ELSEVIER

Journal of Statistical Planning and Inference 87 (2000) 221–239

journal of statistical planning and inference

www.elsevier.com/locate/jspi

Analysis of incomplete data in presence of competing risks

Debasis Kundu[a, *], Sankarshan Basu[b]

[a] Department of Mathematics, Indian Institute of Technology Kanpur, Kanpur, 208016, India

[b] Department of Statistics, London School of Economics, Houghton Street, London WC2A 2AE, UK

Received 26 June 1998; received in revised form 17 August 1999; accepted 3 November 1999

Abstract

In medical studies or in reliability analysis an investigator is often interested in the assessment of a specific risk in the presence of other risk factors. In statistical literature this is known as the analysis of competing risks model. The competing risks model assumes that the data consists of a failure time and an indicator denoting the cause of failure. Several studies have been carried out under this assumption for both the parametric and non-parametric set up. Unfortunately in many situations, the causes of failure are not observed, even if the failure times are observed. Miyawaka (1984, IEEE Trans. Reliability Anal. 33(4), 293–296) obtained some results under the assumption that the failure time distribution is exponential. He obtained the maximum likelihood estimators and the minimum variance unbiased estimators of the unknown parameters. We provide the approximate and asymptotic properties of these estimators. Using the approximate and the asymptotic distributions we compute confidence intervals for different parameters and compare them with the two different bootstrap confidence bounds. We also consider the case when the failure distributions are Weibull. One data set is used to see how different methods work in real-life situations. © 2000 Elsevier Science B.V. All rights reserved.

MSC: 62GO5; 60F15

Keywords: Competing risks; Failure rates; Exponential distribution; Weibull distribution; Maximum likelihood estimators; Bootstrap confidence intervals

1. Introduction

In medical studies or in reliability analysis it is quite common that more than one risk factor may be present at the same time. An investigator is often interested in the assessment of a specific risk in the presence of other risk factors. In statistical literature it is well known as the competing risks model. In analyzing the competing risks model, it is assumed that data consists of a failure time and an indicator denoting the cause of failure. Several studies have been carried out under this assumption for

* Corresponding author.

E-mail address: kundu@iitk.ac.in (D. Kundu).

PII: S0378-3758(99)00193-7

both the parametric and the non-parametric set up. For the parametric set up it is assumed that different lifetime distributions follow some special parametric distribution, namely exponential, gamma or Weibull. Several authors, for example Berkson and Elveback (1960), Cox (1959), David and Moeschberger (1978) considered this problem from the parametric point of view. In the non-parametric set up no specific lifetime distribution is assumed. Kaplan and Meier (1958), Efron (1967) and Peterson (1977) analyzed the non-parametric version of this model. In all of the above cases it is assumed that when the failure times are observed, the causes of failure are also known. However, in certain situations (Dinse, 1982 or Miyawaka, 1982) it is observed that the determination of the cause of failure may be expensive or may be very difficult to observe. Therefore it might occur that the failure time of that item/individual is observed but the corresponding cause of failure is not observed. Miyawaka (1984) considered this model and obtained the maximum likelihood estimators (MLEs) and the uniformly minimum variance unbiased estimators (UMVUEs) of the failure rates of different failure distributions under the assumption that they are of the exponential type. However, he did not provide any distributional properties of these estimators.

In this paper we consider the same model as that of Miyawaka (1984). It is assumed that every member of a certain target population either dies of a particular cause, 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. At the end of the study, we have three types of observation:

(a) Individuals who died of cancer and their lifetimes.
(b) Individuals who died of other causes and their lifetimes.
(c) Individuals whose lifetimes are observed, but causes of death are unknown.

Types a and b will be referred to as complete observations and type c will be referred to as incomplete observations. There is another dimension to this problem. Suppose, a research project is financed for a fixed length of time, say, M. Some individuals could be alive at the end of the project period. For simplicity, we ignore this aspect of the problem and assume that every one in the sample can be monitored until death. We assume that no censoring occurs. The type of data considered here is complete in terms of failure times (not censored) and incomplete only in terms of failure modes. It may be mentioned that David and Moeschberger (1978) considered different cases when the data are censored.

First, we consider the case when the lifetime distributions of the different causes of failure are exponential. It is observed that although the MLEs or the UMVUEs of the hazard rates always exist, the MLEs or the UMVUEs of the mean lifetime of the different causes may not exist always. We propose to use the conditional MLEs of mean lifetimes of different causes. We obtain the exact distribution of the conditional MLEs using the conditional moment generating function approach and also obtain the asymptotic distributions of the MLEs. It is observed that the MLE of the mean lifetime usually over estimates the true parameter, at least for small samples. Similar to the case of a complete sample, in the case of an incomplete sample the bias can also be obtained as the inverse moment of a positive binomial random variable. We provide small tables for the biases of the MLEs for different sample sizes and for different parameter values when 10% and 20% of the failure modes are unobservable (incomplete). Based on the exact and the asymptotic distributions, we propose two approximate confidence intervals of the different parameters of interest. We also use percentile bootstrap and bootstrap-t confidence intervals for the unknown parameters and compare their performances through Monte Carlo simulations.

Since the exponential distribution has constant failure rate it might not be very practical to assume that the lifetime distribution is exponential. The two-parameter Weibull distribution can be used to analyze lifetime data because of its increasing and decreasing failure rates. We consider the model when the underlying lifetime distributions are Weibull. We obtain the maximum likelihood estimators of different parameters and study their properties under this assumption. We also obtain the asymptotic confidence bounds and the bootstrap confidence bounds of the different parameters and compare their performances through Monte Carlo simulations. We consider one data set from Lawless (1982) and see how the different methods work in this practical situation.

The rest of the paper is organized as follows. In Section 2, we describe the model. In Section 3, we consider estimation of the different parameters and also obtain the exact distribution of the MLE of the mean lifetime when the lifetime distribution is exponential. Different confidence intervals of the unknown parameters are considered in Section 4. The Weibull lifetime distribution is considered in Section 5. Numerical results are presented in Section 6. One data set from Lawless (1982) is analyzed in Section 7 and finally we draw conclusions from our work in Section 8.

2. Model description and notations

Without loss of generality, we assume that there are only two causes of failure. We use the following notations:

X_i	lifetime of system i
X_{ji}	lifetime of mode j of system i, $j=1,2$
$F(\cdot)$	cumulative distribution function of X_i
$F_j(\cdot)$	cumulative distribution function of X_{ji}, $j=1,2$
$\bar{F}_j(\cdot)=1-F_j(\cdot)$	
δ_i	indicator variable denoting the cause of failure of system i
$\mathrm{I}[\cdot]$	indicator function of event $[\cdot]$
Gamma (α, λ)	denotes the gamma random variable with density function $(\lambda^\alpha/\Gamma(\alpha))x^{\alpha-1}\mathrm{e}^{-\lambda x}$
Weibull (α, λ)	denotes the Weibull random variable with density function $\alpha\lambda x^{\alpha-1}\mathrm{e}^{-\lambda x^\alpha}$

It is assumed that (X_{1i}, X_{2i}); $i=1,2,\ldots,n$, are n independent identically distributed (i.i.d.) random variables. X_{1i} and X_{2i} are independent for all $i=1,2,\ldots,n$ and

$X_i = \min\{X_{1i}, X_{2i}\}$. With out loss of generality, it is assumed that the first m observations consist of failure times and also causes of failure whereas for the last $(n-m)$ observations we only observe the failure times and not the causes of failure, i.e. the following data are observed: $(X_1,\delta_1),(X_2,\delta_2),\dots,(X_m,\delta_m)$, $(X_{m+1},*),\dots,(X_n,*)$. In order to analyze the incomplete data it is assumed that the failure times are from the same population as the complete data, that is the population remains unchanged irrespective of the cause of failure. We also assume throughout that m is fixed, strictly positive and not random. If m is random then all the results involving m, given in the rest of the sections, are valid on the conditioning argument on m, i.e. for a given m.

