NONPARAMETRIC MAXIMUM LIKELIHOOD ESTIMATION IN CURE-RATE MODELS BASED ON UNCENSORED AND CENSORED DATA

FANG TAN

A Thesis
in
The Department
of
Mathematics and Statistics

Presented in Partial Fulfillment of the Requirements
for the Degree of Master of Science at
Concordia University
Montreal, Quebec, Canada

April 2006

Library and
Archives Canada

Published Heritage
Branch

395 Wellington Street
Ottawa ON K1A 0N4
Canada

Bibliothèque et
Archives Canada

Direction du
Patrimoine de l'édition

395, rue Wellington
Ottawa ON K1A 0N4
Canada

Your file Votre référence
ISBN: 0-494-14234-0
Our file Notre référence
ISBN: 0-494-14234-0

Nonparametric maximum likelihood estimation

in cure-rate models based on uncensored and censored data

Fang Tan

Abstract

In this thesis, we shall attempt to give the NPMLE of the event time distribution and cure-rate based on different types of uncensored and censored data. Cure-mixture model and hidden model are used extensively. We address the non-estimability of the cure-rate when no cures are actually observed, in the uncensored case and some important censoring models. A proof is also given for the almost sure convergence of $\sup_x F(x)$ to $(1-\pi)$, where $\sup_x F(x)$ is the supremum of the MLE of the underlying distribution function, and π is the true underlying cure-rate, for random censoring and interval censoring (case-1). We describe and illustrate the "max-min formula" derived by Groeneboom and Wellner (1992) for interval censoring (case-1), then modify it to get the MLE of the cure-rate under a cure-mixture model, when some cures are observed. We perform a simulation study to give some numerical results as well. Finally, we discuss a probable approach to find the NPMLE in interval censoring (case-2), as a problem for further research.

ACKNOWLEDGEMENTS

I would like to express my deepest gratitude to Dr. Arusharka Sen for his invaluable guidance and encouragement. This paper was improved tremendously by his thoughtful comments and constructive suggestions.

I am thankful to Dr. Yogendra P. Chaubey, Dr. Xiao-Wen Zhou and Dr.Asgharian Masoud (McGill University), from whom I got the valuable knowledge to complete this thesis.

I am grateful to everyone in the Department of Mathematics and Statistics at Concordia University, who helped me getting through tough time last spring. I appreciate their psychological and financial support.

Finally, I wish to thank my family, my friends and my boyfriend, Bin-Qian Song, for being supportive all along.

DEDICATION

I would like to dedicate this thesis to my mother Mrs. Jin-Song Jing and my father Mr. Wei-Zhong Tan.

TABLE OF CONTENTS

Chapter 1 Maximum likelihood method and cure-rate models

1.1 Introduction

Medical data sets offer many challenges. Time-to-event data present themselves in different ways which create special problems in analyzing such data. One peculiar feature, often present in time-to-event data, is known as censoring, which, broadly speaking, occurs when an individual's life length is only known to belong to a certain interval of time. The analysis of survival experiments is usually complicated by the issue of censoring. In this thesis, we shall consider Type-I, Type-II, Random and Interval censoring.

The survival function is perhaps the most important function in medical and health studies, which is the probability of an individual surviving beyond some time point$x \geq 0$. It is defined as $S(x) = P\{X > x\}$, the complement of the cumulative distribution function, i.e., $S(x) = 1 - F(x)$, where $F(x) = P\{X \leq x\}$. Hence, if X is a continuous random variable, the derivative of the survival function is found as $S^{'}(x) = -f(x)$. Note that the survival function is non-negative and non-increasing with $S(0) = 1$ and $S(\infty) = \lim_{x \to \infty} S(x) = 0$. In some cases, we consider $S(\infty) = \pi > 0$, where there are some ultimate survivors.

Survival models that incorporate 'immune' or 'cured' individuals (so called Cure-rate models) have been used in the biostatistical area for several decades. A 'cured' individual means one who is not subject to the event under study; such an event is like death, contraction of a disease, or return to prison, and so on. Areas of interest are such like Recidivism, Market penetration, Engineering reliability, Fisheries research, and Education theory.

We consider cure-mixture model and hidden model to analyze time-to-event data, with a nonparametric maximum likelihood estimation procedure.

1.2 Nonparametric maximum likelihood method

The maximum likelihood method at this point, is by far the most appropriate analysis method for censored data.

There are various categories of censoring, such as right, left and interval censoring. Each type will lead to a different likelihood function which will be the basis for the inference. Though the likelihood function is unique for each type of censoring, there is a common approach to be used in constructing it.

Let X be the time until some specific event, such as death, the appearance of a tumor, the development of some disease, recurrence of a disease, equipment breakdown, and so forth. More precisely, X is a nonnegative random variable from a homogeneous population. Some functions characterizing the distribution of X will be used in this thesis, namely, the survival function, which is the probability of an individual surviving beyond time x, and the probability density function, which is the unconditional probability of the event occurring at time x.

When constructing likelihood functions, a critical assumption is that the lifetimes and censoring times are independent. An observation corresponding to an exact event time provides information on the probability that the event occurs at this time which is proportional to the density function of X at this time. For a right-censored observation, all we know is that the event time is larger than this time, so that the information is the survival function evaluated at the on study time. Similarly, for a left-censored observation, all we know is that the event has already occurred, so that the contribution to the likelihood is the cumulative distribution function evaluate at the on study time. Finally, for interval-censored data, we know only that the event occurred within the interval so that the information is the probability that the event time is in this interval. More specifically, the likelihoods for various types of censoring schemes may all be written by incorporating the following components:

exact lifetimes	-	$f(x)$
right-censored observations	-	$S(C_r)$
left-censored observations	-	$1-S(C_l)$
interval-censored observations	-	$S(L_i)-S(R_i)$

The likelihood may be constructed by factioning all these terms together

$$L \propto \prod_{i\in D} f(x_i) \prod_{i\in R} S(C_r) \prod_{i\in L} (1-S(C_l)) \prod_{i\in I} [S(L_i)-S(R_i)] \qquad (1.2.1)$$

where D is the set of observed death times, R the set of right-censored observations, L the set of left-censored observations, and I the set of interval-censored observations. If there exists a vector x^* such that $L(x^*) \geq L(x)$ for all possible choices of x, then x^* is considered as the maximum likelihood estimator.

1.3 Cure-mixture model and hidden model

1.3.1 Mixture model

Models for survival analysis typically assume that everybody in the study population is susceptible to the event under study and will eventually experience this event if the follow-up is sufficiently long. However, there are situations when a fraction of individuals are not expected to experience the event of interest; that is, those individuals are cured or insusceptible. For example, researchers may be interested in analyzing the recurrence of a disease. Many individuals may never experience a recurrence; therefore, a cured fraction of the population exists.

Survival models that incorporate a surviving fraction are so-called Cure-Rate Models. Historically, cure-rate models have been utilized to estimate the cure fraction. Cure models are survival models which allow for a cured fraction, π, of individuals, i.e., $S(\infty) = \pi$. These models extend the understanding of time-to-event data by allowing for the formulation of more accurate and informative conclusions. These conclusions are otherwise unobtainable from an analysis which fails to account for a cured or insusceptible fraction of the population. If a cured component is not present, the analysis reduces to the standard approach in survival analysis. The use of cure models has been popular for joint modeling of the overall risk of a disease and the age-at-onset distribution of the diseased individuals (e.g. Farewell 1977, 1982; Kuk and Chen 1992).

