Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.12785/jsap/030314

Inferences on the Competing Risk Model

*Bhupendra Singh** *and Sachin Tomer*

Department of Statistics, C.C.S. University, Meerut-250 004, India

Received: 23 Mar. 2014, Revised: 14 Sep. 2014 Accepted: 16 Sep. 2014
Published: 1 Nov. 2014

Abstract: In this article, a competing risk model is analyzed in the presence of complete and censored data when the causes of failures follow different family of failure time distributions. We derive the maximum likelihood and Bayes estimators of the parameters involved in the model and the relative risks. The goodness-of-fit of the competing risks model with the considered failure time distributions to a real data set is also demonstrated.

Key words: Competing risk model, Maximum likelihood estimate, Bayes estimate, Survival function, Censored data.

1. Introduction

The competing risks situation arises when subjects under study are at risk of more than one mutually exclusive event, such as death from different causes, and the occurrence of one precludes the occurrence of the other events. Such problems can occur in many fields, including reliability/survival analysis, demography and actuarial science. In analyzing competing risks data, the data comprises a failure time and an indicator denoting the cause of failure. Suppose, there are k latent failure times, one for each possible type of failure. Let T_j be the time to failure from cause j (j=1, 2,….., k). In the presence of competing risks, we only observe the minimum of the latent failures times (T) and the corresponding cause of failure δ where $\delta = j$ if $T_j=Min(T_1, T_2,\ldots, T_k)$. The latent failure times $T_1, T_2,\ldots, T_k$ are assumed to be independent. Although the assumption of independence seems to be very restrictive, but in case of dependence, the underlying distributions are not identifiable on the basis of (T, δ) [see Tsiatis 1975 and Crowder, 1991 & 1993].

Initially, Daniel Bernoulli (1760) considered the competing risk models to separate the risk of dying from smallpox from other risks. Thereafter, several authors have analyzed competing risks model in different context. A book by Crowder [2001] presents excellent review of

* Corresponding author e-mail: bhupendra.rana@gmail.com

literature on the competing risks model. The studies by Kaplan and Meier [1958] and Peterson [1977] dealt with the non-parametric analysis of the competing risks model. On the other hand, a number of authors like Elveback [1960], David and Moeschberger [1978], Dinse [1982], Miyawaka [1982, 1984], Alwasel [2009], Kundu and Basu [2002], and Sarhan [2007] derived parametric inferences on such models.

In all of the above-mentioned studies, it is assumed that the failure time distributions of all the causes belong to the same family. However, in practice, this assumption is not realistic. Some or all of the causes may follow different failure time distributions. Therefore, the objective of this study is to analyze the competing risks model in the presence of complete and censored data when the causes of failures follow different family of failure time distributions. Also, we assume that every member/unit of a target population either dies/fails due to a particular cause or survived/operative till the end of the experiment. That is, we consider the following three types of observations:

- Individuals/units who died/failed, their lifetimes and cause of failure.
- Individuals/units who died/failed, their lifetimes but not the cause of failure.
- Individuals/units who survived/operative till the end of the experiment.

The rest of the paper is organized as follows: We present the assumption and notations needed for describing the model in Section2. The maximum likelihood estimates (MLEs) and Bayes estimates of the unknown parameters involved in the model are derived in section 3. The relative risk rates due to the causes are obtained in section 4. The goodness-of-fit of the competing risks model with the considered failure time distributions to a real data set is demonstrated in section 5. Finally the conclusion is presented.

2. Model Assumptions and Notations

Without loss of generality we assume that there are only two independent causes of failure. However, the methods developed here can be easily extended to the case k>2. Weassume the following notations:

Notations:

N	: number of individuals on the life test.
T_i	: lifetime of an individuali (i=1, 2,…..N).
T_{ji}	: Lifetime of the i^{th} individual under cause j, j = 1, 2.
F(.)	: Cumulative distribution function of T_i.

$F_j(.)$: Cumulative distribution function of T_{ji}.

$\bar{F}(.)$: $= 1 - F(.)$, the survival function of T_i

$f(.)$: Probability density function of T_i

$f_j(.)$: Probability density function of T_{ji}

$\bar{F}_j(.)$: $= 1 - F_j(.)$, the survival function of T_{ji}

δ_i : Indicator variable denoting the cause of failure of the i^{th} individual.

$I(.)$: Indicator function of event [.]

m : Number of complete failures observed before termination.

Weibull(θ, β) : Weibull distribution with parameter θ and β.

Log-normal(μ, σ^2) : Log-normal distribution with μ and σ^2.

Exponential (λ): Exponential distribution with parameter λ

Assumptions

1. The random vectors T_{ji} ; j = 1, 2 and i = 1, 2,…, N are N independent and identically distributed random vectors.
2. The random variables T_{ji} are independent for all i = 1, 2,……., N and j = 1, 2 and $T = Min\{T_{1i}, T_{2i}\}$.
3. (i) The random variable T_{1i} follows Weibull(θ, β) and T_{2i} follows Log-normal (μ, σ^2) Where i = 1, 2,….., N.
   (ii) The random variable T_{1i} follows Weibull(θ, β) and T_{2i} follows exponential(λ) where i = 1, 2,…..,N.
4. In the first m observations we observe the failure times and also causes of failure. Whereas for the successive (n-m) observations we only observe the failure times and not the causes of failure that is the cause of failure is unknown. In the successive (N-n) observations, the units are still operative at the end the project period.

That is, we observe the following data set $(T_{1,}\delta_1),(T_{2,}\delta_2),....,(T_{m,}\delta_m),(T_{m+1,}*),................$ $(T_n,*),(T_{n+1}*,*),...........(T_N*,*)$. Here, (t,δ) means the unit/individual has failed/expired at time t due to cause δ, (t, *) means the unit/individual has failed/expired at time t but the cause of failure is unknown and $(t*,*)$ means the unit has been tested until time t without failing (censored observations). We denote this set by Ω which can be categorized as a union of three disjoint classes Ω_1,Ω_2 and Ω_3. Where Ω_1 represents the set of data when the cause of unit failure is known, while Ω_2, denotes the set of observation when the cause of unit failure is unknown, Ω_3 denotes the set of censored observations. Further, the set Ω_1 can be divided into two disjoint subsets of observations Ω_{11} and Ω_{12}, where Ω_{1j}representsthe set of all observations when the failure of the unit is due to the cause j, j= 1, 2. We also assume that $|\Omega_i| = r_i$, $|\Omega_{1j}| = r_{1j}$. Namely,

$m = r_1 = r_{11} + r_{12}$, $|\Omega_2| = r_2 = n - m$ and $|\Omega_3| = r_3 = N - n$.

5. m and n are prefixed numbers.

