Biometrical Journal **46** (2004) 2, 165–179 DOI: 10.1002/bimj.200210014

Parameter Estimation for Partially Complete Time and Type of Failure Data

Debasis Kundu*

Department of Mathematics, Indian Institute of Technology Kanpur, Pin 208016, INDIA

Received August 2002, revised July 203, accepted December 2003

Summary

The theory of competing risks has been developed to asses a specific risk in presence of other risk factors. In this paper we consider the parametric estimation of different failure modes under partially complete time and type of failure data using latent failure times and cause specific hazard functions models. Uniformly minimum variance unbiased estimators and maximum likelihood estimators are obtained when latent failure times and cause specific hazard functions are exponentially distributed. We also consider the case when they follow Weibull distributions. One data set is used to illustrate the proposed techniques.

Key words: Competing risks; Incomplete data; Failure distribution; Exponential distribution; Maximum likelihood estimators; Minimum variance unbiased estimator.

1. Introduction

In medical studies an investigator is often concerned with the assessment of a specific risk in presence of other risk factors. In statistical literature it is well known as the competing risks model. In analyzing the competing risks model, ideally data consist of a failure time and an indicator denoting the failure type. But in many situations data may be incomplete, both in failure time and failure type. Moreover, under some circumstances, the failure type can be determined even when the failure time is censored. For example, consider an analysis of prostatic cancer in older man. Since, prostatic cancer is not rapidly lethal and can be diagnosed long before death, a man known to have the prostatic cancer who is alive at the time of analysis contributes complete information on the failure type, but incomplete in failure time. On the other hand, without an autopsy, the time of death provides a complete information on the failure time but no information on the failure type.

Note that in presence of complete information, several studies have been carried out for the last two decades both under the parametric and non-parametric set up. Dinse [3] considered the non-parametric estimation of the survival function of a particular cause for incomplete time and type of failure data. Because of the non-parametric formulation, it was difficult to consider estimations and construction of confidence intervals of several other parameters of interests. Here we adopt the parametric formulation of the problem and consider statistical inference of several important parameters.

We approach the problem in two different ways. First we use the latent failure times model, as suggested by Cox [1]. In the latent failure times model, it is assumed that the competing risks are independent. Secondly, we use the cause specific hazard functions model as suggested by Prentice et al. [9], where the distributions of the competing causes may not be independent. It is assumed that the latent failure time distributions can be either exponential or Weibull. Similarly, under the cause specific hazard functions model, it is assumed that the cause specific hazard functions also can be either exponential or Weibull. Interestingly, in both the formulations, likelihood functions of the observed data are same in each case. Therefore, the estimation procedures of the different unknown

* Corresponding author: e-mail: kundu@iitk.ac.in

parameters and their statistical properties remain unchanged, although the interpretations of the different parameters might be different.

In this paper, we make similar assumptions as of Dinse [3]. It is assumed that every member of a certain target population either dies of a particular disease, say cancer or by other causes. A proportion π of the population die of cancer and the proportion $(1-\pi)$ die due to other causes. Suppose that a subject can experience one of J failure types and let us define X be the time of failure and let δ be the failure type. At the end of the study, we have the following types of observations:

$$1.\ \{X = t, \delta = j\} \qquad 2.\ \{X > t\} \qquad 3.\ \{X > t, \delta = j\} \qquad 4.\ \{X = t\}\,.$$

It may be mentioned that Miyakawa [8] and recently Kundu and Basu [5] considered a similar problem in presence of only type 1 and type 4 observations.

The rest of the paper is organized as follows. In Section 2, we provide the formulation in terms of the latent failure times model and describe different notation, which we are going to use throughout this paper. In Section 3, we consider the estimation of different parameters when the latent failure time distributions are exponential. The corresponding confidence intervals are considered in Section 4. The Weibull latent failure distributions are considered in Section 5. One data set is analyzed in Section 6 for illustration purposes. In Section 7, we formulate the problem using the cause specific hazard functions model and finally we draw conclusions from out work in Section 8.

2. Model Description and Notation

Without loss of generality, we assume that there are only two failures types. We use the following notation;

X_i: lifetime of system i
X_{ji}: lifetime of mode j (= 1 or 2) of system i
$F(.)$: cumulative distribution function of X_i
$f(.)$: density function of X_i
$F_j(.)$: distribution function of X_{ji}
$f_j(.)$: density function of X_{ji}
$\bar{F}_j(.)$: $= 1 - F_j(.)$
δ_i: indicator variable denoting cause of failure of system i
$I[.]$: indicator function of event [.]
Exp (λ): denotes exponential random variable with density function $\lambda\, e^{-\lambda x}$
Gamma (α, λ): denotes gamma random variable with density function $\dfrac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1}\, e^{-\lambda x}$
Weibull (α, λ): denotes Weibull random variable with density function $\alpha\lambda\, x^{\alpha-1}\, e^{-\lambda x^\alpha}$
bin (n, p): denotes binomial random variable with parameters n and p.

In this Section we formulate the problem using the latent failure times model of Cox [1]. It is assumed that (X_{1i}, X_{2i}); $i = 1, 2 \ldots N$ are N independent and identically distributed (i.i.d.) random variables. It is also assumed that X_{1i} and X_{2i} are independent for all $i = 1, 2, \ldots N$ and $X_i = \min\{X_{1i}, X_{2i}\}$. Without loss of generality we assume that there are only two failure types, namely 1 and 2. We assume that we have following observations. The first r_1 observations have complete failure times and corresponding cause of failure is 1 for all of them. We denote this set as I_1, i.e., the observations are of the type (x_i, δ_i); $\delta_i = 1$ for all $i \in I_1$ and the number of elements in I_1, $|I_1|$, is r_1. Similarly the next r_2 observations have complete failure times and the corresponding cause of failure is 2 for all of them. We denote this set as I_2. The next r_3 observations have complete failure times but corresponding causes of failure are unknown and we denote this set as I_3. For the next r_4 observations we know that the failure type will be 1 but failure times have been right censored. We denote this set as I_4. Similarly for the next r_5 observations we have failure type 2, but corresponding failure times are right censored. Finally we have last r_6 right censored observations, where exact failure time and failure type both are unknown and we denote this set as I_6. Therefore, in summary we have following types of observations:

$$\begin{array}{llll} a. & (x_i, 1); & i \in I_1, & |I_1| = r_1, \\ b. & (x_i, 2); & i \in I_2, & |I_2| = r_2, \\ c. & (x_i, *); & i \in I_3, & |I_3| = r_3, \\ d. & (x_i*, 1); & i \in I_4, & |I_4| = r_4, \\ e. & (x_i*, 2); & i \in I_5, & |I_5| = r_5, \\ f. & (x_i*, *); & i \in I_6, & |I_6| = r_6\,. \end{array} \tag{1}$$

We denote $r_1 + r_2 + r_3 = n$, $r_4 + r_5 + r_6 = m$ and therefore, $m + n = N$. In order to analyze incomplete data it is assumed that failure times are from the same population as the complete data. We also assume throughout that $r_1 + r_2 = n_1$, r_3, $r_4 + r_5 = m_1$, r_6 are fixed and not random and $n > 0$. If they are assumed to be random, then all the results are valid conditionally, see Kundu and Basu [5] for details. The likelihood contributions from the observations a, b, c, d, e and f are

$$f_1(x)\,\bar{F}_2(x)\,, \qquad f_2(x)\,\bar{F}_1(x)\,, \qquad f(x)\,, \qquad \int_x^\infty f_1(y)\,\bar{F}_2(y)\,\mathrm{d}y\,, \qquad \int_x^\infty f_2(y)\,\bar{F}_1(y)\,\mathrm{d}y\,, \qquad \bar{F}(x),$$

respectively.