In cure models, we use "cure fraction" and "insusceptible fraction" as interchangeable notions. The population is divided into two sub-populations so that an individual either is cured with probability π, or has a proper survival function $S_0(t)$, with probability $1-\pi$. A model for the distribution of survival times that incorporates a cured fraction is thus given by $S_\pi(t) = \pi + (1-\pi) \cdot S_0(t)$. Traditional cure models assume that those individuals experiencing the event of interest are homogeneous in risk. During the last fifteen years, extensions of cure models were developed in order to allow for heterogeneity among the fraction under risk by using frailty models where the frailty distribution is a mixture of a discrete and a continuous part (e.g. Aalen 1988, 1992; Longini and Halloran 1996).

Suppose that X is a random variable denoting the life time of an individual with the cure-rate π. Let F_0 be the cumulative distribution function for the uncured individuals, and F_π be the cumulative distribution function for the whole population (both cured and uncured individuals). Then we have, for $0 \le \pi < 1$ and $t > 0$:

$$\pi = P\{X = \infty\} = \lim_{t\to\infty} P\{X > t\} = 1 - \lim_{t\to\infty} F_\pi(t) \qquad (1.3.1)$$

$$F_\pi(t) = (1-\pi)F_0(t) \Leftrightarrow F_0(t) = \frac{F_\pi(t)}{1-\pi} = P\{X \le t \mid X < \infty\} \qquad (1.3.2)$$

$$S_\pi(t) = 1 - F_\pi(t) = \pi + (1-\pi)(1-F_0(t)) \qquad (1.3.3)$$

1.3.2 Hidden model

The hidden model, motivated by a biological application, arises as follows (e.g. Chen et al 1999):

Let N be the number of cancer cells, suppose that $N \sim Poisson(\theta)$, i.e., $P\{N = t\} = \dfrac{\theta e^{-\theta}}{t!}, t = 0, 1, \cdots$; X_i be the life time of the ith cancer cell, suppose that $X_i \overset{i.i.d.}{\sim} \exp(\theta)$, i.e., $p(x_i) = \theta e^{-\theta x_i}, x_i > 0$.

Define
$$X = \begin{cases} \min\limits_{1\le i\le N} X_i, & \text{if } N \ge 1 \\ \infty, & \text{if } N = 0 \end{cases}$$

$$\begin{aligned}\Rightarrow P\{X > t\} &= P\{N=0\} + \sum_{n=1}^{\infty} P\{\min_{1\le i\le n} X_i > t, N=n\}, t \ge 0\\ &= e^{-\theta} + \sum_{n=1}^{\infty} e^{-\theta}\cdot\frac{\theta^n}{n!}\cdot(1-F_0(t))^n\\ &= e^{-\theta} + e^{-\theta}\cdot[e^{\theta(1-F_0(t))}-1]\\ &= e^{-\theta\cdot F_0(t)}\end{aligned}$$

which yields $$S_\theta(t) = e^{-\theta\cdot F_0(t)}, t \ge 0. \qquad (1.3.4)$$

This model is suitable for any type of failure data that has a surviving fraction (cure-rate). It can be used in modeling various types of failure time data, such as time to relapse, time to death, time to first infection and so forth.

By taking the first derivative of (1.3.4), we have

$$S_\theta'(t) = -\theta\cdot f_0(t)\cdot e^{-\theta\cdot F_0(t)}, t > 0. \qquad (1.3.5)$$

Note that model (1.3.4) is not a proper survival function, since $S_\theta(\infty) = e^{-\theta}$. This also means that the cure-rate is given by

$$S_\theta(\infty) = P\{N=0\} = e^{-\theta}. \qquad (1.3.6)$$

Meanwhile, in the cure-mixture model, the cure-rate is given by π, hence there is a relationship between the two models:

$$e^{-\theta} = \pi \qquad (1.3.7)$$

$$S_\pi(x) = \pi + (1-\pi)\cdot(1-F_0(x))$$
$$S_\theta(x) = e^{-\theta} + (1-e^{-\theta})(1-F_0(x)) = e^{-\theta\cdot F_0(x)}$$

$$\Rightarrow F_\theta(x) = \frac{1-e^{-\theta\cdot F_0(x)}}{1-e^{-\theta}} \text{ and } S_\theta(x) = \frac{e^{-\theta\cdot F_0(x)}-e^{-\theta}}{1-e^{-\theta}} \qquad (1.3.8)$$

under the hidden model, $$\Delta F_\theta(x_i) = [1-e^{-\theta\cdot F_0(x_i)}]-[1-e^{-\theta\cdot F_0(x_{i-1})}], \qquad (1.3.9)$$

while under the mixture model, $$\Delta F_\pi(x) = (1\text{-}\pi)\cdot\Delta F_0(x). \qquad (1.3.10)$$

Both models are computationally attractive. More importantly, when using the nonparametric maximum likelihood method to estimate the survival function and cure-rate, the cure-mixture model and hidden model are equivalent. In this thesis, we use one or the other model depending on convenience.

Chapter 2 Preliminary results

In this chapter, to better process the maximum likelihood estimation, we apply the hidden model to non-censored data, and the mixture model for censoring models. We can arrive at explicit results for the NPMLE of the cure-rate in each case. All the results presented in the chapter are original work.

2.1 NPMLE of non-censored data, when observing both cures and non-cures, with known distribution

In this uncensored case, we apply the hidden model. Suppose F_0 is known, $X_1,\cdots X_n \overset{i.i.d.}{\sim} S_\theta$, and $S_\theta(t) = e^{-\theta\cdot F_0(t)}, t \geq 0$.

When both cured and non-cured individuals are observed, define the indicator of whether an individual belongs to a cure or not as follows

$$\delta_i = \begin{cases} 1, & \text{if}\, X_i < \infty \\ 0, & \text{if}\, X_i = \infty. \end{cases} \tag{2.1.1}$$

Hence the p.d.f. of X_i is $f_\theta(x_i) = \left[\theta f_0(x_i)\cdot e^{-\theta F_0(x_i)}\right]^{\delta_i} \cdot \left[e^{-\theta}\right]^{1-\delta_i}, x_i > 0.$ (2.1.2)

The likelihood function is given by

$$L = \prod_{i=1}^{n}\left[(\theta f_0(x_i)\cdot e^{-\theta\cdot F_0(x_i)})^{\delta_i} \cdot e^{-\theta\cdot(1-\delta_i)}\right]$$

$$= \prod_{i=1}^{n}[f_0(x_i)]^{\delta_i} \cdot e^{-\theta\cdot\sum_{i=1}^{n}[\delta_i\cdot F_0(x_i)+(1-\delta_i)]} \cdot \theta^{\sum_{i=1}^{n}\delta_i}. \tag{2.1.3}$$

Hence, the log-likelihood function is given by

$$\ln L = \ln C - \theta\cdot\sum_{i=1}^{n}\left[\delta_i\cdot F_0(x_i) + (1-\delta_i)\right] + \sum_{i=1}^{n}\delta_i\cdot\ln\theta, \tag{2.1.4}$$

where $\ln C = \sum_{i=1}^{n} \delta_i \cdot \ln[f_0(x_i)]$

To maximize the log-likelihood function, take the first derivative of (2.1.4) and equate it to zero, then solve for θ:

$$\frac{\partial \ln(L)}{\partial \theta} = -\sum_{i=1}^{n} [\delta_i \cdot F_0(x_i) + (1-\delta_i)] + \sum_{i=1}^{n} \delta_i \cdot \frac{1}{\theta} \tag{2.1.5}$$

$$\frac{\partial \ln(L)}{\partial \theta} = 0 \Rightarrow \hat{\theta} = \frac{\sum_{i=1}^{n} \delta_i}{\sum_{i=1}^{n} [\delta_i \cdot F_0(x_i) + (1-\delta_i)]} \tag{2.1.6}$$