3. The Likelihood Function and Estimation

The likelihood function for the observed data set $(T_{1,}\delta_1),(T_{2,}\delta_2),....,(T_{m,}\delta_m),(T_{m+1,}*),......$ $................(T_n,*),(T_{n+1}*,*),...........(T_N*,*).....$, for the general case, take the form

$$L=\prod_{i=1}^{m}\left\{\left[f_1(t_i)\,\bar{F}_2(t_i)\right]^{I(\delta_i=1)}\left[f_2(t_i)\,\bar{F}_1(t_i)\right]^{I(\delta_i=2)}\right\}\times\prod_{i=m+1}^{n} f(t_i)\prod_{i=n+1}^{N}\bar{F}(t_i) \qquad (1)$$

3.1 Case I

Here, we propose the methods of estimation of the competing risks model's parameters, when cause-1 follows the Weibull distribution and cause-2 follows the Log-normal distribution. Based on the assumption 3(i), for j=1, 2 and i=1, 2……., N, the respective cumulative distribution functions of T_{1i} and T_{2i} are

$$F_1(t) = 1 - e^{-\theta t^{\beta}} \quad ; t > 0 \qquad (2)$$

$$F_2(t) = \Phi\left(\frac{\log t - \mu}{\sigma}\right) \quad ; t > 0 \qquad (3)$$

$$\Phi(z) = \frac{1}{\sqrt{2\pi}}\int_{-\infty}^{z} e^{-\frac{t^2}{2}}dt$$

Therefore, the probability density functions are

$$f_1(t) = \theta\beta t^{\beta-1} e^{-\theta t^{\beta}} \;;\; t > 0 \tag{4}$$

$$f_2(t) = \frac{1}{\sqrt{2\pi}\;\sigma t}\exp\left\{-\frac{1}{2\sigma^2}(\log t-\mu)^2\right\} \;;\; t>0 \tag{5}$$

The survival functions are

$$\bar{F}_1(t) = e^{-\theta t^{\beta}} \tag{6}$$

$$\bar{F}_2(t) = 1-\Phi\left(\frac{\log t-\mu}{\sigma}\right) \tag{7}$$

3.1.1 MLEs

Substituting (4)-(7) into (1), the likelihood functions becomes

$$L = \theta^{r_{11}}\beta^{r_{11}}\left(\frac{1}{\sigma}\right)^{r_{12}} \times \exp\left\{-\sum_{t_i\in\Omega_1} t_i^{\beta}\left[\theta\, I(\delta_i=1)+\theta I(\delta_i=2)\right]\right\} \times \prod_{t_i\in\Omega_{11}}\left[t_i^{\beta-1}\left\{1-\Phi\left(\frac{\log t_i-\mu}{\sigma}\right)\right\}\right]$$

$$\times \prod_{t_i\in\Omega_{12}}\left[\frac{1}{t_i}\exp\left\{-\frac{1}{2\sigma^2}(\log t_i-\mu)^2\right\}\right]$$

$$\times \prod_{t_i\in\Omega_2}\left[\left\{\theta\beta t_i^{\beta-1} e^{-\theta t_i^{\beta}}\left(1-\Phi\left(\frac{\log t_i-\mu}{\sigma}\right)\right)\right\}+\left\{\frac{1}{\sqrt{2\pi}\;\sigma t_i}\exp\left\{-\frac{1}{2\sigma^2}(\log t_i-\mu)^2\right\}\times e^{-\theta t_i^{\beta}}\right\}\right]$$

$$\times \prod_{t_i\in\Omega_3}\left[e^{-\theta t_i^{\beta}}\left(1-\Phi\left(\frac{\log t_i-\mu}{\sigma}\right)\right)\right] \tag{8}$$

Therefore, the log- likelihood function is given by

$$\log L = r_{11}(\log\theta+\log\beta)+r_{12}\log\left(\frac{1}{\sigma}\right)-\theta\sum_{t_i\in\Omega_{11}} t_i^{\beta}-\theta\sum_{t_i\in\Omega_{12}} t_i^{\beta}+(\beta-1)\sum_{t_i\in\Omega_{11}}\log t_i$$

$$+\sum_{t_i\in\Omega_{11}}\log\left\{1-\Phi\left(\frac{\log t_i-\mu}{\sigma}\right)\right\}+\sum_{t_i\in\Omega_{12}}\log\left\{\frac{1}{t_i}\right\}-\frac{1}{2}\sum_{t_i\in\Omega_{12}}\left(\frac{\log t_i-\mu}{\sigma}\right)^2$$

$$+\sum_{t_i\in\Omega_2}\log\left[\left\{\theta\beta t_i^{\beta-1} e^{-\theta t_i^{\beta}}\left(1-\Phi\left(\frac{\log t_i-\mu}{\sigma}\right)\right)\right\}+\left\{\frac{1}{\sqrt{2\pi}\;\sigma t_i}\exp\left\{-\frac{1}{2\sigma^2}(\log t_i-\mu)^2\right\}\times e^{-\theta t_i^{\beta}}\right\}\right]$$

$$-\theta\sum_{t_i\in\Omega_3} t_i^{\beta}+\sum_{t_i\in\Omega_3}\log\left\{1-\Phi\left(\frac{\log t_i-\mu}{\sigma}\right)\right\} \tag{9}$$

Equating the first partial derivates of (9) with respect to θ, β, μ and σ to zeros, we get the likelihood equations as given by (10-13), where

$$\xi_i = \left\{\theta\beta t_i^{\beta-1} e^{-\theta t_i^{\beta}} \left(1-\Phi\left(\frac{\log t_i - \mu}{\sigma}\right)\right)\right\} + \left\{\frac{1}{\sqrt{2\pi}\;\;\sigma\, t_i} \exp\left\{-\frac{1}{2\sigma^2}(\log t_i - \mu)^2\right\} \times e^{-\theta t_i^{\beta}}\right\}$$

$$0 = \frac{r_{11}}{\theta} - \sum_{t_i\in\Omega_{11}} t_i^{\beta} - \sum_{t_i\in\Omega_{12}} t_i^{\beta} - \sum_{t_i\in\Omega_3} t_i^{\beta}$$

$$+ \sum_{t_i\in\Omega_2} \frac{\left[\left(\left\{1-\Phi\left(\frac{\log(t_i)-\mu}{\sigma}\right)\right\}\beta t_i^{\beta-1}\left\{e^{-\theta t_i^{\beta}} - \theta e^{-\theta t_i^{\beta}} t_i^{\beta}\right\}\right) - \left(\frac{1}{\sqrt{2\pi}\;\;\sigma\, t_i}\exp\left\{-\frac{1}{2\sigma^2}(\log t_i - \mu)^2\right\} e^{-\theta t_i^{\beta}} t_i^{\beta}\right)\right]}{\xi_i}$$

$$0 = \frac{r_{11}}{\beta} - \theta \sum_{t_i\in\Omega_{11}} t_i^{\beta}\log t_i - \theta \sum_{t_i\in\Omega_{12}} t_i^{\beta}\log t_i + \sum_{t_i\in\Omega_{11}} \log t_i - \theta \sum_{t_i\in\Omega_3} (t_i \log t_i)$$