3. Exponential Latent Failure Time Distributions: Point Estimation

In this Section, we assume that X_{ji}'s are exponential random variables with parameters λ_j for $j = 1$ and 2 and for $i = 1, \ldots, N$. The distribution function, $F_j(t)$, of X_{ji} has the following form:

$$F_j(t) = (1 - \mathrm{e}^{-\lambda_j t})\,; \quad j = 1, 2\,.$$

In this case the likelihood function of the observed data (1) is

$$\begin{aligned} L &= \lambda_1^{r_1} \exp\left(-(\lambda_1 + \lambda_2) \sum_{i \in I_1} x_i\right) \times \lambda_2^{r_2} \exp\left(-(\lambda_1 + \lambda_2) \sum_{i \in I_2} x_i\right) \\ &\quad \times (\lambda_1 + \lambda_2)^{r_3} \exp\left(-(\lambda_1 + \lambda_2) \sum_{i \in I_3} x_i\right) \times \left(\frac{\lambda_1}{\lambda_1 + \lambda_2}\right)^{r_4} \exp\left(-(\lambda_1 + \lambda_2) \sum_{i \in I_4} x_i\right) \\ &\quad \times \left(\frac{\lambda_2}{\lambda_1 + \lambda_2}\right)^{r_5} \exp\left(-(\lambda_1 + \lambda_2) \sum_{i \in I_5} x_i\right) \times \exp\left(-(\lambda_1 + \lambda_2) \sum_{i \in I_6} x_i\right) \\ &= \lambda_1^{r_1 + r_4} \times \lambda_2^{r_2 + r_5} \times (\lambda_1 + \lambda_2)^{r_3 - r_4 - r_5} \times \exp\left(-(\lambda_1 + \lambda_2) \sum_{i \in I} x_i\right). \end{aligned} \tag{2}$$

Here $I = I_1 \cup I_2 \cup \ldots \cup I_6$. From the likelihood function (2), we obtain MLEs of λ_1 and λ_2 as

$$\hat{\lambda}_1 = \frac{n(r_1 + r_4)}{(n_1 + m_1)\left(\sum_{i \in I} x_i\right)} \quad \text{and} \quad \hat{\lambda}_2 = \frac{n(r_2 + r_4)}{(n_1 + m_1)\left(\sum_{i \in I} x_i\right)}\,.$$

Now to compute mean and variances of $\hat{\lambda}_1$ and $\hat{\lambda}_2$, we observe the following

$$r_1 \sim \text{bin}\left(n_1, \frac{\lambda_1}{\lambda_1 + \lambda_2}\right), \qquad r_2 \sim \text{bin}\left(n_1, \frac{\lambda_2}{\lambda_1 + \lambda_2}\right), \tag{3}$$

$$r_4 \sim \text{bin}\left(m_1, \frac{\lambda_1}{\lambda_1 + \lambda_2}\right), \qquad r_5 \sim \text{bin}\left(m_1, \frac{\lambda_2}{\lambda_1 + \lambda_2}\right), \tag{4}$$

$$\sum_{i=1}^{n+m} X_i \sim \Gamma(n + m,\ \lambda_1 + \lambda_2)\,. \tag{5}$$

Note that using (3)–(5), the exact distributions of $\hat{\lambda}_1$ and $\hat{\lambda}_2$ can be easily calculated. Using (3)–(5), we obtain

$$E(\hat{\lambda}_1) = \frac{n}{n+m-1}\lambda_1\,, \qquad V(\hat{\lambda}_1) = \frac{n^2\lambda_1}{(n+m-1)\,(n+m-2)}\left(\frac{\lambda_1}{n+m-1} + \frac{\lambda_2}{n_1+m_1}\right),$$
$$E(\hat{\lambda}_2) = \frac{n}{n+m-1}\,\lambda_2\,, \qquad V(\hat{\lambda}_2) = \frac{n^2\lambda_2}{(n+m-1)\,(n+m-2)}\left(\frac{\lambda_2}{n+m-1} + \frac{\lambda_1}{n_1+m_1}\right). \tag{6}$$

From (6), it is immediate that

$$\tilde{\lambda}_1 = \frac{(r_1+r_4)\,(n+m-1)}{(n_1+m_1)\left(\sum\limits_{i\in I} x_i\right)} \quad \text{and} \quad \tilde{\lambda}_2 = \frac{(r_2+r_5)\,(n+m-1)}{(n_1+m_1)\left(\sum\limits_{i\in I} x_i\right)}$$

are uniformly minimum variance unbiased estimators (UMVUEs) of λ_1 and λ_2 respectively. Note that these results match with those of Miyakawa [8], when $r_4 = r_5 = r_6 = 0$.

The relative risk rate, π, due to cause 1 is

$$P[X_{1i} < X_{2i}] = \int\limits_0^\infty \lambda_1 \exp\left(-(\lambda_1+\lambda_2)\,x\right) \mathrm{d}x = \frac{\lambda_1}{\lambda_1+\lambda_2}\,.$$

Observe that $\frac{r_1}{n_1}$ and $\frac{r_4}{m_1}$ are unbiased estimators of π, but the MLE of π, $\hat{\pi}$, is

$$\hat{\pi} = \frac{\hat{\lambda}_1}{\hat{\lambda}_1+\hat{\lambda}_2} = \frac{r_1+r_4}{n_1+m_1}\,.$$

Note that the distribution $\hat{\pi}$ is a scaled version of a Binomial distribution and $\hat{\pi}$ is an unbiased estimator of π. Now, we consider the problem of estimating the mean lifetimes $\theta_1 = \frac{1}{\lambda_1}$ and $\theta_2 = \frac{1}{\lambda_2}$ due to cause 1 and cause 2 respectively. Note that although MLEs and UMVUEs of λ_1 and λ_2 always exist but the same is not true for θ_1 and θ_2. UMVUEs of θ_1 and θ_2 do not exist and MLEs of θ_1 and θ_2 exist when $r_1 + r_4 > 0$ and $r_2 + r_5 > 0$ respectively. First we give the results for θ_1. The conditional MLE (conditioning that $r_1 + r_4 > 0$) of θ_1, say $\hat{\theta}_1$, is as follows:

$$\hat{\theta}_1 = \sum_{i\in I} x_i \times \frac{n_1+m_1}{n(r_1+r_4)}\,.$$

Now, we obtain the conditional distribution of $\hat{\theta}_1$, conditioning on $r_1 + r_4 > 0$, which in turn helps us in constructing confidence interval θ_1. We need the following lemma for further development.

Lemma 1: The conditional moment generating function (mgf) of $\hat{\theta}_1$, say $\phi_{\hat{\theta}_1}(t)$, is of the following form:

$$\begin{aligned}\phi_{\hat{\theta}_1}(t) &= E[\exp(t\hat{\theta}_1) \mid r_1+r_4>0] \\ &= \left(1-q^{n_1+m_1}\right)^{-1}\left[\sum_{i=1}^{n_1+m_1} \frac{(n_1+m_1)!}{i!(n_1+m_1-i)!}\left(\frac{\theta_2}{\theta_1+\theta_2}\right)^i \left(\frac{\theta_1}{\theta_1+\theta_2}\right)^{n_1+m_1-i}\right. \\ &\quad \left.\times\left(1-\frac{t(n_1+m_1)}{ni}\,\frac{\theta_1\theta_2}{\theta_1+\theta_2}\right)^{-(n+m)}\right] = \sum_{i=1}^{n_1+m_1} p_i\left(1-\frac{t(n_1+m_1)}{ni}\,\frac{\theta_1\theta_2}{\theta_1+\theta_2}\right)^{-(n+m)}.\end{aligned}$$

Here $q = \frac{\theta_1}{\theta_1+\theta_2}$ and

$$p_i = (1-q^{n_1+m_1})^{-1}\,\frac{(n_1+m_1)!}{i!(n_1+m_1-i)!}\left(\frac{\theta_2}{\theta_1+\theta_2}\right)^i\left(\frac{\theta_1}{\theta_1+\theta_2}\right)^{n_1+m_1-i}.$$

Proof of lemma 1: Note that $\sum_{i=1}^{n+m} X_i$ is Gamma $\left(n+m, \frac{\theta_1+\theta_2}{\theta_1\theta_2}\right)$ and r_1+r_4 is bin $\left(n_1+m_1, \frac{\lambda_1}{\lambda_1+\lambda_2}\right)$ random variables. Therefore,

$$\begin{aligned}\phi_{\hat{\theta}_1}(t) &= E[\exp(t\hat{\theta}_1) \mid r_1+r_4>0] \\ &= \sum_{i=1}^{n_1+m_1} E[\exp(t\hat{\theta}_1) \mid r_1+r_4=i] \times P[r_1+r_4=i \mid r_1+r_4>0]\,.\end{aligned}$$

Note that, $p_i = P[r_1+r_4=i \mid r_1+r_4>0]$ for $i=1,\ldots,n_1+m_1$. Since the moment generating function of Gamma (α, λ) is $\left(1-\frac{t}{\lambda}\right)^{-\alpha}$, the result immediately follows. Therefore, we have the following theorem.