Equivalently, the cure-rate is given by

$$\hat{\pi} = e^{-\hat{\theta}} = \exp\left\{ -\frac{\sum_{i=1}^{n} \delta_i}{\sum_{i=1}^{n} [\delta_i \cdot F_0(x_i) + (1-\delta_i)]} \right\} \tag{2.1.7}$$

2.2 NPMLE of uncensored data, when observing both cures and non-cures, with unknown distribution

When considering the distribution of life time is unknown, the estimation procedure becomes more complicated. Let the indicator of cures be as in (2.1.1) and apply the hidden model

$$F_\theta(t) = 1 - e^{-\theta \cdot F_0(t)}, S_\theta(t) = e^{-\theta \cdot F_0(t)} \tag{2.2.1}$$

Take the derivative of $S_\theta(t)$, then the probability distribution of X_i is

$$f_\theta(x_i, \delta_i) = \left[\theta \cdot f_0(x_i) \cdot e^{-\theta \cdot F_0(x_i)}\right]^{\delta_i} \cdot (e^{-\theta})^{1-\delta_i}, x_i > 0, \tag{2.2.2}$$

where the term $(\theta \cdot f_0(x_i) \cdot e^{-\theta \cdot F_0(x_i)})^{\delta_i}$ is for an observation which is not a cure

($\delta_i = 1$); and the term $(e^{-\theta})^{1-\delta_i}$ is for observation which is a cure ($\delta_i = 0$).

The likelihood function is hence found from (1.3.9) as

$$L(\theta, F_0) = \prod_{i=1}^{n} [\Delta F_\theta(x_i)]^{\delta_i} \cdot [e^{-\theta}]^{1-\delta_i}$$

$$= \prod_{i=1}^{n} [(1-e^{-\theta \cdot F_0(x_i)}) - (1-e^{-\theta \cdot F_0(x_{i-1})})]^{\delta_i} \cdot [e^{-\theta}]^{1-\delta_i} \quad (2.2.3)$$

Let $y_i = F_0(x_i)$

$$\Rightarrow L(\theta, F_0) = L(\theta, y_1, \cdots, y_n)$$

$$= \prod_{i=1}^{n} [(1-e^{-\theta \cdot y_i}) - (1-e^{-\theta \cdot y_{i-1}})]^{\delta_i} \cdot [e^{-\theta}]^{1-\delta_i} \qquad (2.2.4)$$

where $\quad 0 \le y_1 \le \cdots \le y_n \le 1$ and $k = \sum_{i=1}^{n} \delta_i$

$$\Rightarrow \quad 0 \le y_1 \le \cdots \le y_k, y_{k+1} = \cdots = y_n = 1$$

Furthermore, let $z_i = 1 - e^{-\theta \cdot y_i}, 1 \le i \le k$, set $z_0 = 0$. Since $0 \le y_i \le 1$ for any i, we have

$$0 \le z_1 \le z_2 \le \cdots \le z_k \le 1 - e^{-\theta} \qquad (2.2.5)$$

and (2.2.4) becomes

$$L(\theta, z_1, \cdots, z_k) = \prod_{i=1}^{k} (z_i - z_{i-1}) \cdot (e^{-\theta})^{n-k} \qquad (2.2.6)$$

Let $t_i = \frac{z_i}{1-e^{-\theta}}$, set t_0=0 $\Rightarrow$ $z_i = (1-e^{-\theta}) \cdot t_i, i = 0, \cdots, n,$

$$\Rightarrow L(\theta, t_1, \cdots, t_k) = (1-e^{-\theta})^k \cdot (e^{-\theta})^{n-k} \cdot \prod_{i=1}^{k} (t_i - t_{i-1}) \qquad (2.2.7)$$

When θ is fixed, the maximum value of the likelihood function (2.2.7) only depends on the value of $\prod_{i=1}^{k} (t_i - t_{i-1})$. Observe that

$$\prod_{i=1}^{k}(t_i - t_{i-1}) = t_1 \cdot (t_2 - t_1) \cdot (t_3 - t_2) \cdots (t_k - t_{k-1})$$

where $t_0 \overset{def}{=} 0,\ 0 \le t_1 \le t_2 \le \cdots \le t_k \le 1$

Now $$\begin{cases} t_{k-1} = \alpha_{k-1} \cdot t_k \ , & 0 \le \alpha_{k-1} \le 1 \\ t_{k-2} = (\alpha_{k-2} \cdot \alpha_{k-1}) \cdot t_k \ , & 0 \le \alpha_{k-2} \le 1 \\ \quad \vdots & \\ t_2 = (\alpha_2 \cdot \alpha_3 \cdot \alpha_4 \cdots \alpha_{k-1}) \cdot t_k \ , & 0 \le \alpha_2 \le 1 \\ t_1 = (\alpha_1 \cdot \alpha_2 \cdot \alpha_3 \cdots \alpha_{k-1}) \cdot t_k \ , & 0 \le \alpha_1 \le 1 \end{cases}$$

$$\begin{aligned} \Rightarrow \prod_{i=1}^{k}(t_i - t_{i-1}) &= (\alpha_1 \cdot \alpha_2 \cdot \cdots \cdot \alpha_{k-1}) \cdot (\alpha_2 \cdot \alpha_3 \cdot \cdots \cdot \alpha_{k-1}) \cdot (1-\alpha_1) \cdot \\ &\quad (\alpha_3 \cdot \alpha_4 \cdot \cdots \cdot \alpha_{k-1}) \cdot (1-\alpha_2) \cdot (\alpha_4 \cdot \alpha_5 \cdot \cdots \cdot \alpha_{k-1}) \cdot (1-\alpha_3) \cdot \\ &\quad \cdots \cdot \alpha_{k-1} \cdot (1-\alpha_{k-2}) \cdot (1-\alpha_{k-1}) \cdot t_k^{\,k} \\ &= t_k^{\,k} \cdot \left[\prod_{i=1}^{k-1}(1-\alpha_i)\right] \cdot \left[\prod_{i=1}^{k-1}\alpha_i^{\,i}\right] \\ &= t_k^{\,k} \cdot \left[\prod_{i=1}^{k-1}(1-\alpha_i) \cdot \alpha_i^{\,i}\right] \end{aligned} \tag{2.2.8}$$

where $0 \le \alpha_i \le 1$ for $1 \le i \le k-1$.

