

A Comparative Study of Shared Frailty Models for Kidney Infection Data with Generalized Exponential Baseline Distribution

David D. Hanagal* and Alok D. Dabade
University of Pune

Abstract: Shared frailty models are often used to model heterogeneity in survival analysis. The most common shared frailty model is a model in which hazard function is a product of random factor (frailty) and baseline hazard function which is common to all individuals. There are certain assumptions about the baseline distribution and distribution of frailty. Mostly assumption of gamma distribution is considered for frailty distribution. To compare the results with gamma frailty model, we introduce three shared frailty models with generalized exponential as baseline distribution. The other three shared frailty models are inverse Gaussian shared frailty model, compound Poisson shared frailty model and compound negative binomial shared frailty model. We fit these models to a real life bivariate survival data set of McGilchrist and Aisbett (1991) related to kidney infection using Markov Chain Monte Carlo (MCMC) technique. Model comparison is made using Bayesian model selection criteria and a better model is suggested for the data.

Key words: Bayesian model comparison, compound negative binomial distribution, compound Poisson distribution, gamma distribution, generalized exponential distribution, inverse Gaussian distribution, likelihood ratio test, MCMC, predictive density, shared frailty.

1. Introduction

Survival models have been extensively used in medical research during last several years. Especially the Cox proportional hazards model. In fact, Cox proportional hazards model become workhorse of regression analysis for censored data. In Cox proportional hazards model, we include explanatory variables or covariates to study the effect of covariates on distribution of survival times. Unfortunately, in many of the cases it is not possible to include all the relevant

*Corresponding author.

covariates related to disease. Sometimes because we do not know the values of the factor for each individual or sometimes we are not aware that there exists factors that we ought to include. For example, a genetic factor as we do not know all possible genes having influence on survival. This unknown or unobservable risk factor of the hazard function is often termed as the heterogeneity or frailty. To measure heterogeneity caused by unobserved covariates it is necessary to include random effect term or frailty into the model. Thus the frailty model is a random effect model for time to event data which is an extension of the Cox proportional hazards model.

Frailty models are substantially promoted by its applications to multivariate survival data in a seminal paper by Clayton (1978) without using the notion frailty. The term frailty was first coined by Vaupel *et al.* (1979).

The frailty model is usually modeled as an unobserved random variable acting multiplicatively on the baseline hazard function and it is assumed that frailty random variable follows one of the parametric distribution such as gamma, log-normal, positive stable, inverse Gaussian, power variance function etc. Let a continuous random variable T be life time of an individual and the random variable Z be frailty variable. The conditional hazard function for a given frailty variable $Z = z$ at time $t > 0$ is,

$$h(t | z) = zh_0(t)e^{\underline{X}\underline{\beta}}, \quad (1.1)$$

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

$$S(t | z) = e^{-\int_0^t h(x|z)dx} = e^{-zH_0(t)e^{\underline{X}\underline{\beta}}}, \quad (1.2)$$

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

$$S(t) = \int_0^\infty S(t | z)f(z)dz = L_Z(H_0(t)e^{\underline{X}\underline{\beta}}), \quad (1.3)$$

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

To model life times, Weibull distribution is mostly used in the literature. For example, Ibrahim *et al.* (2001) and references their in, Sahu *et al.* (1997), Boneg (2001), Yu (2006) and Santos *et al.* (2010). All of the above references considered shared gamma frailty model with baseline as Weibull distribution.

Hazard rate for Weibull distribution is a monotone function, which increases with time to infinity when shape parameter α is greater than one and it decreases up to the the value zero for $\alpha < 1$. A near zero failure rate implies that almost no failure will occur which is hardly feasible in real life. Generalized exponential distribution, suggested by Gupta and Kundu (1999), can be used effectively in analyzing many life time data sets particularly in place of gamma and Weibull. For generalized exponential distribution, hazard rate increases from zero to a finite constant, when shape parameter α increases and hazard rate decreases from infinity to a finite number when α is less than one. A nearly constant rate after a certain time period implies that the occurrence of failure is purely random and is independent of past life; this is a property of the failure rate of an exponential distribution which has been extensively used in reliability studies. So we have considered generalized exponential as baseline distribution. The aim of our study is to compare different shared frailty models with baseline as generalized exponential distribution.

In this paper, we consider shared frailty model with baseline as generalized exponential distribution and four frailty distributions, gamma, inverse Gaussian, compound Poisson and compound negative binomial distribution. We are using Markov Chain Monte Carlo (MCMC) technique to fit these four models. We apply our models to bivariate survival data set of McGilchrist and Aisbett (1991) related to Kidney infection and compare these four models using Bayesian comparison techniques such as AIC, BIC, DIC, Bayes' factor, predicted model choice criteria etc.

The remainder of the paper is organized as follows. In Section 2, we give introduction of general shared frailty model. In Sections 3, 4, 5 and 6 we introduce four shared frailty models, gamma, inverse Gaussian, compound Poisson and compound negative binomial shared frailty models. In Section 7, we introduce shared frailty model with generalized exponential as baseline distribution. In the same Section we propose the four frailty models, gamma, inverse Gaussian, compound Poisson and compound negative binomial shared frailty models with generalized exponential baseline distribution respectively. Section 8 gives an outline of model fitting and model comparison using Bayesian approach. Section 9 is devoted to the analysis of kidney infection data. Finally in Section 10, the paper conclude with discussions.

2. General Shared Frailty Model

The shared frailty model is relevant to event time of related individuals, similar organs and repeated measurements for example, if the timing of failure of paired organs like kidneys, lungs, eyes, ears, dental implants etc. are considered. In this model individuals from a group share common covariates. For monozy-

gotic twins, examples are sex and any other genetically based covariates. Both monozygotic and dizygotic twins share date of birth and common pre birth environment. Also for human lifetime, natural disasters and accidents lead to the death of several persons at the same time or in the infectious diseases, two or more family members might visit an infected person and all of them become infected.

For the shared frailty model it is assumed that survival times are conditionally independent, for given shared frailty. That means dependence between survival times is only due to unobservable covariates or frailty. When there is no variability in the distribution of frailty variable Z that implies Z has a degenerate distribution and when the distribution of Z is not degenerate the dependence is positive.

Suppose n individuals are observed for the study and let a bivariate random variable (T_{1j}, T_{2j}) be represents first and second survival times of j^{th} individual ($j = 1, 2, 3, \dots, n$). Also suppose that there are k observed covariates collected in a vector $\underline{X}_j = (X_{1j}, \dots, X_{kj})$ for j^{th} individual, where X_{aj} ($a = 1, 2, 3, \dots, k$) represents the value of a^{th} observed covariate for j^{th} individual. Here we assume that the first and second survival times for each individual share the same value of the covariates. Let Z_j be represents shared frailty for j^{th} individual. Assuming that the frailties are acting multiplicatively on the baseline hazard function and both the survival times of individuals are conditionally independent for given frailty, the conditional survival function for j^{th} individual at i^{th} survival time $t_{ij} > 0$ for given frailty $Z_j = z_j$ from (1.2) is given by,

$$S(t_{ij} | z_j, \underline{X}_j) = e^{-z_j H_0(t_{ij}) \eta_j}, \quad (2.1)$$

where $\eta_j = e^{\underline{X}_j \underline{\beta}}$ and $H_0(t_{ij})$ is cumulative baseline hazard function at time $t_{ij} > 0$. Under the assumption of independence, bivariate conditional survival function for given frailty $Z_j = z_j$ at time $t_{1j} > 0$ and $t_{2j} > 0$ is,

$$\begin{aligned} S(t_{1j}, t_{2j} | z_j, \underline{X}_j) &= S(t_{1j} | z_j, \underline{X}_j) S(t_{2j} | z_j, \underline{X}_j) \\ &= e^{-z_j (H_{01}(t_{1j}) + H_{02}(t_{2j})) \eta_j}. \end{aligned} \quad (2.2)$$

Unconditional bivariate survival function at time $t_{1j} > 0$ and $t_{2j} > 0$ can be obtained by integrating over frailty variable Z_j having the probability function $f(z_j)$, for j^{th} individual

$$S(t_{1j}, t_{2j} | \underline{X}_j) = L_{Z_j}[(H_{01}(t_{1j}) + H_{02}(t_{2j})) \eta_j]. \quad (2.3)$$

Here onwards we represent $S(t_{1j}, t_{2j} | \underline{X}_j)$ as $S(t_{1j}, t_{2j})$.

Once we have unconditional survival function of bivariate random variable (T_{1j}, T_{2j}) we can obtain likelihood function and estimate the parameters of the model.

3. Gamma Frailty Model

We consider first frailty distribution as gamma distribution because the gamma distribution fits well to failure data from a computational and analytical point of view and it is easy to derive the closed form expression of survival and hazard function. Here the cross ratio function (Clayton, 1978) is constant and is independent of the lifetimes. Let a continuous random variable Z follows a gamma distribution with shape parameter λ and scale parameter α then density function of Z is,

$$f(z) = \begin{cases} \frac{\alpha^\lambda}{\Gamma(\lambda)} z^{\lambda-1} e^{-\alpha z}, & z > 0, \alpha > 0, \lambda > 0, \\ 0, & \text{otherwise,} \end{cases} \quad (3.1)$$

and Laplace transform is, $L_Z(s) = (1 + s/\alpha)^{-\lambda}$.

In order to solve the problem of non-identifiability, we assume Z has expected value equal to one which introduces restriction on the parameters $\alpha = \lambda$. Under this restriction, density function and Laplace transformation of a gamma distribution reduces to,

$$f(z) = \begin{cases} \frac{\frac{1}{\sigma^2}}{\Gamma(\frac{1}{\sigma^2})} z^{\frac{1}{\sigma^2}-1} e^{-\frac{z}{\sigma^2}}, & z > 0, \sigma^2 > 0, \\ 0, & \text{otherwise,} \end{cases} \quad (3.2)$$

and $L_Z(s) = (1 + \sigma^2 s)^{-1/\sigma^2}$ with variance of Z is $\sigma^2 = \lambda/\alpha^2 = 1/\alpha = 1/\lambda$.

Replacing Laplace transformation in (2.3), we get the unconditional bivariate survival function for j^{th} individual at time $t_{1j} > 0$ and $t_{2j} > 0$ as,

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

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

4. Inverse Gaussian Frailty Model

The gamma distribution is most commonly used frailty distribution because of its mathematical convenience. However, it has drawbacks (see Kheiri *et al.*, 2007) for example it may weaken the effect of covariates. Alternative to the gamma distribution Hougaard (1984) introduced inverse Gaussian as a frailty distribution. The inverse Gaussian distribution have many similarities to standard Gaussian distribution (see Chikkara and Folks, 1986). Furthermore, it provides much flexibility in modeling, when early occurrences of failures are dominant in a life time distribution and its failure rate is expected to be non-monotonic. In

such situations the inverse Gaussian distribution might provide a suitable choice for the lifetime model. Also inverse Gaussian is almost an increasing failure rate distribution when it is slightly skewed and hence is also applicable to describe lifetime distribution which is not dominated by early failures. Secondly, for the inverse Gaussian distribution the surviving population becomes more homogeneous with respect to time, where as for gamma distribution the relative heterogeneity is constant. The inverse Gaussian distribution has unimodal density and is the member of exponential family. While its shape resembles the other skewed density functions, such as lognormal and gamma. These properties of inverse Gaussian distribution motivates us to use inverse Gaussian as frailty distribution.