Theorem 1: The conditional probability density function (pdf) of $\hat{\theta}_1$, $f_{\hat{\theta}_1}(x)$, and the conditional probability distribution function, $F_{\hat{\theta}_1}(x)$, becomes

$$f_{\hat{\theta}_1}(x) = \sum_{i=1}^{n_1+m_1} p_i g_i(x)\,, \qquad F_{\hat{\theta}_1}(x) = \sum_{i=1}^{n_1+m_1} p_i G_i(x)\,.$$

Here $g_i(x)$ and $G_i(x)$ denote the density function and the distribution function respectively of a gamma random variable with shape parameter $(n+m)$ and scale parameter $\frac{ni(\theta_1+\theta_2)}{(n_1+m_1)(\theta_1\theta_2)}$ for $i=1,\ldots,n_1+m_1$.

Proof of theorem 1: The proof follows immediately from lemma 1.

From theorem 1, it is clear that the conditional distribution of $\hat{\theta}_1$ is a mixture of gamma random variables. It is easy to compute different conditional moments of $\hat{\theta}_1$. We provide the first and the second conditional moments, others can be computed along the same line. We denote two conditional moments as $E(\hat{\theta}_1)$ and $E(\hat{\theta}_1^2)$ respectively. We have

$$E(\hat{\theta}_1) = \left(\sum_{i=1}^{n_1+m_1} \frac{p_i}{i}\right) \frac{1}{n}\,(n+m)\,(n_1+m_1)\left(\frac{\theta_1\theta_2}{\theta_1+\theta_2}\right) \tag{7}$$

and

$$E(\hat{\theta}_1^2) = \left(\sum_{i=1}^{n_1+m_1} \frac{p_i}{i^2}\right) \frac{1}{n^2}\,(n+m)\,(n+m+1)\,(n_1+m_1)^2\left(\frac{(\theta_1\theta_2)^2}{(\theta_1+\theta_2)^2}\right).$$

Note that $\hat{\theta}_1$ is not an unbiased estimator of θ_1 and since the summation sign denotes inverse moments of positive binomial random variables, it is not possible to give exact expressions. However, it is known (Mendenhall and Lehmann [7]) that if Z is a binomial random variable with parameters N and p, then for large N, $E\left[\frac{1}{Z} \mid Z>0\right] \approx \frac{1}{E(Z)} = \frac{1}{Np}$ and $E\left[\frac{1}{Z^2} \mid Z>0\right] \approx \frac{1}{(E(Z))^2} = \frac{1}{(Np)^2}$. Using these approximations, we have $E(\hat{\theta}_1) \approx \frac{n+m}{n}\,\theta_1$ and $E(\hat{\theta}_1^2) \approx \frac{(n+m)\,(n+m+1)}{n^2}\,\theta_1^2$. Therefore, when $\frac{m}{n}$ tends to zero, then $\hat{\theta}_1$ is an asymptotically unbiased and consistent estimator of θ_1. For fixed m and n (large), the expression (7) of $E(\hat{\theta}_1)$ can be used for biased correction. Same results are true for $\hat{\theta}_2$ also and they can be obtained simply by interchanging the roles of θ_1 and θ_2.

4. Exponential Latent Failure Time Distributions: Confidence Intervals

In this Section we obtain confidence bounds of λ_1 and λ_2 using the asymptotic distributions of $\hat{\lambda}_1$ and $\hat{\lambda}_2$. The 2×2 Fisher Information of λ_1 and λ_2 is $\boldsymbol{I}(\lambda_1, \lambda_2) = (I_{ij}(\lambda_1,\lambda_2))$ for $i, j = 1$ and 2. Here

$$I_{ij}(\lambda_1, \lambda_2) = -E\left(\frac{\partial^2 \ln(\lambda_1, \lambda_2)}{\partial\lambda_i\,\partial\lambda_j}\right)$$

and

$$I_{11}(\lambda_1,\, \lambda_2) = \frac{n\lambda_1 + (n_1 + m_1)\,\lambda_2}{\lambda_1(\lambda_1 + \lambda_2)^2}\,,$$

$$I_{12}(\lambda_1,\, \lambda_2) = I_{21}(\lambda_1,\, \lambda_2) = \frac{(n - m_1) - n_1}{(\lambda_1 + \lambda_2)^2} \quad \text{and}$$

$$I_{22}(\lambda_1,\, \lambda_2) = \frac{n\lambda_2 + (n_1 + m_1)\,\lambda_1}{\lambda_2(\lambda_1 + \lambda_2)^2}\,.$$

Therefore, if $\boldsymbol{\lambda} = (\lambda_1,\, \lambda_2)$ and $\hat{\boldsymbol{\lambda}} = (\hat{\lambda}_1,\, \hat{\lambda}_2)$, then we have,

$$(\hat{\boldsymbol{\lambda}} - \boldsymbol{\lambda}) \to N_2\big(\mathbf{0},\, \boldsymbol{I}^{-1}(\lambda_1,\, \lambda_2)\big)\,,$$

where $\boldsymbol{I}^{-1}(\lambda_1,\, \lambda_2) = (I_{ij}^{-1}(\lambda_1,\, \lambda_2))$ and

$$I_{11}^{-1}(\lambda_1,\, \lambda_2) = \frac{\lambda_1(n\lambda_2 + (n_1 + m_1)\,\lambda_1)}{n(n_1 + m_1)}\,,$$

$$I_{12}^{-1}(\lambda_1,\, \lambda_2) = I_{21}^{-1}(\lambda_1,\, \lambda_2) = \frac{\lambda_1\lambda_2(-(n - m_1) + n_1)}{n(n_1 + m_1)}\,,$$

$$I_{22}^{-1}(\lambda_1,\, \lambda_2) = \frac{\lambda_2(n\lambda_1 + (n_1 + m_1)\,\lambda_2)}{n(n_1 + m_1)}\,.$$

Similarly, we can have the Fisher Information matrix, $\boldsymbol{I}(\theta_1,\, \theta_2)$, for θ_1 and θ_2 by using simply Jacobian transformation.

Some other important parameters are, say

[1] $\pi = \dfrac{\lambda_1}{\lambda_1 + \lambda_2} = \dfrac{\theta_2}{\theta_1 + \theta_2}$: Probability of death due to cause 1.

[2] $\tau_1 = \pi\theta_1 + (1 - \pi)\,\theta_2 = \pi\lambda_1^{-1} + (1 - \pi)\,\lambda_2^{-1}$: The expected lifetime of all individuals.

[3] $\tau_2 = \pi\tau_1^{-1}$: The cause (1) specific mortality index.

[4] $\tau_3 = \exp\,(-\lambda_1 x)$: The survival function due to cause 1.

The MLEs of π, τ_1, τ_2 and τ_3 can be obtained very easily by using the in-variance property of the MLEs and their asymptotic distributions can be obtained using the δ-method.