$$+ \sum_{t_i\in\Omega_2} \frac{\left[\left(\theta\left\{1-\Phi\left(\frac{\log t_i - \mu}{\sigma}\right)\right\}\left\{t_i^{\beta-1} e^{-\theta t_i^{\beta}}\left\{(1+\beta\log t_i) - \theta\beta t_i^{\beta}\log t_i\right\}\right\}\right) - \left(\frac{1}{\sqrt{2\pi}\;\;\sigma\, t_i}\exp\left\{-\frac{1}{2\sigma^2}(\log t_i - \mu)^2\right\} e^{-\theta t_i^{\beta}} \theta t_i^{\beta}\log t_i\right)\right]}{\xi_i}$$

$$0 = \sum_{t_i\in\Omega_{11}} \frac{\varphi\left(\frac{\log t_i - \mu}{\sigma}\right)\left(\frac{1}{\sigma}\right)}{1-\Phi\left(\frac{\log t_i - \mu}{\sigma}\right)} + \sum_{t_i\in\Omega_{12}} \left(\frac{\log t_i - \mu}{\sigma}\right)\left(\frac{1}{\sigma}\right) + \sum_{t_i\in\Omega_3} \frac{\varphi\left(\frac{\log t_i - \mu}{\sigma}\right)\left(\frac{1}{\sigma}\right)}{1-\Phi\left(\frac{\log t_i - \mu}{\sigma}\right)}$$

$$+ \sum_{t_i\in\Omega_2} \frac{\left[\left\{\theta\beta t_i^{\beta-1} e^{-\theta t_i^{\beta}} \varphi\left(\frac{\log t_i - \mu}{\sigma}\right)\left(\frac{1}{\sigma}\right)\right\} + \left\{\frac{1}{\sqrt{2\pi}\;\;\sigma^2\, t_i} e^{-\theta t_i^{\beta}} \exp\left\{-\frac{1}{2\sigma^2}(\log t_i - \mu)^2\right\}\left(\frac{\log t_i - \mu}{\sigma}\right)\right\}\right]}{\xi_i}$$

$$0 = -\frac{r_{12}}{\sigma} + \sum_{t_i\in\Omega_{11}} \frac{\varphi\left(\frac{\log t_i - \mu}{\sigma}\right)\left(\frac{\log t_i - \mu}{\sigma^2}\right)}{1-\Phi\left(\frac{\log t_i - \mu}{\sigma}\right)} + \sum_{t_i\in\Omega_{12}} \left(\frac{\log t_i - \mu}{\sigma}\right)\left(\frac{\log t_i - \mu}{\sigma^2}\right)$$

$$+ \sum_{t_i\in\Omega_3} \frac{\varphi\left(\frac{\log t_i - \mu}{\sigma}\right)\left(\frac{\log t_i - \mu}{\sigma^2}\right)}{1-\Phi\left(\frac{\log t_i - \mu}{\sigma}\right)}$$

$$+\sum_{t_i\in\Omega_2}\frac{\left[\left\{\theta\beta t_i^{\beta-1}e^{-\theta t_i^{\beta}}\varphi\left(\frac{\log t_i-\mu}{\sigma}\right)\left(\frac{\log t_i-\mu}{\sigma^2}\right)\right\}+\left\{\frac{1}{\sqrt{2\pi}\ \ \sigma^2\ t_i}e^{-\theta t_i^{\beta}}\exp\left\{-\frac{1}{2\sigma^2}(\log t_i-\mu)^2\right\}\left(\left(\frac{\log t_i-\mu}{\sigma}\right)^2-1\right)\right\}\right]}{\xi_i}$$

(10-13)

As it seems, the system of non-linear equations (10-13) has no closed form solution in θ, β, μ and σ. So, a numerical method technique such as Newton-Raphson method is required for computing the MLEs of the parameters θ, β, μ and σ.

3.1.2Bayesian Estimation

In practice, it is observed that the life-testing experiments are very time consuming as such the parameters involved in the lifetime model cannot be remained static throughout the testing period. Therefore, it seems logical to treat the parameters as random variables instead of fixed constants. In lieu of this, the present sub-section proposes Bayesian estimation procedure by assuming the parameters of the Weibull and log normal distributions as random variables. The prior distributions of θ, β, μ and σ are considered as non informative:

$$g_1(\theta,\beta)=1 \quad ;(\theta,\beta)>0 \tag{14}$$

and

$$g_2(\mu,\sigma)=1 \quad ;-\infty<\mu<\infty,\sigma>0 \tag{15}$$

Using likelihood function in (8) and prior distributions in (14) and (15), the joint posterior distribution of θ, β, μ and σ can be written as

$$\prod(\theta,\beta,\mu,\sigma\,|\,\underset{\sim}{t}) = L(\underset{\sim}{t}\,|\,\theta,\beta,\mu,\sigma)\,g_1(\theta,\beta)\,g_2(\mu,\sigma)$$

$$=\theta^{r_{11}}\beta^{r_{11}}\left(\frac{1}{\sigma}\right)^{r_{12}}\exp\left\{-\sum_{t_i\in\Omega_1}t_i^{\beta}\left[\theta\, I(\delta_i=1)+\theta I(\delta_i=2)\right]\right\}\times\prod_{t_i\in\Omega_{11}}\left[t_i^{\beta-1}\left\{1-\Phi\left(\frac{\log t_i-\mu}{\sigma}\right)\right\}\right]$$

$$\times\prod_{t_i\in\Omega_{12}}\left[\frac{1}{t_i}\exp\left\{-\frac{1}{2\sigma^2}(\log t_i-\mu)^2\right\}\right]$$

$$\times\prod_{t_i\in\Omega_2}\left[\left\{\theta\beta t_i^{\beta-1}e^{-\theta t_i^{\beta}}\left(1-\Phi\left(\frac{\log t_i-\mu}{\sigma}\right)\right)\right\}+\left\{\frac{1}{\sqrt{2\pi}\ \ \sigma\, t_i}\exp\left\{-\frac{1}{2\sigma^2}(\log t_i-\mu)^2\right\}\times e^{-\theta t_i^{\beta}}\right\}\right]$$

$$\times\prod_{t_i\in\Omega_3}\left[e^{-\theta t_i^{\beta}}\left(1-\Phi\left(\frac{\log t_i-\mu}{\sigma}\right)\right)\right] \tag{16}$$

From (16), it is apparent that one cannot obtain the closed form solutions for Bayes estimates of the parameters. Therefore, for computing Bayes estimates of the competing risk parameters, the MCMC techniques such as Metropolis-Hastings and Gibbs sampling algorithms have been utilized. For implementing Gibbs sampling procedure, the full conditional posterior distributions of θ, β, μ and σ are given in the appendix *A1*.