Let a continuous random variable Z follows inverse Gaussian distribution with parameters μ and α then density function of Z is,

$$f(z) = \begin{cases} \left[\frac{\alpha}{2\pi}\right]^{\frac{1}{2}} z^{-\frac{3}{2}} e^{-\frac{\alpha(z-\mu)^2}{2z\mu^2}}, & z > 0, \alpha > 0, \mu > 0, \\ 0, & \text{otherwise,} \end{cases}$$

and Laplace transform is,

$$L_Z(s) = \exp\left[\frac{\alpha}{\mu} - \left(\frac{\alpha^2}{\mu^2} + 2\alpha s\right)^{\frac{1}{2}}\right].$$

The mean and variance of frailty variable are $E(Z) = \mu$ and $V(Z) = \mu^3/\alpha$.

For identifiability, we assume Z has expected value equal to one i.e., $\mu = 1$. Under the restriction density function and Laplace transformation of inverse Gaussian distribution reduces to,

$$f(z) = \begin{cases} \left[\frac{1}{2\pi\sigma^2}\right]^{\frac{1}{2}} z^{-\frac{3}{2}} e^{-\frac{(z-1)^2}{2z\sigma^2}}, & z > 0, \sigma^2 > 0, \\ 0, & \text{otherwise,} \end{cases} \quad (4.1)$$

and

$$L_Z(s) = \exp\left[\frac{1 - (1 + 2\sigma^2 s)^{\frac{1}{2}}}{\sigma^2}\right],$$

with variance of Z is $\sigma^2 = 1/\alpha$. Note that there is heterogeneity if $\sigma^2 > 0$.

Replacing Laplace transformation in (2.3), we get the unconditional bivariate survival function for j^{th} individual at time $t_{1j} > 0$ and $t_{2j} > 0$ as,

$$S(t_{1j}, t_{2j}) = \exp\left\{\frac{1 - [1 + 2\sigma^2(H_{01}(t_{1j}) + H_{02}(t_{2j}))\eta_j]^{\frac{1}{2}}}{\sigma^2}\right\}, \quad (4.2)$$

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

5. Compound Poisson Frailty Model

In shared frailty models, we assume that, each individual from a pair shares common frailty but sometimes it may be possible that some individuals are immune to a particular event i.e., they are non-susceptible or they have zero susceptibility. For example, some cancer patients survive their cancer. In medicine, there are several examples of diseases primarily attacking people with particular susceptibility, for instance, a genetic kind, other people having virtually zero susceptibility of getting the disease. Another example is fertility, some couples are unable to conceive children so that the time to have first child birth for them have zero susceptibility. In case of marriages, some people never marry, some marriages are not prone to dissolve so that time to divorce for such couples have zero susceptibility. In such type of data, compound Poisson distribution is convenient since it has an explicit Laplace transform and it deals with the feature that some people may have zero susceptibility. This motivates us to use compound Poisson distribution as third frailty distribution. Aalen (1992), Aalen and Tretli (1999), Moger and Aalen (2005), Hanagal (2010a,b,c) have considered compound Poisson frailty models.

The compound Poisson distribution is defined as follows,

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

where N is Poisson distributed with mean ρ , while X_1, X_2, \cdots , are independent and gamma distributed with scale parameter ν and shape parameter γ .

The distribution of Z consists of two parts; a discrete part which corresponds to the probability of zero susceptibility, and a continuous part on the positive real line. The discrete part is, $P(Z = 0) = e^{-\rho}$ which decreases as ρ increases. The distribution of the continuous part can be found by conditioning N and using the fact that the X 's are gamma distributed. It can be written as

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

The parameter set for the compound Poisson distribution is $\rho > 0, \nu > 0, \gamma > 0$. The expectation and variance are given by

$$E(Z) = \frac{\rho\gamma}{\nu}, \quad \text{Var}(Z) = \frac{\rho\gamma(\gamma + 1)}{\nu^2}.$$

Laplace transform of compound Poisson distribution is given by,

$$L_Z(s) = \exp\left\{-\rho\left[1 - \left(\frac{\nu}{\nu + s}\right)^\gamma\right]\right\}. \quad (5.1)$$

For identifiability, we assume Z has expected value equal to one i.e., $\nu = \rho\gamma$. Under the restriction Laplace transformation of inverse Gaussian distribution reduces to,

$$L_Z(s) = \exp\{-\rho[1 - (1 + \frac{s}{\rho\gamma})^{-\gamma}]\},$$

and frailty variance is given by $\sigma^2 = (\gamma + 1)/\rho\gamma$.

Replacing Laplace transformation in (2.3), we get the unconditional bivariate survival function for j^{th} individual at time $t_{1j} > 0$ and $t_{2j} > 0$ as,

$$S(t_{1j}, t_{2j}) = \exp\{-\rho[1 - (1 + (H_{01}(t_{1j}) + H_{02}(t_{2j}))\eta_j\rho\gamma)^{-\gamma}]\}, \quad (5.2)$$

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

6. Compound Negative Binomial Frailty Model

In non-susceptible or zero susceptibility type of data a convenient frailty distribution is compound Poisson distribution. As defined in Section 5, the compound Poisson variate (Z) is defined as,

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

where N is a Poisson variate and X_i 's are gamma distributed random variables which represents length of i^{th} failure. If $N > 0$ then we can interpret Z as aggregate heterogeneity due to failures before we get first success or in general r^{th} success. For example, Aalen and Tretli (1999) analyzed testis cancer data and they represented X_i as size of the damage at i^{th} occasion and N be the number of damages occurred. A man receives damages during a critical period of their fetal development which may develop testis cancer after the hormonal process of puberty has started. The damage may be a result of the mother's exposure to environmental factors, for example an excessive estrogen burden, and may also interact with genetic factors. So Z is now cumulative effect of damages before testis cancer is observed with a man. Some other examples can be given as, in case of marriage data, Z may represents heterogeneity due to difficulties in finding a marriageable partner before individual meet first suitable partner. In case of fertility, Z may be heterogeneity due to miss-carriages observed or unable to conceive a child with a couple before they have their first child or second child. Some mothers would like to deliver babies until she delivers a baby boy or two baby boys. Politicians go on contesting elections until they win once or twice and so on. Individuals go on changing the jobs until he/she gets a suitable job. In such situations negative binomial distribution or geometric distribution is a suitable

choice of distribution for variate N . In risk model also, another parallel choice to compound Poisson model is compound negative binomial model or compound geometric model. In fact, there is a relation between compound negative binomial distribution and compound Poisson distribution. If Z_1 and Z_2 are respectively follows compound Poisson and compound negative binomial distribution with $Z_1 = X_1 + X_2 + \dots + X_{N_1}$ and $Z_2 = X_1 + X_2 + \dots + X_{N_2}$, where N_1 is Poisson random variate with intensity λ and N_2 is and negative binomial random variate with parameters, number of successes (r) and probability of success (p) then Z_1 and Z_2 are identically distributed if $\lambda = -r \log p$. When the number of successes is equal to one, the compound negative binomial distribution reduces to compound geometric distribution. This motivates us to use compound negative binomial distribution to model zero susceptibility.

The compound negative binomial distribution is defined as follows,

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

where N is negative binomial variate with parameters r and p ; r and p denotes respectively, the number of successes and the probability of success. The distribution of Z consists of two parts; a discrete part which corresponds to the probability of zero susceptibility, and a continuous part on the positive real line. The discrete part is, $P(Z = 0) = p^r$. The distribution of the continuous part can be found by conditioning N and using the fact that the X 's are gamma distributed. It can be written as

$$f(z) = \begin{cases} p^r \frac{1}{z} e^{-\nu z} \sum_{N=1}^{\infty} N^{r-1} C_N q^N \frac{(\nu z)^{N\gamma}}{\Gamma(N\gamma)}, & z > 0, \nu > 0, \gamma > 0, 0 < p < 1, \\ 0, & \text{otherwise.} \end{cases}$$

$q = 1 - p, r = 1, 2, \dots$

The parameter set for the compound negative binomial distribution is $0 < p < 1$, $\nu > 0$ and $\gamma > 0$. The Laplace transform, mean and variance of compound negative binomial distribution are given by,

$$L_Z(s) = \left\{ \frac{p}{1 - q[1 + \frac{s}{\nu}]^{-\gamma}} \right\}^r, \quad (6.1)$$

$$E(Z) = \mu = \frac{rq\gamma}{p\nu}, \quad (6.2)$$

$$Var(Z) = \sigma^2 = \frac{rq\gamma(p + \gamma)}{p^2\nu^2}. \quad (6.3)$$

For identifiability, we assume Z has expected value equal to one, therefore from (6.2) we have

$$\nu = \frac{rq\gamma}{p}. \quad (6.4)$$

Under the restriction, (6.4), variance is given by, $\sigma^2 = (p+\gamma)/rq\gamma$ and Laplace transformation of compound negative binomial distribution reduces to,

$$L_Z(s) = \left\{ \frac{p}{1 - q\left[1 + \frac{ps}{rq\gamma}\right]^{-\gamma}} \right\}^r.$$

Replacing Laplace transformation in (2.3), we get the unconditional bivariate survival function for j^{th} individual at time $t_{1j} > 0$ and $t_{2j} > 0$ as,

$$S(t_{1j}, t_{2j}) = \left\{ \frac{p}{1 - q\left[1 + \frac{p(H_{01}(t_{1j}) + H_{02}(t_{2j}))\eta_j}{rq\gamma}\right]^{-\gamma}} \right\}^r, \quad (6.5)$$

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

7. Generalized Exponential Baseline Model

A continuous random variable T is said to follow generalized exponential distribution if its survival function is,

$$S_0(t) = \begin{cases} 1 - (1 - e^{-\lambda t})^\alpha, & t > 0, \alpha > 0, \lambda > 0, \\ 1, & \text{otherwise,} \end{cases}$$

where λ and α are respectively scale and shape parameters of the distribution. The hazard function and cumulative hazard function are respectively,

$$h_0(t) = \begin{cases} \frac{\alpha\lambda e^{-\lambda t}(1 - e^{-\lambda t})^{\alpha-1}}{1 - (1 - e^{-\lambda t})^\alpha}, & t > 0, \alpha > 0, \lambda > 0, \\ 0, & \text{otherwise,} \end{cases}$$

$$H_0(t) = \begin{cases} -\ln[1 - (1 - e^{-\lambda t})^\alpha], & t > 0, \alpha > 0, \lambda > 0, \\ 0, & \text{otherwise.} \end{cases}$$

For $\alpha = 1$ distribution reduces to one parameter exponential distribution and have constant failure rate λ . When $\alpha > 1$ the hazard function is increasing function of time and for $\alpha < 1$ hazard function is decreasing function of time.