5. Weibull Latent Failure Time Distributions

5.1 Estimation of parameters

In this Section we assume that X_{ji}s are Weibull random variables with parameters $(\alpha,\, \lambda_j)$ for $j = 1, 2$ and for $i = 1, \ldots, n$. The distribution function $F_j(.)$ of X_{ji} has the following form:

$$F_j(t) = 1 - \exp\,(-\lambda_j t^{\alpha})\,.$$

We assume that the lifetime distributions of the different causes follow Weibull distributions with different scale parameters but they have the same shape parameter. The hazard rates due to cause 1 and cause 2 are given by

$$\frac{\partial F_1(t)}{\bar{F}_1(t)} = \alpha\lambda_1 t^{\alpha-1} \quad \text{and} \quad \frac{\partial F_2(t)}{\bar{F}_2(t)} = \alpha\lambda_2 t^{\alpha-1}$$

respectively. The mean lifetimes due to cause 1 and cause 2 are

$$E(X_{1i}) = \frac{1}{\lambda_1^{\frac{1}{\alpha}}}\,\Gamma\left(1 + \frac{1}{\alpha}\right) \quad \text{and} \quad E(X_{2i}) = \frac{1}{\lambda_2^{\frac{1}{\alpha}}}\,\Gamma\left(1 + \frac{1}{\alpha}\right)$$

respectively. The relative risk rate, π, due to cause 1 is

$$\pi = P[X_{1i} < X_{2i}] = \int_0^\infty \alpha\lambda_1 x^{\alpha-1} \exp\left(-(x^\alpha(\lambda_1+\lambda_2)\right) \mathrm{d}x = \frac{\lambda_1}{\lambda_1+\lambda_2}$$

and it is independent of the shape parameter. The log-likelihood of the observed data is

$$\begin{aligned}\ln(L) = {} & (r_1+r_4)\ln(\lambda_1) + (r_2+r_5)\ln(\lambda_2) + (r_3-r_4-r_5)\ln(\lambda_1+\lambda_2) \\ & -(\lambda_1+\lambda_2)\sum_{i\in I} x_i^\alpha + (\alpha-1)\sum_{i\in I_1\cup I_2\cup I_3}\ln(x_i) + (r_1+r_2+r_3)\ln(\alpha)\,.\end{aligned}$$

Taking derivatives of $\ln(L)$ with respect to λ_1, λ_2 and α and making them equal to zero, give

$$\hat{\lambda}_1(\alpha) = \frac{n(r_1+r_4)}{\sum_{i\in I} x_i^\alpha(n_1+m_1)} \quad \text{and} \quad \hat{\lambda}_2(\alpha) = \frac{n(r_2+r_5)}{\sum_{i\in I} x_i^\alpha(n_1+m_1)}\,.$$

Putting values of $\hat{\lambda}_1(\alpha)$ and $\hat{\lambda}_2(\alpha)$ in $\ln(L)$, we obtain

$$\ln(L(\alpha)) = K + (r_1+r_2+r_3)\left[\ln(\alpha) - \ln\left(\sum_{i\in I} x_i^\alpha\right)\right] + (\alpha-1)\sum_{i\in I_1\cup I_2\cup I_3}\ln(x_i)\,, \tag{8}$$

here K is a constant independent of α. Now differentiating (8) with respect to α and equating it to zero, we obtain

$$\frac{d}{d\alpha}\ln(L(\alpha)) = -(r_1+r_2+r_3)\frac{\sum_{i\in I} x_i^\alpha \ln(x_i)}{\sum_{i\in I} x_i^\alpha} + \sum_{i\in I_1\cup I_2\cup I_3}\ln(x_i) + \frac{r_1+r_2+r_3}{\alpha} = 0\,. \tag{9}$$

The equation (9) can be written equivalently as

$$\alpha = \left[\frac{\sum_{i\in I} x_i^\alpha \ln(x_i)}{\sum_{i\in I} x_i^\alpha} - \frac{\sum_{i\in I_1\cup I_2\cup I_3}\ln(x_i)}{r_1+r_2+r_3}\right]^{-1} = h(\alpha)\ (\text{say})\,. \tag{10}$$

Therefore, a simple iterative scheme may be used to solve (10). From the ith iterate $\alpha_{(i)}$, $\alpha_{(i+1)}$ can be obtained as $h(\alpha_{(i+1)})$. Once we obtain $\hat{\alpha}$, $\hat{\lambda}_1(\hat{\alpha})$ and $\hat{\lambda}_2(\hat{\alpha})$ can be obtained as MLEs of λ_1 and λ_2 respectively.

Alternatively, we can use EM algorithm of Dempster, Laird and Rubin [2] to compute MLEs of unknown parameters. The EM algorithm consists of two steps. The first one is the estimation (E) step, where complete observations are left intact and 'pseudo observations' are framed. In this case we consider the E-step as follows. All data belong to category a, b, d, e, and f are left intact, only if the observation x belongs to category c, we form 'pseudo observation' by fractioning x to two partially complete 'pseudo observation' of the form $(x,\, w_1(x;\,\gamma))$, $(x,\, w_2(x;\,\gamma))$, where $\gamma = (\lambda_1,\, \lambda_2,\, \alpha)$. Specifically, the fractional mass, $(w_j(x;\,\gamma))$, assigned to this 'pseudo observation' x, is the conditional probability that the individual died from risk j given that the individual had died at time point point x. Therefore,

$$w_1(x;\,\gamma) = \frac{\lambda_1}{\lambda_1+\lambda_2} \quad \text{and} \quad w_2(x;\,\gamma) = \frac{\lambda_2}{\lambda_1+\lambda_2}\,.$$

From now on we denote $w_1(x;\,\gamma)$, $w_2(x;\,\gamma)$ as w_1 and w_2 respectively. The log-likelihood function of the 'pseudo data' can be written as $\ln(L_S(\lambda_1,\, \lambda_2,\, \alpha))$, where

$$\begin{aligned}\ln(L_S(\lambda_1,\lambda_2,\alpha)) = {} & (r_1+r_2+r_3)\ln(\alpha) + (r_1+w_1r_3)\ln(\lambda_1) + (r_2+w_2r_3)\ln(\lambda_2) \\ & + (\alpha-1)\sum_{i\in I_1\cup I_2\cup I_3}\ln(x_i) - (\lambda_1+\lambda_2)\sum_{i\in I} x_i^\alpha\,.\end{aligned} \tag{11}$$

The maximization step (M) involves the maximization of the log-likelihood function of the 'pseudo data' (11) with respect to λ_1, λ_2 and α. Assuming w_js to be known, we need to maximize $L_S(\lambda_1, \lambda_2, \alpha)$. It has to be performed iteratively. If $(\lambda_1^{(i)}, \lambda_2^{(i)}, \alpha^{(i)})$ is the ith iterate, then $(i+1)$th iterate, $(\lambda_1^{(i+1)}, \lambda_2^{(i+1)}, \alpha^{(i+1)})$ can be obtained as

$$\lambda_1^{(i+1)} = \frac{r_1 + r_3 w_1}{\sum\limits_{i \in I} x_i^{\alpha^{(i)}}} \quad \text{and} \quad \lambda_2^{(i+1)} = \frac{r_2 + r_3 w_2}{\sum\limits_{i \in I} x_i^{\alpha^{(i)}}} .$$

Finally $\alpha^{(i+1)}$ can be obtained as follows;

$$\alpha^{(i+1)} = \arg\max g(\alpha) ,$$

where

$$\begin{aligned} g(\alpha) = {} & (r_1 + r_2 + r_3) \ln(\alpha) + (r_1 + w_1 r_3) \ln(\lambda_1^{(i)}) + (r_2 + w_2 r_3) \ln(\lambda_2^{(i)}) \\ & + (\alpha - 1) \sum_{i \in I_1 \cup I_2 \cup I_3} \ln(x_i) - (\lambda_1^{(i)} + \lambda_2^{(i)}) \sum_{i \in I} x_i^{\alpha} . \end{aligned}$$

Equating $g'(\alpha) = 0$, we obtain

$$\alpha = \frac{r_1 + r_2 + r_3}{(\lambda_1^{(i)} + \lambda_2^{(i)}) \sum\limits_{i \in I} x_i^{\alpha} \ln(x_i) - \sum\limits_{i \in I_1 \cup I_2 \cup I_3} \ln(x_i)} = v(\alpha) \text{ (say)} . \tag{12}$$

Note that (12) also can be solved similarly as (10).

