the survival analysis of related individuals is a class of frailty models which are based on the modifying the hazard function of individuals by introducing multiplicative effect on the baseline hazard function. Thus the frailty model is a random effect model for time to event data which is an extension of the Cox's proportional hazards model. Vaupel (1979), Keyfitz and Littman (1979) showed that ignoring individual heterogeneity lead to an incorrect conclusions.

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

$$h(t|z) = zh_0(t) \ ; \ t>0 \qquad (1.1)$$

where $h_0(t)$ is called the baseline hazard function. Then the conditional survival function is given by,

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

where $H(t) = \int_0^t h_0(u)du$ is the cumulative baseline hazard.

And the marginal survival function

$$S(t) = E[S(t|z)] = E\left[e^{-zH(t)}\right] = L_Z(H(t)) \quad ; t>0 \tag{1.3}$$

where $L_Z(.)$ is the Laplace transform of the frailty distribution.

In this paper, we consider the shared frailty model with the compound Poisson distribution as a frailty distribution and the Pareto and linear failure rate distributions as baseline distributions. The remainder of the paper is organized as follows. In Section 2, we give the properties of the general shared frailty models. In Section 3, we introduce the shared compound Poisson frailty model. The baseline distributions are given in Section 4. In Section 5, we propose two different compound Poisson frailty models. In Section 6, we discuss Bayesian method which is used to estimate the parameters of the proposed models. We discuss the different model selection criterion by Bayesian approach in Section 7. In Section 8, we present the simulation study. We present analysis of kidney infection data set and suggest a better model from these two proposed models in Section 9. Finally in Section 10, we discuss the conclusion of the study.

2. 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.

Bivariate survival data arise when each subject under study experiences two events. For e.g., failure times of paired organs like kidneys, eyes, ears or any other paired organs of an individual, recurrences of a given disease. Both monozygotic and dizygotic twins share date of birth and common pre-birth environment. 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 the illustrations of bivariate survival data with the shared frailty.

In order to build the shared frailty model for such kind of the data it is assumed that survival times are conditionally independent, for given shared frailty. i.e. there is an association between survival times only due to frailty.

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

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

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 $\underline{\beta}$ is the vector of regression coefficients with k components and $\underline{X_j}$ is the vector of observed covariates having k components.

Integrating the hazard function $h_i(t_{ij}|z_j, \underline{X_j})$ we get, the conditional cumulative hazard function for the j[th] individual at the i[th] component survival time $t_{ij} > 0$ for the given frailty $Z_j = z_j$ is,

$$H_i(t_{ij}|z_j, \underline{X_j}) = z_j H_{0i}(t_{ij})\eta_j \tag{2.2}$$

where $\eta_j = \exp(\underline{X_j}'\underline{\beta})$ and $H_{0i}(t_{ij})$ is the cumulative baseline hazard function at time $t_{ij} > 0$. The conditional survival function for the j[th] individual at the i[th] component survival time $t_{ij} > 0$ for the given frailty $Z_j = z_j$ is,

$$\begin{aligned} S(t_{ij}|z_j, \underline{X_j}) &= e^{-H_i(t_{ij}|z_j, \underline{X_j})} \\ &= e^{-z_j H_{0i}(t_{ij})\eta_j} \end{aligned} \tag{2.3}$$

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

$$S(t_{1j}, t_{2j}|z_j, \underline{X_j}) = e^{-z_j(H_{01}(t_{1j}) + H_{02}(t_{2j}))\eta_j} \tag{2.4}$$

where $H_{01}(t_{1j})$ and $H_{02}(t_{2j})$ be the cumulative baseline hazard functions of the first and the second components 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 the j[th] individual in order to obtain the unconditional bivariate survival function at time $t_{ij} > 0$.

$$\begin{aligned} S(t_{1j}, t_{2j}|\underline{X_j}) &= \int_{Z_j} e^{-z_j(H_{01}(t_{1j}) + H_{02}(t_{2j}))\eta_j} f(z_j)\, dz_j \\ &= L_{Z_j}\left(\left(H_{01}(t_{1j}) + H_{02}(t_{2j})\right)\eta_j\right) \end{aligned} \tag{2.5}$$

where $L_{Z_j}(.)$ is the Laplace transform of frailty the variable of Z_j for the j[th] individual.

3. Shared compound Poisson frailty model

Aalen (1988, 1992) introduced compound Poisson distribution as a mixing distribution in survival models. In many situations hazard rates or intensities are raising at the start, reaching a maximum value and then declining, that's why the intensity has a unimodal shape with finite mode. For e.g., death rates for cancer patients, divorce rates, etc. The reason to start decline in the population intensity is that the high risk individuals have already died or been divorced in case of above examples.

Also in is often seen that the total integral under the intensity or hazard rate is to be finite. It occurs due to the distribution is defective. It means that some individuals have zero susceptibility; they will survive forever. In the case of above examples some patients survive their cancer, some people never marry, some marriages are not prone to

be dissolved. In such kind of data compound Poisson distribution plays an important role of mixing distribution. The distribution arises as a sum of a random number of independent gamma variables, where the number of terms in the sum is Poisson distributed.

The Compound Poisson variable Z is defined as

$$Z = \begin{cases} X_1 + X_2 + \cdots + X_N & ; N > 0 \\ 0 & ; N = 0 \end{cases}$$

where N is Poisson distributed with mean ρ while $X_1, X_2, \ldots, X_N$ are independent gamma distributed with scale parameter ν and shape parameter ξ.