3.2. Case II

In this case, we consider that cause-1 follows the Weibull distribution and cause-2 follows the exponential distribution. Based on the assumption 3(ii), for j=1,2 and i=1,2 ,………, N, the cumulative distribution functions of T_{1i} and T_{2i} are respectively given by

$$F_1(t) = 1 - e^{-\theta t^{\beta}} \quad ; t > 0 \tag{17}$$

$$F_2(t) = 1 - e^{-\lambda t} \quad ; t > 0 \tag{18}$$

Therefore, the probability density functions are

$$f_1(t) = \theta\beta t^{\beta-1} e^{-\theta t^{\beta}} \quad ; t > 0 \tag{19}$$

$$f_2(t) = \lambda\, e^{-\lambda t} \quad ; t > 0 \tag{20}$$

The survival functions are

$$\bar{F}_1(t) = e^{-\theta t^{\beta}} \tag{21}$$

$$\bar{F}_2(t) = e^{-\lambda t} \tag{22}$$

3.2.1. MLEs

Substituting (19)-(22) into (1), the likelihood functions becomes

$$L = \theta^{r_{11}} \beta^{r_{11}} \lambda^{r_{12}} \exp\left\{ - \sum_{t_i \in \Omega_1} t_i \left[\lambda\, I(\delta_i = 1) + \lambda I(\delta_i = 2) \right] \right\} \times \prod_{t_i \in \Omega_{11}} \left[t_i^{\beta-1} e^{\left\{-\theta t_i^{\beta}\right\}} \right]$$
$$\times \prod_{t_i \in \Omega_{12}} \left[\exp\left\{ -\theta t_i^{\beta} \right\} \right]$$
$$\times \prod_{t_i \in \Omega_2} \left[\left\{ \theta\beta t_i^{\beta-1} e^{-\theta t_i^{\beta}} e^{-\lambda t_i} \right\} + \left\{ \lambda e^{-\lambda t_i} \times e^{-\theta t_i^{\beta}} \right\} \right] \times \prod_{t_i \in \Omega_3} \left[e^{-\theta t_i^{\beta}} e^{-\lambda t_i} \right] \tag{23}$$

The log- likelihood function is given by

$$\log L = r_{11} \left(\log\theta + \log\beta \right) + r_{12} \log\lambda - \lambda \sum_{t_i \in \Omega_{11}} t_i - \lambda \sum_{t_i \in \Omega_{12}} t_i$$
$$+ \sum_{t_i \in \Omega_{11}} \left\{ (\beta - 1) \log t_i - \theta t_i^{\beta} \right\} - \sum_{t_i \in \Omega_{12}} \theta t_i^{\beta}$$
$$+ \sum_{t_i \in \Omega_2} \log \left[\left\{ \theta\beta t_i^{\beta-1} e^{-\theta t_i^{\beta}} e^{-\lambda t_i} \right\} + \left\{ \lambda e^{-\lambda t_i} \times e^{-\theta t_i^{\beta}} \right\} \right] + \sum_{t_i \in \Omega_3} \left(-\theta t_i^{\beta} - \lambda t_i \right) \tag{24}$$

Equating the first partial derivates of (24) with respect to θ, β and λ to zeros, we get the likelihood equations as given by (25-27), where

$$\psi_i = \left\{\theta\beta t_i^{\beta-1} e^{-\theta t_i^{\beta}} e^{-\lambda t_i}\right\} + \left\{\lambda e^{-\lambda t_i} \times e^{-\theta t_i^{\beta}}\right\}$$

$$0 = \frac{r_{11}}{\theta} - \sum_{t_i \in \Omega_{11}} t_i^{\beta} - \sum_{t_i \in \Omega_{12}} t_i^{\beta} + \sum_{t_i \in \Omega_2} \frac{\left[\left(\beta t_i^{\beta-1} e^{-\lambda t_i} e^{-\theta t_i^{\beta}} \left\{1-\theta t_i^{\beta}\right\}\right) - \left(\lambda e^{-\lambda t_i} e^{-\theta t_i^{\beta}} t_i^{\beta}\right)\right]}{\psi_i} - \sum_{t_i \in \Omega_3} t_i^{\beta}$$

$$0 = \frac{r_{11}}{\beta} - \theta \sum_{t_i \in \Omega_{11}} t_i^{\beta} \log t_i - \theta \sum_{t_i \in \Omega_{12}} t_i^{\beta} \log t_i + \sum_{t_i \in \Omega_{11}} \log t_i - \theta \sum_{t_i \in \Omega_3} \left(\theta t_i^{\beta} \log t_i\right)$$
$$+ \sum_{t_i \in \Omega_2} \frac{\left[\left(\theta e^{-\lambda t_i} t_i^{\beta-1} e^{-\theta t_i^{\beta}} \left\{(1+\beta \log t_i) - \left(\theta\beta t_i^{\beta} \log t_i\right)\right\}\right) - \left(\lambda e^{-\lambda t_i} e^{-\theta t_i^{\beta}} \theta t_i^{\beta} \log t_i\right)\right]}{\psi_i}$$

$$0 = \frac{r_{12}}{\lambda} - \sum_{t_i \in \Omega_{11}} t_i - \sum_{t_i \in \Omega_{12}} t_i + \sum_{t_i \in \Omega_2} \frac{\left[\left(-\theta\beta t_i^{\beta-1} e^{-\theta t_i^{\beta}} e^{-\lambda t_i}\right) + e^{-\theta t_i^{\beta}} e^{-\lambda t_i} (1-\lambda t_i)\right]}{\psi_i} - \sum_{t_i \in \Omega_3} t_i$$

(25-27)

The equations (25-27) can be solved using any suitable iterative procedure such as Newton-Raphson method to get the MLEs of the parameters θ, β and λ.

3.2.2Bayesian Estimation

For performing Bayesian estimation procedure, the prior distributions of θ, β and λ are again considered as non informative:

$$g_1(\theta, \beta) = 1 \quad ; (\theta, \beta) > 0 \tag{28}$$

and

$$g_2(\lambda) = 1 \quad ; \lambda > 0 \tag{29}$$

Using likelihood function in (23) and prior distributions in (28) and (29), the joint posterior distribution of θ, β and λ can be written as

$$\Pi(\theta,\beta,\lambda \mid \underset{\sim}{t}) = L(\underset{\sim}{t} \mid \theta,\beta,\lambda)\, g_1(\theta,\beta)\, g_2(\lambda)$$

$$= \theta^{r_{11}} \beta^{r_{11}} \lambda^{r_{12}} \exp\left\{-\sum_{t_i \in \Omega_1} t_i \left[\lambda I(\delta_i = 1) + \lambda I(\delta_i = 2)\right]\right\} \times \prod_{t_i \in \Omega_{11}} \left[t_i^{\beta-1} e^{\left\{-\theta t_i^{\beta}\right\}}\right] \times \prod_{t_i \in \Omega_{12}} \left[\exp\left\{-\theta t_i^{\beta}\right\}\right]$$

$$\times \prod_{t_i \in \Omega_2} \left[\left\{\theta\beta t_i^{\beta-1} e^{-\theta t_i^{\beta}} e^{-\lambda t_i}\right\} + \left\{\lambda e^{-\lambda t_i} \times e^{-\theta t_i^{\beta}}\right\}\right] \times \prod_{t_i \in \Omega_3} \left[e^{-\theta t_i^{\beta}} e^{-\lambda t_i}\right] \tag{30}$$

For implementing Gibbs sampling procedure, the full conditional posterior distributions of θ, β and λ are given in the appendix *A2*.