5.2 Confidence intervals

In this Section we provide confidence intervals of different parameters. Since it is not possible to obtain the exact distribution of MLEs, we use the asymptotic method. The asymptotic results can be stated as follows:

$$(\hat{\lambda}_1 - \lambda_1, \hat{\lambda}_2 - \lambda_2, \hat{\alpha} - \alpha) \to N_3(\mathbf{0}, \boldsymbol{I}^{-1}(\lambda_1, \lambda_2, \alpha)) . \tag{13}$$

Here $\boldsymbol{I}(\lambda_1, \lambda_2, \alpha)$ is the Fisher Information matrix for the unknown parameters $(\lambda_1, \lambda_2, \alpha)$. The elements of the 3×3 matrix $\boldsymbol{I} = ((I_{ij}))$ are as follows:

$$\begin{aligned} I_{11}(\lambda_1, \lambda_2, \alpha) &= \frac{n\lambda_1 + (n_1 + m_1)\lambda_2}{\lambda_1 (\lambda_1 + \lambda_2)^2} , \\ I_{22}(\lambda_1, \lambda_2, \alpha) &= \frac{n\lambda_2 + (n_1 + m_1)\lambda_1}{\lambda_2 (\lambda_1 + \lambda_2)^2} , \\ I_{12}(\lambda_1, \lambda_2, \alpha) &= \frac{n - (n_1 + m_1)}{(\lambda_1 + \lambda_2)^2} = I_{21}(\alpha, \lambda_1, \lambda_2) , \\ I_{33}(\lambda_1, \lambda_2, \alpha) &= \frac{n}{\alpha^2} + V(n + m)(\lambda_1 + \lambda_2) , \\ I_{13}(\lambda_1, \lambda_2, \alpha) &= (n + m)\, U = I_{31}(\lambda_1, \lambda_2, \alpha) = I_{23}(\alpha, \lambda_1, \lambda_2) = I_{32}(\lambda_1, \lambda_2, \alpha) . \end{aligned}$$

Here $U = E(X^{\alpha} \ln(X))$ and $V = E(X^{\alpha} (\ln(X))^2)$, where X is distributed as Weibull $(\alpha, \lambda_1 + \lambda_2)$. Note that U and V can be written as

$$\begin{aligned} U &= \frac{1}{\alpha(\lambda_1 + \lambda_2)^2} \left[\psi(2) - \ln(\lambda_1 + \lambda_2)\right] \\ V &= \frac{1}{\alpha^2(\lambda_1 + \lambda_2)^2} \left[\psi'(2) + \psi(2) (\ln(\lambda_1 + \lambda_2))^2 - 2\psi(2) \ln(\lambda_1 + \lambda_2)\right] . \end{aligned}$$

Here $\psi(.)$ and $\psi'(.)$ are the digamma and polygamma functions respectively. We can use (13) to compute asymptotic confidence intervals of all unknown parameters.

Since it is observed that in many cases (see Efron and Hinkley [4]), it is more appropriate to use the observed information matrix than the expected information matrix, we provide the observed information matrix also. It is possible to obtain the observed information matrix when the EM algorithm is used (Louis [6]). Using the same notation of Louis [6], the observed information matrix, $\hat{\boldsymbol{I}}$, takes the form

$$\hat{\boldsymbol{I}} = \boldsymbol{B} - \boldsymbol{S}\boldsymbol{S}'.$$

Here $\boldsymbol{B}$ is the 3×3 negative of the second derivative matrix and $\boldsymbol{S}$ is the 3×1 gradient vector. They are as follows:

$$S(1) = \frac{r_1 + w_1 r_3}{\lambda_1} - \sum_{i \in I} x_I^{\alpha}, \qquad S(2) = \frac{r_2 + w_2 r_3}{\lambda_2} - \sum_{i \in I} x_I^{\alpha},$$

$$S(3) = \frac{r_1 + r_2 + r_3}{\alpha} + \sum_{i \in I_1 \cup I_2 \cup I_3} \ln(x_i) - (\lambda_1 + \lambda_2) \sum_{i \in I} x_i^{\alpha} \ln(x_i).$$

and

$$B(1,1) = -\frac{r_1 + w_1 r_3}{\lambda_1^2}, \qquad B(2,2) = -\frac{r_2 + w_2 r_3}{\lambda_2^2}, \qquad B(1,2) = B(2,1) = 0,$$

$$B(1,3) = B(3,1) = B(2,3) = B(3,2) = -\sum_{I \in I} x_i^{\alpha} \ln(x_i),$$

$$B(3,3) = -\frac{r_1 + r_2 + r_3}{\alpha^2} - (\lambda_1 + \lambda_2) \sum_{i \in I} x_i^{\alpha} (\ln(x_i))^2.$$

6. Data Analysis

We illustrate the proposed parametric estimation technique using one cancer clinical trials. The data was originally analyzed by Dinse [3] using the non-parametric approach.

Data Set: This example is concerned with the relationship between time until progression and an indicator function of patient status at the time of progression for patients with glioblastoma, a cancer of the brain. Define X as the time until disease progression, as measured in month from the date of randomization. Suppose, δ indicates whether a patient is non ambulatory ($\delta = 1$) or ambulatory ($\delta = 2$) at the time of their most recent examination within the month before progression.

If a patient does not experience a progression by the time of analysis, the patient belongs to category f. For a patient with a known progression time, if there is no record of his/her ambulatory status within the month of preceding progression falls into category c. For detailed discussions on the data set the readers are referred to Dinse [3]. The data set presents times and indicators for 172 patients from a clinical trial. The summary of the data set is as follows: $r_1 = 41$, $r_2 = 17$, $r_3 = 31$, $r_4 = r_5 = 0$, $r_6 = 83$, $n_1 = 58$, $n = 89$, $m_1 = 0$ and $m = 83$. Also $\sum_{i \in I_1} x_i = 312$, $\sum_{i \in I_2} x_i = 150$, $\sum_{i \in I_3} = 276$, $\sum_{i \in I_6} x_i = 885$ and $\sum_{i \in I} x_i = 1639$.

Before fitting our model to the complete data set we would like to test the validity of the proposed parametric models. Since the closeness between the empirical survival function and the estimated survival function is a good measure for model validity for uncensored data, we consider a subset of the data which has only a, b or c types of failure. Assuming that the latent failure time distributions are exponential we obtain the MLEs of the parameters of the restricted data as $\hat{\lambda}_{1r} = 0.085\,249$ and $\hat{\lambda}_{2r} = 0.035\,347$. The corresponding UMVUEs are $\tilde{\lambda}_{1r} = 0.084\,291$ and $\tilde{\lambda}_{2r} = 0.034\,950$ respectively. In this case MLEs and UMVUEs are quite close to each other and we consider MLEs only. The Kolmogorov-Smirnov (K-S) distance between the empirical distribution function of the restricted data and the estimated distribution function using the exponential model is 0.1912 and the corresponding p