Hence the MLEs are

$$\hat{t}_k = 1 \ \text{ and } \ \hat{\alpha}_i = \frac{i}{i+1}, \text{ for } 1 \le i \le k-1 \tag{2.2.9}$$

Consequently,

$\hat{t}_{k-1} = \frac{k-1}{k},\ \hat{t}_{k-2} = \frac{k-1}{k} \cdot \frac{k-2}{k-1} = \frac{k-2}{k},\ \cdots\cdots,\ \hat{t}_1 = \frac{1}{k}$, for $k = 1, \cdots, n$, i.e.,

$$\hat{t}_i = \frac{i}{k}, \text{ for } i = 1, 2, \cdots, k \tag{2.2.10}$$

With the estimator of θ,

$$\hat{z}_i = (1 - e^{-\hat{\theta}}) \cdot \frac{i}{k} \ , \text{ for } i = 1, 2, \cdots, k \tag{2.2.11}$$

And $z_i = 1 - e^{-\theta \cdot y_i},\ 1 \le i \le k$

$$\Rightarrow F_0(x_i)=\hat{y}_i = -\frac{1}{\hat{\theta}}\cdot \ln\left[1-\frac{i}{k}\cdot\left(1-e^{-\hat{\theta}}\right)\right],\quad i=1,2,\cdots,k \tag{2.2.12}$$

Notice that

$$S_{\hat{\theta}}(\infty) = e^{-\hat{\theta}\cdot F_0(\infty)} = e^{-\hat{\theta}},$$

implying

$$P(non-cured) = 1-e^{-\hat{\theta}}. \tag{2.2.13}$$

Also

$$P(non-cured) = P(X<\infty) = \frac{k}{n}, \tag{2.2.14}$$

hence

$$1-e^{-\hat{\theta}} = \frac{k}{n} = \frac{1}{n}\cdot\sum_{i=1}^{n}\delta_i,$$

which yields

$$\hat{\theta} = -\ln(1-\sum_{i=1}^{n}\delta_i/n). \tag{2.2.15}$$

Equivalently, the NPMLE of the cure-rate is

$$\hat{\pi} = 1-\sum_{i=1}^{n}\delta_i/n. \tag{2.2.16}$$

2.3 NPMLE of Type-I censoring with no cures

For convenience, we apply mixture model in this section and the following one. We will consider Type-I censoring where the event is observed only if it occurs prior to some pre-specified time. These censoring times may vary from individual to individual.

Data from experiments involving right censoring can be conveniently represented by pairs of random variables (T,δ), where δ indicates whether the lifetime X is observed $(\delta=1)$ or not $(\delta=0)$, and T is equal to X if the lifetime is observed and to C_r if it is right-censored, i.e., $T=\min(X,C_r)$.

Details of constructing the likelihood function for Type-I censoring are as follows.

For $\delta = 0$, it can be seen that

$$\begin{aligned} P\{T, \delta = 0\} &= P\{T = C_r \mid \delta = 0\} \cdot P\{\delta = 0\} \\ &= P\{\delta = 0\} = P\{X > C_r\} = S(C_r), \end{aligned}$$

Also, for $\delta = 1$,

$$\begin{aligned} P\{T, \delta = 1\} &= P\{T = X \mid \delta = 1\} \cdot P\{\delta = 1\} \\ &= P\{X = T \mid X \le C_r\} \cdot P\{X \le C_r\} \\ &= \left[\frac{f(t)}{1 - S(C_r)}\right] \cdot [1 - S(C_r)] = f(t). \end{aligned}$$

Combining these expressions, when we have a random sample of pairs (T_i, δ_i), $i = 1, \cdots, n$, the likelihood function is

$$L = \prod_{i=1}^{n} P\{t_i, \delta_i\} = \prod_{i=1}^{n} [f(t_i)]^{\delta_i} [S(t_i)]^{1-\delta_i}. \qquad (2.3.1)$$

Define the indicator $\delta_i = 1\{X_i \le C_r\}$, $X_1, \ldots, X_n \overset{i.i.d.}{\sim} F_\pi(x)$, where $F_\pi(t) = (1 - \pi) \cdot F_0(t)$ (under mixture model), $F_0(t)$ is some appropriate distribution function and X_i is observed if and only if $X_i \le C_r$, where C_r is some pre-specified number. Let $T_i = X_i \wedge C_r$, if $X_i > C_r$, then the individual is a survivor, whose event time is censored at C_r.

As stated before, the likelihood function is

$$\begin{aligned} L(\pi, F_0) &= \prod_{i=1}^{n} [\Delta F_\pi(x_i)]^{\delta_i} \cdot [1 - F_\pi(C_r)]^{1-\delta_i} \\ &= \prod_{i=1}^{n} [\Delta F_\pi(t_i)]^{\delta_i} \cdot [1 - F_\pi(t_i)]^{1-\delta_i} \end{aligned} \qquad (2.3.2)$$

$$\Rightarrow L(\pi, F_o) = \prod_{i=1}^{n} \left[(1 - \pi) \cdot \left(F_0(t_i) - F_0(t_{i-1})\right)\right]^{\delta_i} \cdot \left[1 - (1 - \pi) \cdot F_0(t_i)\right]^{1-\delta_i} \qquad (2.3.3)$$

Put $y_i = F_0(t_i)$, $z_i = (1 - \pi) \cdot y_i$, set $z_0 = 0$. Since $T_1, \cdots, T_n$ are ordered statistics, and $0 \le F_0(t_i) \le 1$, we have $0 \le z_1 \le \cdots \le z_n \le 1 - \pi$

$$\Rightarrow L(\pi,\tilde{z}) = \prod_{i=1}^{n} (z_i - z_{i-1})^{\delta_i} \cdot (1-z_i)^{1-\delta_i} \tag{2.3.4}$$

Let $X_1,\ldots,X_n$ be ordered statistics, where *k* of them are not censored, i.e., $T_i = X_i$, for *i=1,2,...,k*; and *n-k* of them are censored, i.e., $T_{k+1} = \cdots = T_n = C_r$; implying that $k = \sum_{i=1}^{n} \delta_i$, and $z_{k+1} = \cdots z_n = (1-\pi)\cdot F_0(C_r)$.

$$\Rightarrow L(\pi,\tilde{z}) = \prod_{i=1}^{k} (z_i - z_{i-1}) \cdot (1-z_{k+1})^{n-k} \tag{2.3.5}$$

where define $z_0 = 0$.

Firstly, to maximize the product $\prod_{i=1}^{k}(z_i - z_{i-1})$, follow the same approach as shown in Section 2.2, we get the MLEs as $\hat{\alpha}_i = \frac{i}{i+1}$, for $1 \le i \le k-1$, i.e.,

$$\hat{z}_i = \frac{i}{k}\cdot \hat{z}_k, \text{ for } 1 \le i \le k-1 \tag{2.3.6}$$

Equation (2.3.5) becomes

$$L(\pi,\tilde{z}) = \frac{1}{k^k}\cdot {z_k}^k \cdot \left(1-z_{k+1}\right)^{n-k} \tag{2.3.7}$$

$$\begin{aligned}
&\underset{\substack{0\le\pi\le1,\\ 0\le z_1\le\cdots\le z_n\le 1-\pi}}{Max} \left\{ {z_k}^k \cdot \left(1-z_{k+1}\right)^{n-k} \right\} \\
&\le \underset{\substack{0\le\pi\le1,\\ 0\le z_1\le\cdots\le z_n\le 1-\pi}}{Max} \left\{ {z_k}^k \cdot \left(1-z_{k}\right)^{n-k} \right\} \\
&= Max \begin{cases} \underset{0\le\pi\le1}{Max}\left\{ \left(\frac{k}{n}\right)^k \cdot \left(1-\frac{k}{n}\right)^{n-k} \right\}, & \text{if } \pi \le 1-\frac{k}{n} \\ \underset{0\le\pi\le1}{Max}\left\{ (1-\pi)^k \cdot \pi^{n-k} \right\}, & \text{if } \pi > 1-\frac{k}{n} \end{cases} \\
&= Max \begin{cases} \left(\frac{k}{n}\right)^k \cdot \left(1-\frac{k}{n}\right)^{n-k}, & \text{if } \pi \le 1-\frac{k}{n} \\ \left(\frac{k}{n}-\varepsilon\right)^k \cdot \left(1-\frac{k}{n}+\varepsilon\right)^{n-k}, & \text{with arbitrary } \varepsilon>0, \ \text{if } \pi > 1-\frac{k}{n} \end{cases}
\end{aligned}$$

$$=\left(\frac{k}{n}\right)^{k}\cdot\left(1-\frac{k}{n}\right)^{n-k}, \text{ with } \hat{\pi}=1-\frac{k}{n} \text{ and } \hat{z}_k=\frac{k}{n}$$