Substituting cumulative hazard function for Weibull distribution in (3.3), (4.2), (5.2) and (6.5) we get the unconditional bivariate survival functions as,

$$S(t_{1j}, t_{2j}) = [1 - \sigma^2(\ln[1 - G_1(t_{1j})] + \ln[1 - G_2(t_{2j})])\eta_j]^{-\frac{1}{\sigma^2}}, \quad (7.1)$$

$$S(t_{1j}, t_{2j}) = \exp\left\{\frac{1 - [1 - 2\sigma^2((\ln[1 - G_1(t_{1j})] + \ln[1 - G_2(t_{2j})])\eta_j)]^{\frac{1}{2}}}{\sigma^2}\right\}, \quad (7.2)$$

$$S(t_{1j}, t_{2j}) = \exp\left\{-\rho\left[1 - \left(1 - \frac{(\ln[1 - G_1(t_{1j})] + \ln[1 - G_2(t_{2j})])\eta_j}{\rho\gamma}\right)^{-\gamma}\right]\right\}, \quad (7.3)$$

$$S(t_{1j}, t_{2j}) = \left\{ \frac{p}{1 - q\left[1 - \frac{p(\ln[1 - G_1(t_{1j})] + \ln[1 - G_2(t_{2j})])\eta_j}{rq\gamma}\right]^{-\gamma}} \right\}^r, \quad (7.4)$$

where

$$G_1(t_{1j}) = (1 - e^{-\lambda_1 t_{1j}})^{\alpha_1} \quad \text{and} \quad G_2(t_{2j}) = (1 - e^{-\lambda_2 t_{2j}})^{\alpha_2}.$$

Here onwards we call (7.1), (7.2), (7.3) and (7.4) as model-I, model-II, model-III and model-IV respectively.

To fit the model, Bayesian approach is now popularly used, because computation of the Bayesian analysis become feasible due to advances in computing technology. Several authors have discussed Bayesian approach for the frailty models. Some of them are, Ibrahim *et al.* (2001) and references therein, Santos *et al.* (2010). Santos *et al.* (2010) considered parametric models with baseline distribution as Weibull and generalized gamma distribution and gamma, log-normal as frailty distributions. Ibrahim *et al.* (2001) and references therein considered Weibull model and piecewise exponential model with gamma frailty. They also considered positive stable frailty models. Kheiri *et al.* (2007) considered inverse Gaussian correlated frailty model.

8. Model Fitting and Comparison Using Bayesian Approach

8.1 Model Fitting

Suppose there are n individuals under study whose first and second observed failure times are represented by (t_{1j}, t_{2j}) . Let c_{1j} and c_{2j} be the observed censoring times for j^{th} individual ($j = 1, 2, 3, \dots, n$) for first and second recurrence times respectively. We assume independence between censoring scheme and life times of individuals. One of the following censoring situations can happen for each data point (T_{1j}, T_{2j}) .

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

The contribution of bivariate life time random variable of j^{th} individual in likelihood function is given by,

$$L_j(t_{1j}, t_{2j}) = \begin{cases} f_1(t_{1j}, t_{2j}), & t_{1j} < c_{1j}, t_{2j} < c_{2j}, \\ f_2(t_{1j}, c_{2j}), & t_{1j} < c_{1j}, t_{2j} > c_{2j}, \\ f_3(c_{1j}, t_{2j}), & t_{1j} > c_{1j}, t_{2j} < c_{2j}, \\ f_4(c_{1j}, c_{2j}), & t_{1j} > c_{1j}, t_{2j} > c_{2j}, \end{cases}$$

and likelihood function is,

$$L(\underline{\theta}, \underline{\beta}, \underline{\tau}) = \prod_{j=1}^{n_1} f_1(t_{1j}, t_{2j}) \prod_{j=1}^{n_2} f_2(t_{1j}, c_{2j}) \prod_{j=1}^{n_3} f_3(c_{1j}, t_{2j}) \prod_{j=1}^{n_4} f_4(c_{1j}, c_{2j}), \quad (8.1)$$

where $\underline{\theta}$, $\underline{\beta}$ and $\underline{\tau}$ are respectively vector of baseline parameters, vector of regression coefficients and vector of frailty parameters. $\underline{\theta} = (\alpha_1, \lambda_1, \alpha_2, \lambda_2)$ for all the four models and $\underline{\tau} = \sigma^2$ for model-I and model-II, for model-III, $\underline{\tau} = (\rho, \gamma)$ and for model-IV, $\underline{\tau} = (p, \gamma)$. Let n_1, n_2, n_3 and n_4 be the number of pairs for which first and second failure times (t_{1j}, t_{2j}) lies 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

$$\begin{aligned} f_1(t_{1j}, t_{2j}) &= \frac{\partial^2 S(t_{1j}, t_{2j})}{\partial t_{1j} \partial t_{2j}}, \\ f_2(t_{1j}, c_{2j}) &= -\frac{\partial S(t_{1j}, c_{2j})}{\partial t_{1j}}, \\ f_3(c_{1j}, t_{2j}) &= -\frac{\partial S(c_{1j}, t_{2j})}{\partial t_{2j}}, \\ f_4(c_{1j}, c_{2j}) &= S(c_{1j}, c_{2j}), \end{aligned}$$

differentiating survival function for each of the model given in (7.1) to (7.4) we get,

(for model-I)

$$\begin{aligned} f_1(t_{1j}, t_{2j}) &= (1 + \sigma^2) \frac{\alpha_1 \lambda_1 e^{-\lambda_1 t_{1j}} F_1(t_{1j})}{R_1(t_{1j})} \frac{\alpha_2 \lambda_2 e^{-\lambda_2 t_{2j}} F_2(t_{2j})}{R_2(t_{2j})} S(t_{1j}, t_{2j})^{1+2\sigma^2} \eta_j^2, \\ f_2(t_{1j}, c_{2j}) &= \frac{\alpha_1 \lambda_1 e^{-\lambda_1 t_{1j}} F_1(t_{1j})}{R_1(t_{1j})} S(t_{1j}, c_{2j})^{1+\sigma^2} \eta_j, \\ f_3(c_{1j}, t_{2j}) &= \frac{\alpha_2 \lambda_2 e^{-\lambda_2 t_{2j}} F_2(t_{2j})}{R_2(t_{2j})} S(c_{1j}, t_{2j})^{1+\sigma^2} \eta_j, \\ f_4(c_{1j}, c_{2j}) &= S(c_{1j}, c_{2j}), \end{aligned} \tag{8.2}$$

where $S(a_j, b_j)$ is given by (7.1);

(for model-II)

$$\begin{aligned} f_1(t_{1j}, t_{2j}) &= \frac{\alpha_1 \lambda_1 e^{-\lambda_1 t_{1j}} F_1(t_{1j})}{R_1(t_{1j})} \frac{\alpha_2 \lambda_2 e^{-\lambda_2 t_{2j}} F_2(t_{2j})}{R_2(t_{2j})} \frac{S(t_{1j}, t_{2j}) \phi_1(t_{1j}, t_{2j})}{[\phi_2(t_{1j}, t_{2j})]^{\frac{3}{2}}} \eta_j^2, \\ f_2(t_{1j}, c_{2j}) &= \frac{\alpha_1 \lambda_1 e^{-\lambda_1 t_{1j}} F_1(t_{1j})}{R_1(t_{1j})} \frac{S(t_{1j}, c_{2j})}{[\phi_2(t_{1j}, c_{2j})]^{\frac{1}{2}}} \eta_j, \\ f_3(c_{1j}, t_{2j}) &= \frac{\alpha_2 \lambda_2 e^{-\lambda_2 t_{2j}} F_2(t_{2j})}{R_2(t_{2j})} \frac{S(c_{1j}, t_{2j})}{[\phi_2(c_{1j}, t_{2j})]^{\frac{1}{2}}} \eta_j, \\ f_4(c_{1j}, c_{2j}) &= S(c_{1j}, c_{2j}), \end{aligned} \tag{8.3}$$

where

$$\begin{aligned}\phi_1(t_{1j}, t_{2j}) &= 1 + \sigma^2(1 - \ln S(t_{1j}, t_{2j})), \\ \phi_2(t_{1j}, t_{2j}) &= 1 - 2\sigma^2(\ln R_1(t_{1j}) + \ln R_2(t_{2j}))\eta_j,\end{aligned}$$

and $S(a_j, b_j)$ is given by (7.2);

(for model-III)

$$\begin{aligned}f_1(t_{1j}, t_{2j}) &= \frac{\alpha_1 \lambda_1 e^{-\lambda_1 t_{1j}} F_1(t_{1j})}{R_1(t_{1j})} \frac{\alpha_2 \lambda_2 e^{-\lambda_2 t_{2j}} F_2(t_{2j})}{R_2(t_{2j})} \frac{S(t_{1j}, t_{2j}) \phi_1(t_{1j}, t_{2j})}{[\phi(t_{1j}, t_{2j})]^{\gamma+2}} \eta_j^2, \\ f_2(t_{1j}, c_{2j}) &= \frac{\alpha_1 \lambda_1 e^{-\lambda_1 t_{1j}} F_1(t_{1j})}{R_1(t_{1j})} \frac{S(t_{1j}, c_{2j})}{[\phi(t_{1j}, c_{2j})]^{\gamma+1}} \eta_j, \\ f_3(c_{1j}, t_{2j}) &= \frac{\alpha_2 \lambda_2 e^{-\lambda_2 t_{2j}} F_2(t_{2j})}{R_2(t_{2j})} \frac{S(c_{1j}, t_{2j})}{[\phi(c_{1j}, t_{2j})]^{\gamma+1}} \eta_j, \\ f_4(c_{1j}, c_{2j}) &= S(c_{1j}, c_{2j}),\end{aligned}\tag{8.4}$$

where

$$\phi_1(a_j, b_j) = \frac{\gamma + 1}{\rho\gamma} + \phi(a_j, b_j)^{-\gamma}, \quad \phi(a_j, b_j) = 1 - \frac{[(\ln R_1(a_j) + \ln R_2(b_j))\eta_j]}{\rho\gamma},$$

and $S(a_j, b_j)$ is given by (7.3);