The distribution of Z can be partitioned into two parts, one a discrete part corresponding to the probability of zero susceptibility, and second is due to continuous part on the positive real line.
The discrete part is given by

$$P(Z = 0) = \exp(-\rho)$$

which decreases as ρ increases.

The distribution of the continuous part is given by

$$f(z) = \begin{cases} \exp[-(\rho + \nu z)] \frac{1}{z} \sum_{n=1}^{\infty} \frac{\rho^n (\nu z)^{n\xi}}{\Gamma(n\xi) n!} & ; z > 0, \rho > 0, \nu > 0, \xi > 0 \\ 0 & ; \quad otherwise. \end{cases} \tag{3.1}$$

The expectation and variance of Z are
$E(Z) = \rho\xi/\nu$ and $var(Z) = \sigma^2 = \rho\xi(\xi + 1)/\nu^2$.

The Laplace transformation of Z is given by,

$$\begin{aligned} L_Z(s) &= E(e^{-sZ}) = E\left(e^{-s(X_1 + X_2 + \cdots + X_N)}\right) \\ &= E(L_X(s)^N) \\ &= L_N\left(-ln\left(L_X(s)\right)\right) \\ &= exp\left\{-\rho + \rho\left[\frac{\nu}{\nu + s}\right]^{\xi}\right\} \end{aligned}$$

where $L_X(s) = \left[\frac{\nu}{\nu+s}\right]^{\xi}$.

…

In order to solve the non-identifiability problem, we take $E(z) = 1$ which leads to $\nu = \rho\xi$ and $var(Z) = \sigma^2 = (\xi + 1)/\rho\xi$. Thus, we have the following form of Laplace transformation.

$$\begin{aligned} L_Z(s) &= exp\left\{-\rho + \rho\left[\frac{\rho\xi}{\rho\xi + s}\right]^{\xi}\right\} \\ &= exp\left\{-\rho + \rho\left[\frac{\nu}{\nu + s}\right]^{\xi}\right\} \\ &= exp\left\{-\rho\left[1 - \left(1 + \frac{s}{\rho\xi}\right)^{-\xi}\right]\right\} \end{aligned} \tag{3.2}$$

The unconditional bivariate survival function for the jth individual at the time $t_{1j} > 0$ and $t_{2j} > 0$ using equations (2.5) and (3.2), we have,

$$S(t_{1j},t_{2j}) = exp\left\{-\rho\left[1-\left(1+\frac{\left(H_{01}(t_{1j})+H_{02}(t_{2j})\right)\eta_j}{\rho\xi}\right)^{-\xi}\right]\right\} \quad (3.3)$$

In this paper, we consider the two baseline distributions namely Pareto and linear failure rate distribution which yield two compound Poisson frailty models.

4. 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,

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

and the hazard function and the cumulative hazard function as

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

$$H_0(t) = \begin{cases}\alpha\log(\lambda\,t+1) & ; \quad t>0,\lambda>0,\alpha>0\\ 0 & ; \quad otherwise\end{cases} \quad (4.3)$$

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

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

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 therein in this connection.
The hazard function and the cumulative hazard function of linear failure rate distribution are respectively as stated below:

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

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

5. Proposed models

Here we present the two compound Poisson 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).

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

$$S(t_{1j,}t_{2j}) = exp\left\{-\rho\left[1-\left(1+\frac{\left(\left(\alpha_1 t_{1j}+\lambda_1\frac{t_{1j}^2}{2}\right)+\left(\alpha_2 t_{2j}+\lambda_2\frac{t_{2j}^2}{2}\right)\right)\eta_j}{\rho\xi}\right)^{-\xi}\right]\right\} \quad ;\ t_{1j}>0, t_{2j}>0 \tag{5.2}$$

Here onwards, equations (5.1) and (5.2) as Model I and Model II which correspond to compound Poisson frailty models with baseline Pareto and linear failure rate distributions respectively.

6. Estimation of parameters by Bayesian approach

For the bivariate life time distribution, we use the bivariate censoring scheme given by Hanagal and Dabade (2013).

Suppose that there are *n* independent pairs of components under study and the j[th] 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 j[th] individual is given by,

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

where c_{1j} and c_{2j} be the observed censoring times for j[th] 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,

$$L(\underline{\theta},\underline{\beta},\sigma^2) = \left(\prod_{j=1}^{n_1} f_1(t_{1j},t_{2j})\right)\left(\prod_{j=1}^{n_2} f_2(t_{1j},c_{2j})\right)\left(\prod_{j=1}^{n_3} f_3(c_{1j},t_{2j})\right)\left(\prod_{j=1}^{n_4} f_4(c_{1j},c_{2j})\right) \tag{6.1}$$

where $\underline{\theta},\underline{\beta},\sigma^2$ 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

$$f_1(t_{1j},t_{2j}) = \frac{\partial^2 S(t_{1j},t_{2j})}{\partial t_{1j}\partial t_{2j}} = \frac{h_{01}(t_{1j})\,h_{02}(t_{2j})\,S(t_{1j},t_{2j})\,\phi_1(t_{1j},t_{2j})\,\eta_j^2}{\left(\phi(t_{1j},t_{2j})\right)^{2+\xi}}$$