Which implies $$\hat{z}_i=\begin{cases}\frac{i}{n}, & i=1,\cdots,k-1\\ \frac{k}{n}, & i=k,\cdots,n\end{cases} \qquad (2.3.8)$$

and $$\hat{y}_i=F_0(t_i)=\frac{\hat{z}_i}{1-\hat{\pi}}=\begin{cases}\frac{i}{k}, & i=1,\cdots,k-1\\ 1, & i=k,\cdots,n\end{cases} \qquad (2.3.9)$$

The MLE of the cure-rate is $$\hat{\pi}=1-\frac{k}{n} \qquad (2.3.10)$$

2.4 NPMLE of Type-II censoring with no cures

A second type of right censoring is Type-II censoring in which the study continues until the failure of the first r individuals, where r is some predetermined integer ($r<n$). It is true that the statistical treatment of Type-II censored data is simpler because the data consists of the r smallest lifetimes in a random sample of n lifetimes, so that the theory of order statistics is directly applicable in determining the likelihood and any inferential technique employed. Here it should be noted that r, the number of failures and n-r, the number of censored observations are fixed integers and the censoring time $T_{(r)}$, the rth ordered lifetimes, is random.

For Type-II censoring, the data consists of the rth smallest lifetimes $X_{(1)}\le X_{(2)}\le\cdots\le X_{(r)}$ out of a random sample of n lifetimes $X_1,\ldots,X_n$ from the assumed life distribution. Assuming $X_1,\ldots,X_n$ are i.i.d, and have a continuous distribution with p.d.f. f and survival function S, it follows from the likelihood is (cf. David, 1981)

$$L_{Type-II} = \frac{n!}{(n-r)!}\left[\prod_{i=1}^{r} f(x_{(i)})\right]\left[S(x_{(r)})\right]^{n-r} \tag{2.4.1}$$

The study continues until the failure of the first r individuals, where r is a predetermined integer ($r<n$). Let the event times $X_1,\ldots,X_n$ be ordered statistics, so the last n-r of them, i.e., $X_{r+1},\ldots,X_n$ are censored observations, and they are all censored at a random life time $X_{(r)}$. If $X_1,\ldots,X_n \overset{i.i.d.}{\sim} F_\pi(x)$, $F_\pi(x) = (1-\pi)\cdot F_0(x)$, we have the likelihood function

$$\begin{aligned} L(\pi, F_0) &= \frac{n!}{(n-r)!}\cdot\left[\prod_{i=1}^{r} f_\pi(x_i)\right]\cdot\left[S_\pi(x_{(r)})\right]^{n-r} \\ &= \frac{n!}{(n-r)!}\cdot\left[\prod_{i=1}^{r} \Delta F_\pi(x_{(i)})\right]\cdot\left[1-F_\pi(x_{(r)})\right]^{n-r} \\ &= \frac{n!}{(n-r)!}\cdot\left[1-(1-\pi)\cdot F_0(x_r)\right]^{n-r}\cdot\prod_{i=1}^{r}\left[(1-\pi)\cdot\left(F_0(x_i)-F_0(x_{i-1})\right)\right] \end{aligned} \tag{2.4.2}$$

Making the same substitution as in Section 2.3, (2.4.2) becomes

$$L(\pi, F_0) = \frac{n!}{(n-r)!}\cdot\left[1-(1-\pi)\cdot F_0(x_r)\right]^{n-r}\cdot\prod_{i=1}^{r}\left(z_i - z_{i-1}\right) \tag{2.4.3}$$

We consequently get the MLEs of the underlying distribution function and cure-rate by replacing k by r in the estimators given by (2.3.9) and (2.3.10).

Chapter 3 Non-estimability of cure-rate when no cures are observed

In this chapter we adopt the estimate in random censoring considered by Laska and Meisner (1992) and the result given by Groeneboom and Wellner (1992) to show that, when no cures are actually observed, the likelihood function is independent of θ in the uncensored case, and that in some important censoring models, the non-parametric maximum-likelihood method produces only the trivial estimator $\hat{\pi} = 0$ of the cure-rate. This leads us to consider NPMLE with a non-zero number of cures in the next chapter.

However, as part of our original work, we are able to show that in the random censoring and interval censoring (case-1) models, the NPMLE of $F(\cdot)$ at the largest observation (i.e., $\hat{F}(x_n)$, where $0 \le x_1 \le \cdots \le x_n$ are the observed data) is a consistent estimator of $(1-\pi)$.

3.1 Non-censored data, with no cures

Given $X_i < \infty$, under the hidden model, the conditional probability distribution function of X_i is given by

$$f_\theta(x_i) = \frac{d}{d\partial x}\left[\frac{1-e^{-\theta \cdot F_0(x)}}{1-e^{-\theta}}\right] = \frac{\theta \cdot f_0(x) \cdot e^{-\theta \cdot F_0(x)}}{1-e^{-\theta}} \qquad (3.1.1)$$

The likelihood function is then

$$L(\theta, F_0) = \prod_{i=1}^{n}[\Delta F_\theta(x_i)]$$

$$=\prod_{i=1}^{n}\left[\frac{1-e^{-\theta\cdot F_0(x_i)}}{1-e^{-\theta}}-\frac{1-e^{-\theta\cdot F_0(x_{i-1})}}{1-e^{-\theta}}\right]$$

$$=(1-e^{-\theta})^{-n}\cdot\prod_{i=1}^{n}\left[(1-e^{-\theta\cdot F_0(x_i)})-(1-e^{-\theta\cdot F_0(x_{i-1})})\right] \quad (3.1.2)$$

Let $y_i=F_0(x_i)$, set x_0=0, $0\le y_i\le 1$, $i=1,\cdots,n$

$z_i=1-e^{-\theta\cdot y_i}$, set z_0=0, $0\le z_i\le 1-e^{-\theta}$, $i=1,\cdots,n$

$$\Rightarrow \mathrm{L}(\theta,z_1,\cdots,z_n)\text{=}(1-e^{-\theta})^{\text{-n}}\cdot\prod_{i=1}^{n}\left(z_i-z_{i-1}\right) \quad (3.1.3)$$

Put $t_i=\dfrac{z_i}{1-e^{-\theta}}$, set t_0=0, $0\le t_i\le 1$, $i=1,\cdots,n$

$$\begin{aligned}\Rightarrow \mathrm{L}(\theta,t_1,\cdots,\mathrm{t}_n)&\text{=}(1-e^{-\theta})^{-n}\cdot\prod_{i=1}^{n}(1-e^{-\theta})\cdot\left(t_i-t_{i-1}\right)\\&=\prod_{i=1}^{n}\left(t_i-t_{i-1}\right)\end{aligned} \quad (3.1.4)$$