4. The Relative Risk Rates

4.1 Case I

Here, we derive the relative risk rates due the causes 1 and 2, when cause-1 follows the Weibull distribution and cause-2 follows the Log-normal distribution. The relative risk rate π_1, due to cause-1 is

$$\pi_1 = P\left[T_{1i} < T_{2i}\right]$$

$$= \int_0^\infty \left(1 - e^{-\theta t^\beta}\right) \frac{1}{\sqrt{2\pi} \quad \sigma t} \exp\left\{-\frac{1}{2\sigma^2}\left(\log t - \mu\right)^2\right\} dt$$

And the relative risk π_2, due to cause-2 is

$$\pi_2 = P\left[T_{2i} < T_{1i}\right]$$

$$=1- \pi_1$$

4.2 Case II

In this case, we derive the relative risk rates, when cause-1 follows the Weibull distribution and cause-2 follows the Exponential distribution. The relative risk rate π_1, due to cause-1 is

$$\pi_1 = P\left[T_{1i} < T_{2i}\right]$$

$$= \int_0^\infty \left(1 - e^{-\theta t^\beta}\right) \lambda e^{-\lambda t} dt$$

And the relative risk π_2, due to cause-2 is

$$\pi_2 = P\left[T_{2i} < T_{1i}\right]$$

$$=1- \pi_1$$

5. Real Data Analysis

In this section, a real data set from Boag (1949) is analyzed under the competing risks model with the following assumptions:

- Both the causes of failures follow exponential distributions.
- First cause of failures follow exponential distribution and second cause of failures follow Weibull distribution.
- First cause of failures follow Weibull distribution and second cause of failures follow exponential distribution.

- First cause of failures follow Weibull distribution and second cause of failures follow log-normal distribution.
- First cause of failures follow log-normal distribution and second cause of failures follow Weibull distribution.
- Both the causes of failures follow Weibull distributions.
- Both the causes of failures follow log-normal distributions.

Note that though we consider seven combinations of failure time distributions of causes however, the theoretical developments are provided only for two cases.

The data consists of survival times (in months) for 121 breast cancer patients. It comes from the clinical records of one hospital from the years 1929 to 1938. The causes of death are cancer (1) and others (2). Our aim is to test the goodness-of-fit of this data set to the suitable competing risks model. Further, we want to test whether the cancer occurs earlier compared to the other risks. In this data set, out of 121 breast cancer patients, total death due to cancer is observed to be as 78 and that due to others causes is 18 and 25 patients are survived till the end of the experiment. Here, it is to be noted that there is no observation whose cause of failure is not known.

Here, $N = 121,\ r_{11} = 78,\ r_{12} = 18,$ i.e. $r_1 = r_{11} + r_{12} = 96,\ r_2 = 0$ and $r_3 = 25$

Now, with the above information, we want to test which one of the pairs of the considered distributions to the different failures of causes are reasonable. For this, we compute the MLEs and Bayes estimates of the unknown parameters and compare Kolmogrov-Simrnov (K-S) distances under the considered competing risks models, which are summarized in Table 1. To see the goodness-of-fit of the data with the considered competing risk models, the fitted survival functions (with ML and Bayes methods)and empirical survival function have also been plotted [Fig-1-7].The corresponding relative risks are also estimated with both classical and Bayesian methods and the same are listed in Table 1. Note that in Table 1, KS-1 and KS-2 respectively stand for Kolmogrov-Simrnov distances computed with MLEs and Bayes estimates. For numerical computations, the programs are developed in R-software and are available with the authors.

6. Conclusion

From Table-1, it is observed that the relative risk due to cancer is

- 81.25 (with ML method) and 81.89 (with Bayes method) when both the causes follow exponential distribution.
- 80.47 (with ML method) and 78.90 (with Bayes method) when cause-1 follow

exponential distribution and cause-2 follows the Weibull distribution.

- 80.23 (with ML method) and 79.72 (with Bayes method) when cause-1 follow Weibull distribution and cause-2 follows the exponential distribution.
- 55.65 (with ML method) and 54.51 (with Bayes method) when cause-1 follow Weibull distribution and cause-2 follows the log normal distribution.
- 90.25 (with ML method) and 88.25 (with Bayes method) when cause-1 follow log normal distribution and cause-2 follows the Weibull distribution.
- 79.11 (with ML method) and 77.87 (with Bayes method) when cause-1 follow Weibull distribution and cause-2 follows the Weibull distribution.
- 35.25 (with ML method) and 31.46 (with Bayes method) when cause-1 follow log normal distribution and cause-2 follows the log normal distribution.

The values of the K-S distances suggest that three competing risks models as

- Weibull-Exponential
- Weibull-log normal
- Weibull-Weibull

are more suitable models to the considered data set as they are having least distances. However, among these three, Weibull-Exponential competing risks model is observed to be best fitted model for analyzing the given data set in the presence of the two causes of failures (cancer and others). The plots of the fitted survival functions are also provided the same evidences. Further, from Table-1, it is observed that in all the cases (excluding log normal-log normal competing risk model), the relative risks due to cause-1(cancer) are higher than the relative risks due to cause-2(others). Thus, we can conclude that cancer is the cause which occurs earlier to the other risks.

Acknowledgment

The authors thankfully acknowledge the critical suggestions and comments from the learned referee which greatly helped us in the improvement of the paper.

Appendix

A1. *The full conditional posterior distributions of θ, β, μ and σ are given as follows*

$$\pi_1(\theta\,|\,\underset{\sim}{t},\beta,\mu,\sigma) \propto \theta^{r_{11}}\exp\left\{-\sum_{t_i\in\Omega_1} t_i^{\beta}\left[\theta I(\delta_i=1)+\theta I(\delta_i=2)\right]\right\}$$

$$\times \prod_{t_i\in\Omega_2}\left[\left\{\theta\beta t_i^{\beta-1}e^{-\theta t_i^{\beta}}\left(1-\Phi\left(\frac{\log t_i-\mu}{\sigma}\right)\right)\right\}+\left\{\frac{1}{\sqrt{2\pi}\;\;\sigma\, t_i}\exp\left\{-\frac{1}{2\sigma^2}(\log t_i-\mu)^2\right\}\times e^{-\theta t_i^{\beta}}\right\}\right]\times\prod_{t_i\in\Omega_3}\left[e^{-\theta t_i^{\beta}}\right]$$

$$\pi_2(\beta\,|\,\underset{\sim}{t},\theta,\mu,\sigma) \propto \beta^{r_{11}}\exp\left\{-\sum_{t_i\in\Omega_1} t_i^{\beta}\left[\theta\, I(\delta_i=1)+\theta I(\delta_i=2)\right]\right\}\times\prod_{t_i\in\Omega_{11}} t_i^{\beta-1}$$