$$f_2(t_{1j},c_{2j}) = -\frac{\partial S(t_{1j},c_{2j})}{\partial t_{1j}} = \frac{h_{01}(t_{1j})\,S(t_{1j},c_{2j})\,\eta_j}{\left[\phi(t_{1j},c_{2j})\right]^{1+\xi}}$$

$$f_3(c_{1j},t_{2j}) = -\frac{\partial S(c_{1j},t_{2j})}{\partial t_{2j}} = \frac{h_{02}(t_{2j})\,S(c_{1j},t_{2j})\,\eta_j}{\left[\phi(c_{1j},t_{2j})\right]^{1+\xi}}$$

$$f_4(c_{1j},c_{2j}) = S(c_{1j},c_{2j}) \tag{6.2}$$

where

$$\phi(a_i,b_j) = 1 + \frac{\left(H_{01}(t_{1j}) + H_{02}(t_{2j})\right)\eta_j}{\rho\xi}$$

and

$$\phi_1(a_i,b_j) = \frac{1}{\left(\phi(a_i,b_j)\right)^{\xi}} + \frac{1+\xi}{\xi\rho}$$

Thus, we get the two likelihood functions for the two proposed compound Poisson 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 is obtained for the Model I as:

$$f_1(t_{1j},t_{2j}) = \frac{\frac{\alpha_1\lambda_1}{(1+\lambda_1 t_{1j})}\frac{\alpha_2\lambda_2}{(1+\lambda_2 t_{2j})}\,S(t_{1j},t_{2j})\,\phi_1(t_{1j},t_{2j})\,\eta_j^2}{\left(\phi(t_{1j},t_{2j})\right)^{2+\xi}}$$

$$f_2(t_{1j}, c_{2j}) = \frac{\frac{\alpha_1\lambda_1}{(1+\lambda_1 t_{1j})}\, S(t_{1j}, c_{2j})\, \eta_j}{\left[\phi(t_{1j}, c_{2j})\right]^{1+\xi}}$$
$$f_3(c_{1j}, t_{2j}) = \frac{\frac{\alpha_2\lambda_2}{(1+\lambda_2 t_{2j})}\, S(c_{1j}, t_{2j})\, \eta_j}{\left[\phi(c_{1j}, t_{2j})\right]^{1+\xi}}$$
$$f_4(c_{1j}, c_{2j}) = S(c_{1j}, c_{2j}). \tag{6.3}$$

For Model II as:

$$f_1(t_{1j}, t_{2j}) = \frac{(\alpha_1+\lambda_1 t_{1j})(\alpha_2+\lambda_2 t_{2j})\, S(t_{1j}, t_{2j})\, \phi_1(t_{1j}, t_{2j})\, \eta_j^2}{\left(\phi(t_{1j}, t_{2j})\right)^{2+\xi}}$$
$$f_2(t_{1j}, c_{2j}) = \frac{(\alpha_1+\lambda_1 t_{1j})\, S(t_{1j}, c_{2j})\, \eta_j}{\left[\phi(t_{1j}, c_{2j})\right]^{1+\xi}}$$
$$f_3(c_{1j}, t_{2j}) = \frac{(\alpha_2+\lambda_2 t_{2j})\, S(c_{1j}, t_{2j})\, \eta_j}{\left[\phi(c_{1j}, t_{2j})\right]^{1+\xi}}$$
$$f_4(c_{1j}, c_{2j}) = S(c_{1j}, c_{2j}). \tag{6.4}$$

Each of the proposed models consists of eleven parameters, computing the maximum likelihood estimators (MLEs) involve solving a eleven dimensional optimization problem in these models. 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) 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.

$L(\theta|y)$ 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,

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

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

$$\pi\left(\underline{\theta} = (\alpha_1, \lambda_1, \alpha_2, \lambda_2, \sigma^2, \underline{\beta} \,\middle|\, (t_{1j},\ t_{2j})\right) \propto$$
$$L\left(\underline{\theta} = (\alpha_1, \lambda_1, \alpha_2, \lambda_2, \sigma^2, \underline{\beta} \,\middle|\, (t_{1j},\ t_{2j})\right) * \left(g_1(\alpha_1)\, g_2(\lambda_1)\, g_3(\alpha_2)\, g_4(\lambda_2)\, g_5(\sigma^2)\right)$$
$$* \prod_{i=1}^{k} p_i(\beta_i) \tag{6.6}$$

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 of 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.

7. 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

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

where p represents number of parameters of the model. $D(\hat{\underline{\theta}})$ represents an estimate of the deviance evaluated at the posterior mean, $\hat{\underline{\theta}} = E(\underline{\theta}\,|data)$. The deviance is given by, $(\underline{\theta}) = -2 \log L(\underline{\theta})$, 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

$$BIC = D(\hat{\underline{\theta}}) + p \log (n) \tag{7.2}$$

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;

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

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

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

Kadane and Lazar (2004) reviewed 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

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

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), 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,

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

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.