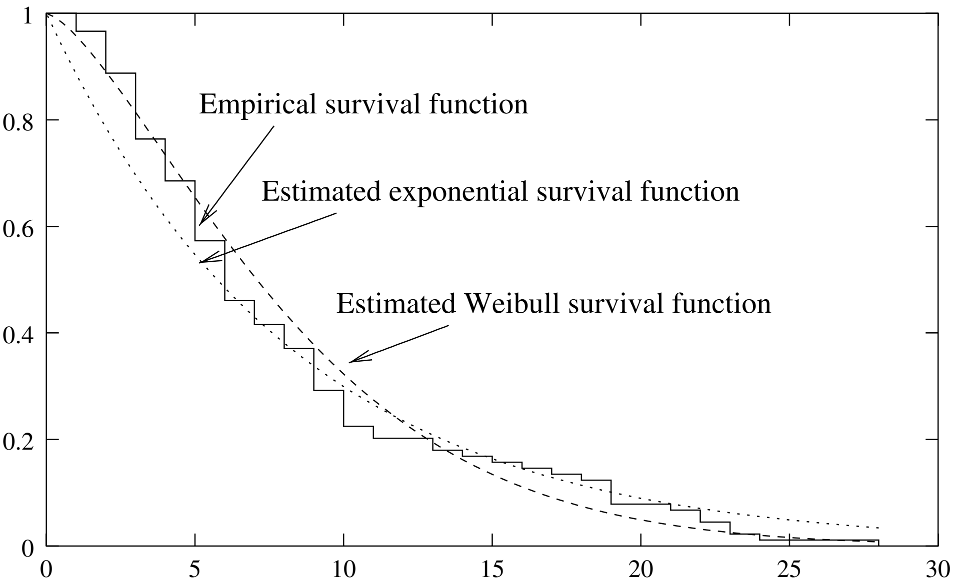

Fig. 1 Empirical survival function and the estimated survival functions using exponential and Weibull models.

value is 0.00298. Similarly assuming that the latent failure time distributions are Weibull, we obtain the MLEs of the parameters of the restricted data as $\hat{\lambda}_{1r} = 0.030678$, $\hat{\lambda}_{2r} = 0.012719$ and $\hat{\alpha}_r = 1.415506$. In this case the K-S distance between the empirical distribution function and the estimated distribution function using the Weibull model is 0.1173 and the corresponding p value is 0.17254. Since the p value corresponding to the Weibull model is quite high, therefore the preliminary data analysis suggests that the proposed Weibull model can be used to analyze this data set. We plot the empirical survival function and the estimated survival functions using exponential model and

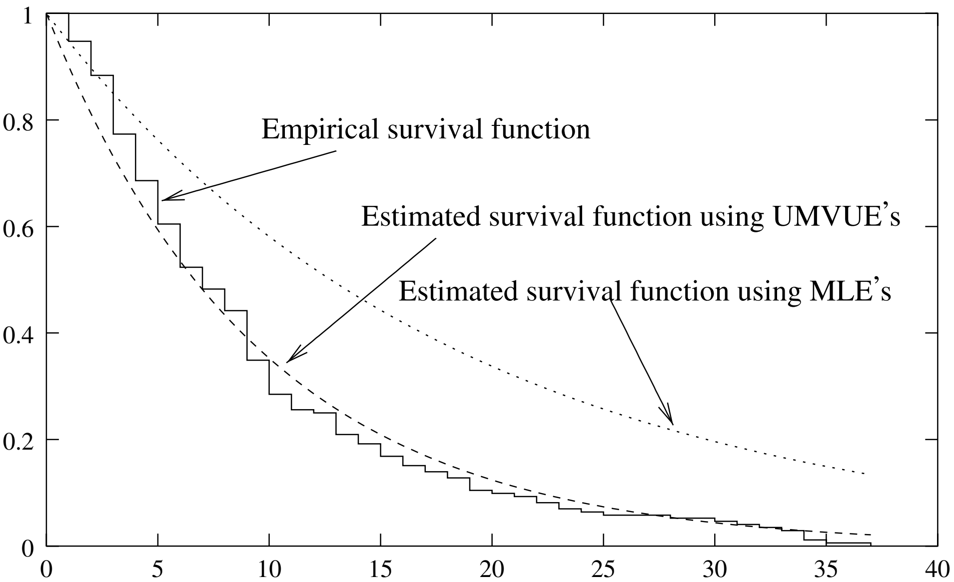

Fig. 2 Empirical survival function and estimated survival functions using MLE's and UMVUE's. Here lifetime distributions are exponential.

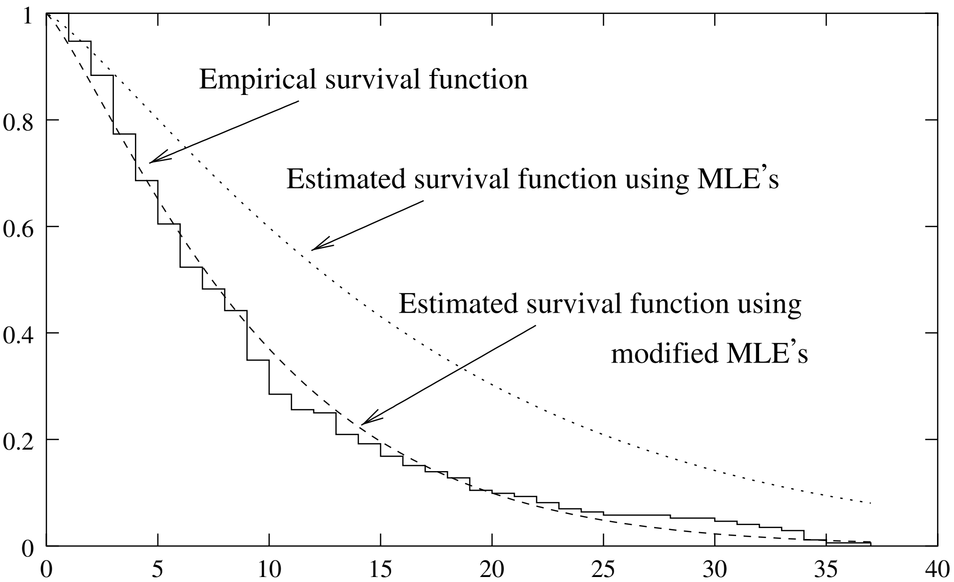

Fig. 3 Empirical survival function and the estimated survival functions using MLE's and modified MLE's. Here the lifetime distributions are Weibull.

Weibull model in Figure 1. Since it is assumed that the incomplete data are the coming from the same population as the complete data, therefore it is not unreasonable to use the proposed models to analyze this data set.

Now we consider the full data set. Under the assumption that the latent failure times are exponentially distributed, MLEs of λ_1 and λ_2 are $\hat{\lambda}_1 = 0.0384$ and $\hat{\lambda}_2 = 0.0159$ respectively, the log-likelihood value is -383.3599. Also $V(\hat{\lambda}_1) = 5.22 \times 10^{-6}$ and $V(\hat{\lambda}_1) = 3.27 \times 10^{-6}$. Moreover, UMVUEs of λ_1 and λ_2 are $\tilde{\lambda}_1 = 0.0738$ and $\tilde{\lambda}_2 = 0.0305$ respectively, with $V(\tilde{\lambda}_1) = 1.93 \times 10^{-5}$ and $V(\tilde{\lambda}_2) = 1.21 \times 10^{-5}$. Clearly MLEs and UMVUEs are quite different and it is mainly due to heavy censoring. Estimated survival function obtained using the MLEs is $e^{-(0.0384 + 0.0159)\,x}$ and the corresponding 95% confidence bounds are $e^{-(0.0384 + 0.0159)\,x} \pm 0.0113 \times e^{-(0.0384 + 0.0159)\,x}$. The Kolmogorov-Smirnov (K-S) distance between the empirical distribution function and the estimated distribution function is 0.296 and the corresponding p value is almost zero. Similarly the estimated survival function using the UMVUEs is $e^{-(0.0738 + 0.0305)\,x}$ and the corresponding 95% confidence bands are $e^{-(0.0738 + 0.0305)x} \pm 0.0416 \times e^{-(0.0738 + 0.0305)\,x}$. In this case the K-S distance is 0.152 and the corresponding p value is 0.007. We provide the empirical survival function and the estimated survival functions using MLEs and UMVUEs in the Figure 2. Clearly, UMVUEs provide a much better fit to the empirical survival function but still its p value is quite small. Instead of MLEs we prefer to use the bias corrected UMVUEs in this case.