$$\times \prod_{t_i\in\Omega_2}\left[\left\{\theta\beta t_i^{\beta-1}e^{-\theta t_i^{\beta}}\left(1-\Phi\left(\frac{\log t_i-\mu}{\sigma}\right)\right)\right\}+\left\{\frac{1}{\sqrt{2\pi}\;\;\sigma\, t_i}\exp\left\{-\frac{1}{2\sigma^2}(\log t_i-\mu)^2\right\}\times e^{-\theta t_i^{\beta}}\right\}\right]\times\prod_{t_i\in\Omega_3}\left\{e^{-\theta t_i^{\beta}}\right\}$$

$$\pi_3(\mu\,|\,\underset{\sim}{t},\theta,\beta,\sigma) \propto \prod_{t_i\in\Omega_{11}}\left\{1-\Phi\left(\frac{\log t_i-\mu}{\sigma}\right)\right\}\times\prod_{t_i\in\Omega_{12}}\left[\frac{1}{t_i}\exp\left\{-\frac{1}{2\sigma^2}(\log t_i-\mu)^2\right\}\right]\times\prod_{t_i\in\Omega_3}\left(1-\Phi\left(\frac{\log t_i-\mu}{\sigma}\right)\right)$$

$$\times \prod_{t_i\in\Omega_2}\left[\left\{\theta\beta t_i^{\beta-1}e^{-\theta t_i^{\beta}}\left(1-\Phi\left(\frac{\log t_i-\mu}{\sigma}\right)\right)\right\}+\left\{\frac{1}{\sqrt{2\pi}\;\;\sigma\, t_i}\exp\left\{-\frac{1}{2\sigma^2}(\log t_i-\mu)^2\right\}\times e^{-\theta t_i^{\beta}}\right\}\right]$$

$$\pi_4(\sigma\,|\,\underset{\sim}{t},\theta,\beta,\mu) \propto \left(\frac{1}{\sigma}\right)^{r_{12}}\times\prod_{t_i\in\Omega_{11}}\left\{1-\Phi\left(\frac{\log t_i-\mu}{\sigma}\right)\right\}\times\prod_{t_i\in\Omega_{12}}\left[\frac{1}{t_i}\exp\left\{-\frac{1}{2\sigma^2}(\log t_i-\mu)^2\right\}\right]$$

$$\times \prod_{t_i\in\Omega_2}\left[\left\{\theta\beta t_i^{\beta-1}e^{-\theta t_i^{\beta}}\left(1-\Phi\left(\frac{\log t_i-\mu}{\sigma}\right)\right)\right\}+\left\{\frac{1}{\sqrt{2\pi}\;\;\sigma\, t_i}\exp\left\{-\frac{1}{2\sigma^2}(\log t_i-\mu)^2\right\}\times e^{-\theta t_i^{\beta}}\right\}\right]$$

$$\times \prod_{t_i\in\Omega_3}\left(1-\Phi\left(\frac{\log t_i-\mu}{\sigma}\right)\right)$$

A2. *The full conditional posterior distributions θ,* β and λ *are given as follows*

$$\pi_1(\theta\,|\,\underset{\sim}{t},\beta,\lambda) \propto \theta^{r_{11}}\prod_{t_i\in\Omega_{11}} e^{-\theta t_i^{\beta}}\times\prod_{t_i\in\Omega_{12}} e^{-\theta t_i^{\beta}}\times\prod_{t_i\in\Omega_2}\left[\left\{\theta\beta t_i^{\beta-1}e^{-\theta t_i^{\beta}}e^{-\lambda t_i}\right\}+\left\{\lambda e^{-\lambda t_i}\times e^{-\theta t_i^{\beta}}\right\}\right]\times\prod_{t_i\in\Omega_3} e^{-\theta t_i^{\beta}}$$

$$\pi_2(\beta\,|\,\underset{\sim}{t},\theta,\lambda) \propto \beta^{r_{11}}\times\prod_{t_i\in\Omega_{11}}\left[t_i^{\beta-1}e^{-\theta t_i^{\beta}}\right]\times\prod_{t_i\in\Omega_{12}}\left[\exp\left\{-\theta t_i^{\beta}\right\}\right]\times\prod_{t_i\in\Omega_2}\left[\left\{\theta\beta t_i^{\beta-1}e^{-\theta t_i^{\beta}}e^{-\lambda t_i}\right\}+\left\{\lambda e^{-\lambda t_i}\times e^{-\theta t_i^{\beta}}\right\}\right]\times\prod_{t_i\in\Omega_3}\left[e^{-\theta t_i^{\beta}}\right]$$

$$\pi_3(\lambda\,|\,\underset{\sim}{t},\theta,\beta) \propto \lambda^{r_{12}}\exp\left\{-\sum_{t_i\in\Omega_1} t_i\left[\lambda\, I(\delta_i=1)+\lambda I(\delta_i=2)\right]\right\}\times\prod_{t_i\in\Omega_2}\left[\left\{\theta\beta t_i^{\beta-1}e^{-\theta t_i^{\beta}}e^{-\lambda t_i}\right\}+\left\{\lambda e^{-\lambda t_i}\times e^{-\theta t_i^{\beta}}\right\}\right]\times\prod_{t_i\in\Omega_3}\left[e^{-\lambda t_i}\right]$$