(for model-IV)

$$\begin{aligned}f_1(t_{1j}, t_{2j}) &= \frac{p^2 \eta^2}{r q \gamma} \frac{\Phi_1(t_{1j}, t_{2j}) S(t_{1j}, t_{2j})}{\Phi(t_{1j}, t_{2j})^{2(\gamma+1)} [1 - q\Phi(t_{1j}, t_{2j})^{-\gamma}]^2} \frac{\alpha_1 \lambda_1 e^{-\lambda_1 t_{1j}} F_1(t_{1j})}{R_1(t_{1j})} \\ &\quad \cdot \frac{\alpha_2 \lambda_2 e^{-\lambda_2 t_{2j}} F_2(t_{2j})}{R_2(t_{2j})}, \\ f_2(t_{1j}, c_{2j}) &= p \eta \frac{S(t_{1j}, c_{2j})}{\Phi(t_{1j}, c_{2j})^{\gamma+1} [1 - q\Phi(t_{1j}, c_{2j})^{-\gamma}]} \frac{\alpha_1 \lambda_1 e^{-\lambda_1 t_{1j}} F_1(t_{1j})}{R_1(t_{1j})}, \\ f_3(c_{1j}, t_{2j}) &= p \eta \frac{S(c_{1j}, t_{2j})}{\Phi(c_{1j}, t_{2j})^{\gamma+1} [1 - q\Phi(c_{1j}, t_{2j})^{-\gamma}]} \frac{\alpha_2 \lambda_2 e^{-\lambda_2 t_{2j}} F_2(t_{2j})}{R_2(t_{2j})}, \\ f_4(c_{1j}, c_{2j}) &= S(c_{1j}, c_{2j}),\end{aligned}\tag{8.5}$$

where

$$\begin{aligned}\Phi(t_{1j}, t_{2j}) &= [1 - \frac{p(\ln[1 - (1 - e^{-\lambda_1 t_{1j}})^{\alpha_1}] + \ln[1 - (1 - e^{-\lambda_2 t_{2j}})^{\alpha_2}])\eta_j}{r q \gamma}], \\ \Phi_1(t_{1j}, t_{2j}) &= q\gamma(r + 1) + (\gamma + 1)\Phi(t_{1j}, t_{2j})^\gamma [1 - q\Phi(t_{1j}, t_{2j})^{-\gamma}],\end{aligned}$$

and

$$\begin{aligned} F_1(t_{1j}) &= (1 - e^{-\lambda_1 t_{1j}})^{\alpha_1 - 1}, & F_2(t_{2j}) &= (1 - e^{-\lambda_2 t_{2j}})^{\alpha_2 - 1}, \\ R_1(t_{1j}) &= 1 - (1 - e^{-\lambda_1 t_{1j}})^{\alpha_1}, & R_2(t_{2j}) &= 1 - (1 - e^{-\lambda_2 t_{2j}})^{\alpha_2}, \end{aligned}$$

and $S(a_j, b_j)$ is given by (7.4).

The joint posterior density function of a parameters for given failure times is given by,

$$\begin{aligned} \pi(\alpha_1, \lambda_1, \alpha_2, \lambda_2, \underline{\tau}, \underline{\beta}) &\propto L(\alpha_1, \lambda_1, \alpha_2, \lambda_2, \underline{\tau}, \underline{\beta}) \cdot g_1(\alpha_1)g_2(\lambda_1)g_3(\alpha_2)g_4(\lambda_2) \\ &\cdot \prod_{i=1}^k \varepsilon_i(\tau_i) \prod_{i=1}^5 p_i(\beta_i), \end{aligned} \quad (8.6)$$

where $g_i(\cdot)$ ($i = 1, 2, \dots, 4$) indicates the prior density function with known hyper parameters of corresponding argument for baseline parameters. $\varepsilon_i(\cdot)$ and $p_i(\cdot)$ are prior density functions for frailty parameters τ_i , $i = 1, 2, \dots, l$ and regression coefficient β_i , $i = 1, 2, \dots, 5$. The likelihood function $L(\cdot)$ is given by (8.1). Here we assume that all the parameters are independently distributed.

A widely used prior for frailty parameter is the gamma distribution with mean one and large variance, $G(\phi, \phi)$, say with a small choice of ϕ and the prior for regression coefficient is the normal with mean zero and large variance say ϵ^2 . Similar types of prior distributions are used in Ibrahim *et al.* (2001), Sahu *et al.* (1997) and Santos *et al.* (2010). So in our study also we use same noninformative prior for frailty parameters and regression coefficients. 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. We consider two different noninformative prior distributions for baseline parameters, one is $G(a_1, a_2)$ and another is $U(b_1, b_2)$. All the hyper-parameters $\phi, \epsilon^2, a_1, a_2, b_1$ and b_2 are known. Here $G(a_1, a_2)$ is gamma distribution with shape parameter a_1 and scale parameter a_2 and $U(b_1, b_2)$ represents uniform distribution over the interval b_1 to b_2 . We set hyper-parameters $\phi = 0.0001$, $\epsilon^2 = 1000$, $a_1 = 1$, $a_2 = 0.0001$, $b_1 = 0$ and $b_2 = 100$.

The Bayesian model with the above prior density functions and likelihood function (8.1) can be fitted and the posterior distributions of the parameters of interest can be computed using the MCMC methods such as, Metropolis-Hastings algorithm. We generate two parallel Markov chains for all the models using two sets of prior distributions with different starting points using single component Metropolis-Hastings algorithm based on normal transition kernel. We iterate both the chains for N times. Convergence of chains to stationary distribution is monitored by Gelman-Rubin scale reduction factor and Geweke test. Burn in period is decided by using coupling from the past plot and sample autocorrelation function plots are used to decide autocorrelation lag of the chain. If burn in period

is B and autocorrelation lag is k then every k^{th} value of chain after burn in period is taken as sample value. Thus a sample of size n is formed and posterior summary is obtained.

8.2 Model Comparison

In order to compare proposed models we use several Bayesian model selection criteria, which we outline the concepts in this section.

Information criteria, Akaike Information Criteria (AIC), Bayesian Information Criteria (BIC) and Deviance Information Criteria (DIC) are respectively defined as,

$$\text{AIC} = D(\hat{\theta}) + 2p, \quad (8.7)$$

$$\text{BIC} = D(\hat{\theta}) + p \cdot \log(n), \quad (8.8)$$

$$\text{DIC} = D(\hat{\theta}) + 2 \cdot p_D, \quad (8.9)$$

where p represents number of parameters of the model and n represents number of data points. $D(\hat{\theta})$ represents an estimate of the deviance evaluated at the posterior mean $\hat{\theta} = E(\theta \mid \text{data})$. The deviance is defined by, $D(\theta) = -2 \cdot \log L(\theta)$, where θ is a vector of unknown parameters of the model and $L(\theta)$ is the likelihood function of the model. p_D is the difference between the posterior mean of the deviance and the deviance of the posterior mean of parameters of interest, that is, $p_D = \bar{D} - D(\hat{\theta})$, where $\bar{D} = E(D(\theta) \mid \text{data})$.

The Bayes factor B_{uv} for a model M_u against M_v for given data $D = (t_{1j}, t_{2j})$, $j = 1, 2, \dots, n$ is

$$B_{uv} = \frac{P(D \mid M_u)}{P(D \mid M_v)}, \quad (8.10)$$

where

$$P(D \mid M_k) = \int_S P(D \mid \theta_k, M_k) \cdot \pi(\theta_k \mid M_k) d\theta_k, \quad k = 1, 2, \dots, m.$$

θ_k is the vector of unknown parameters of model M_k , $\pi(\theta_k \mid M_k)$ is prior density and S is the support of the parameter θ_k . Here m represents number of models. The Kass and Raftery (1995) given the categories for the decision against the v^{th} model.

To compute Bayes factor we need to obtain $I_k = P(D \mid M_k)$ which is not easy to calculate analytically in our case. So to compute Bayes factor, we consider one of the approach given in Kass and Raftery (1995), a MCMC estimate of I_k given by,

$$\hat{I}_k = \left\{ \frac{\sum_{i=1}^N P(D \mid \theta^i)^{-1}}{N} \right\}^{-1},$$

which is harmonic mean of the likelihood values. Here N represents the posterior sample size.

Bayesian model examination for adequacy and model comparison can be proceeds by predictive distribution $\pi(y | y_{obs})$. By marginalizing $\pi(y | y_{obs})$, we obtain the posterior predictive density of a single observation y_r , $r = 1, 2, \dots, n$ as follows,

$$\pi(y_r | y_{obs}) = \int \pi(y_r | \theta) \pi(\theta | y_{obs}) d\theta. \quad (8.11)$$

A simple checking for assessment of model is predictive interval. Suppose we generate a sample $y_{r1}, y_{r2}, \dots, y_{rB}$ from the predictive density (8.11) for r^{th} observation and create the $100(1 - \alpha)\%$ equal tailed credible interval also known as predictive interval, then the model under consideration would be an adequate model for the data if $100(1 - \alpha)\%$ of the $y_{r,obs}$ to fall in their respective interval. To draw a random sample from predictive density (8.11), suppose we have θ_j^* ($j = 1, 2, \dots, B$), B samples from posterior density $\pi(\theta | Y)$ possibly using one of the MCMC methods. Then a random sample y_r^j drawn from $\pi(y_r | \theta_j^*)$ is a sample from the predictive density (8.11).

Another approach for model selection is based on cross-validation predictive density. The cross-validation predictive densities are the set $\{f(y_r | Y_{(r)}); r = 1, 2, \dots, n\}$ where $Y_{(r)}$ denotes all elements of data set Y except observation y_r and

$$f(y_r | Y_{(r)}) = \int f(y_r | \theta, Y_{(r)}) f(\theta | Y_{(r)}) d\theta, \quad (8.12)$$

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

Monte Carlo estimate of CPO, is given by,

$$\frac{B}{\sum_{j=1}^B \frac{1}{f(y_{r,obs} | \theta_j^*)}}, \quad (8.13)$$

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

Recently Gelfand and Ghosh (1998) have proposed a model choice criterion

$$D'_\omega = \sum_{r=1}^n \sigma_r^2 + \frac{\omega}{\omega + 1} \sum_{r=1}^n (\mu_r - y_{r,obs})^2, \quad (8.14)$$

where μ_r and σ_r^2 are Monte Carlo estimates of the posterior predictive mean and variance of y_r under the density (8.11) based on the sample y_r^j , $j = 1, 2, \dots, B$, and $\omega > 0$ is constant. The first term is a penalty term which penalizes both under-fitted and over-fitted models, since the predictive variances in such cases will tend to be larger. The second term without the factor involving ω is a goodness-of-fit measure. Model selection using D'_ω is usually not sensitive to ω . For censored data the criterion must be modified because $y_{r,obs}$ is not available for censored cases. The modified criterion is,

$$D_\omega = \sum_{r=1}^n \sigma_r^2 + \frac{\omega}{\omega + 1} \sum_{r=1}^n (\mu_r - v_r)^2, \quad (8.15)$$

where $v_r = y_{r,obs}$ if the r^{th} observation is a failure time and $v_r = \max(\mu_r, c_r)$ if the r^{th} observation is censored at c_r . A model with minimum value of D_ω is selected as the best model among all the models considered.

9. Analysis of Kidney Infection Data

The kidney infection data of McGilchrist and Aisbett (1991) is related to 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 (0 for censoring and 1 for occurrence of infection). 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 consist 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. The infection times from each patient share the same value of the covariates. Let T_1 and T_2 be represents first and second recurrence time to infection. Five covariates age, sex and presence or absence of disease type GN, AN and PKD are represented by $X_1, X_2, X_3, X_4,$ and X_5 . First we check goodness of fit of

the data for generalized exponential baseline distributions and then apply the Bayesian estimation procedure.