3. Exponential failure distributions, estimation

In this section we assume that X_{ji}'s are exponential random variables with parameters λ_j for $i=1,2,\dots,n$ and for $j=1$ and 2. The distribution function $F_j(\cdot)$ of X_{ji} has the following form:

$$F_j(t) = (1 - e^{-\lambda_j t}) \tag{3.1}$$

for $j=1$ and 2. The likelihood function of the observed data $(x_1,\delta_1),\dots(x_m,\delta_m),(x_{m+1},*)$, $\dots,(x_n,*)$ for the general case takes the following form:

$$L = \prod_{i=1}^{m} [dF_1(x_i)\bar{F}_2(x_i)]^{I(\delta_i=1)}[dF_2(x_i)\bar{F}_1(x_i)]^{I(\delta_i=2)} \prod_{i=m+1}^{n} dF(x_i). \tag{3.2}$$

Therefore for the particular case when F_1 and F_2 are exponentials with parameters λ_1 and λ_2, respectively, the likelihood function becomes

$$L = \lambda_1^{r_1}\lambda_2^{r_2}(\lambda_1+\lambda_2)^{n-m}\exp\left(-(\lambda_1+\lambda_2)\sum_{i=1}^{n} x_i\right). \tag{3.3}$$

Here r_1 and r_2 denote the number of failures due to modes 1 and 2, respectively. Taking the logarithm of (3.3) and equating the partial derivatives to be zeros, we get the MLE of λ_1 as

$$\hat{\lambda}_1 = \frac{nr_1}{m\sum_{i=1}^{n} x_i}$$

and the UMVUE of λ_1 as

$$\lambda_1^* = \frac{(n-1)r_1}{m\sum_{i=1}^{n} x_i}, \tag{3.4}$$

see Miyawaka (1984). Therefore, the MLE of the survival function due to cause 1 (say cancer) is

$$\hat{\bar{F}}_j(x) = e^{-\hat{\lambda}_1 x}, \tag{3.5}$$

and the MLE of the hazard rate or the instantaneous death rate due to cause 1 is given by

$$\frac{\widehat{dF_1(x)}}{\hat{\bar{F}}_1(x)} = \hat{\lambda}_1.$$

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

$$P[X_{1i} < X_{2i}] = \int_0^\infty \lambda_1 e^{-\lambda_1 x} e^{-\lambda_2 x}\,dx = \frac{\lambda_1}{\lambda_1 + \lambda_2}$$

and because of the invariance property of the MLE, the relative risk rate $\hat{\pi}$ due to cause 1 is

$$\hat{\pi} = \frac{\hat{\lambda}_1}{\hat{\lambda}_1 + \hat{\lambda}_2} = \frac{r_1}{m}. \tag{3.6}$$

For the exponential lifetime distribution as in (3.1), λ_1 represents the hazard rate and $\theta_1 = 1/\lambda_1$ denotes the mean lifetime due to cause 1. Although the MLE and the UMVUE of λ_1 always exist the same is not true for θ_1. The UMVUE of θ_1 does not exist and the MLE of θ_1 exists only when $r_1 > 0$. The conditional MLE of θ_1, say $\hat{\theta}_1$, when $r_1 > 0$ is as follows:

$$\hat{\theta}_1 = \frac{m \sum_{i=1}^n x_i}{n r_1}. \tag{3.7}$$

If $r_1 = 0$, $\hat{\theta}_1$ does not exist. Now, we obtain the conditional distribution of the MLE of θ_1, conditioning on $r_1 > 0$. Our approach to produce the confidence bound for θ_1 is based on the distribution of $\hat{\theta}_1$ and is similar to Chen and Bhattacharya (1988) and Gupta and Kundu (1998). Moreover, use of the MLE provides a safeguard against any obvious loss of information and ensures asymptotic optimality of the present method. Consider the following lemma:

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

$$\phi_{\hat{\theta}_1}(t) = E[e^{t\hat{\theta}_1} | r_1 > 0]$$

$$= (1 - q^m)^{-1} \left[\sum_{i=1}^m \frac{m!}{i!(m-i)!} \left(\frac{\theta_2}{\theta_1 + \theta_2} \right)^i \left(\frac{\theta_1}{\theta_1 + \theta_2} \right)^{m-i} \left(1 - \frac{tm\theta_1\theta_2}{ni(\theta_1 + \theta_2)} \right)^{-n} \right]$$

$$= \left[\sum_{i=1}^m p_i \left(1 - \frac{tm\theta_1\theta_2}{ni(\theta_1 + \theta_2)} \right)^{-n} \right].$$

Here

$$q = \frac{\theta_1}{\theta_1 + \theta_2} \quad \text{and} \quad p_i = (1 - q^m)^{-1} \frac{m!}{i!(m-i)!} \left(\frac{\theta_2}{\theta_1 + \theta_2} \right)^i \left(\frac{\theta_1}{\theta_1 + \theta_2} \right)^{m-i},$$

for $i = 1, 2, \ldots, m$.

Proof. Note that $\sum_{i=1}^n X_i$ is a Gamma $(n, \lambda_1 + \lambda_2)$ random variable and r_1 is a binomial random variable with parameters m and $\lambda_1/(\lambda_1 + \lambda_2)$

$$\phi_{\hat{\theta}_1}(t) = E[e^{t\hat{\theta}_1} | r_1 > 0]$$

$$= \sum_{i=1}^n E[e^{t\hat{\theta}_1} | r_1 = i] P(r_1 = i | r_1 > 0).$$

Using that $p_i = P(r_1 = i | r_1 > 0)$ for $i = 1, 2, \ldots, m$ and the moment generating function of Gamma (α, λ) is $(1 - t/\lambda)^{-\alpha}$, the result follows immediately. □

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

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

where $g_i(x)$ and $G_i(x)$ denote the density function and the distribution function, respectively, of a gamma random variable with shape parameter n and scale parameter $ni(\theta_1 + \theta_2)/m\theta_1\theta_2$ for $i = 1, 2, \ldots, m$.

Proof. Obvious from Lemma 1. □

Therefore, the conditional distribution of $\hat{\theta}_1$ is a mixture of m gamma random variables. It may be noted also that when $m = n$, we get the distribution of $\hat{\theta}_1$ in the competing risk model when the lifetime distributions of the different causes are exponential and there is no censoring or incomplete data (see David and Moeschberger, 1978). With the help of Theorem 1, we can obtain different conditional moments of $\hat{\theta}_1$. We give the first and the second conditional moments of $\hat{\theta}_1$. For brevity we denote them as $E(\hat{\theta}_1)$ and $E(\hat{\theta}_1^2)$, respectively:

$$E(\hat{\theta}_1) = \frac{m\theta_1\theta_2}{(\theta_1 + \theta_2)} \sum_{i=1}^{m} \frac{p_i}{i} \tag{3.8}$$

and

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

Since in both the cases the quantities within the summation sign denote inverse moments of positive binomial random variables, it is not possible to give exact expressions. Therefore it is difficult to obtain the exact bias from (3.8). The tabulated values of the first moment of the inverse positive binomial random variable are available in Edwin and Savage (1954). These may be used to make some bias correction. However, this is not pursued here. If Z is a binomial random variable with parameters N and P, then for large N, $E(1/Z | Z > 0) \approx 1/E(Z) = 1/NP$ and $E(1/Z^2 | Z > 0) \approx 1/(E(Z))^2 = 1/(NP)^2$ (see Mendenhall and Lehmann, 1960). Using these approximations, we can say that for large m and n, $E(\hat{\theta}_1) \approx \theta_1$ and $E(\hat{\theta}_1^2) \approx \theta_1^2$. We can also show that asymptotically $\hat{\theta}_1$ is an unbiased estimator of θ_1. The variance of $\hat{\theta}_1 \approx 0$ for large m and n, therefore $\hat{\theta}_1$ is a consistent estimator of θ_1 also.