8. 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 compound Poisson distribution with variance 3.6. Life times (T1j , T2j) for the j[th] 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 T_{1j} and T_{2j}. But here we are giving procedure for sample generation of only one sample size, say, n = 20. Samples are generated using the following procedure:

1. Generate a random sample of size 20 from the compound Poisson distribution with shape ν and the scale parameter ξ as the shared frailties (z_j) for the j[th] (j = 1, 2, …, 20) individual.
To generate random observation from compound Poisson distribution, we firstly generate a random observation $N = n$ from Poisson distribution with mean ρ.
Then we consider the following two cases:
(i) If $n = 0$; take frailty $Z = 0$.
(ii) If $n > 0$; generate n gamma variates say $X_1, X_2, \ldots, X_N$ and then frailty is taken as $Z = X_1 + X_2 + \cdots + X_N$.
2. Generate 20 covariate values for X_1 from the normal distribution.
3. Compute $\eta = e^{-X_1\beta}$ with the regression coefficient $\beta = 0.5$.
4. Generate 20 pairs of the lifetimes (t_{1j} , t_{2j}) for the given frailty (z_j) using the following generators, for Model-I and Model-II respectively,

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

$$t_{1j} = \frac{1}{\lambda_1}\left(-\alpha_1 - \sqrt{\alpha_1^2 - 4\left(\frac{\lambda_1}{2}\right)A_{1j}}\right) \qquad t_{2j} = \frac{1}{\lambda_1}\left(-\alpha_1 - \sqrt{\alpha_1^2 - 4\left(\frac{\lambda_1}{2}\right)A_{1j}}\right) \tag{8.2}$$

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).
5. Generate the censoring times (c_{1j} and c_{2j}) from the exponential distribution.
6. Observe the i[th] survival time $t_{ij} = \min(t_{ij}, c_{ij})$ and the censoring indicator δ_{ij} for the j[th] 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 non-informative prior for the frailty parameter σ^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. Similar types of prior distributions are used in Ibrahim et al. (2001), Sahu et al. (1997) and Santos and Achcar (2010). We set hyper-parameters ϕ = 0.0001, ϵ^2=1000. We consider the non informative prior distribution for baseline parameters as the Gamma (1, 0.0001).

For both the chains the results were somewhat similar so we present here the analysis for only one chain (i.e. chain 1). Also due to lack of space we are not providing graphs. Simulated values of the parameters have the autocorrelation of lag k, so every k^{th} 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.

parameter	α_1	λ_1	α_2	λ_2	ρ	ξ	β_1
True values	3.2	1.1	3.2	1.1	4.5	1.5	0.5
sample size=20							
burn in period = 750					autocorrelation lag= 10		
estimates	3.2569	1.10123	3.2442	1.10395	4.5020	1.47995	0.49655
S. E.	0.2253	0.23108	0.2239	0.23030	0.2272	0.22740	0.16230
L.C.L.	2.8311	0.72174	2.8299	0.72154	4.1218	1.11922	0.22007
U.C.L.	3.5853	1.47830	3.5836	1.48025	4.8785	1.87523	0.78130
sample size=40							
burn in period=780					autocorrelation lag=7		
estimates	3.2338	1.09978	3.1595	1.10617	4.5100	1.47827	0.49872
S.E.	0.2249	0.22946	0.2226	0.22670	0.2258	0.22594	0.16153
L.C.L.	2.8287	0.72022	2.8179	0.72304	4.1244	1.11917	0.22028
U.C.L.	3.5816	1.48192	3.5697	1.47860	4.8798	1.87495	0.77905

sample size=60							
burn in period = 580					autocorrelation lag=7		
estimates	3.2108	1.09906	3.1859	1.09807	4.4957	1.47044	0.50077
S.E.	0.2212	0.22800	0.2218	0.22587	0.2250	0.22471	0.16334
L.C.L.	2.8261	0.72132	2.8221	0.72240	4.1239	1.11797	0.22018
L.C.L.	3.5787	1.47971	3.5720	1.47970	4.8761	1.87431	0.78089

Table 1: Posterior summary for simulation study of Model-I.

S.E. – Standard error; L.C.L. – Lower credible limit; U.C.L. – Upper credible limit

parameter	α_1	λ_1	α_2	λ_2	ρ	ξ	β_1
True values	2.5	2.1	2.1	2.5	2.1	5.5	0.5
sample size=20							
burn in period = 520					autocorrelation lag= 7		
estimates	2.4938	2.09359	2.52369	2.12208	2.11617	5.49133	0.49827
S.E.	0.2251	0.22494	0.22413	0.22165	0.21996	0.22365	0.16199
L.C.L.	2.1216	1.72143	2.12833	1.72966	1.72703	5.12215	0.22038
L.C.L.	2.8766	2.47516	2.87985	2.47994	2.47824	5.87457	0.78080
sample size=40							
burn in period=400					autocorrelation lag=9		
estimates	2.4749	2.07149	2.49050	2.08436	2.18290	5.49683	0.49741
S.E.	0.2228	0.22183	0.22266	0.21914	0.20789	0.21923	0.16074
L.C.L.	2.1194	1.71871	2.12103	1.72054	1.75204	5.12292	0.21994
L.C.L.	2.8704	2.47217	2.87352	2.47210	2.48550	5.87374	0.77857
sample size=60							
burn in period = 1800					autocorrelation lag=11		
estimates	2.4830	2.08799	2.46551	2.06507	2.20391	5.49060	0.49954
S.E.	0.2218	0.22291	0.22108	0.22028	0.20060	0.22032	0.16193
L.C.L.	2.1217	1.72137	2.12033	1.71908	1.76875	5.12248	0.21906
L.C.L.	2.8733	2.47332	2.87195	2.46895	2.48563	5.87301	0.77909

Table 2: Posterior summary for simulation study of Model-II.