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

To check goodness of fit of kidney data set, firstly we consider graphical procedure to check appropriateness of the model. To assess model graphically we linearize survival function $\bar{F}(t)$. This means that some transform of $\bar{F}(t)$ (say $\Psi_1(\bar{F}(t))$) is a linear function of t or of some function of t (say $\Psi_2(t)$), that is, $\Psi_1(\bar{F}(t)) = \beta_0 + \beta_1 \Psi_2(t)$ for some functions $\Psi_1(\cdot)$ and $\Psi_2(\cdot)$. The idea is then to plot $\Psi_1(\bar{F}(t))$ versus $\Psi_2(t)$, if the resulted plot is roughly a straight line then we can say that underline parametric model is appropriate. This procedure is useful only when $\bar{F}(t)$ can be expressed as linear function of time and $\Psi_1(\bar{F}(t))$ is independent of unknown parameters. For generalized exponential distribution, $\Psi_1(\bar{F}(t)) = -\ln[1 - (1 - \bar{F}(t))^{1/\alpha}]$, $\beta_0 = 0$, $\beta_1 = \lambda$, $\Psi_2(t) = t$. In this case $\Psi_1(\bar{F}(t))$ depends on the unknown parameter α , so we substitute estimate of α in $\Psi_1(\bar{F}(t))$ and plot against $\Psi_2(t)$. The maximum likelihood estimates of α_1 and α_2 are $\hat{\alpha}_1 = 0.6176655$ and $\hat{\alpha}_2 = 0.9214646$ respectively. Figure 1 represents goodness of fit plots for generalized exponential for first and second recurrence times respectively. From the figure we can observe that many of the points are nearly on the straight line.

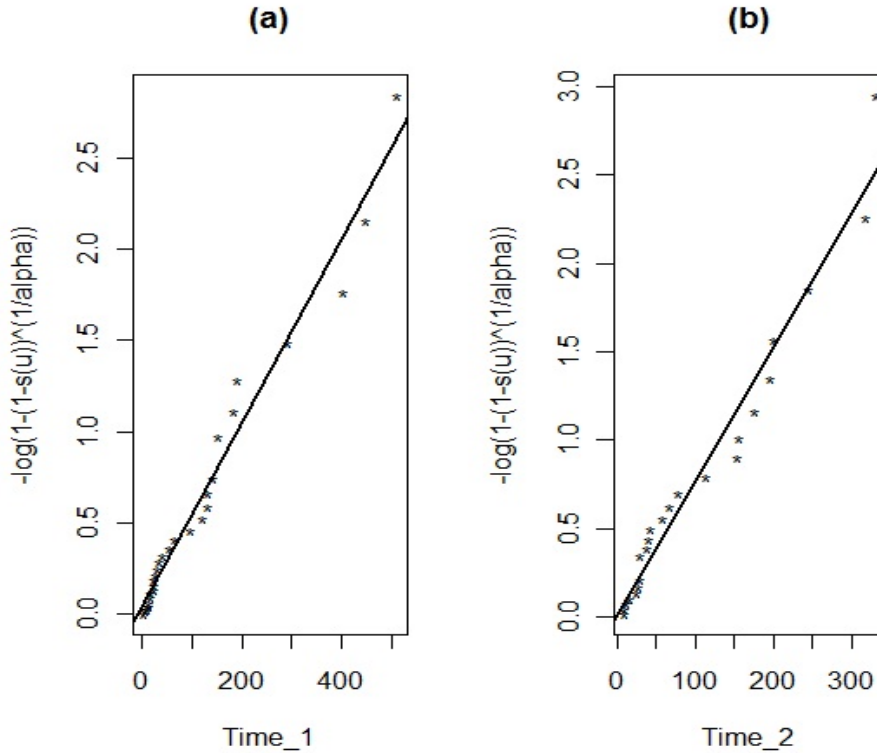


Figure 1: Graph for goodness of fit for kidney infection data for generalized exponential distribution (a) recurrence time 1; (b) recurrence time 2

Now we consider likelihood ratio statistic proposed by Turnbull and Weiss (1978) and discussed in Lawless (2003) to check goodness of fit of kidney data set. Test procedure is developed for testing goodness of fit with grouped data which are subject to random right censoring. Procedure is discussed in the Appendix. Here we assume that survival times of individuals are conditionally independent so we apply likelihood ratio statistic to the survival times of both the individuals separately. We classify the survival times into five groups for first and second recurrence times. Table 1 gives the classification for first and second recurrence times respectively. To illustrate how censoring times across the interval are obtained, consider class interval $[4, 26)$ of second recurrence time. In the interval $[4, 26)$ censoring times are 4, 5, 8, 8, 13, 16, 24 and 25. Out of which 4 is exactly and 5 is nearly at the beginning of the interval. So assuming that individuals are censored at the beginning of the interval, 24 and 25 are more or less at the end of the interval. So assuming that individuals are survived for complete interval and 8, 8, 13, and 16 are more or less at half the interval. So number of censoring times across the interval are $w_j = 4$. The p -values of likelihood ratio test for generalized exponential distribution for first and second recurrence times are

respectively, 0.2564597 and 0.2194112. Thus from goodness of fit graph (Figure 1) and p -value of likelihood ratio test we can say there is no statistical evidence to reject the hypothesis that first and second recurrence times follows generalized exponential distribution.

Table 1: Classification of survival times for first and second recurrence times of kidney data set

class interval	censoring		death	risk
	at the beginning or at the end	across the interval		
first recurrence time				
[2, 13)	0	2	5	38
[13, 23)	1	0	5	31
[23, 54)	0	0	7	25
[54, 149)	1	1	6	17
[149, 537)	1	0	9	9
second recurrence time				
[4, 26)	4	4	5	36
[26, 38)	0	0	5	25
[38, 108)	0	2	6	20
[108, 201)	1	1	5	11
[201, 563)	0	0	5	5

The proportional hazard model for j^{th} individual at i^{th} survival time with frailty term for all the four models we have considered in our study is,

$$h(t_{1j}, t_{2j} | z_j, \underline{X}_j) = z_j \frac{\alpha_1 \lambda_1 e^{-\lambda_1 t_{1j}} (1 - e^{-\lambda_1 t_{1j}})^{\alpha_1 - 1}}{1 - (1 - e^{-\lambda_1 t_{1j}})^{\alpha_1}} \frac{\alpha_2 \lambda_2 e^{-\lambda_2 t_{2j}} (1 - e^{-\lambda_2 t_{2j}})^{\alpha_2 - 1}}{1 - (1 - e^{-\lambda_2 t_{2j}})^{\alpha_2}} \cdot e^{\underline{X}_j \underline{\beta}},$$

where $e^{\underline{X}_j \underline{\beta}} = e^{\beta_1 X_{1j} + \beta_2 X_{2j} + \beta_3 X_{3j} + \beta_4 X_{4j} + \beta_5 X_{5j}}$, β_a is a^{th} regression coefficient. Covariates X_{1j} and X_{2j} represents respectively age and sex of j^{th} patient. X_{2j} takes the value 1 if j^{th} patient is female and zero otherwise. X_{3j} , X_{4j} , X_{5j} are indicator variables taking the values 1 or 0 if disease type GN, AN and PKD are present or absent for j^{th} individual.

For model fitting, we run two parallel chains for all four models using two sets of prior distributions with the different starting points using Metropolis-Hastings algorithm and Gibbs sampler based on normal transition kernels. We iterate

both the chains for 99000 times. We got nearly same estimates of parameters for both the set of priors, so estimates are not dependent on the different prior distributions. Convergence rate of Gibbs sampler for both the prior sets is also not greatly different. Also both the chains shows somewhat similar results, so we present here the analysis for only one chain with $G(a_1, a_2)$ as prior for baseline parameters, for all the four models.

Trace plots, coupling from past plots and sample autocorrelation plots for baseline parameter β_1 and frailty parameters σ^2 for model-I are presented in Figure 2 and Figure 3 respectively. These plots are presented only for β_1 and σ^2 because sample autocorrelation plot for β_1 is larger amongst all other parameters for the model-I and σ^2 is main interest. For all other parameters graphs have similar pattern so due to lack of space we are not presenting graphs for other parameters. Trace plots for all the parameter shows zigzag pattern which indicate the parameters move more freely and appropriate.

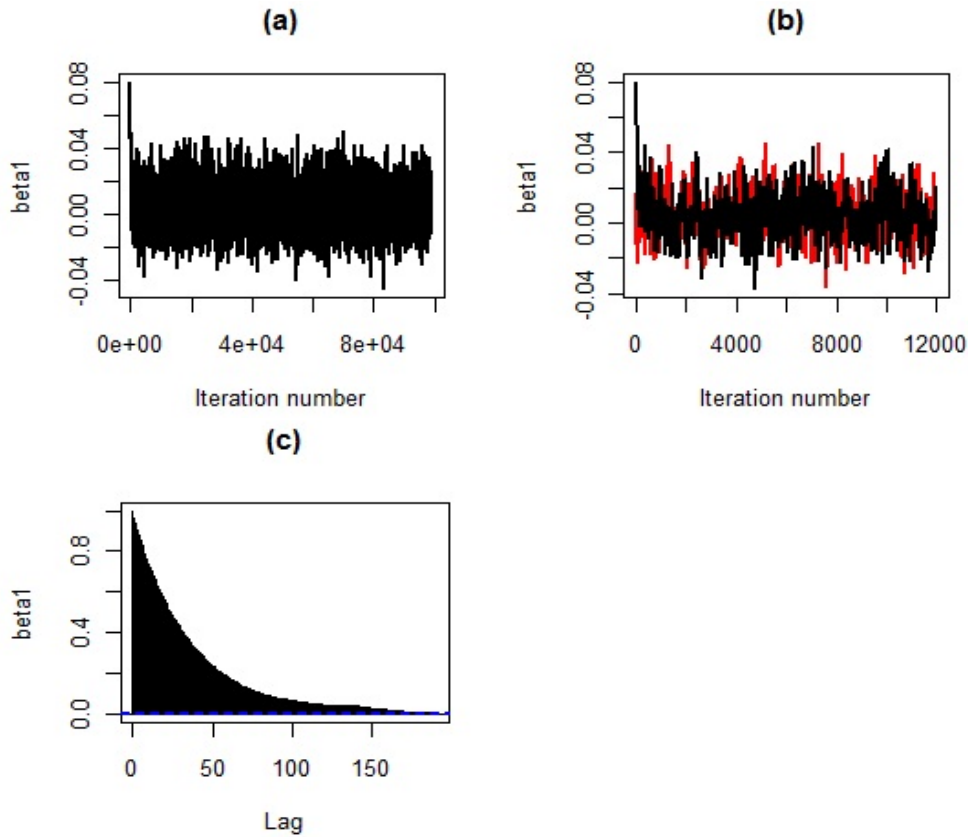


Figure 2: (a) Trace plot; (b) Coupling with past plot and (c) sample autocorrelation function plot for the parameter β_1 of model with gamma as frailty distribution