We present two small tables which give the numerical values of the exact biases of $\hat{\theta}_1$ for different sample sizes and for different values of θ_2. In all these calculations we have kept $\theta_1 = 1$. Table 1 represents the negative value of the biases when 10% of the data are incomplete and Table 2 represents the negative value of the biases when 20% of the data are incomplete.

Table 1
$\theta_1 - E(\hat{\theta}_1)$ when 10% of the date are incomplete

n	$\theta_2 = 1.25$	$\theta_2 = 1.50$	$\theta_2 = 1.75$	$\theta_2 = 2.00$	$\theta_2 = 2.25$	$\theta_2 = 2.50$
10	0.1284	0.1041	0.0868	0.0741	0.0646	0.0572
20	0.0528	0.0431	0.0363	0.0315	0.0277	0.0248
30	0.0330	0.0271	0.0231	0.0201	0.0177	0.0159
40	0.0241	0.0198	0.0169	0.0147	0.0130	0.0117
50	0.0189	0.0156	0.0133	0.0116	0.0103	0.0093

Table 2
$\theta_1 - E(\hat{\theta}_1)$ when 20% of the date are incomplete

n	$\theta_2 = 1.25$	$\theta_2 = 1.50$	$\theta_2 = 1.75$	$\theta_2 = 2.00$	$\theta_2 = 2.25$	$\theta_2 = 2.50$
10	0.1469	0.1207	0.1011	0.0867	0.0756	0.0669
20	0.0610	0.0495	0.0417	0.0360	0.0317	0.0284
30	0.0377	0.0309	0.0263	0.0228	0.0202	0.0181
40	0.0273	0.0225	0.0192	0.0167	0.0148	0.0133
50	0.0214	0.0177	0.0151	0.0132	0.0117	0.0105

From Tables 1 and 2 it is observed that the bias of the MLE of $\hat{\theta}_1$ is negative in all cases considered. It means $\hat{\theta}_1$ always over estimates θ_1, although we could not prove this theoretically. As the sample size increases the bias also decreases. If we have more incomplete data there is also more bias. Since $E(\hat{\theta}_1/\theta_1)$ is a function of θ_2/θ_1 only, Tables 1 or 2 can be used to compute the bias for any value of θ_1. For example, if $\theta_1 = 50$, and $\theta_2/\theta_1 = 1.50$ and only 90% of the data are complete for a sample of size 40, then the bias of $\hat{\theta}_1$ can be obtained from the Table 1 as $50 \times 0.0198 = 0.990$. Therefore, it is not a very serious bias.

4. Exponential failure distributions, confidence intervals

In this section we propose four different confidence intervals for θ_1. The first one is based on the distribution of $\hat{\theta}_1$ under similar assumptions as those of Chen and Bhattacharya (1988) or Gupta and Kundu (1998). The second one is based on the asymptotic distribution of $\hat{\theta}_1$. We also propose to use the percentile bootstrap confidence interval and the bootstrap-t confidence interval and give their implementation procedures in this section.

4.1. Approximate confidence bound

First let us assume that θ_2 is known. Suppose $P_{\theta_1}[\hat{\theta}_1 \geqslant b]$ is a monotonically increasing function of θ_1, and let $b(\theta)$ be a function such that $\alpha/2 = P_{\theta_1}[\hat{\theta}_1 \geqslant b(\theta_1)]$. Then for $\theta_1 < \theta_1'$, we have

$$\frac{\alpha}{2} = P_{\theta_1'}[\hat{\theta}_1 \geqslant b(\theta_1')] = P_{\theta_1}[\hat{\theta}_1 \geqslant b(\theta_1)] \leqslant P_{\theta_1'}[\hat{\theta}_1 \geqslant b(\theta_1)].$$

Table 3
The tabulated values of $P_{\theta_1}[\hat{\theta}_1 \geqslant b]$, for $\theta_2 = 2,\ b = 1$ and when 10% of the data are incomplete

n	θ_1								
	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00
10	0.475	0.685	0.815	0.890	0.933	0.958	0.973	0.982	0.987
20	0.483	0.766	0.905	0.962	0.985	0.994	0.997	0.999	0.999
30	0.486	0.818	0.948	0.986	0.996	0.999	1.00	1.00	1.00
40	0.488	0.855	0.971	0.995	0.999	1.00	1.00	1.00	1.00
50	0.489	0.883	0.983	0.998	1.00	1.00	1.00	1.00	1.00

Table 4
The tabulated values of $P_{\theta_1}[\hat{\theta}_1 \geqslant b]$, for $\theta_2 = 2, b = 2$ and when 10% of the data are incomplete

n	θ_1								
	1.00	1.25	1.50	1.75	2.00	2.25	2.50	2.75	3.00
10	0.045	0.130	0.246	0.371	0.486	0.585	0.667	0.732	0.784
20	0.009	0.058	0.172	0.329	0.491	0.630	0.738	0.817	0.873
30	0.002	0.028	0.125	0.297	0.492	0.661	0.785	0.868	0.920
40	0.000	0.014	0.093	0.271	0.493	0.686	0.821	0.903	0.949
50	0.000	0.007	0.070	0.249	0.494	0.708	0.849	0.927	0.966

This implies that $b(\theta_1) < b(\theta_1')$, that is $b(\theta)$ is an increasing function. Thus $b^{-1}(\theta)$ exists and is also an increasing function. So we get $1-\alpha/2 = P_{\theta_1}[b^{-1}(\hat{\theta}_1) \leqslant \theta_1]$, which implies $b^{-1}(\hat{\theta}_1)$ is the lower bound of the $(1-\alpha)100\%$ confidence bound of θ_1. If $\hat{\theta}_{\text{obs}}$ denotes the observed value of $\hat{\theta}_1$ find $\theta_{\text{L}} = b^{-1}(\hat{\theta}_{\text{obs}})$ such that $\alpha/2 = P_{\theta_{\text{L}}}(\hat{\theta}_1 \geqslant \hat{\theta}_{\text{obs}})$, which is equivalent to finding θ_{L} such that $1 - \alpha/2 = P_{\theta_{\text{L}}}(\hat{\theta}_1 \leqslant \hat{\theta}_{\text{obs}})$. As $P_{\theta_1}[\hat{\theta}_1 \geqslant b]$ is a monotonically increasing function of θ_1, therefore $P_{\theta_1}[\hat{\theta}_1 \leqslant c]$ is a monotonically decreasing function of θ_1. Let $c(\theta)$ be a function such that $\alpha/2 = P_{\theta_1}[\hat{\theta}_1 \leqslant c(\theta_1)]$. It can be shown exactly as before that $c(\theta)$ is a decreasing function and therefore $c^{-1}(\theta)$ exists. Next, we need to find $\theta_{\text{U}} = c^{-1}(\hat{\theta}_{\text{obs}})$, such that $\alpha/2 = P_{\theta_{\text{U}}}(\hat{\theta}_1 \leqslant \hat{\theta}_{\text{obs}})$. Since it is not possible to obtain closed-form expressions of the functions $b(\theta)$ or $c(\theta)$, we need to use some iterative technique to get θ_{L} and θ_{U}. Note that we can get an exact $(1-\alpha)100\%$ confidence bound of θ_1, if we know θ_2. Since θ_2 is usually unknown, we need to estimate θ_2 and we will get an approximate $(1-\alpha)100\%$ confidence bound of θ_1. The construction of the confidence bound of θ_1 is based on the assumption that $P_{\theta_1}[\hat{\theta}_1 \geqslant b]$ is a monotonically increasing function of θ_1. We could not prove this assumption because of the complicated nature of the function. Although, some heuristic justification can be given as follows. Since $\hat{\theta}_1$ is the MLE of the mean life θ_1 due to cause 1, for fixed θ_2, the larger the parameter θ_1 is, the more probable it will be for its MLE to exceed a given value. Numerical values of $P_{\theta_1}[\hat{\theta}_1 \geqslant b]$ for various values of θ_1 and b support this conjecture.