S.E. – Standard error; L.C.L. – Lower credible limit; U.C.L. – Upper credible limit

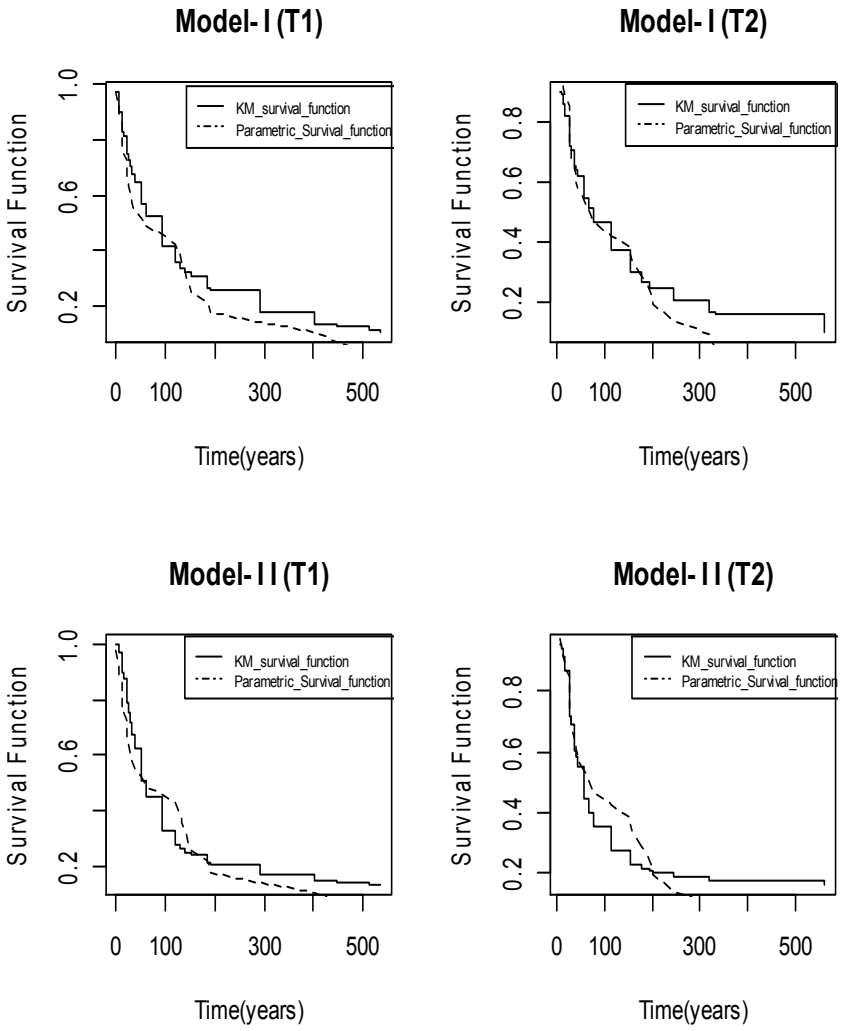
9. Analysis of kidney infection data

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, the first and the 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 the 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 X1, X2, X3, X4, and X5. First we check goodness of fit of the T1 and T2. If marginal distributions of T1 and T2 for two proposed distributions fit well then the bivariate distribution of T1 and T2 may be fit well for the same. We used Kolmogorov-Smirnov Goodness-of-Fit Test. Thus from p-values of K-S test we can say that there is no statistical evidence to reject the hypothesis that data are from proposed models in the univariate case and we assume that the models also fit for the bivariate case. The Table 3 gives the p-values of Kolmogorov-Smirnov test for the proposed models. Figure 1 shows the parametric versus non-parametric plots.

Model	K-S statistic		p-value	
	T1	T2	T1	T2
Model I	0.1085134	0.1144341	0.8482	0.80126
Model II	0.1445132	0.159587	0.53265	0.40831

Table 3: Goodness-of-Fit Test: p-values K-S statistic for kidney infection data.



F i g u r e 1 : Survival function plots for (K-M survival and parametric survival)

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. 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.

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. Due to lack of space we are presenting trace plots, coupling from the past plots and sample autocorrelation plots for the parameters of Model I only as shown in Figure 2-4. 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 k^{th} 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 4. The AIC, BIC and DIC values for both the models are given in Table 5. The Bayes factor for the proposed models is also computed.

…

From the Table 4, we can observe that for these two models, the value zero is not a credible value for the credible interval of the regression coefficient X_2, that is, sex variable is significant. Negative value of β_2 indicates that the female patients have a slightly lower risk of infection. The covariate X_1 , age is insignificant for both the models. It is also observed that X_5 i.e. disease type PKD is significant for Model I and insignificant in Model II. The covariate X_3 i.e. disease type GN and covariate X_4 i.e. disease type AN are insignificant for both the Models.

The estimate of σ^2 for Model I and Model II are 0.0067 and 0.8116 respectively. The estimate of σ^2 from Model 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 done using AIC, BIC and DIC values given in Table 5. It is observed that both Model I and Model II have AIC, BIC except DIC values are near about same. An alternative way to take the decision about the best model between the proposed Model I and Model II, we use the Bayes factor which is defined as,

$$M_{ij} = 2\, log\left(\frac{I_i}{I_j}\right) \tag{9.1}$$

Where I_i is as defined as in equation (7.5). Using Equation (9.1) we computed Bayes factor M_{12} which is 3.31550 for our proposed models.