Following the same approach as in Section 2.2, with a certain $\hat{\theta}$, we have that the MLEs are as follows

$$\begin{cases}\hat{t}_i=\dfrac{i}{n} \quad , \quad i=1,\cdots,n\\ \hat{z}_i=\dfrac{i}{n}\cdot(1-e^{-\hat{\theta}}) \quad , \quad i=1,\cdots,n\\ \hat{y}_i=-\dfrac{1}{\hat{\theta}}\cdot\ln\left[1-\dfrac{i}{n}\cdot(1-e^{-\hat{\theta}})\right] \quad , \quad i=1,\cdots,n\end{cases} \quad (3.1.5)$$

Going back to (3.1.2), the likelihood function becomes

$$\begin{aligned}L(\theta,\hat{y}_1,\cdots,\hat{y}_n)&=(1-e^{-\theta})^{-n}\cdot\prod_{i=1}^{n}\left[e^{-\theta\cdot\hat{y}_{i-1}}-e^{-\theta\cdot\hat{y}_i}\right]\\&=(1-e^{-\theta})^{-n}\cdot\prod_{i=1}^{n}\left\{\left[1-\frac{i-1}{n}\cdot\left(1-e^{-\theta}\right)\right]-\left[1-\frac{i}{n}\cdot\left(1-e^{-\theta}\right)\right]\right\}\\&=(1-e^{-\theta})^{-n}\cdot\prod_{i=1}^{n}\left[\frac{1}{n}\cdot\left(1-e^{-\theta}\right)\right]\end{aligned}$$

$$=\left(\frac{1}{n}\right)^n$$

Thus the likelihood function is independent of θ, hence θ is non-estimable.

Comment: As illustrated in Chapter 1, the relationship between the cure-mixture model and the hidden model is that

$$P(non-cured)=1-\pi=1-e^{-\theta}$$

Implying that π is equivalent to $e^{-\theta}$. In Section 2.2, when there are (*n-k*) cures observed, we get

$$1-e^{-\hat{\theta}}=\frac{k}{n} \quad \Rightarrow \quad \hat{\theta}=\ln(\frac{n}{n-k})$$

$$\Leftrightarrow \ \hat{\pi}=1-\frac{k}{n}.$$

When *k*=*n*, i.e., no cure is observed, the MLE of θ is then

$$\hat{\theta}=\lim_{k\to n}\ln(\frac{n}{n-k})=\infty.$$

The estimator of cure-rate is hence given by

$$\hat{\pi}=\lim_{k\to n}(1-\frac{k}{n})\text{=}0,$$

which also implies the non-estimability of the cure-rate when there are no observed cures.

3.2 Random censoring with no cures

Sometimes, individuals will experience some other competing event which causes them to be removed from the study. In such cases the event of interest is not observable. Some events which cause the individual to be randomly censored, with respect to the event of interest, are accidental deaths, migration of human populations, death due to some cause other than the one of interest, patient withdrawal from a clinical trial, and so forth. If the distribution of random censoring times contains no parameters common with *S*(*t*), then, the estimators of such parameters may be obtained in the usual fashion for generalized Type-I censoring. However, the distribution of such estimators may be influenced by the distribution of the random censoring times.

Here we consider one particular instance encountered frequently, in which there is no complication. This random censoring process is one in which each subject has a lifetime X and a censoring time C_r, X and C_r being independent random variables with the usual notation for the probability density and survival function of X and the p.d.f. and survival function of C_r are denoted by g and G, respectively. Furthermore, let $T=min(X, C_r)$ and δ indicates whether the lifetime X is censored ($\delta = 0$) or not ($\delta = 1$). The data from a sample of n subjects consist of the pairs (t_i, δ_i), $i=1,...,n$. The density function of the pair may be obtained from the joint density function of X and C_r, $f(x, c_r)$, as

$$\begin{aligned} P\{T_i = t, \delta = 0\} &= P\{C_{r,i} = t, X_i > C_{r,i}\} \\ &= \frac{d}{dt}\int_0^t \int_v^\infty f(u,v)dudv \end{aligned}$$

When X and C_r are independent of marginal densities of f and g, respectively, the above probability becomes

$$\begin{aligned}&= \frac{d}{dt}\int_0^t \int_v^\infty f(u)g(v)du dv \\ &= \frac{d}{dt}\int_0^t S(v)g(v)dv \\ &= S(t)g(t)\end{aligned} \qquad (3.2.1)$$

and similarly,

$$P\{T_i = t, \delta = 1\} = P\{X_i = t, X_i < C_r, i\} = f(t)G(t) \qquad (3.2.2)$$

Therefore, the likelihood function is given by

$$\begin{aligned}L &= \prod_{i=1}^{n}[f(t_i)G(t_i)]^{\delta_i}[g(t_i)S(t_i)]^{1-\delta_i} \\ &= \left\{\prod_{i=1}^{n} G(t_i)^{\delta_i} g(t_i)^{1-\delta_i}\right\}\left\{\prod_{i=1}^{n} f(t_i)^{\delta_i} S(t_i)^{1-\delta_i}\right\}.\end{aligned}$$

If the distribution of the censoring times, as alluded to earlier, does not depend upon the parameters of interest, then, the first term will be a constant with respect to the parameters of interest and the likelihood function takes the form of (1.2.1):

$$L \propto \prod_{i=1}^{n}[f(t_i)]^{\delta_i}[S(t_i)]^{1-\delta_i} \qquad (3.2.3)$$

Under the cure-mixture model, the likelihood function for random censoring without observing any cure is given by

$$L(F_\pi) = \prod_{i=1}^{n}[\Delta F_\pi(t_i)]^{\delta_i} \cdot [1 - F_\pi(t_i)]^{1-\delta_i} \qquad (3.2.4)$$

$$\Leftrightarrow L(\pi, F_0) = \prod_{i=1}^{n}[(1-\pi)\cdot F_0(t_i) - (1-\pi)\cdot F_0(t_{i-1})]^{\delta_i} \cdot [\pi + (1-\pi)\cdot S_0(t_i)]^{1-\delta_i}.$$

Now put $y_i = (1-\pi)\cdot F_0(t_i)$, set z_0=0, hence y_0=0, $0 \le y_1 \le \cdots \le y_n \le 1-\pi$

$$\Rightarrow L(\pi,\tilde{y}) = \prod_{i=1}^{n} [y_i - y_{i-1}]^{\delta_i} \cdot [1 - y_i]^{1-\delta_i} \tag{3.2.5}$$

Since

$$\max_{0 \le y_1 \le \cdots \le y_n \le 1-\pi} \left\{ \prod_{i=1}^{n} [y_i - y_{i-1}]^{\delta_i} \cdot [1 - y_i]^{1-\delta_i} \right\}$$
$$\le \max_{0 \le y_1 \le \cdots \le y_n \le 1} \left\{ \prod_{i=1}^{n} [y_i - y_{i-1}]^{\delta_i} \cdot [1 - y_i]^{1-\delta_i} \right\}$$