Conditional estimates of mean times until progression of patients with non ambulatory status and ambulatory status are $\tilde{\theta}_1 = 13.4801$ and $\tilde{\theta}_2 = 32.5113$ respectively. The corresponding 95% confidence intervals are (4.8404, 22.1199) and (26.4388, 38.5834) respectively. Clearly the mean progression time of the patients with ambulatory status is significantly more than the patients with non ambulatory status. The biased corrected estimate of the mean progression time of all patients, say $\tilde{\tau}_1 = 19.0586$ and the corresponding 95% asymptotic confidence interval is (15.0924, 23.0249). The estimate of the relative risk rate of the non ambulatory status is 0.7072 and the corresponding 95% asymptotic confidence interval is (0.6076, 0.8076). Clearly it is siginificantly greater than 0.5.

The estimates of the survival functions of the time until progression of the patients with non-ambulatory and ambulatory status at the time point x are $e^{-0.0738x}$ and $e^{-0.0305x}$ respectively. The corresponding 95% confidence intervals are $e^{-0.0738x} \pm 0.0196x\, e^{-0.0738x}$ and $e^{-0.0305x} \pm 0.0138x\, e^{-0.0305x}$ respectively.

Now we analyze the data under the assumptions that the latent failure time distributions are Weibull. In this case, we obtain the MLEs of the unknown parameters using the methods proposed in the Section 5. Both the methods converge to the same solutions and they are as follows, $\hat{\alpha} = 1.2145$, $\hat{\lambda}_1 = 0.0222$ and $\hat{\lambda}_2 = 0.0092$, the corresponding log-likelihood value is -378.1220. We obtain the empirical survival function and the estimated survival function using the MLEs as $e^{-(0.0222+0.0092)x^{1.2145}}$. We provide it in the Figure 3. Unfortunately it does not match very well. The K-S distance between the empirical distribution function and the estimated distribution function is 0.313 and the corresponding p value is almost zero. Similar things happen even in the exponential case also. We feel that the scale parameters are highly biased. We make the bias correction as follows. Assuming that the shape parameter is known to be 1.2145, we transform the date as X^{α} and then estimate the scale parameters by UMVUEs. It gives the improved estimates of λ_1 and λ_2, say $\tilde{\lambda}_1 = 0.0429$ and $\tilde{\lambda}_2 = 0.0178$ respectively. We obtain the empirical survival function and the estimated survival function using the improved estimates as $e^{-(0.0429+0.0178)x^{1.2145}}$. It is also plotted in Figure 3. The K-S distance between the empirical survival function and the estimated survival function in this case is 0.089 and the corresponding p value is 0.127. We will be using these new estimates from now on.

The asymptotic variance covariance matrix of $(\hat{\lambda}_1, \hat{\lambda}_2, \hat{\alpha})$ obtained from the observed information matrix using the EM algorithm is

$$\begin{bmatrix} 2.7261 \times 10^{-6} & -2.1173 \times 10^{-6} & -1.1 \quad \times 10^{-5} \\ -2.1173 \times 10^{-6} & 2.3672 \times 10^{-6} & -0.4943 \times 10^{-5} \\ -1.1 \quad \times 10^{-5} & -0.4943 \times 10^{-5} & 16.469 \quad \times 10^{-4} \end{bmatrix}.$$

Therefore, by proper multiplication we obtain the variance covariance matrix of $(\tilde{\lambda}_1, \tilde{\lambda}_2, \tilde{\alpha})$ as

$$\begin{bmatrix} 1.0180 \times 10^{-5} & -4.0915 \times 10^{-6} & -2.1257 \times 10^{-5} \\ -4.0915 \times 10^{-6} & 8.8398 \times 10^{-6} & -0.9552 \times 10^{-5} \\ -2.1257 \times 10^{-5} & -0.9552 \times 10^{-5} & 16.469 \quad \times 10^{-4} \end{bmatrix}.$$

The asymptotic 95% confidence bounds for λ_1, λ_2 and α are (0.0366, 0.0492), (0.0119, 0.0236) and (1.1350, 1.2940) respectively. Since the 95% confidence interval of α does not include 1, therefore it is reasonable to use the Weibull model rather than the exponential model for this data set. Even the log-likelihood values also suggest (performing χ^2 testing) that the Weibull model is a preferred model than the exponential model in this case. The estimates of the mean lifetime of the patients with non-ambulatory and ambulatory status are 12.5349 and 25.8634 and the corresponding 95% confidence bounds are (5.0299, 20.0399) and (14.5047, 37.2221) respectively. The estimate of the relative risk rate of the patients with non-ambulatory status is 0.7068 and the corresponding asymptotic 95% confidence band is (0.5885, 0.8251). The estimates of the survival functions of the time until progression of the patients with non-ambulatory and ambulatory status at the time point x are $e^{-0.0429x^{1.2145}}$ and $e^{-0.0178x^{1.2145}}$ respectively. The corresponding 95% asymptotic confidence bounds are $e^{-0.0429x^{1.2145}} \pm e^{-0.0429x^{1.2145}} x^{1.2145} \left(19.9528 - 3.5527 \ln(x) + 5.9570 (\ln(x))^2\right)^{\frac{1}{2}} \times 10^{-3}$ and $e^{-0.0178x^{1.2145}} \pm e^{-0.0178x^{1.2145}} x^{1.2145} \left(17.3260 - 0.6665 \ln(x) + 1.0227 (\ln(x))^2\right)^{\frac{1}{2}} \times 10^{-3}$. They are plotted in Figure 4 and it is clear from the figure that the two survival functions are significantly different.

For illustrations purposes we assume that 10% of the category f data have known status of the patients. We randomly choose 10% of the data from the category f and assign status 1 and status 2 with probability 0.7 ($\approx \pi$) and 0.3 respectively. Now the summary of the new data set is as follows: $r_1 = 41$, $r_2 = 17$, $r_3 = 31$, $r_4 = 7$, $r_5 = 5$, $r_6 = 71$, $n_1 = 58$, $n = 89$, $m_1 = 12$ and $m = 83$. Also $\sum_{i \in I_1} x_i = 312$, $\sum_{i \in I_2} x_i = 150$, $\sum_{i \in I_3} = 276$, $\sum_{i \in I_4} = 109$, $\sum_{i \in I_5} = 76$, $\sum_{i \in I_6} x_i = 700$ and $\sum_{i \in I} x_i = 1639$. Assuming that the lifetime distributions are exponential, the estimates of the scale parameters are as follows: $\hat{\lambda}_1 = 0.0372$, $\hat{\lambda}_2 = 0.0171$, $\tilde{\lambda}_1 = 0.0714$ and $\tilde{\lambda}_2 = 0.0329$. Similarly assuming that the lifetime distri-