We can observe that the Bayes factor of Model I vs Model II is 3.31550 implies that, there is positive evidence against Model-II, so Model-I is better than Model-II. Thus, Model I is best model of the proposed compound Poisson frailty models. Now, we are in a position to say that, we have suggested a new shared compound Poisson frailty model with Pareto distribution as the baseline distribution is best for modeling of kidney infection data. For simulation study and to analysis the kidney infection data we used R software.

parameter	estimates	Standard errors	Credible Intervals Lower	Upper
Model-I burn in period =4200 ; autocorrelation lag = 280				
α_1	22.94962	0.591966	22.05117	23.93646
λ_1	0.001675696	0.0004387207	0.001038241	0.00261299
α_2	18.43275	0.5489014	17.50093	19.35814
λ_2	0.001421859	0.0003851145	0.0007945032	0.002169576
ρ	509.9641	0.4618561	509.1203	510.8436
ξ	0.4148859	0.04469093	0.3359841	0.5011177
β_1	0.00307309	0.00471257	-0.005322972	0.01244158
β_2	-1.7413	0.2948506	-2.308226	-1.19979
β_3	-0.0586276	0.3759451	-0.7994152	0.618691
β_4	0.4188102	0.3505457	-0.2530626	1.101939
β_5	-1.426814	0.5169103	-2.467788	-0.5105207
Model-II burn in period = 4200 ; autocorrelation lag = 320				
α_1	0.02870925	0.008430826	0.01755664	0.05003709
λ_1	0.0001891125	5.170476e-05	0.0001079208	0.0002929388
α_2	0.006640035	0.001476604	0.00355642	0.008814052
λ_2	0.0002226568	4.880152e-05	0.0001180898	0.0002981948
ρ	24.05729	1.25025	21.71634	26.57135
ξ	0.05403221	0.0141685	0.02843128	0.08225272
β_1	0.008055036	0.006446512	-0.004089688	0.01867487
β_2	-2.090539	0.4304029	-2.881527	-1.237484
β_3	0.2705803	0.3413715	-0.377473	0.9121983
β_4	0.909617	0.5441327	-0.02651981	1.918116
β_5	-0.5736833	0.7265549	-2.109769	0.7390602

Table 4: Posterior summary for kidney infection data

Model	Distribution	AIC values	BIC values	DIC values
I	Pareto	683.0215	701.035	673.8789
II	linear failure rate	686.7364	704.7499	675.5379

Table 5: AIC, BIC and DIC values for kidney infection data.