We see that $$\hat{\pi} = 0 \tag{3.2.6}$$

Putting $p_i = y_i - y_{i-1} \Rightarrow y_i = \sum_{j=1}^{i} p_j$, then define $y_0 = 0$, (3.2.5) becomes

$$L(\pi,\tilde{p}) = \prod_{i=1}^{n} p_i^{\delta_i} \cdot \left(1 - \sum_{j=1}^{i} p_j \right)^{1-\delta_i} . \tag{3.2.7}$$

Introducing the conditional probability (c.f., Laska and Meisner, 1992)

$$\lambda_i = \frac{p_i}{1 - \sum_{j=1}^{i-1} p_j}, i = 2, \cdots n \tag{3.2.8}$$

and $\lambda_1 = p_1$,

then $$1 - \lambda_i = \frac{1 - \sum_{j=1}^{i} p_j}{1 - \sum_{j=1}^{i-1} p_j}, i = 2 \cdots, n \tag{3.2.9}$$

$$\prod_{j=1}^{i} (1 - \lambda_j) = 1 - \sum_{j=1}^{i} p_j \tag{3.2.10}$$

$$\Rightarrow L(\pi,\tilde{\lambda}) = \prod_{i=1}^{n} \lambda_i^{\delta_i} \cdot (1-\lambda_i)^{1-\delta_i} \cdot \prod_{i=1}^{n} \left(1 - \sum_{j=1}^{i-1} p_j \right)$$
$$= \prod_{i=1}^{n} \lambda_i^{\delta_i} \cdot (1-\lambda_i)^{1-\delta_i} \cdot \prod_{i=1}^{n} \left(\prod_{j=1}^{i-1} (1-\lambda_j) \right)$$
$$= \prod_{i=1}^{n} \lambda_i^{\delta_i} \cdot (1-\lambda_i)^{1-\delta_i} \cdot \prod_{i=1}^{n} (1-\lambda_i)^{n-i}$$

$$=\prod_{i=1}^{n}\lambda_i^{\delta_i}\cdot(1-\lambda_i)^{n-i+1-\delta_i} \tag{3.2.11}$$

Consequently, we have that the MLE of λ_i is

$$\hat{\lambda}_i=\frac{\delta_i}{n-i+1},\ \text{for } i=1,\ldots,n \tag{3.2.12}$$

Hence $$\hat{\lambda}_i=\frac{\delta_i}{n-i+1}=\frac{\hat{y}_i-\hat{y}_{i-1}}{1-\hat{y}_{i-1}},$$

with $\hat{y}_0=0$.

$$\Rightarrow 1-\frac{\delta_i}{n-i+1}=1-\frac{\hat{y}_i-\hat{y}_{i-1}}{1-\hat{y}_{i-1}}=\frac{1-\hat{y}_i}{1-\hat{y}_{i-1}}$$

Thus we have

$$1-\hat{y}_i=\prod_{j=1}^{i}\left(1-\frac{\delta_j}{n-j+1}\right),\ \text{for } i=1,\ldots,n \tag{3.2.13}$$

In particular,

$$1-\hat{y}_n=\prod_{j=1}^{n}\left(1-\frac{\delta_j}{n-j+1}\right)=\left(1-\frac{\delta_1}{n}\right)\cdot\left(1-\frac{\delta_2}{n-1}\right)\cdots\cdot\left(1-\frac{\delta_{n-1}}{2}\right)\cdot(1-\delta_n) \tag{3.2.14}$$

Suppose that $\delta_n=0$, i.e., the last observation is censored, then $1-\delta_n>0$ and so $1-\hat{y}_n>0$, implying that $\hat{y}_n<1$.

Convergence of $(1-\hat{y}_n)$ to π.

Now we turn our attention to the convergence property of the NPMLE. We will show that the above maximum likelihood estimator of the cure-rate converges to the true parameter value.

$$\begin{aligned}\hat{\pi} &= 1-\hat{y}_n \\ &= \prod_{j=1}^{n}\left(1-\frac{\delta_j}{n-j+1}\right) \\ &= \prod_{j=1}^{n}\left\{1-\frac{\delta_j}{n\cdot\left[1-H_n(t_j)\right]+1}\right\},\end{aligned} \qquad (3.2.15)$$

where $H_n(t)=\frac{1}{n}\cdot\sum_{i=1}^{n}1\{T_i \le t\}$, and $E\left[H_n(t)\right]=P\{T\le t\}=H(t)$. Hence,

$$\begin{aligned}\hat{\pi} &= \exp\left\{\ln\left\{\prod_{j=1}^{n}\left\{1-\frac{\delta_j}{n\cdot\left[1-H_n(t_j)\right]+1}\right\}\right\}\right\} \\ &= \exp\left\{\sum_{j=1}^{n}\ln\left\{1-\frac{\delta_j}{n\cdot\left[1-H_n(t_j)\right]+1}\right\}\right\} \\ &\approx \exp\left\{-\sum_{j=1}^{n}\left[\frac{\delta_j}{n\cdot\left[1-H_n(t_j)\right]+1}\right]\right\}\end{aligned}$$

which converges to

$$\begin{aligned}\exp\left\{-E\left\{\frac{\delta}{1-H(t)}\right\}\right\} &= \exp\left\{-E\left\{\frac{1\{X\le C_r\}}{1-H(t)}\right\}\right\} \qquad (T=X\wedge C_r) \\ &= \exp\left\{-\iint\frac{1\{X\le C_r\}\cdot f(x)\cdot g(c_r)}{(1-F(x))\cdot(1-G(c_r))}dxdc_r\right\} \\ &= \exp\left\{-\int\frac{f(x)}{1-F(x)}dx\right\} \\ &= e^{-\theta} \\ &= \pi\end{aligned}$$

Hence we have proved the convergence.

We shall also address the convergence property for interval censoring-case 1, which is illustrated in the following section.

3.3 Interval censoring (case-1) with no cures

By interval-censored data, we mean that a random variable of interest is known only to lie in an interval, instead of being observed exactly. In such cases, the only information we have for each individual is that their event time falls in an interval, but the exact time is unknown. Interval censoring occurs when patients in a clinical trial or longitudinal study have periodic follow-up and the patient's event time is only known to fall in an interval $(L_i, R_i]$ (L for left endpoint and R for right endpoint of the censoring interval).

Let (X_1, T_1), ... , (X_n, T_n) be a sample of random variables in R_+^2, where X_i represents the life time and T_i represents the observation time of each individual. X_i and T_i are independent (non-negative) random variables with distribution functions F_0 and G, respectively, where F_0 is a right-continuous distribution function.

The only available observations are (T_i, δ_i), for $i = 1, \cdots, n$. Notice that here T_i is the i^{th} order statistic of $T_1, \cdots, T_n$, and δ_i is the corresponding indicator, i.e., if $T_j = T_{(i)}$, then $\delta_i = 1\{X_j \le T_j\}$.