Table-1:The MLEs, Bayes Estimates, RelativeRisk and K-S distances

Family of Distribution	MLE	BAYES	RELATIVE RISK	K-S
Cause 1~Exponential (θ_1) Cause 2~Exponential (θ_2)	$\theta_1 = 0.1174$ $\theta_2 = 0.0271$	$\theta_1^* = 0.1200$ $\theta_2^* = 0.0283$	$\pi_{12} = 0.8125$ $\pi_{21} = 0.1875$ $\pi_{12}^* = 0.8089$ $\pi_{21}^* = 0.1911$	KS-1= 0.1605 KS-2= 0.1558
Cause 1~exponential (λ) Cause 2~ Weibull (θ,β)	$\lambda = 0.1174$ $\hat{\theta} = 0.0181$ $\hat{\beta} = 1.1901$	$\lambda^* = 0.1189$ $\theta^* = 0.0244$ $\beta^* = 1.1128$	$\pi_{12} = 0.8047$ $\pi_{21} = 0.1953$ $\pi_{12}^* = 0.7890$ $\pi_{21}^* = 0.2110$	KS-1= 0.1696 KS-2= 0.1604
Cause 1~Weibull (θ,β) Cause 2~exponential (λ)	$\hat{\theta} = 0.1616$ $\hat{\beta} = 0.8422$ $\lambda = 0.0271$	$\theta^* = 0.1683$ $\beta^* = 0.8341$ $\lambda^* = 0.0287$	$\pi_{12} = 0.8023$ $\pi_{21} = 0.1977$ $\pi_{12}^* = 0.7972$ $\pi_{21}^* = 0.2028$	KS-1= 0.1202 KS-2= 0.1118
Cause 1~Weibull (θ,β) Cause 2~Log-normal (μ,σ)	$\hat{\theta} = 0.1616$ $\hat{\beta} = 0.8422$ $\mu = 1.8839$ $\sigma = 1.1736$	$\theta^* = 0.1674$ $\beta^* = 0.8355$ $\mu^* = 1.8133$ $\sigma^* = 0.9023$	$\pi_{12} = 0.5565$ $\pi_{21} = 0.4435$ $\pi_{12}^* = 0.5451$ $\pi_{21}^* = 0.4549$	KS-1= 0.1450 KS-2= 0.1449
Cause 1~Log-normal (μ,σ) Cause2~Weibull (θ,β)	$\mu = 0.9515$ $\sigma = 1.0198$ $\hat{\theta} = 0.0181$ $\hat{\beta} = 1.1902$	$\mu^* = 1.0824$ $\sigma^* = 0.9297$ $\theta^* = 0.0249$ $\beta^* = 1.0977$	$\pi_{12} = 0.9025$ $\pi_{21} = 0.0975$ $\pi_{12}^* = 0.8825$ $\pi_{21}^* = 0.1174$	KS-1= 0.2929 KS-2= 0.3014
Cause 1~Weibull (θ_1,β_1) Cause2~Weibull (θ_2,β_2)	$\hat{\theta}_1 = 0.1616$ $\hat{\beta}_1 = 0.8422$ $\hat{\theta}_2 = 0.0181$ $\hat{\beta}_2 = 1.1902$	$\theta_1^* = 0.1718$ $\beta_1^* = 0.8223$ $\theta_2^* = 0.0240$ $\beta_2^* = 1.1163$	$\pi_{12} = 0.7911$ $\pi_{21} = 0.2089$ $\pi_{12}^* = 0.7787$ $\pi_{21}^* = 0.2213$	KS-1= 0.1289 KS-2= 0.1129
Cause1~Log-normal (μ_1,σ_1) Cause2~Log-normal (μ_2,σ_2)	$\mu_1 = 0.9515$ $\sigma_1 = 1.0198$ $\mu_2 = 1.8839$ $\sigma_2 = 1.1736$	$\mu_1^* = 1.0867$ $\sigma_1^* = 0.9268$ $\mu_2^* = 1.9566$ $\sigma_2^* = 0.9864$	$\pi_{12} = 0.3525$ $\pi_{21} = 0.6475$ $\pi_{12}^* = 0.3146$ $\pi_{21}^* = 0.6854$	KS-1= 0.3123 KS-2= 0.3376