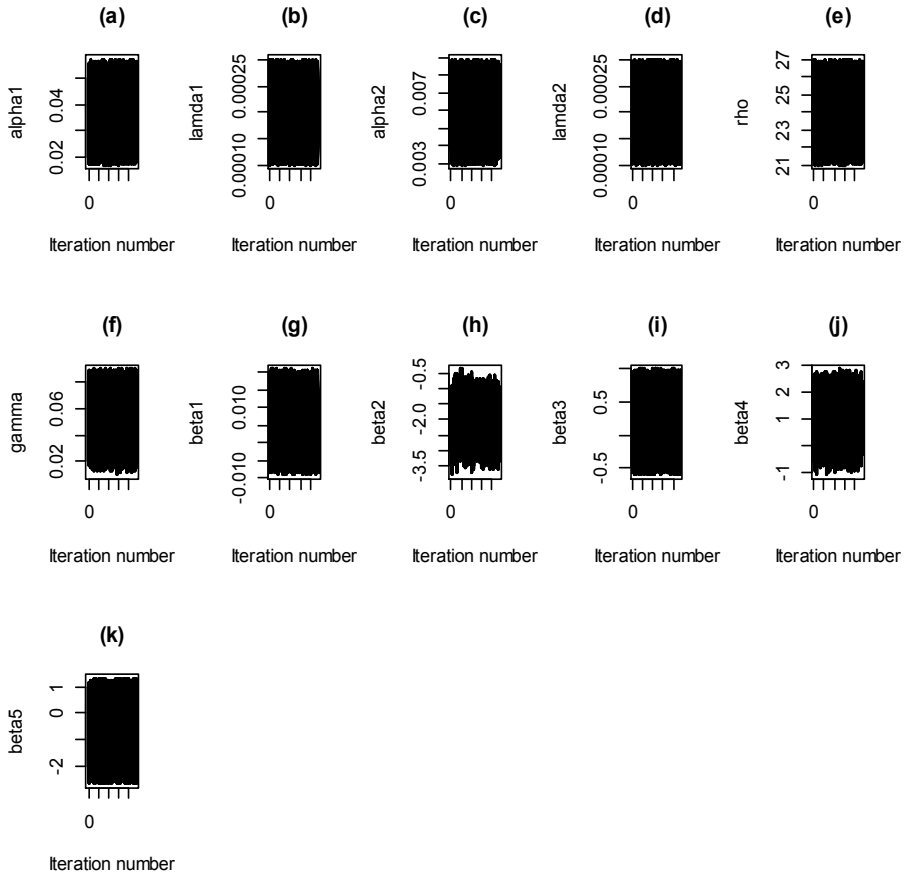


Figure2: Trace plot (Chain I) for Model I

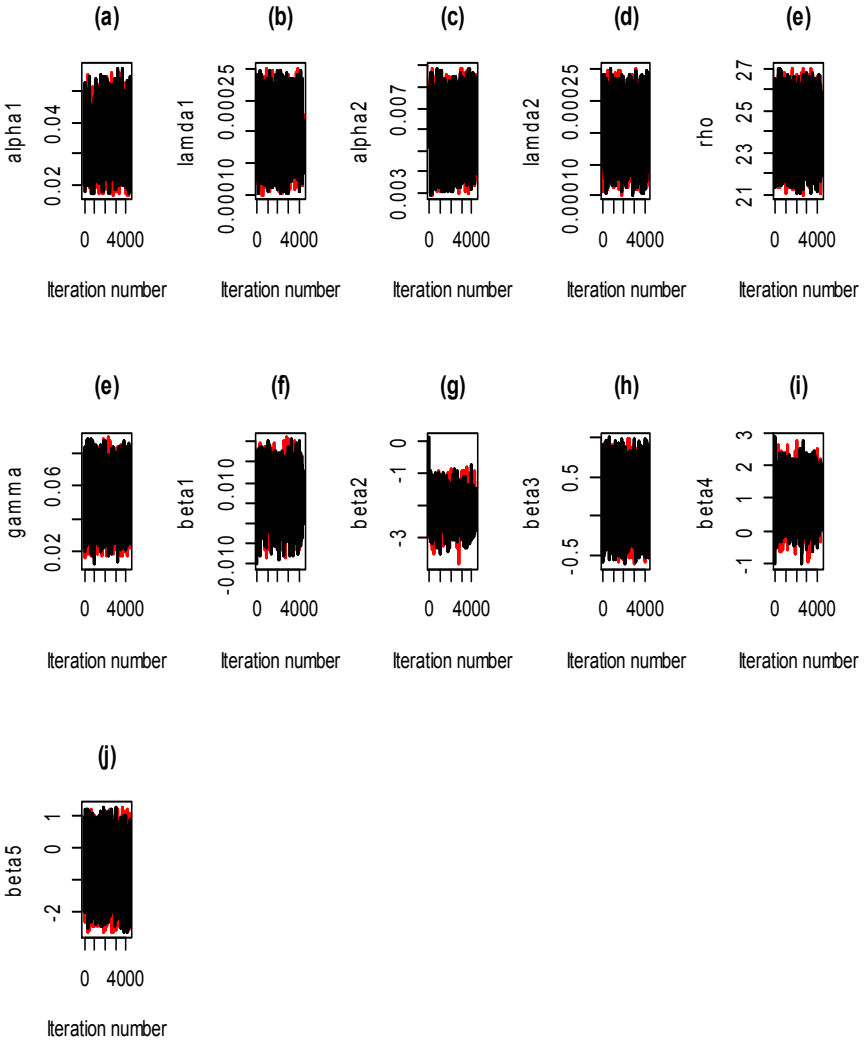


Figure3: Coupling from the past plots for Model I

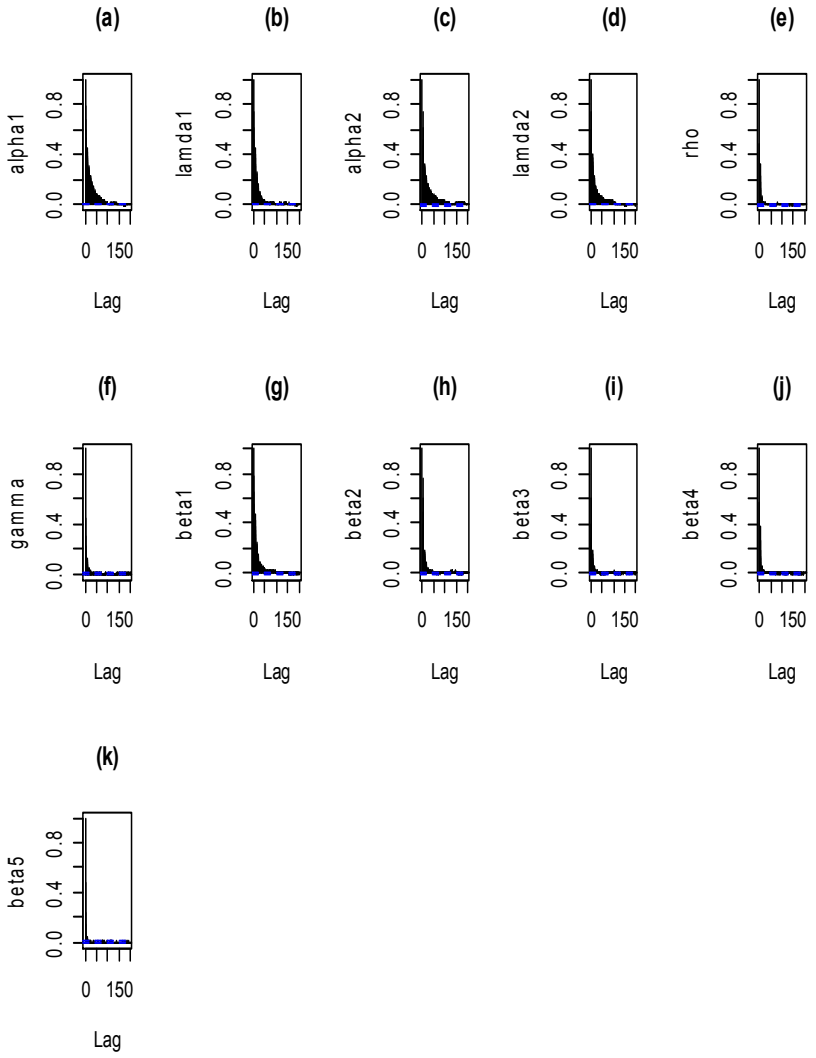


Figure4: Autocorrelation Graphs for Model I

10. Conclusions

In the present study we discuss results for the two proposed models of compound Poisson frailty namely Pareto, linear failure rate distribution as the baseline distributions. Our aim is to find the model which fit best between proposed models. For maximum likelihood estimate, likelihood equations do not converge and method of maximum likelihood fails to estimate the parameters so we use Bayesian approach. Using the Bayesian approach we perform simulation study and analyze kidney infection data. The estimate of frailty variance from Model II shows that there is a strong evidence of high degree of heterogeneity in the population of patients. The covariate