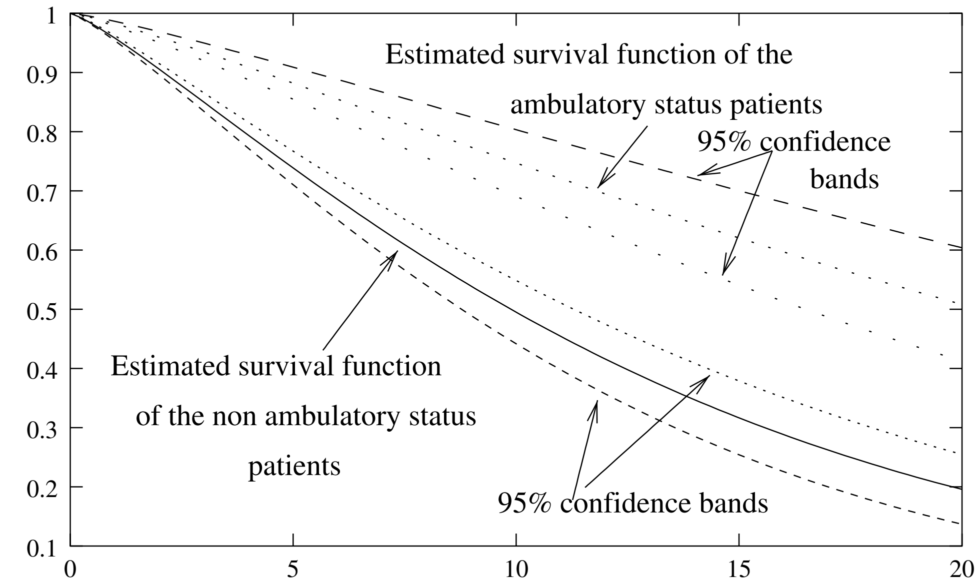

Fig. 4 Estimated survival functions and the corresponding 95% confidence bounds for the non-ambulatory and the ambulatory status patients. The lifetime distributions are assumed to be Weibull.

butions are Weibull the estimates are $\hat{\alpha} = 1.2145$, $\hat{\lambda}_1 = 0.0214$, $\hat{\lambda}_2 = 0.0099$, $\tilde{\lambda}_1 = 0.0413$ and $\tilde{\lambda}_2 = 0.0189$. Therefore the estimates of the other parameters can be obtained similarly as before, using these estimates.

Now we compare ours with Dinse's [3] findings. First of all the non-parametric estimate of $\bar{F}(.)$ as obtained by Dinse [3] is quite close to the parametric estimate of the survival function, particularly when the latent failure time distributions are assumed to be Weibull (see Figure 3). This indicates that both analysis should provide comparable results. Dinse [3] also mentioned that the non-parametric estimates of $\delta_j(t) = P[\delta = j \mid X = t]$ are very erratic and therefore, he has provided the smoothed estimates of $\delta_j(t)$. In our formulation, $\delta_j(t)$ is constant for all t. Therefore, the parametric estimate of $\delta_j(t)$ can be considered as the smoothed version of the non-parametric estimates. In our case using Weibull distributions, we obtain $\hat{\delta}_1(t) \approx 0.70$ and $\hat{\delta}_2(t) \approx 0.30$. From Figure 1 of Dinse [3], it is clear that the non-parametric smoothed estimate of $\delta_2(t)$ is close to 0.30. Therefore, both methods provide similar results.

7. Cause Specific Hazard Functions Model

In this Section, we formulate the problem, using the cause specific hazard functions, as originally proposed by Prentice et al. [9]. We assume cause specific hazard functions to be exponential or Weibull. We use $\lambda(t)$ as the overall hazard function and $\lambda_j(t)$ as the cause specific hazard function for $j = 1$ and 2. It is known that $\lambda(t) = \lambda_1(t) + \lambda_2(t)$.

Suppose X, δ and $\bar{F}(.)$ are same as defined in Sections 1 and 2, the likelihood contributions of the observations a, b, c, d, e and f are

$$\lambda_1(t)\,\bar{F}(t), \qquad \lambda_2(t)\,\bar{F}(t), \qquad \lambda(t)\,\bar{F}(t), \qquad \int_t^\infty \lambda_1(u)\,\bar{F}(u)\,\mathrm{d}u, \qquad \int_t^\infty \lambda_2(u)\,\bar{F}(u)\,\mathrm{d}u,$$
$$\int_t^\infty \lambda(u)\,\bar{F}(u)\,\mathrm{d}u\,,$$

respectively.

7.1 Exponential cause specific hazard functions

In this subSection, it is assumed

$$\lambda_1(t) = \lambda_1 \quad \text{and} \quad \lambda_2(t) = \lambda_2 .$$

In this case the likelihood contributions from the observations a, b, c, d, e and f are

$$\lambda_1 \, e^{-(\lambda_1+\lambda_2)\,t}, \qquad \lambda_2 \, e^{-(\lambda_1+\lambda_2)\,t}, \qquad (\lambda_1+\lambda_2)\, e^{-(\lambda_1+\lambda_2)\,t}, \qquad \frac{\lambda_1}{\lambda_1+\lambda_2}\, e^{-(\lambda_1+\lambda_2)\,t},$$
$$\frac{\lambda_2}{\lambda_1+\lambda_2}\, e^{-(\lambda_1+\lambda_2)\,t}, \qquad e^{-(\lambda_1+\lambda_2)\,t}$$

respectively. Therefore, under this formulation, the likelihood function of the observed data (1) is same as (2).

7.2 Weibull cause specific hazard functions

In this subSection, it is assumed,

$$\lambda_1(t) = \alpha\lambda_1 t^{\alpha-1} \quad \text{and} \quad \lambda_2(t) = \alpha\lambda_2 t^{\alpha-1} .$$

Therefore,

$$\bar{F}(t) = e^{-(\lambda_1+\lambda_2)\,t^{\alpha}} .$$

In this case the likelihood contributions from the observations a, b, c, d, e and f are

$$\alpha\lambda_1 t^{\alpha-1}\, e^{-(\lambda_1+\lambda_2)\,t^{\alpha}}, \qquad \alpha\lambda_2 t^{\alpha-1}\, e^{-(\lambda_1+\lambda_2)\,t^{\alpha}}, \qquad \alpha(\lambda_1+\lambda_2)\, t^{\alpha-1}\, e^{-(\lambda_1+\lambda_2)\,t^{\alpha}},$$
$$\frac{\lambda_1}{\lambda_1+\lambda_2}\, e^{-(\lambda_1+\lambda_2)\,t^{\alpha}}, \qquad \frac{\lambda_2}{\lambda_1+\lambda_2}\, e^{-(\lambda_1+\lambda_2)\,t^{\alpha}}, \qquad e^{-(\lambda_1+\lambda_2)\,t^{\alpha}},$$

respectively. In this case also the likelihood function of the data is same as the likelihood function obtained using independent Weibull latent failure times distributions in Section 5. Since the likelihood functions are equal in both cases, the estimation procedures of the different unknown parameters, namely (λ_1, λ_2) or $(\lambda_1, \lambda_2, \alpha)$, and the statistical properties of these estimates are same in both the formulation.

8. Conclusions

In this paper we consider the analysis of the partially complete time and type of failure data in presence of competing risks. We formulate the problem in two different ways and it is observed that both of them lead to the same likelihood function. It raises the important question of the identifiability problem of competing risks. Our analysis shows, as expected, that in the above situation it is not possible to identify from the given data whether the competing causes are independent or not without the presence of covariates.

References

Cox, D. R. (1959). The analysis of exponentially distributed lifetimes with two types of failures. *Journal of the Royal Statistical Society, Series B* **21**, 411–421.

Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977). Maximum likelihood estimation from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B* **39**, 1–22.

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

Efron, B. and Hinkley, D. V. (1978). The observed versus expected information. *Biometrika* **65**, 457–487.

Kundu, D. and Basu, S. (2000). Analysis of incomplete data in presence of competing risks. *Journal of Statistical Planning and Inference* **87**, 221–239.

Louis, T. A. (1982). Finding the observed information matrix when using the EM algorithm. *Journal of the Royal Statistical Society, Series B* **44**, 2, 226–233.

Mendenhall, W. and Lehmann, E. H. Jr. (1960). An approximation of the negative moments of the positive binomial useful in life testing. *Technometrics* **2**, 227–242.

Miyakawa, M. (1984). Analysis of incomplete data in competing risks model. *IEEE Transactions on Reliability Analysis* **33**, 293–296.

Prentice, R. L., Kalbfleisch, J. D., Peterson, Jr., A. V., Flournoy, N., Farewell, V. T. and Breslow, N. E. (1978). The analysis of failure times in the presence of competing risks. *Biometrics* **34**, 541–554.