Tables 3 and 4 indicate that $P_{\theta_1}[\hat{\theta}_1 \geqslant b]$ is an increasing function of θ_1.

4.2. Asymptotic confidence bound

In this subsection we first obtain the asymptotic distribution of $\hat{\lambda}_1$ and $\hat{\lambda}_2$, and using that we obtain the asymptotic distribution of $\hat{\theta}_1$ and $\hat{\theta}_2$. The Fisher information matrix of the parameters 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

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

So

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

$$I_{12}(\lambda_1,\lambda_2)=I_{21}(\lambda_1,\lambda_2)=\frac{n-m}{(\lambda_1+\lambda_2)^2},$$

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

The Fisher information matrix of θ_1 and θ_2, $\boldsymbol{I}(\theta_1,\theta_2)$, can be obtained easily from $\boldsymbol{I}(\lambda_1,\lambda_2)$ by the Jacobian transformation. The elements of the Fisher information matrix, $\boldsymbol{I}(\theta_1,\theta_2)=(I_{ij}(\theta_1,\theta_2))$ for $i,j=1$ and 2, are as follows:

$$I_{11}(\theta_1,\theta_2)=\frac{\theta_2(m\theta_1+n\theta_2)}{\theta_1^2(\theta_1+\theta_2)^2},$$

$$I_{12}(\theta_1,\theta_2)=I_{21}(\theta_1,\theta_2)=\frac{(n-m)}{(\theta_1+\theta_2)^2},$$

$$I_{22}(\theta_1,\theta_2)=\frac{\theta_1(m\theta_2+n\theta_1)}{\theta_2^2(\theta_1+\theta_2)^2}.$$

Therefore if $\theta=(\theta_1,\theta_2)$ and $\hat{\theta}=(\hat{\theta}_1,\hat{\theta}_2)$, then from asymptotic theory of the MLEs, see Miller (1981), we have

$$(\hat{\theta}-\theta)\rightarrow \mathrm{N}_2(\mathbf{0},\boldsymbol{I}^{-1}(\theta_1,\theta_2))$$

where $\boldsymbol{I}^{-1}(\theta_1,\theta_2)=(I_{ij}^{-1}(\theta_1,\theta_2))$ for $i,j=1$ and 2 and

$$I_{11}^{-1}(\theta_1,\theta_2)=\frac{\theta_1^2(m\theta_2+n\theta_1)}{mn\theta_2},$$

$$I_{12}^{-1}(\theta_1,\theta_2)=I_{21}^{-1}(\theta_1,\theta_2)=\frac{\theta_1\theta_2(n-m)}{mn},$$

$$I_{22}^{-1}(\theta_1,\theta_2)=\frac{\theta_2^2(m\theta_1+n\theta_2)}{mn\theta_1}.$$

Note that here $\boldsymbol{I}(\theta_1,\theta_2)$ is the Fisher information matrix for the whole sample. To obtain the confidence interval on θ_1, we substitute the true values of the parameters by the corresponding MLEs in the expression of $\boldsymbol{I}(\theta_1,\theta_2)$.

4.3. Bootstrap confidence intervals

In this subsection we use two bootstrap confidence intervals, the percentile bootstrap confidence intervals suggested by Efron (1982) and the bootstrap-t confidence intervals suggested by Hall (1988). Hall (1988) showed that the bootstrap-t confidence interval is better than the percentile bootstrap confidence intervals from an asymptotic point of view, although the finite sample properties are not yet known.

We use the following percentile bootstrap confidence interval of Efron (1982):

(1) From $(x_1,\delta_1),\dots,(x_m,\delta_m)$ obtain the bootstrap sample $(x_1^*,\delta_1^*),\dots,(x_m^*,\delta_m^*)$ by re-sampling with replacement and from the sample $(x_{m+1},*),\dots,(x_n,*)$ obtain the bootstrap sample $(x_{m+1}^*,*),\dots,(x_n^*,*)$ again by re-sampling with replacement.
(2) From the bootstrap sample $(x_1^*,\delta_1^*)\dots,(x_m^*,\delta_m^*)$, $(x_{m+1},*),\dots,(x_n,*)$ obtain the bootstrap estimates of all the unknown parameters. For any unknown parameter, say θ, denote the bootstrap estimate of θ as $\hat{\theta}^*$.
(3) Repeat the process 1–2 NBOOT times.
(4) Let $\widehat{\mathrm{CDF}}(t)=P_*(\hat{\theta}^*\leqslant t)$ be the cumulative distribution of $\hat{\theta}^*$, the bootstrap estimator of the parameter θ, then from the NBOOT $\hat{\theta}^*$ obtain the upper bound and the lower bound of the $100(1-2\alpha)\%$ bootstrap confidence bound for θ as follows. For a given α define $\hat{\theta}_{\text{boot}}(\alpha)=\widehat{\mathrm{CDF}}^{-1}(\alpha)$, then the approximate $100(1-2\alpha)\%$ confidence interval for θ is given by

$$(\hat{\theta}_{\text{boot}}(\alpha),\hat{\theta}_{\text{boot}}(1-\alpha)).$$

We use the following bootstrap-t confidence interval of Hall (1988):

(1) From $(x_1,\delta_1),\dots,(x_m,\delta_m)$ obtain the bootstrap sample $(x_1^*,\delta_1^*),\dots,(x_m^*,\delta_m^*)$ by re-sampling with replacement and from the sample $(x_{m+1},*),\dots,(x_n,*)$ obtain the bootstrap sample $(x_{m+1}^*,*),\dots,(x_n^*,*)$ again by re-sampling with replacement.
(2) From the bootstrap sample $(x_1^*,\delta_1^*),\dots,(x_m^*,\delta_m^*)$, $(x_{m+1},*),\dots,(x_n,*)$ obtain the bootstrap estimates of all the unknown parameters. For any unknown parameter, say θ, denote the bootstrap estimate of θ as $\hat{\theta}^*$.
(3) For any unknown parameter θ, compute

$$T^*=\frac{\sqrt{n}(\hat{\theta}^*-\hat{\theta})}{\hat{\sigma}^*},$$

where $\hat{\theta}$ is the MLE of θ and $\hat{\sigma}^*$ is the estimated standard error of $\hat{\theta}^*$.
(4) Repeat process 1–3 NBOOT times.
(5) From the NBOOT T^* obtain the upper bound and the lower bound of the $100(1-2\alpha)\%$ bootstrap-t confidence bound for θ as follows. Let $\widehat{\mathrm{CDFN}}(t)=P_*(\hat{\theta}^*\leqslant t)$ be the cumulative distribution of T^*, then for a given α define

$$\hat{\theta}_{\text{boot}-\text{t}}(\alpha)=\hat{\theta}+n^{-\frac{1}{2}}\hat{\sigma}\widehat{\mathrm{CDFN}}^{-1}(\alpha).$$

The approximate $100(1-2\alpha)\%$ confidence interval for θ is given by

$$(\hat{\theta}_{\text{boot}-\text{t}}(\alpha),\hat{\theta}_{\text{boot}-\text{t}}(1-\alpha)).$$

5. Weibull failure distributions

5.1. Estimation of the parameters

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

$$F_j(t) = (1 - \mathrm{e}^{-\lambda_j t^{\alpha}}). \tag{5.1}$$

We assume that the lifetime distribution of the different causes follow Weibull distributions with different scale parameters, but that they have the same shape parameter, which is quite a practical assumption, see for example Rao et al. (1991). We introduce one more parameter in the model; this gives more flexibility in the hazard rate unlike it being constant as in case of the exponential distribution. The hazard rate or the instantaneous death rate due to cause 1 is given by

$$\frac{\mathrm{d}F_1(t)}{\bar{F}_1(t)} = \alpha\lambda_1 t^{\alpha-1}$$

and the mean lifetime due to cause 1 is

$$E(X_{1i}) = \frac{1}{\lambda_1^{1/\alpha}}\Gamma\left(1 + \frac{1}{\alpha}\right).$$