sex is significant for both models. Negative value of the regression coefficient (β_2) of covariate sex indicates that the female patients have a slightly lower the risk of infection. Negative value of the regression coefficient (β_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 in Model I only. On the basis of AIC, BIC, DIC and Bayes factor, the Model I, the shared compound Poisson frailty with Pareto distribution is best model for kidney infection.

References

1. Aalen, O. O. (1992). Modelling heterogeneity in survival analysis by the compound Poisson distribution, Annals of Applied Probability, 2, p.951-72.
2. Aalen, O.O. (1988). Heterogeneity in survival analysis, Statistics in Medicine, 7, p. 1121-37.
3. 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, p.267-281.
4. Bain, L. J. (1974). Analysis for the Linear Failure-Rate Life-Testing Distribution, Technometrics, 16(4), p. 551 - 559.
5. Barndorff-Nielsen, O. E. (1994). A note on electrical networks. Advances in Applied Probability, 26, p.63-67.
6. Boneg, Y. T. (2001). Weibull frailty for modelling heterogeneity, Unpublished Ph.D. thesis, University of Gueiph.
7. Chhikara, R.S., Folks, J.L. (1989). The Inverse Gaussian Distribution, Marcel Dekker: New York.
8. Deshpande, J. V. and Purohit, S. G. (2005). Life Time Data: Statistical Models and Methods, World Scientific, New Jersey.
9. Gacula, M. C. Jr. and Kubala, J. J. (1975). Statistical models for shelf life failures, Journal of Food Science, 40, p.404-409.
10. Ghitany, M. E. and Kotz, S. (2007). Reliability properties of extended linear failure-rate distributions, Probability in the Engineering and Information Sciences, 21, p. 441 - 450.
11. Hanagal, D. D. and Dabade (2013). Modeling inverse Gaussian frailty model for bivariate survival data, Communications in Statistic - Theory and Methods, 42(20), p. 3744-69.
12. Ibrahim, J.G, Ming-Hui Chen and Sinha, D. (2001): Bayesian Survival Analysis, Springer, Verlag.
13. Jeffreys, H. (1961). Theory of Probability, 3rd edn., Oxford: Oxford University Press.
14. Kadane, J. and Lazar, N. (2004). Methods and criteria for model selection, Journal of the American Statistical Association, 99(465), p. 279-90.
15. Kass, R. A. and Raftery, A. E. (1995). Bayes' factor, Journal of the American Statistical Association, 90, p. 773-795.
16. Katz, R. W. (1981). On some criteria for estimating the order of a Markov chain, Technometrics , 23, p.243-249.
17. Keyfitz, N. and Littman, G. (1979). Mortality in a heterogeneous population, Population Studies, 33, p. 333-342.

18. Kheiri, S., Kimber, A. and Meshkani, M. R. (2007). Bayesian analysis of an inverse Gaussian correlated frailty model, Computational Statistics and Data Analysis, 51, p. 5317-5326.
19. 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, p. 575 - 590.
20. McGilchrist, C. A. (1993). REML estimation for survival models with frailty, Biometrics, 49, p. 221-225.
21. McGilchrist, C. A. and Aisbett, C. W. (1991). Regression with frailty in survival analysis, Biometrics, 47, p. 461-466.
22. Raftery, A. E. (1994). Approximate Bayes factors and accounting for model uncertainty in generalized linear models, Biometrika, 83(2), p.251-266.
23. 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, p.123-137.
24. 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, p. 233-253.
25. Schwarz, G. E. (1978). Estimating the dimension of a model, Annals of Statistics, 6(2), p. 461-464.
26. Sen, A. and Bhattacharya, G. K. (1995). Inference procedure for the linear failure rate model, Journal of Statistical Planning and Inference, 46, p. 59 - 76.
27. Seshadri, V. (1999). The Inverse Gaussian Distribution: Statistical Theory and Applications, Springer: New York.
28. Shibata, R. (1976). Selection of the order of an autoregressive model by Akaikes information criterion, Biometrika, 63, p. 117-126.
29. Spiegelhalter, D., Best, N., Carlin, B. and Linde, A. V. D. (2002). Bayesian measures of model complexity and fit, J. Roy. Statist. Soc. Ser. B, 64, p. 583-639.
30. Vaupel, J. W., Manton, K. G. and Stallaed, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality, Demography, 16, p. 439-454.
31. Vilmann, H., Kirkeby, S., and Kronberg, D. (1990). Histomorphometrical analysis of the inuence of soft diet on masticatory muscle development in the muscular dystrophic mouse, Archives of Oral Biology, 35(1), p. 37-42.
32. Yu, B. (2006). Estimation of shared gamma frailty models by a modified EM algorithm, Computational Statistics and Data Analysis, 50, p. 463-474.