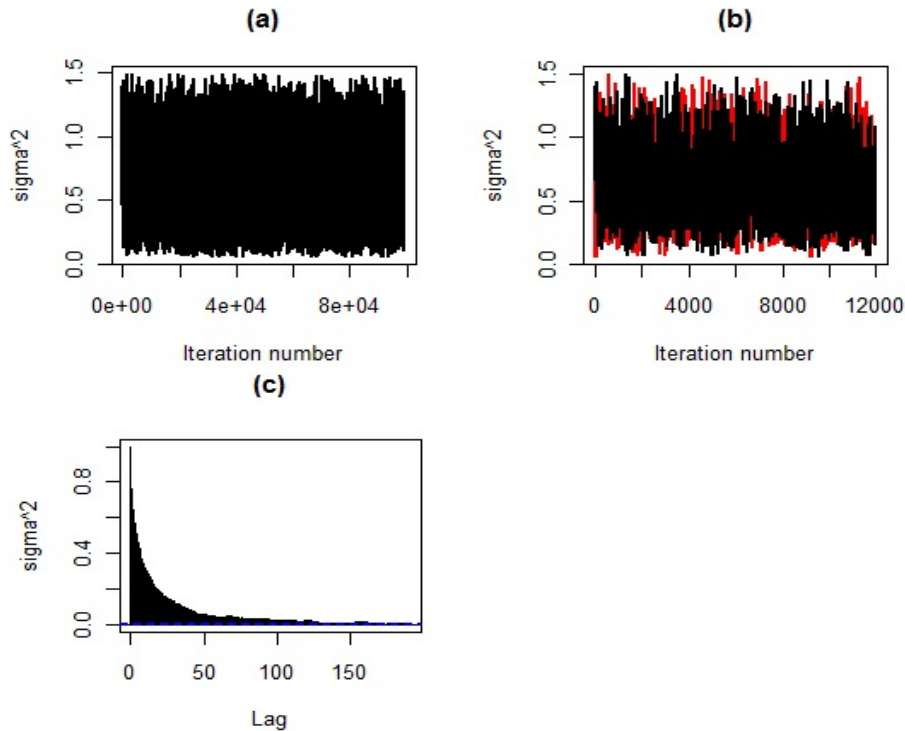


Figure 3: (a) Trace plot; (b) Coupling with past plot and (c) sample autocorrelation function plot for the parameter σ^2 of model with gamma as frailty distribution

Tables 2, 3, 4 and 5 present posterior summary for models I, II, III and IV respectively. 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 attains stationary distribution as posterior distribution. Simulated values of parameters have autocorrelation of lag k so every k^{th} iteration is selected as sample. Autocorrelation lag is different for different parameters, to have single value we have taken maximum of all the lag value as k . From the posterior summary, we can observe that regression coefficients for all the four models are more or less same also these estimates of regression coefficients in our models are qualitatively similar to those found by McGilchrist and Aisbett (1991). Also for all these four models, the value zero is not a credible value for the only credible interval of the regression coefficient X_2 , so X_2 that is sex variable seems only significant. Negative value of β_2 indicates that the female patients have a lower risk for infection as compared to male patients. The estimate of variance frailty σ^2 from different models (model-I = 0.4601; model-II = 0.9428; model-III = 0.4047 and model-IV = 1.2158) shows that there is a strong evidence of high degree of heterogeneity in the population

of patients. Some patients are expected to be vary 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.

Table 2: Posterior summary for the model-I fitted to kidney infection data set

sample size = 457;		burn in period = 12000;		autocorrelation lag = 190			
parameter	Gelman-Rubin factor	Geweke test		estimates	standard error	Credible limits	
		test value	p -value			lower	upper
α_1	1.0001	-0.0087	0.4965	1.6271	0.4666	0.8315	2.7253
λ_1	1.0001	-0.0087	0.4965	0.0485	0.0120	0.0252	0.0668
α_2	1.0001	-0.0007	0.4997	2.0834	0.4963	1.1888	2.9466
λ_2	1.0002	-0.0021	0.4992	0.0404	0.0112	0.0199	0.0615
σ^2	1.0008	0.0019	0.5008	0.0523	0.2710	0.0904	1.1396
β_1	1.0020	0.0021	0.5008	0.0052	0.0128	-0.0198	0.0310
β_2	1.0003	0.0023	0.5009	-2.0162	0.4456	-2.9035	-1.1826
β_3	1.0025	0.0009	0.5004	0.1683	0.5765	-0.9550	1.3242
β_4	1.0006	-0.0004	0.4998	0.5370	0.5596	-0.5467	1.6324
β_5	1.0005	-0.0006	0.4998	-0.8764	0.9642	-2.6354	0.9944

Table 3: Posterior summary for the model-II fitted to kidney infection data set

sample size = 289;		burn in period = 11000;		autocorrelation lag = 290			
parameter	Gelman-Rubin factor	Geweke test		estimates	standard error	Credible limits	
		test value	p -value			lower	upper
α_1	1.0007	0.0039	0.5016	1.6085	0.4872	0.7805	2.6558
λ_1	1.0009	0.0060	0.5024	0.0429	0.0096	0.0222	0.0572
α_2	1.0034	0.0018	0.5007	2.1389	0.5464	1.1784	3.1392
λ_2	1.0008	-0.0023	0.4991	0.0367	0.0094	0.0167	0.0548
σ^2	1.0009	0.0004	0.5001	0.9428	0.6003	0.1051	2.3756
β_1	1.0000	-0.0102	0.4959	0.0094	0.0133	-0.0137	0.0404
β_2	1.0000	0.0044	0.5017	-1.9056	0.4826	-2.9078	-1.0237
β_3	1.0006	0.0144	0.5057	0.0680	0.6247	-1.2991	1.3727
β_4	1.0002	0.0160	0.5064	0.4759	0.5854	-0.6686	1.6687
β_5	1.0004	0.0149	0.5060	-1.1654	0.8767	-2.7474	0.4783

Table 4: Posterior summary for the model-III fitted to kidney infection data set

sample size = 366;		burn in period = 7500;		autocorrelation lag = 250			
parameter	Gelman-Rubin factor	Geweke test		estimates	standard error	Credible limits	
		test value	<i>p</i> -value			lower	upper
α_1	1.0016	0.0001	0.5000	1.5137	0.4181	0.8017	2.4265
λ_1	1.0017	0.0040	0.5016	0.0424	0.0094	0.0232	0.0569
α_2	1.0042	-0.0061	0.4976	2.0856	0.5564	1.1131	3.2762
λ_2	1.0003	0.0020	0.5008	0.0372	0.0089	0.0215	0.0554
ρ	1.0003	0.0022	0.5009	3.5844	0.7400	2.1020	4.9108
γ	1.0040	-0.0099	0.4960	2.2192	1.2392	0.4339	5.1329
β_1	1.0070	-0.0089	0.4964	0.0012	0.0133	-0.0240	0.0284
β_2	1.0154	0.0038	0.5015	-1.9800	0.5095	-2.9818	-1.0048
β_3	1.0030	0.0122	0.5049	0.3353	0.5876	-0.7619	1.5083
β_4	1.0096	0.0079	0.5032	0.7007	0.5888	-0.5476	1.8307
β_5	1.0006	0.0011	0.5044	0.3476	2.0948	-3.8703	4.0066

Table 5: Posterior summary for the model-IV fitted to kidney infection data set

sample size = 94;		burn in period = 13000;		autocorrelation lag = 910			
parameter	Gelman-Rubin factor	Geweke test		estimates	standard error	Credible limits	
		test value	<i>p</i> -value			lower	upper
α_1	1.0000	0.0030	0.5012	1.8241	0.4986	0.9591	2.8201
λ_1	1.0006	-0.0051	0.4980	0.0406	0.0097	0.0226	0.0570
α_2	1.0014	-0.0014	0.4994	2.3908	0.5127	1.4216	3.5155
λ_2	1.0001	-0.0078	0.4969	0.0361	0.0105	0.0170	0.0558
p	1.0000	0.0028	0.5011	0.0674	0.0226	0.0405	0.1186
γ	1.0014	0.0002	0.5001	0.5037	0.0514	0.4054	0.5981
β_1	1.0058	0.0109	0.5044	0.0199	0.0157	-0.0141	0.0467
β_2	1.0004	-0.0047	0.4981	-2.2746	0.5687	-3.2459	-1.1672
β_3	1.0036	-0.0077	0.4969	0.1034	0.6406	-1.2087	1.4417
β_4	1.0050	-0.0050	0.4980	0.4040	0.7624	-1.0926	1.9591
β_5	1.0066	-0.0132	0.4948	-1.1188	1.2073	-3.3837	1.5444

To check the adequacy of the models firstly we have constructed 99%, 95%, 90%, 75% and 50% equal tailed predictive intervals of generated random sample from predictive distribution and counted the total number of intervals in which r^{th} observation falls in their respective intervals. For 99% predictive intervals almost all intervals contains r^{th} observation. Under the model-I, 73, 72, 64 and 52 out of 76 observations were contained in the 95%, 90%, 75% and 50% predictive intervals where as under model-II, model-III and model-IV the respective number of observations were 73, 71, 61 and 53; 74, 73, 66 and 51; 74, 71, 63 and 56. For each of the observation at least $100(1 - \alpha)\%$ predictive interval contains their respective observation. This shows that all four models are adequate for the kidney infection data.

To compare four models we firstly use AIC, BIC and DIC values which are given in Table 6. The AIC, BIC and DIC values for model-I and model-II are comparatively lesser than for model-III and model-IV, so model-III and model-IV are not the better models than model-I and model-II. The difference between AIC, BIC and DIC values for model-I and model-II is very small, so AIC, BIC and DIC values are not worthy to take decision between the model-I and model-II. Note that log likelihood (see Table 6) is larger for model-I and model-II than other two models also log likelihood for model-I and model-II are more or less equal. To take decision about better model between model-I and model-II, we now use Bayes factor and pseudo-Bayes factor which are provided in Tables 7 and 8 respectively.

Table 6: AIC, BIC and DIC values for all the four models fitted to kidney data set

	model-I	model-II	model-III	model-IV
AIC	677.0561	677.4435	682.2805	690.0620
BIC	693.4320	693.8193	700.2940	708.0755
DIC	675.4083	675.2130	692.3332	687.7289
log likelihood	-328.5280	-328.7220	-330.1400	-334.0310

Table 7: $D_{uv} = 2 \log B_{uv}$ values under Bayes factor for models fitted to kidney data set

numerator model	denominator model		
	inverse Gaussian	compound Poisson	compound negative binomial
gamma	$D_{I,II} = -0.4107$	$D_{I,III} = 43.0520$	$D_{I,IV} = 13.7281$
inverse Gaussian	-	$D_{II,III} = 43.4627$	$D_{II,IV} = 14.1388$
compound Poisson	-	-	$D_{III,IV} = -29.3240$

Table 8: $D_{uv} = 2 \log B_{uv}$ values under pseudo-Bayes factor for models fitted to kidney data set

numerator model	denominator model		
	inverse Gaussian	compound Poisson	compound negative binomial
gamma	$D_{I,II} = -1.7617$	$D_{I,III} = 29.3788$	$D_{I,IV} = 8.8751$
inverse Gaussian	-	$D_{II,III} = 31.1405$	$D_{II,IV} = 10.6368$
compound Poisson	-	-	$D_{III,IV} = -20.5038$

From the Table 7, we can observe that, except $D_{I,II}$ all other D_{uv} values are greater than two. $D_{I,III}$, $D_{I,IV}$, $D_{II,III}$ and $D_{II,IV}$ are positive which shows the positive evidence against denominator models where as $D_{III,IV}$ is negative indicating numerator model is better. Thus compound Poisson model is better than compound negative binomial model but both these models are worst than gamma and inverse Gaussian models. Same conclusion is found with pseudo-Bayes factor.