It is well known that the hazard rate can be increasing or decreasing depending on whether $\alpha > 1$ or $\alpha < 1$. For $\alpha=1$, we have a constant hazard rate, that is an exponential distribution. The relative risk rate, π, due to cause 1 is

$$\pi = P(X_{1i} < X_{2i}) = \int_0^{\infty} \alpha_1\lambda_1 x^{\alpha-1}\mathrm{e}^{-\lambda_1 x^{\alpha}}\mathrm{e}^{-\lambda_2 x^{\alpha}}\,\mathrm{d}x = \frac{\lambda_1}{\lambda_1 + \lambda_2},$$

which is independent of the shape parameter α and is the same as the exponential case. The log likelihood function of the observed data as given in Section 3, becomes

$$\ln(L) = n\ln(\alpha) + r_1\ln(\lambda_1) + r_2\ln(\lambda_2) + (\alpha-1)\sum_{i=1}^{n}\ln(x_i) - (\lambda_1+\lambda_2)\sum_{i=1}^{n}x_i^{\alpha} + (n-m)\ln(\lambda_1+\lambda_2).$$

Then taking the derivatives with respect to the unknown parameters α, λ_1 and λ_2 and equating them to zeros, we get

$$\hat{\lambda}_1(\alpha) = \frac{n}{m}\frac{r_1}{\sum_{i=1}^{n}x_i^{\alpha}} \qquad \hat{\lambda}_2(\alpha) = \frac{n}{m}\frac{r_2}{\sum_{i=1}^{n}x_i^{\alpha}}.$$

We put $\hat{\lambda}_1$ and $\hat{\lambda}_2$ in the expression of $\ln(L)$ above and maximize with respect to α. We do not have any explicit expression for $\hat{\alpha}$. We obtain $\hat{\alpha}$ by maximizing

$$\ln[L(\alpha)] = n\ln(\alpha) + r_1\ln(\hat{\lambda}_1(\alpha)) + r_2\ln(\hat{\lambda}_2(\alpha)) + (\alpha-1)\sum_{i=1}^{n}\ln(x_i) - (\hat{\lambda}_1(\alpha)+\hat{\lambda}_2(\alpha))\sum_{i=1}^{n}x_i^{\alpha} + (n-m)\ln(\hat{\lambda}_1(\alpha)+\hat{\lambda}_2(\alpha))$$

with respect to α. Once we obtain $\hat{\alpha}$, we obtain the maximum likelihood estimators of λ_1 and λ_2 as $\hat{\lambda}_1(\hat{\alpha})$ and $\hat{\lambda}_2(\hat{\alpha})$, respectively. From the invariance principle of the MLEs we can say that the MLE of the relative risk rate due to cause 1 is

$$\hat{\pi} = \frac{\hat{\lambda}_1(\hat{\alpha})}{\hat{\lambda}_1(\hat{\alpha}) + \hat{\lambda}_2(\hat{\alpha})}$$

and also the MLE of the mean lifetime due to cause 1 is

$$\hat{\tau}_1 = \frac{1}{\hat{\lambda}_1^{1/\hat{\alpha}}} \Gamma\left(1 + \frac{1}{\hat{\alpha}}\right).$$

For known α the distribution of $\hat{\lambda}_1(\alpha)$ or $\hat{\lambda}_2(\alpha)$ can be obtained similarly as Section 3 simply by transforming the data. But if α is unknown then the exact distribution of $\hat{\lambda}_1(\hat{\alpha})$ or $\hat{\lambda}_2(\hat{\alpha})$ is not possible to obtain, so we have to rely on the asymptotic distribution only.

5.2. Confidence intervals

In this section we provide confidence intervals for the different parameters. Since exact confidence intervals are not possible to obtain when the shape parameter is unknown, we propose the asymptotic confidence intervals and also two different bootstrap confidence intervals. The asymptotic result can be stated as follows:

$$(\hat{\alpha} - \alpha, \hat{\lambda}_1 - \lambda_1, \hat{\lambda}_2 - \lambda_2) \rightarrow \mathrm{N}_3(\mathbf{0}, \boldsymbol{I}^{-1}(\alpha, \lambda_1, \lambda_2)).$$

Here $\boldsymbol{I}(\alpha, \lambda_1, \lambda_2)$ is the Fisher information matrix for the parameters $(\alpha, \lambda_1, \lambda_2)$. The matrix $\boldsymbol{I} = ((I_{ij}))$ for $i, j = 1, 2$ and 3 are as follows:

$$I_{11}(\alpha, \lambda_1, \lambda_2) = n\left[\frac{1}{\alpha^2} + (\lambda_1 + \lambda_2)V\right],$$

$$I_{12}(\alpha, \lambda_1, \lambda_2) = nU = I_{21}(\alpha, \lambda_1, \lambda_2),$$

$$I_{13}(\alpha, \lambda_1, \lambda_2) = nU = I_{31}(\alpha, \lambda_1, \lambda_2),$$

$$I_{22}(\alpha, \lambda_1, \lambda_2) = \frac{m\lambda_2 + n\lambda_1}{\lambda_1(\lambda_1 + \lambda_2)^2},$$

$$I_{33}(\alpha, \lambda_1, \lambda_2) = \frac{m\lambda_1 + n\lambda_2}{\lambda_2(\lambda_1 + \lambda_2)^2},$$

$$I_{23}(\alpha, \lambda_1, \lambda_2) = \frac{(n - m)}{(\lambda_1 + \lambda_2)^2} = I_{32}(\alpha, \lambda_1, \lambda_2).$$

Here $U = E(X^\alpha \ln(X))$ and $V = E(X^\alpha \ln(X).\ln(X))$, where X is distributed as Weibull $(\alpha, (\lambda_1 + \lambda_2))$. We propose to use bootstrap confidence intervals similar to the ones described in Section 4.3. Note that another way to obtain the approximate confidence interval for θ_1 or θ_2 is to use the exact distribution of $\hat{\theta}_1$ or $\hat{\theta}_2$ for known α and then replace α by its estimate. This is not pursued here.

6. Numerical experiments

In this section we present some numerical results to see how the different methods behave for small sample sizes and also for different parametric values. All these numerical works were performed at the University of New Brunswick using HP workstations. We use a method of Press et al. (1986) for random deviate generation. We consider cases when the lifetimes are exponential and also when the lifetimes are Weibull. We mainly observe the behavior of the MLEs in terms of their biases and in terms of their variances. We also compare the performances of the different proposed confidence intervals in terms of the coverage percentages and also in terms of their average confidence lengths.

6.1. Exponential case

In this subsection we present results when the lifetimes are exponential. Since for the exponential lifetime, the biases and the asymptotic variances are functions of θ_2/θ_1 only, we keep $\theta_1 = 1$ fixed, and consider different values of $\theta_2 = 1.25, 1.50, \ldots, 2.50$. We take sample sizes $n = 10, 20, 30, 40$ and 50. In all cases we assume 10% of the data are incomplete. We draw random samples for different values of n and θ_2 and compute the MLEs of θ_1 and θ_2. We also compute four different 95% confidence intervals, namely (1) asymptotic (Asymp.) (2) approximate (Approx.) (3) percentile Bootstrap (Boot-p) and (4) Bootstrap-t (Boot-t) confidence intervals. We replicate the process one thousand times and compute average values of the MLEs, the variances, the biases and the absolute biases ($|E(\hat{\theta}) - \theta|$). For different confidence intervals we compute the coverage percentages and also the average confidence lengths. The results of θ_1 and θ_2 are quite similar in nature so we present the results only for θ_1. The average values of $\hat{\theta}_1$, the variances, the absolute biases and the negative biases are reported in Table 5. The average error bounds (half of the confidence length) and the corresponding coverage percentages are reported within brackets for different methods in Table 6.