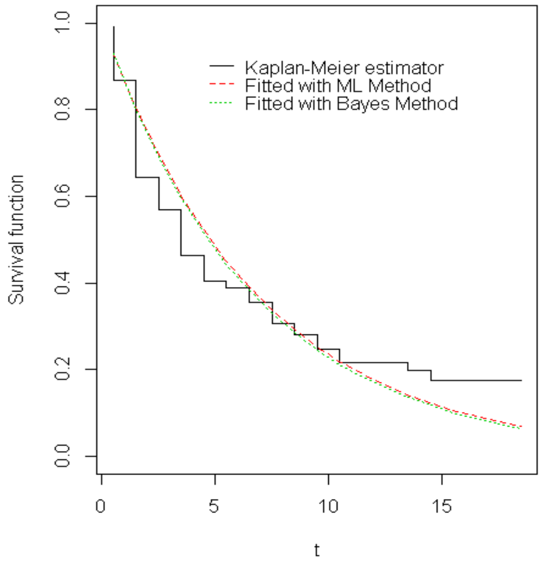

Fig-1: Fitted Survival Function when
Cause 1~Exponential and Cause2~Exponential

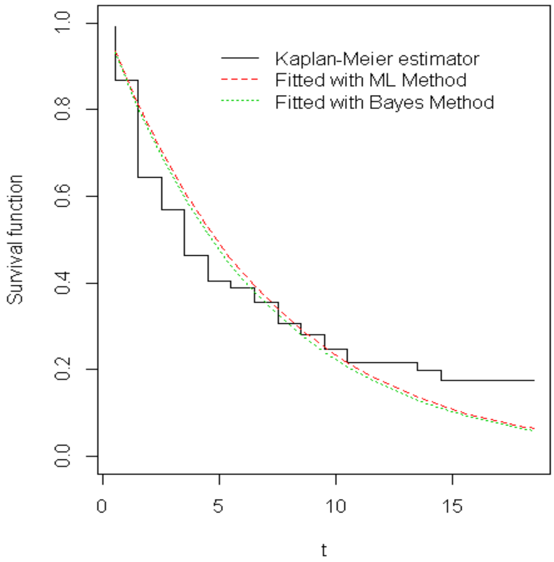

Fig-2: Fitted Survival Function when
Cause 1~Exponential and Cause2~Weibull

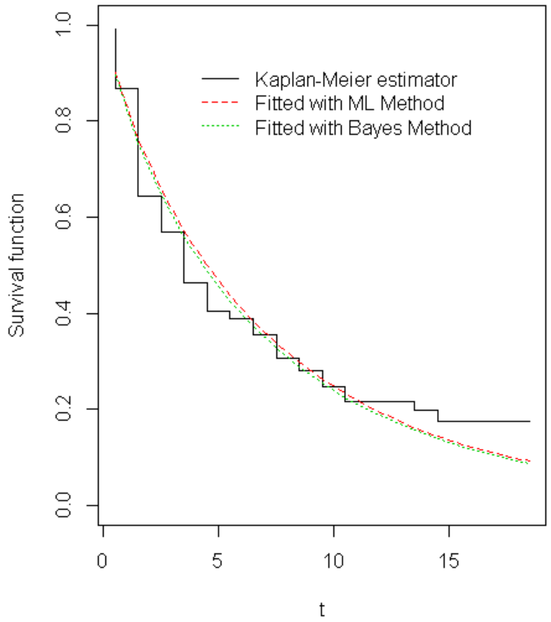

Fig-3: Fitted Survival Function when
Cause 1~Weibull and Cause2~Exponential

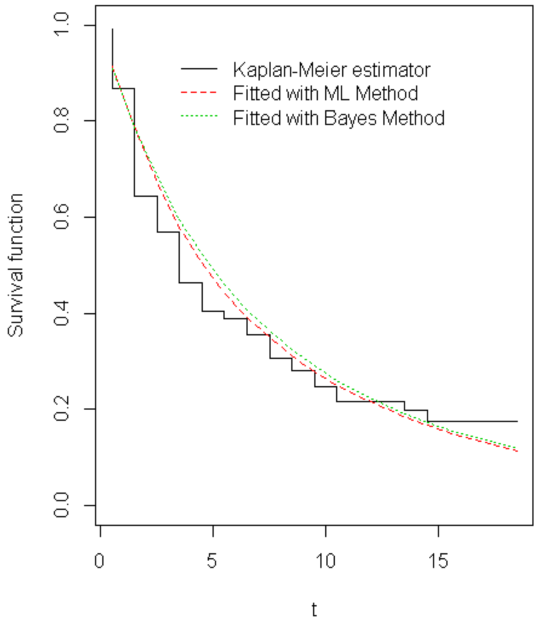

Fig-4: Fitted Survival Function when
Cause 1~Weibull and Cause2~Log-normal

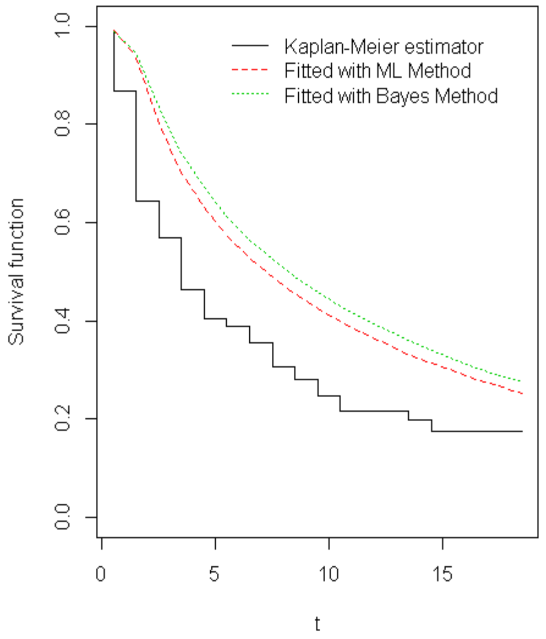

Fig-5: Fitted Survival Function when
Cause 1~Log-normal and Cause2~Weibull

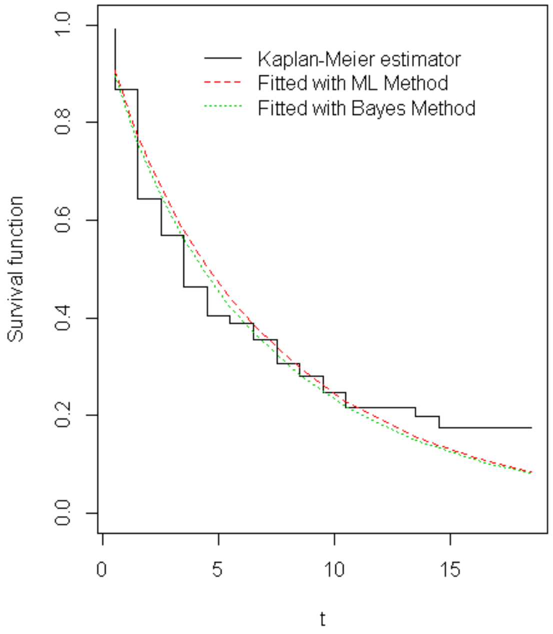

Fig-6: Fitted Survival Function when
Cause 1~Weibull and Cause2~Weibull

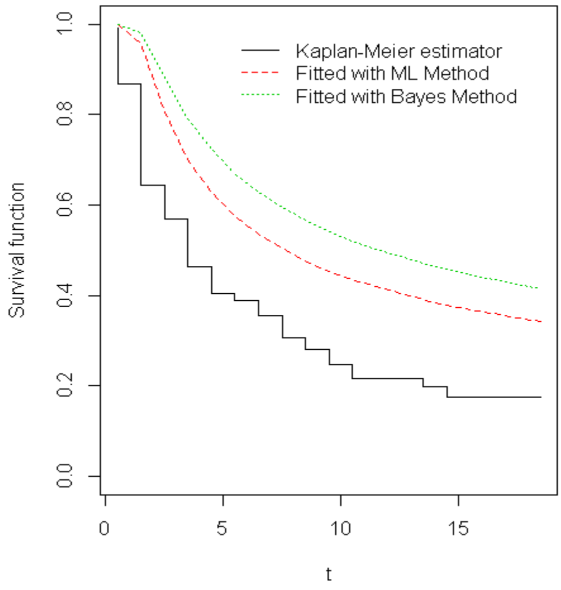

Fig-7: Fitted Survival Function when
Cause 1~Log-normal and Cause2~Log-normal

References

[1] Tsiatis, A. (1975), "A non-identifiability aspect in the problem of competing risks",Proceedings of National Academy of Science, USA 72, 20-22.

[2] Alwasel, I. A. (2009), "Parameter estimation for the modified Weibull Distributionbased on grouped and censored data". International Journal ofApplied Mathematics &Statistics, Vol. 12, 78-87.

[3] Berkson, J. and Elveback, L. (1960), "Competing exponential risks with particular inference to the study of smoking lung cancer", Journal of American Statistical Association, 55, 415-428.

[4] Boag's. J. W. (1949), "Maximum Likelihood Estimates of the Proportion of Patients Cured by Cancer Therapy", Journal of the Royal Statistical Society. Series B(Methodological),Vol. 11, No. 1 , pp. 15-53.

[5] Crowder, M.J. (2001), "Classical competing risks". Chapman & Hall.

[6] Crowder, M.J. (1993), " Identifiability crisis in competing risks analysis", Int. Statist. Rev., 62, 379-391.

[7] Crowder, M.J., Kimber, A.C., Smith, L.L. and Sweeting, T.J. (1991), "Statistical Analysis of Reliability Data", Chapman and Hall.

[9] D. Bernoulli, R_eflexions sur les avantages de l'inoculation, Mercure deFrance (1760) 173. (Reprinted in L.P.Bouckaert, B.L. van der Waerden (Eds), Die Werke von Daniel Bernoulli, Bd. 2 Analysis und Wahrscheinlich-keitsrechnung, Birkhaqauser, Basel, 1982, p. 268).

[10] David, H.A. and Moeschberger, M.L. (1978), "The Theory of Competing Risks", Griffin, London.

[11] Dinse, G.E. (1982), "Non-parametric estimation of partially incomplete time and types of failure data", Biometrics 38, 417-431.

[12] Kaplan, E. L. and Meier, P.(1958),"On the identifiability crisis in competing risks analysis"Jourmal of American Statistical Association53, 457-481.

[13] Kundu, D. and Basu, S. (2002), "Analysis of incomplete sata in the presenceof competing risks", Journal of Statistical Planning and Inference, 87,221-239.

[14] Miyakawa, M. (1982), "Statistical analysis of incomplete data in competing risks model", Journal of Japanese Soc. Quality Control 12, 49-52.

[15] Miyakawa, M. (1984), "Analysis of incomplete data in competing risks model", IEEE Trans. Reliability Anal. 33 (4), 293-296.

[16] Peterson Jr., A.P. (1977), "Expressing the Kaplan-Meier estimator as a function of empirical survival functions, Journal of American Statistical Association", 72, 854-858.

[17] Sarhan, A.M. (2007), "Analysis of incomplete, censored data in competing risks models with generalized exponential distributions", IEEE Transactions on Reliability, 56(1), 132-138.