Another diagnostic we now use is CPO plot. We have plotted difference of log of CPO values for pair of models. Figure 4 represents these plots. In figure part (a), (b) and (c) considers pair of model-I with models II, III and IV, respectively. Part (d), (e) and (f) considers pair of model-II with models III, IV and model-III with model-IV. From the part (a) we can observe that larger the points are in negative side indicating model-II is better than model-I. In the parts (c) and (e) hardly some points are tending to negative side so we can say model-III is worst model than model-I and model-II. In the parts (b), (d) and (f) almost all points get clustered at zero, only some points are to the positive side in part (b), (d) and to the negative side in part (f) so we can say according to CPO model-III is somewhat similar to model-I and model-II also model-III is similar to model-IV.

Finally, the values relating to the model choice criteria D_ω (8.15) for all for models are shown in Table 9. First and second column of the Table 9 gives penalty and goodness-of-fit term, next four columns gives D_ω values for different values of $\omega = 1, 5, 10$ and ∞ . Penalty term and goodness-of-fit term are minimum for model-I, also for all the values of ω , D_ω is minimum for model-I, so this criteria also suggest model-I is better model. Thus model-I that is gamma frailty model with baseline as generalized exponential distribution is better model amongst all other model for modeling kidney infection data. One thing is to be noted that, model-III comes out to be better model than model-II in this situation.

10. Discussion

In the present paper, Bayesian shared frailty model is adopted to analyze kidney infection data. A generalized exponential distribution is considered as

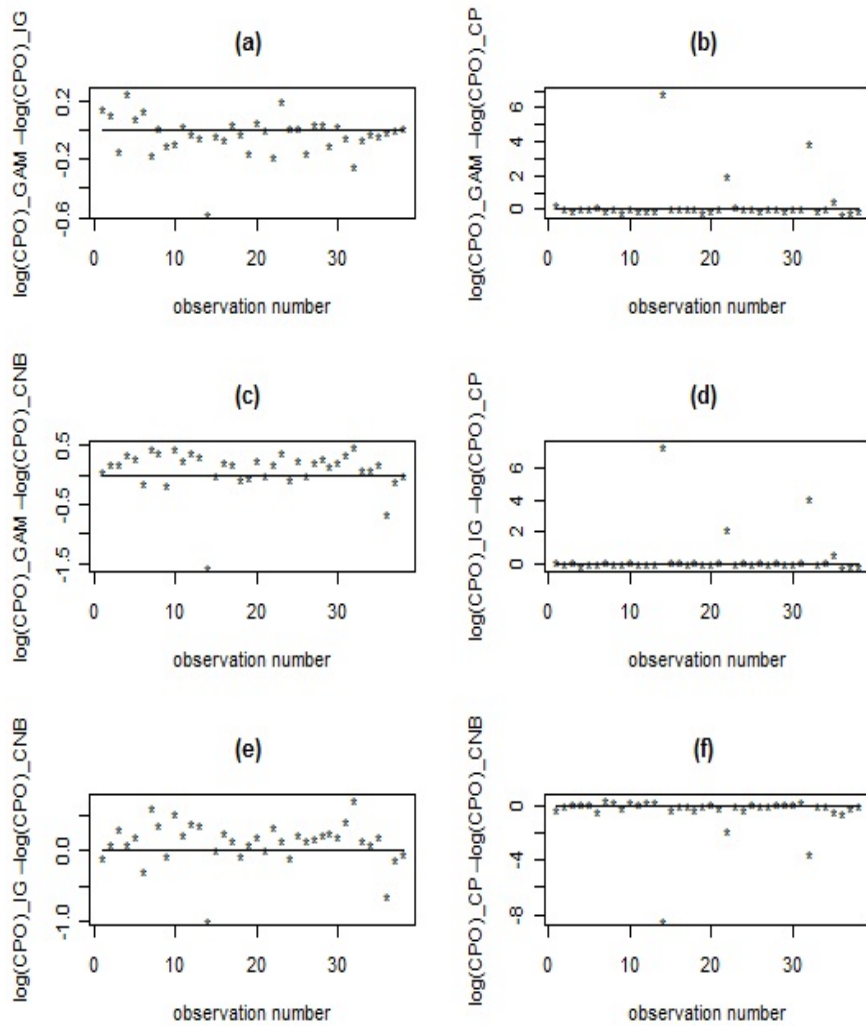


Figure 4: Plot of log CPO values for models fitted to kidney infection data

Table 9: Model selection criteria $(8.15) \times 10^{-5}$ for all the models fitted to kidney data set

frailty distribution	Penalty term	Goodness-of-fit term	D_1	D_5	D_{10}	D_∞
gamma	24.7503	17.1280	33.3143	39.0236	40.3211	47.8783
inverse Gaussian	27.2831	19.0657	36.8160	43.1712	44.6155	46.3488
compound poisson	27.2543	17.8312	36.1699	42.1136	43.4644	45.0854
compound negative binomial	38.7892	20.9833	49.2802	56.2752	57.8649	59.7725

baseline distribution, because for generalized exponential distribution, hazard rate increases from zero to a finite constant, when shape parameter α increases and hazard rate decreases from infinity to a finite number when α is less than one. A nearly constant rate after a certain time period implies that the occurrence of failure is purely random and is independent of past life; this is a property of the failure rate of an exponential distribution which has been extensively used in reliability studies. Also Gupta and Kundu (1999) suggested that, generalized exponential distribution can be used effectively in analyzing many life time data sets particularly in place of Weibull which is mostly used in the literature for modeling life times.

In the literature, gamma distribution is mostly used as frailty distribution because of its simplicity. So our aim of this study is to compare different shared frailty models to analyze kidney infection data with generalized exponential as baseline distribution. As discussed in the Section 4, Hougaard (1984) introduced inverse Gaussian distribution as a better frailty distribution. So we selected inverse Gaussian distribution as one of the frailty distribution. To consider non-susceptibility or zero susceptibility we considered two other frailty distributions, compound Poisson and compound negative binomial distribution. We used Metropolis-Hastings algorithm and Gibbs sampler to fit all the four models and a comparison is made between the models using Bayesian model comparison criteria such as, AIC, BIC, DIC, Bayes factor, CPO plot, prediction model choice criteria. We used R statistical software to perform programs to analyze kidney infection data.

From the graphical plot and likelihood ratio test we can say that, recurrence times are well fitted by the generalized exponential distribution. The posterior summaries for all the regression coefficients for all the four models are more or less same also these estimates of regression coefficients in our models are qualitatively similar to those found by McGilchrist and Aisbett (1991). For all the four models, the value zero is not a credible value for the only credible interval of the regression coefficient X_2 , so X_2 that is sex variable seems only significant. Negative value of β_2 indicates that the female patients have a lower risk for infection as compared to male patients. Also the estimates of frailty variance for all the models say that there is strong evidence of high degree of heterogeneity in the population of patients.

The AIC, BIC, DIC values, Bayes factor, pseudo-Bayes factor and log likelihood all these criteria suggest that model-I and model-II are better than model-III and model-IV and model-I and model-II are similar to fit kidney infection data. Only loss function criteria (8.15) suggest that model-I is better than all other models whereas CPO plot suggest model-II is better than model-I. Thus, by referring all the above analysis, now we are in a position to conclude that, gamma

and inverse Gaussian shared frailty models with generalized exponential as baseline distribution are equivalent models and are better fit for modeling of kidney infection data than other models considered in this study. One thing is to be noted here that, there is no any physical reason for choosing frailty distribution it is just based on mathematical convenience.

Appendix: Likelihood Ratio Test

Suppose we have n subjects. Let T_i and c_i be respectively represents actual life time and random censoring time for i^{th} subject. We observe event time t_i^* and censoring indicator δ_i where, $t_i^* = \text{minimum}(T_i, c_i)$ and

$$\delta_i = \begin{cases} 1, & T_i < c_i, \\ 0, & T_i > c_i. \end{cases}$$

We wish to test the hypothesis that, life time of a subject T_i follows a distribution having survival function $\bar{F}_0(t, \underline{\theta})$, where $\bar{F}_0(t, \underline{\theta})$ is some specified function except for vector of unknown parameters say $\underline{\theta}$. Mathematically we can write hypothesis as,

$$H_0 : \bar{F}(t, \underline{\theta}) = \bar{F}_0(t, \underline{\theta}), \quad \forall t \in \mathfrak{R}, \text{ against,}$$

$$H_1 : \bar{F}(t, \underline{\theta}) \neq \bar{F}_0(t, \underline{\theta}), \quad \text{for at least one } t \in \mathfrak{R}.$$

Suppose the observations $t_1^*, t_2^*, \dots, t_n^*$ are grouped and the life times recorded belongs to one and only one of the k interval $I_j = [a_{j-1}, a_j)$ with $a_0 = 0$ and $a_k = T$ as an upper limit of observations. We assume that any right censored observation at T_i occur immediately after any observed deaths. Suppose an interval I_j has n_j subjects at risk at the beginning of an interval, d_j number of deaths and w_j number of censored subjects. If censoring occurs at the beginning or at the end of the interval then there is no requirement to add censoring in the likelihood function because it get adjusted in the risk factor but if censoring occurs anywhere across the interval then some adjustment is required for example, an ad-hoc procedure discussed in Lawless (2003). In ad-hoc procedure we suppose that a censored individual is at risk for half the interval and we define an effective number of individuals at risk for the interval I_j as $n'_j = n_j - w_j/2$. This adjustment is somewhat arbitrary but sensible in many situations when the intervals are not too long and censoring is not too heavy. Of course its appropriateness depends on the failure and censoring process.