Some points are very clear from these experiments. From Table 5, it is clear that as sample size increases, the biases and the variances decrease. This suggests that the MLEs are asymptotically unbiased and consistent estimators of the corresponding parameters as it was indicated in Section 3. From Tables 5 and 3, it is observed that theoretical biases and simulated biases are quite close to each other. For fixed θ_1, as θ_2 increases, the biases and variances of $\hat{\theta}_1$ decrease and the corresponding biases and variances of $\hat{\theta}_2$ increase (not reported here). This is not very surprising, because as θ_2/θ_1 increases the mean life due to cause 1 decreases compared to the mean life due to cause 2. It is expected that as θ_2/θ_1 increases, the sample consists of more deaths due to cause 1 than due to cause 2. Therefore the sample has more information about θ_1 than θ_2.

From Table 6 it is clear that as sample size increases or θ_2/θ_1 increases, the average confidence lengths of θ_1 decrease for all four methods. In the case of θ_2 it is observed (not reported here) that as sample size increases the average confidence lengths

Table 5
Average values of $\hat{\theta}_1$, Variance of $\hat{\theta}_1$, bias and absolute bias of the LSEs when 10% of the data are incomplete. The lifetimes are exponential

n	Average	$\theta_2 = 1.25$	$\theta_2 = 1.50$	$\theta_2 = 1.75$	$\theta_2 = 2.00$	$\theta_2 = 2.25$	$\theta_2 = 2.50$
10	$\hat{\theta}_1$	1.1261	1.1018	1.0947	1.0741	1.0776	1.0615
	V($\hat{\theta}_1$)	0.4107	0.3459	0.2872	0.2409	0.2227	0.1902
	Bias	0.1261	0.1018	0.0947	0.0741	0.0776	0.0615
	|Bias|	0.4107	0.3801	0.3575	0.3343	0.3235	0.3056
20	$\hat{\theta}_1$	1.0489	1.0411	1.0340	1.0327	1.0303	1.0291
	V($\hat{\theta}_1$)	0.1226	0.1410	0.1010	0.0944	0.0894	0.0850
	Bias	0.0489	0.0411	0.0340	0.0327	0.0303	0.0291
	|Bias|	0.2602	0.2510	0.2409	0.2310	0.2273	0.2223
30	$\hat{\theta}_1$	1.0286	1.0266	1.0243	1.0228	1.0183	1.0173
	V($\hat{\theta}_1$)	0.0782	0.0680	0.0636	0.0581	0.0560	0.0523
	Bias	0.0286	0.0266	0.0243	0.0228	0.0183	0.0173
	|Bias|	0.2100	0.1986	0.1953	0.1854	0.1834	0.1782
40	$\hat{\theta}_1$	1.0264	1.0215	1.0134	1.0175	1.0095	1.0097
	V($\hat{\theta}_1$)	0.0564	0.0500	0.0440	0.0437	0.0402	0.0373
	Bias	0.0264	0.0215	0.0134	0.0175	0.0095	0.0097
	|Bias|	0.1816	0.1727	0.1647	0.1625	0.1558	0.1522
50	$\hat{\theta}_1$	1.0192	1.0151	1.0153	1.0116	1.0086	1.0078
	V($\hat{\theta}_1$)	0.0425	0.0388	0.0362	0.0339	0.0326	0.0298
	Bias	0.0192	0.0151	0.0153	0.0116	0.0086	0.0078
	|Bias|	0.1588	0.1534	0.1483	0.1438	0.1423	0.1360

decrease but as θ_2/θ_1 increases, the average confidence lengths increase. This is not very surprising as has been mentioned previously. Among different methods, it is clear that all methods work quite well if the sample size is large, and all of them are able to keep the nominal coverage percentage, although for small sample sizes, mainly for $n = 10$, all methods cannot maintain the nominal coverage percentage. It is observed that for all methods except the approximate one, the coverage percentages are slightly lower than the nominal level, whereas for the approximate method the coverage percentages are slightly higher than the nominal levels. Between the approximate and the asymptotic methods, the average lengths of confidence intervals using the asymptotic method are slightly lower than that of the approximate method. Comparing the percentile bootstrap method and the bootstrap-t method, it is observed that bootstrap-t is preferred in terms of the confidence lengths although their coverage percentages are almost equal in all the cases considered. Both the methods fail to maintain the coverage percentages at least for small samples. Now comparing the computational complexities, the asymptotic method is the easiest to implement. Approximate confidence intervals can be obtained by equating two nonlinear equations whereas the bootstrap confidence intervals can be obtained by re-sampling from the original sample. It is observed that bootstrap methods take longer times than the approximate method. Considering all these points it is recommended that for small sample sizes, approximate confidence

Table 6
Different confidence intervals of θ_1, when 10% of the date are incomplete. The lifetimes are exponential. The average length of the error bound and the corresponding coverage probability (within parenthesis) are reported. The nominal level is 95%

n	Methods	$\theta_2 = 1.25$	$\theta_2 = 1.50$	$\theta_2 = 1.75$	$\theta_2 = 2.00$	$\theta_2 = 2.25$	$\theta_2 = 2.50$
10	Asymp.	1.089 (0.917)	0.979 (0.900)	0.949 (0.898)	0.908 (0.897)	0.892 (0.867)	0.865 (0.847)
	Approx.	0.703 (0.928)	0.707 (0.938)	0.701 (0.925)	0.697 (0.931)	0.694 (0.910)	0.692 (0.909)
	Boot-p	1.528 (0.901)	1.442 (0.894)	1.298 (0.863)	1.219 (0.870)	1.160 (0.861)	1.166 (0.873)
	Boot-t	1.140 (0.898)	1.076 (0.891)	0.985 (0.873)	0.948 (0.867)	0.888 (0.862)	0.919 (0.868)
20	Asymp.	0.656 (0.917)	0.627 (0.923)	0.596 (0.908)	0.576 (0.911)	0.559 (0.940)	0.552 (0.918)
	Approx.	0.625 (0.952)	0.601 (0.962)	0.605 (0.940)	0.590 (0.962)	0.522 (0.956)	0.523 (0.955)
	Boot-p	0.969 (0.933)	0.841 (0.939)	0.774 (0.923)	0.683 (0.935)	0.691 (0.931)	0.668 (0.922)
	Boot-t	0.794 (0.935)	0.719 (0.931)	0.683 (0.926)	0.617 (0.925)	0.622 (0.928)	0.608 (0.924)
30	Asymp.	0.516 (0.913)	0.492 (0.918)	0.471 (0.928)	0.462 (0.933)	0.454 (0.934)	0.441 (0.938)
	Approx.	0.547 (0.952)	0.529 (0.952)	0.508 (0.960)	0.494 (0.956)	0.485 (0.948)	0.483 (0.960)
	Boot-p	0.645 (0.940)	0.604 (0.951)	0.561 (0.944)	0.535 (0.932)	0.514 (0.944)	0.502 (0.938)
	Boot-t	0.585 (0.939)	0.558 (0.951)	0.528 (0.934)	0.506 (0.928)	0.490 (0.945)	0.482 (0.940)
40	Asymp.	0.444 (0.928)	0.427 (0.944)	0.408 (0.924)	0.400 (0.946)	0.392 (0.940)	0.377 (0.928)
	Approx.	0.473 (0.950)	0.449 (0.964)	0.430 (0.953)	0.424 (0.963)	0.417 (0.953)	0.408 (0.956)
	Boot-p	0.524 (0.944)	0.503 (0.938)	0.466 (0.945)	0.452 (0.945)	0.435 (0.950)	0.431 (0.944)
	Boot-t	0.492 (0.935)	0.477 (0.941)	0.447 (0.941)	0.436 (0.939)	0.421 (0.944)	0.419 (0.942)
50	Asymp.	0.399 (0.946)	0.378 (0.931)	0.362 (0.948)	0.352 (0.933)	0.345 (0.931)	0.340 (0.939)
	Approx.	0.424 (0.966)	0.399 (0.963)	0.384 (0.949)	0.372 (0.948)	0.365 (0.957)	0.359 (0.967)
	Boot-p	0.466 (0.962)	0.438 (0.948)	0.415 (0.960)	0.401 (0.953)	0.393 (0.949)	0.386 (0.948)
	Boot-t	0.445 (0.965)	0.421 (0.947)	0.404 (0.961)	0.389 (0.955)	0.383 (0.949)	0.369 (0.948)

interval can be used whereas for large sample the asymptotic confidence bound is preferred.

6.2. Weibull case

For the Weibull lifetime distribution, we mainly consider the MLEs of the α's and λ's for different sample sizes. We consider $\alpha=1$, $\lambda_1=1$, $\lambda_2^{-1}=1.2$, 1.50, 1.75, 2.00, 2.25 and 2.50. We take $n = 10, 20, 30, 40$ and 50. For a particular choice of α, n and λ_2 we draw a random sample from the Weibull lifetime distribution and compute the MLEs of λ_1, λ_2 and α when 10% of the data are incomplete. We also compute three different 95% confidence bounds namely the asymptotic one, the percentile bootstrap and the bootstrap-t confidence intervals. We replicate the process one thousand times and compute the average biases, average absolute biases and the average variances of the MLEs over one thousand replications. The average variances, the negative biases and the absolute biases of λ_1 are reported in Table 7. The average error bounds and the coverage percentages (within brackets) are reported in Table 8. Since the results of λ_2 and α are quite similar to that of λ_1 they are not reported here.

Some points are very clear from this experiment. From Table 7 it is clear that as sample size increases the variances, biases and the absolute biases all decrease which

Table 7
Average values of $\hat{\lambda}_1$, Variance of $\hat{\lambda}_1$, bias and absolute bias of the LSEs when 10% of the date are incomplete. The lifetimes are Weibull

n	Average	$\lambda_2^{-1}=1.25$	$\lambda_2^{-1}=1.50$	$\lambda_2^{-1}=1.75$	$\lambda_2^{-1}=2.00$	$\lambda_2^{-1}=2.25$	$\lambda_2^{-1}=2.50$
10	$\hat{\lambda}_1$	1.2356	1.2138	1.1926	1.1994	1.1925	1.1385
	$V(\hat{\lambda}_1)$	0.8067	0.5599	0.6956	0.5110	0.8079	0.3775
	Bias	0.2356	0.2138	0.4926	0.1994	0.1925	0.1385
	|Bias|	0.5006	0.4530	0.4334	0.4315	0.4260	0.3742
20	$\hat{\lambda}_1$	1.0996	1.1116	1.0749	1.0844	1.0972	1.0781
	$V(\hat{\lambda}_1)$	0.1553	0.1318	0.1250	0.1278	0.1077	0.1241
	Bias	0.0996	0.1116	0.0794	0.0844	0.0972	0.0781
	|Bias|	0.2859	0.2769	0.2696	0.2649	0.2499	0.2556
30	$\hat{\lambda}_1$	1.0499	1.0426	1.0448	1.0488	1.0591	1.0357
	$V(\hat{\lambda}_1)$	0.0793	0.0713	0.0623	0.0661	0.0764	0.0627
	Bias	0.0499	0.0426	0.0448	0.0488	0.0591	0.0357
	|Bias|	0.2117	0.2030	0.1962	0.1982	0.2059	0.1925
40	$\hat{\lambda}_1$	1.0425	1.0429	1.0452	1.0312	1.0293	1.0250
	$V(\hat{\lambda}_1)$	0.0542	0.0502	0.0470	0.0428	0.0457	0.0410
	Bias	0.0425	0.0429	0.0452	0.0312	0.0293	0.0250
	|Bias|	0.1819	0.1755	0.1670	0.1608	0.1672	0.1591
50	$\hat{\lambda}_1$	1.0199	1.0274	1.0273	1.0267	1.0230	1.0180
	$V(\hat{\lambda}_1)$	0.0404	0.0408	0.0370	0.0372	0.0324	0.0310
	Bias	0.0199	0.0274	0.0273	0.0267	0.0230	0.0180
	|Bias|	0.1595	0.1577	0.1518	0.1502	0.1414	0.1412

Table 8
Different confidence intervals of λ_1, when 10% of the data are incomplete. The length of the confidence intervals and the corresponding coverage probability (within parenthesis) are reported. The lifetime distributions are Weibull

n	Methods	$\lambda_2^{-1}=1.25$	$\lambda_2^{-1}=1.50$	$\lambda_2^{-1}=1.75$	$\lambda_2^{-1}=2.00$	$\lambda_2^{-1}=2.25$	$\lambda_2^{-1}=2.50$
10	Asymp.	1.131 (0.929)	1.063 (0.932)	1.012 (0.925)	0.993 (0.931)	0.973 (0.893)	0.893 (0.890)
	Boot-p	1.282 (0.875)	1.273 (0.854)	1.176 (0.848)	1.136 (0.836)	1.087 (0.841)	1.026 (0.838)
	Boot-t	1.152 (0.865)	1.121 (0.840)	1.052 (0.834)	1.047 (0.832)	0.972 (0.831)	0.892 (0.806)
20	Asymp.	0.675 (0.952)	0.648 (0.969)	0.610 (0.932)	0.600 (0.947)	0.593 (0.948)	0.575 (0.939)
	Boot-p	1.057 (0.920)	0.926 (0.924)	0.896 (0.919)	0.882 (0.939)	0.824 (0.923)	0.802 (0.925)
	Boot-t	0.767 (0.902)	0.730 (0.920)	0.699 (0.911)	0.685 (0.947)	0.657 (0.921)	0.609 (0.913)
30	Asymp.	0.521 (0.938)	0.496 (0.941)	0.479 (0.956)	0.471 (0.953)	0.467 (0.937)	0.449 (0.942)
	Boot-p	0.694 (0.935)	0.634 (0.951)	0.605 (0.945)	0.598 (0.931)	0.572 (0.950)	0.569 (0.939)
	Boot-t	0.602 (0.938)	0.571 (0.943)	0.544 (0.938)	0.540 (0.943)	0.529 (0.944)	0.514 (0.941)
40	Asymp.	0.446 (0.951)	0.427 (0.952)	0.416 (0.952)	0.400 (0.948)	0.392 (0.947)	0.383 (0.943)
	Boot-p	0.554 (0.934)	0.527 (0.940)	0.497 (0.942)	0.493 (0.941)	0.476 (0.947)	0.462 (0.940)
	Boot-t	0.516 (0.929)	0.485 (0.942)	0.474 (0.937)	0.455 (0.943)	0.453 (0.939)	0.437 (0.945)
50	Asymp.	0.392 (0.957)	0.377 (0.952)	0.365 (0.943)	0.356 (0.941)	0.348 (0.950)	0.341 (0.959)
	Boot-p	0.470 (0.963)	0.453 (0.944)	0.428 (0.959)	0.423 (0.952)	0.407 (0.948)	0.406 (0.950)
	Boot-t	0.454 (0.942)	0.429 (0.940)	0.417 (0.939)	0.416 (0.946)	0.399 (0.947)	0.391 (0.948)

indicates the consistency of the MLEs even in the Weibull case. Comparing Tables 5 and 7 it is observed that the biases, variances and the absolute biases of $\hat{\theta}_1$ are less than the corresponding biases, variances and the absolute biases of $\hat{\lambda}_1$, although both are estimating the same parameter $\lambda_1 = \theta_1^{-1} = 1$ in this case. This is not very surprising because in the exponential case there are only two parameters to estimate whereas for the Weibull case there are three unknown parameters.