Probability of surviving beyond an interval $I_j (j = 1, 2, \dots, k-1)$ i.e., beyond time a_j is, $\bar{F}_j = P(T > a_j) = \bar{F}(a_j)$ and for k^{th} interval $P(T > a_k) = 0$. Probability of surviving beyond any interval I_j given that subject is alive at the

beginning of an interval is,

$$s_j = P(T > a_j | T > a_{j-1}) = \frac{\bar{F}_j}{\bar{F}_{j-1}}, \quad j = 1, 2, \dots, k-1.$$

Probability that subject dies in interval I_j but alive at the beginning of the interval is,

$$q_j = P(T < a_j | T \geq a_{j-1}) = 1 - s_j.$$

Thus for j^{th} interval I_j ($j = 1, 2, \dots, k-1$) there are d_j deaths with probability $(1 - s_j)$ and $(n'_j - d_j)$ number of survivors with probability s_j , therefore contribution of j^{th} interval in likelihood function is,

$$(1 - s_j)^{d_j} \cdot s_j^{(n'_j - d_j)},$$

and for k^{th} interval there are d_k deaths with probability $(1 - s_k)$ and no one survives beyond the interval therefore contribution of k^{th} interval in likelihood function is, $(1 - s_k)^{d_k}$ therefore likelihood function is,

$$L(s_1, s_2, \dots, s_k) = \prod_{j=1}^{k-1} (1 - s_j)^{d_j} \cdot s_j^{(n'_j - d_j)} \cdot (1 - s_k)^{d_k},$$

where

$$n'_j = n_j - \frac{w_j}{2} = n - \sum_{i=1}^{j-1} (d_i + w_i) - \frac{w_j}{2},$$

is the effective number of individuals at risk for j^{th} interval. In terms of s_j the null hypothesis H_0 can also be rewritten as,

$$H_0 : s_j = s_{j0}, \quad j = 1, 2, \dots, k, \text{ against,}$$

$$H_1 : s_j \neq s_{j0}, \quad \text{for at least one } j = 1, 2, \dots, k.$$

From generalized likelihood ratio test,

$$\Lambda = \frac{\max_{s \in \Omega_0} L(s)}{\max_{s \in \Omega} L(s)},$$

where $\Omega = \{s_i; 0 \leq s_i \leq 1 \text{ and } \sum_{i=1}^k s_i = 1\}$ is general parameter space and Ω_0 is some sub-space of Ω under H_0 . Under H_0 , $-2 \log \Lambda$ is asymptotically distributed as χ_{k-p}^2 , where p represents number of estimated parameters under H_0 and k

number of intervals. Hence test criteria is, reject H_0 against H_1 if $-2 \log \Lambda$ exceeds the $(1-\alpha)\%$ quantile of the χ_{k-p}^2 .

Likelihood function under general parameter space Ω can be maximized using Kaplan-Meier estimator of survival function. So MLE of s_j is given by $\hat{s}_j = (n'_j - d_j)/n'_j$. Therefore under general parameter space maximum log likelihood is,

$$\begin{aligned} \log L(\hat{s}_1, \hat{s}_2, \dots, \hat{s}_k) &= \sum_{j=1}^k d_j \log d_j + \sum_{j=1}^{k-1} (n'_j - d_j) \log(n'_j - d_j) \\ &\quad - \sum_{j=1}^{k-1} n'_j \log n'_j - d_k \log n_k. \end{aligned}$$

Let $\tilde{\theta}$ be denote MLE of θ under H_0 and therefore, $\tilde{s}_j = \tilde{F}_j(\tilde{\theta})/\tilde{F}_{j-1}(\tilde{\theta})$ will maximize likelihood function under H_0 . Therefore,

$$\begin{aligned} \log \max_{\underline{s} \in \Omega_0} L(\underline{s}) &= \log L(\tilde{s}_1, \tilde{s}_2, \dots, \tilde{s}_k) \\ &= \sum_{j=1}^k d_j \log[\tilde{F}_{j-1} - \tilde{F}_j] + \sum_{j=1}^{k-1} (n'_j - d_j) \log \tilde{F}_j - \sum_{j=1}^{k-1} n'_j \log \tilde{F}_{j-1} - d_k \log \tilde{F}_{k-1}. \end{aligned}$$

Here we use Newton-Raphson iterative procedure to estimate the parameters under H_0 because first order partial derivatives of the log likelihood are not easy to solve. Suppose $\tilde{\lambda}$ and $\tilde{\alpha}$ are estimators of the parameters λ and α respectively then $\log L(\tilde{s}_1, \tilde{s}_2, \dots, \tilde{s}_k)$ is given by,

$$\sum_{j=1}^k d_j \log[e^{-\tilde{\lambda} a_{j-1}^{\tilde{\alpha}}} - e^{-\tilde{\lambda} a_j^{\tilde{\alpha}}}] - \tilde{\lambda} \sum_{j=1}^{k-1} (n'_j - d_j) a_j^{\tilde{\alpha}} + \tilde{\lambda} \sum_{j=1}^{k-1} n'_j a_{j-1}^{\tilde{\alpha}} + \tilde{\lambda} d_k a_{k-1}^{\tilde{\alpha}},$$

for Weibull distribution,

$$\begin{aligned} &\sum_{j=1}^k d_j \log[(1 - e^{-\tilde{\lambda} a_j^{\tilde{\alpha}}})^{\tilde{\alpha}} - (1 - e^{-\tilde{\lambda} a_{j-1}^{\tilde{\alpha}}})^{\tilde{\alpha}}] + \sum_{j=1}^{k-1} (n'_j - d_j) \log[1 - (1 - e^{-\tilde{\lambda} a_j^{\tilde{\alpha}}})^{\tilde{\alpha}}] \\ &- \sum_{j=1}^{k-1} n'_j \log[1 - (1 - e^{-\tilde{\lambda} a_{j-1}^{\tilde{\alpha}}})^{\tilde{\alpha}}] - d_k \log[1 - (1 - e^{-\tilde{\lambda} a_{k-1}^{\tilde{\alpha}}})^{\tilde{\alpha}}], \end{aligned}$$

for generalized exponential distribution and for exponential power distribution,

$$\begin{aligned} &\sum_{j=1}^k d_j \log\{e^{[1 - e^{-\tilde{\lambda} a_{j-1}^{\tilde{\alpha}}}] - e^{[1 - e^{-\tilde{\lambda} a_j^{\tilde{\alpha}}}]}\} + \sum_{j=1}^{k-1} (n'_j - d_j) [1 - e^{-\tilde{\lambda} a_j^{\tilde{\alpha}}}] \\ &- \sum_{j=1}^{k-1} n'_j [1 - e^{-\tilde{\lambda} a_{j-1}^{\tilde{\alpha}}}] - d_k [1 - e^{-\tilde{\lambda} a_{k-1}^{\tilde{\alpha}}}]. \end{aligned}$$

Acknowledgements

We thank the referee for the useful suggestions and comments which improved the earlier version of the manuscript.

References

- Aalen, O. O. (1992). Modelling heterogeneity in survival analysis by the compound Poisson distribution. *Annals of Applied Probability* **2**, 951-972.
- Aalen, O. O. and Tretli, S. (1999). Analyzing incidence of testis cancer by means of a frailty model. *Cancer Causes and Control* **10**, 285-292.
- Bolstad, W. M. (2010). *Understanding Computational Bayesian Statistics*. Wiley John & Sons, New York.
- Boneg, Y. T. (2001). Weibull frailty for modelling heterogeneity. Unpublished Ph.D. Dissertation, University of Gueiph.
- Carlin, B. P. and Louis, T. A. (2009). *Bayesian Methods for Data Analysis*, 3rd edition. Chapman and Hall/CRC Press, New York.
- 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**, 141-151.
- Chhikara, R. S. and Folks, J. L. (1986). *The Inverse Gaussian Distribution*. Marcel Dekker, New York.
- Gamerman, D. (1997). *Markov Chain Monte Carlo: Stochastic Simulation for Bayesian Inference*. Chapman and Hall, London.
- Gelfand, A. E. (1996). Model determination using sampling based methods. In *Markov Chain Monte Carlo in Practice* (Edited by W. R. Gilks, S. Richardson and D. J. Spiegelhalter), 145-161. Chapman and Hall, London.
- Gelfand, A. E. and Ghosh, S. K. (1998). Model choice: a minimum posterior predictive loss approach. *Biometrika* **85**, 1-11.
- Gelman, A., Carlin, J. B., Stern, H. S. and Rubin, D. B. (2003). *Bayesian Data Analysis*, 2nd edition. Chapman and Hall, New York.
- Gupta, R. D. and Kundu, D. (1999). Generalized exponential distributions. *Australian and New Zealand Journal of Statistics* **41**, 173-188.

-
- Hanagal, D. D. (2010a). Correlated compound Poisson frailty model for the bivariate survival data. *International Journal of Statistics and Management Systems* **5**, 127-140.
- Hanagal, D. D. (2010b). Modeling heterogeneity for bivariate survival data by compound Poisson distribution. *Model Assisted Statistics and Applications* **5**, 1-9.
- Hanagal, D. D. (2010c). Modeling heterogeneity for bivariate survival data by compound Poisson distribution with random scale. *Statistics and Probability Letters* **80**, 1781-1790.
- Hougaard, P. (1984). Life table methods for heterogeneous populations. *Biometrika* **71**, 75-83.
- Hougaard, P. (2000). *Analysis of Multivariate Survival Data*. Springer, New York.
- Ibrahim, J. G., Chen, M. H. and Sinha, D. (2001). *Bayesian Survival Analysis*. Springer, New York.
- Kass, R. E. and Raftery, A. E. (1995). Bayes factor. *Journal of the American Statistical Association* **90**, 773-795.
- 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.
- Lawless, J. F. (2003). *Statistical Models and Methods for Lifetime Data*, 2nd edition. John Wiley & Sons, New York.
- McGilchrist, C. A. (1993). REML estimation for survival models with frailty. *Biometrics* **49**, 221-225.
- McGilchrist, C. A. and Aisbett, C. W. (1991). Regression with frailty in survival analysis. *Biometrics* **47**, 461-466.
- Moger, T. A. and Aalen, O. O. (2005). A distribution for multivariate frailty based on the compound Poisson distribution with random scale. *Lifetime Data analysis* **11**, 41-59.
- Sahu, S. K., Dey, D. K., Aslanidou, H. and Sinha, D. (1997). A Weibull regression model with gamma frailties for multivariate survival data. *Lifetime Data Analysis* **3**, 123-137.

- Santos dos, 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.
- Spiegelhalter, D. J., Best, N. G., Carlin, B. P. and Van der Linde, A. (2002). Bayesian measure of model complexity and fit (with discussion). *Journal of the Royal Statistical Society, Series B* **64**, 583-639.
- Turnbull, B. W. and Weiss, L. (1978). A likelihood ratio statistic for testing goodness of fit with randomly censored data. *Biometrics* **34**, 367-375.
- Vaupel, J. W., Manton, K. G. and Stallaed, E. (1979). The impact of heterogeneity in individual frailty on the dynamics of mortality. *Demography* **16**, 439-454.
- Yu, B. (2006). Estimation of shared gamma frailty models by a modified EM algorithm. *Computational Statistics and Data Analysis* **50**, 463-474.

Received June 12, 2012; accepted August 23, 2012.

David D. Hanagal
Department of Statistics
University of Pune
Ganeshkhind Road, Pune-411007, Maharashtra, India.
david_hanagal@yahoo.co.in

Alok D. Dabade
Department of Statistics
University of Pune
Ganeshkhind Road, Pune-411007, Maharashtra, India.
addabade@yahoo.co.in