From Table 8, it is clear that none of the methods are able to maintain the coverage percentages for small sample sizes, although for large sample sizes all the three methods behave reasonably well. As far as the confidence lengths are concerned, the confidence lengths due to the asymptotic method have marginally smaller size than the other two methods and the coverage percentages are close to the nominal value, at least for large sample sizes. Therefore, for moderate or large sample sizes the asymptotic method can be used.

7. Data analysis

In this section we consider one real-life data set from Lawless (1982, p. 491). It consists of failure or censoring times for 36 appliances subjected to an automatic life test. Failures were classified into 18 different modes, though among 33 observed failures only 7 modes are present and only modes 6 and 9 appear more than once. We are mainly interested in failure mode 9. The data consist of two causes of failure, $\delta = 1$ (failure mode 9) and $\delta = 2$ (all other failure modes). The data are given below, consisting of the failure times and the cause of failure if available.

Data Set. (11,2), (35,2), (49,2), (170,2), (329,2), (381,2), (708,2), (958,2), (1062,2), (1167,1), (1594,2), (1925,1), (1990,1), (2223,1), (2327,2), (2400,1), (2451,2), (2471,1), (2551,1), (2565,*), (2568,1), (2694,1), (2702,2), (2761,2), (2831,2), (3034,1), (3059,2), (3112,1), (3214,1), (3478,1), (3504,1), (4329,1), (6367,*), (6976,1), (7846,1), (13403,*)

Here we have $n = 36$, $m = 33$, $r_1 = 17$, $r_2 = 16$, $\sum_{i=1}^{36} x_i = 99245$. Therefore using the exponential lifetime distribution, the ML estimators $\hat{\theta}_1 = 5351.45$ and $\hat{\theta}_2 = 5685.91$. The estimates of the mean life due to cause 1 and cause 2 become 5351.45 and 5685.91, respectively. The ML estimators of the relative risk rate due to cause 1 is $\hat{\pi} = 0.5152$ and due to cause 2 is $1 - \hat{\pi} = 0.4848$. The following 95% confidence intervals are obtained for θ_1 and θ_2 by using different methods.

Using the Weibull lifetime distribution, we obtain the estimates of $\hat{\theta}_1 = \hat{\lambda}_1^{-1} = 6980.22$, $\hat{\theta}_2 = \hat{\lambda}_2^{-1} = 7416.49$ and $\hat{\alpha} = 1.0321$. The 95% confidence band for α becomes (0.7625, 1.3016). Since this interval contains one, it may not be unreasonable to assume that the lifetimes are exponential. Since the simulation results indicate that the approximate method works quite well for the exponential case, we use the approximate confidence bands for the parameters θ_1 and θ_2, respectively, which are given in Table 9.

Table 9

Methods	θ_1		θ_2	
	LB	UB	LB	UB
Asymp.	2862.73	7840.16	2956.68	8415.14
Approx.	4451.45	6251.28	4785.91	6585.22
Boot-p	3683.07	7407.33	2464.83	9406.65
Boot-t	3534.64	6992.28	3421.01	7103.94

8. Conclusions

In this paper we consider estimation of the parameters of the competing risks model when the data may not be complete. We consider two different lifetime distributions of the competing causes, namely exponential and Weibull. We obtain the exact distribution of the MLEs of the mean lifetime when the lifetime distributions are assumed to be exponential. We propose approximate confidence bands for the mean lifetime and compare their performances with the asymptotic confidence bands and two other bootstrap confidence bands. It is observed that approximate confidence bands work quite well for the exponential case. When the lifetime distributions are Weibull, it is observed that MLEs behave reasonably well and as in the exponential case they also seem to provide consistent estimates of the unknown parameters. To obtain the confidence bounds for the unknown parameters the asymptotic results can be used for moderate or large sample sizes but for small sample sizes more work is needed.

Another important aspect which is not addressed here is the analysis when the competing risks may not be independent and when some of the causes of failure are not known. It is a difficult problem. One way it can be handled is through mixture model formulation as was suggested by Babu et al. (1992) or Kundu et al. (1992). It will be reported else where.

Acknowledgements

Part of the work was done when the first author was visiting the University of New Brunswick as a visiting associate professor. He would like to thank the members of the department for their kind hospitality. The authors would also like to thank the referees, the associate editor and Professor Nandini Kannan of the University of Texas at San Antonio, for very useful and constructive comments.

References

Babu, G.J., Rao, C.R., Rao, M.B., 1992. Non-parametric estimation of specific exposure rate in risk and survival analysis. J. Amer. Statist. Assoc. 87, 84–89.

Berkson, J., Elveback, L., 1960. Competing exponential risks with particular inference to the study of smoking lung cancer. J. Amer. Statist. Assoc. 55, 415–428.

Chen, S.M., Bhattacharya, G.K., 1988. Exact confidence bound for an exponential parameter hybrid censoring. Comm. Statist. Ser. A 17 (6), 1858–1870.
Cox, D.R., 1959. The analysis of exponentially distributed lifetimes with two types of failures. J. Roy. Statist B 21, 411–421.
David, H.A., Moeschberger, M.L., 1978. The Theory of Competing Risks. Griffin, London.
Dinse, G.E., 1982. Non-parametric estimation of partially incomplete time and types of failure data. Biometrics 38, 417–431.
Edwin, G.L., Savage, R.I., 1954. Tables of expected value of $1/X$ for positive Bernoulli and Poisson variables. J. Amer. Statist. Assoc. 49, 169–177.
Efron, B., 1967. The two sample problem with censored data. Proceedings of the Fifth Berkeley Symposium Math. Statist. Prob., pp. 831–853.
Efron, B., 1982. The Jackknife, the Bootstrap and Other Resampling Plans. CBMS-NSF Regional Conference Series in Applied Mathematics, Vol. 38, SIAM, Philadelphia.
Gupta, R.D., Kundu, D., 1998. Exact confidence bound for an exponential hybrid censoring. Comm. Statist. Ser. A 17 (6), 1858–1870.
Hall, P., 1988. Theoretical comparison of Bootstrap confidence intervals. Ann. Statist. 16 (3), 927–953.
Kaplan, E.L., Meier, P., 1958. Non-parametric estimation from incomplete observation. J. Amer. Statist. Assoc. 53, 457–481.
Kundu, D., Kannan, N., Mazumdar, M., 1992. Inference on risk rates based on mortality data under censoring and competing risks using parametric models. Biometrical J. 34 (3), 315–328.
Lawless, J.F., 1982. Statistical Models and Methods for Lifetime Data. Wiley, New York.
Mendenhall, W., Lehmann Jr., E.H., 1960. An approximation of the negative moments of the positive binomial useful in life testing. Technometrics 2, 227–242.
Miller Jr., R.G., 1981. Survival Analysis. Wiley, New York.
Miyawaka, M., 1982. Statistical analysis of incomplete data in competing risks model. J. Japanese Soc. Quality Control 12, 49–52.
Miyawaka, M., 1984. Analysis of incomplete data in competing risks model. IEEE Trans. Reliability Anal. 33 (4), 293–296.
Peterson Jr., A.P., 1977. Expressing the Kaplan–Meier estimator as a function of empirical survival functions. J. Amer. Statist. Assoc. 72, 854–858.
Press, W.H., Flannery, B.P., Teukolsky, S.A., Vetterling, W.T., 1986. Numerical Recipes; The Art of Scientific Computing. Cambridge University Press, Cambridge, U.K.
Rao, B.R., Talwalker, S., Kundu, D., 1991. Confidence intervals for the relative risk ratio parameters from survival data under a random epidemiologic study. Biometrical J. 33 (8), 959–984.