3.3.1 Non-estimablity when no cure is observed

Applying the cure-mixture model in (1.3.2) and (1.3.3), with the assumption that the cure-rate is π:

$$\begin{cases} F_\pi(t) = (1-\pi)F_0(t) \\ S_\pi(t) = 1 - F_\pi(t) = \pi + (1-\pi)\cdot S_0(t) \end{cases} \tag{3.3.1}$$

When there are no cures observed, the likelihood function is given by:

$$\begin{aligned} L\,(\pi, F_0) &= \prod_{i=1}^{n} [F_\pi(t_i)]^{\delta_i} \cdot [1 - F_\pi(t_i)]^{1-\delta_i} \\ &= \prod_{i=1}^{n} [(1-\pi)\cdot F_0(t_i)]^{\delta_i} \cdot [\pi + (1-\pi)\cdot(1 - F_0(t_i))]^{1-\delta_i} \end{aligned} \tag{3.3.2}$$

Put $x_i = (1-\pi)\cdot F_0(t_i)$, obviously $0 \le x_1 \le \cdots \le x_n \le 1-\pi$.

$$\Rightarrow L(\pi, \tilde{x}) = \prod_{i=1}^{n} x_i^{\delta_i} \cdot (1-x_i)^{1-\delta_i}$$

Observe that

$$\max_{0 \le x_1 \le \cdots \le x_n \le 1-\pi} \left\{ \prod_{i=1}^{n} x_i^{\delta_i} \cdot (1-x_i)^{1-\delta_i} \right\} \le \max_{0 \le x_1 \le \cdots \le x_n \le 1} \left\{ \prod_{i=1}^{n} x_i^{\delta_i} \cdot (1-x_i)^{1-\delta_i} \right\}$$

implying that $$\hat{\pi} = 0 \qquad (3.3.3)$$

3.3.2 Convergence property of the MLE of the cure-rate

In a general case, where we do not apply the cure-mixture model and no cure is observed, the likelihood function is given by

$$L\,(F_0) = \prod_{i=1}^{n} F_0(t_i)^{\delta_i} \cdot \left(1 - F_0(t_i)\right)^{1-\delta_i} \qquad (3.3.4)$$

Take logarithms and make the replacements $x_i = F_0(t_i)$, for *i=1,...,n*, to get the log likelihood function

$$\ln L(F_0) = \sum_{i=1}^{n} \left\{ \delta_i \cdot \ln x_i + (1-\delta_i) \cdot \ln(1-x_i) \right\} \qquad (3.3.5)$$

where $0 \le x_1 \le \cdots \le x_n \le 1$.

Groeneboom and Wellner (1992) have given the so-called "max-min formula" for the solution of the above general maximization problem (3.3.5) of interval censoring Case-1. In this section, we prove that the MLE of the cure-rate based on the "max-min formula" converges to the real value of the cure-rate, if there exists such a cure-rate.

Since $$y_m = \max_{i \le m} \min_{k \ge m} \frac{\sum_{i \le j \le k} \delta_{(j)}}{k-i+1}, \qquad (3.3.6)$$

we have that
$$y_n = \max_{i \le n} \frac{\sum_{j=i}^{n} \delta_{(j)}}{n-i+1}. \tag{3.3.7}$$

Define
$$G_n(x) = \frac{1}{n} \cdot \sum_{i=1}^{n} 1\{T_i < x\}, \tag{3.3.8}$$

$$\bar{G}_n(x) = 1 - G_n(x) = \frac{1}{n} \cdot \sum_{i=1}^{n} 1\{T_i \ge x\}, \tag{3.3.9}$$

$$G_{1n}(x) = \frac{1}{n} \cdot \sum_{i=1}^{n} \delta_{(i)} \cdot 1\{T_i < x\}, \tag{3.3.10}$$

$$\bar{G}_{1n}(x) = \sum_{i=1}^{n} \delta_{(i)} - G_{1n}(x) = \frac{1}{n} \cdot \sum_{i=1}^{n} \delta_{(i)} \cdot 1\{T_i \ge x\} \tag{3.3.11}$$

Hence y_n can be written as

$$\begin{aligned} y_n &= \max_{i \le n} \frac{\sum_{j=1}^{n} \delta_{(j)} \cdot 1\{T_j \ge T_i\}}{n \cdot \bar{G}_n(T_i)} \\ &= \max_{x} \frac{n \cdot \bar{G}_{1n}(x)}{n \cdot \bar{G}_n(x)} \\ &= \max_{x} \frac{\bar{G}_{1n}(x)}{\bar{G}_n(x)} \end{aligned} \tag{3.3.12}$$

Observe that

$$\begin{aligned} \lim_{n \to \infty} \frac{\bar{G}_{1n}(x)}{\bar{G}_n(x)} &= \lim_{n \to \infty} \frac{\frac{1}{n} \cdot \sum_{i=1}^{n} \delta_{(i)} \cdot 1\{T_i \ge x\}}{\frac{1}{n} \cdot \sum_{i=1}^{n} 1\{T_i \ge x\}} \\ &= \frac{\int_x^{\infty} F(z) dG(z)}{\int_x^{\infty} dG(z)} \\ &= \frac{\bar{G}_1(x)}{\bar{G}(x)}, \text{ say,} \end{aligned} \tag{3.3.13}$$

where $F(x) = F_\pi(x)$ represents the distribution function of the event time X.

Further,

$$\begin{aligned}
\frac{d}{dx}\left(\frac{\bar{G}_1(x)}{\bar{G}(x)}\right) &= \frac{1}{\left[\bar{G}(x)\right]^2}\cdot\left\{\bar{G}(x)\cdot\left[F(x)\cdot(-g(x))\right]-\bar{G}_1(x)\cdot\left[-g(x)\right]\right\} \\
&= \frac{g(x)}{\left[\bar{G}(x)\right]^2}\cdot\left\{\bar{G}_1(x)-\bar{G}(x)\cdot F(x)\right\} \\
&= \frac{g(x)}{\left[\bar{G}(x)\right]^2}\cdot\left\{\int_x^\infty F(y)dG(y)-\int_x^\infty F(x)dG(y)\right\} \\
&= \frac{g(x)}{\left[\bar{G}(x)\right]^2}\cdot\int_x^\infty\left[F(y)-F(x)\right]dG(y) \\
&\geq 0. \qquad (3.3.14)
\end{aligned}$$

hence

$$\begin{aligned}
\max_x\frac{\bar{G}_1(x)}{\bar{G}(x)} &= \lim_{x\to\infty}\frac{\bar{G}_1(x)}{\bar{G}(x)} \\
&= \lim_{x\to\infty}\frac{\frac{d\bar{G}_1(x)}{dx}}{\frac{d\bar{G}(x)}{dx}} \\
&= \lim_{x\to\infty.}\frac{-F(x)g(x)}{-g(x)} \\
&= \lim_{x\to\infty}F(x) \\
&= 1-\pi, \qquad (3.3.15)
\end{aligned}$$

where π is supposed to be the cure-rate of our interest.

Now

$$\begin{aligned}
|y_n - (1-\pi)| &= (1-\pi) - y_n \\
&= (1-y_n) - \pi \\
&= \min_{1\le i\le n} \left[1 - \frac{\frac{1}{n}\cdot\sum_{j=i}^{n}\delta_{(j)}}{n-i+1} - \pi \right] \\
&= \min_{1\le i\le n} \left[1 - \pi - \frac{\bar{G}_{1n}(y_i)}{\bar{G}_n(y_i)} \right] \\
&\le \max_{1\le i\le n} \left[1 - \pi - \frac{\bar{G}_{1n}(y_i)}{\bar{G}_n(y_i)} \right] \\
&\le \max_{x} \left| \frac{\bar{G}_{1n}(x)}{\bar{G}_n(x)} - (1-\pi) \right| \\
&\to 0 \quad \text{with probability 1, as } n \to \infty,
\end{aligned}$$

as in the Glivenko-Cantelli Theorem.

This we prove the convergence of the cure-rate MLE.

Chapter 4 Maximum likelihood estimation with observed cures and censored data

4.1 NPMLE for Type-I censoring with observing some cures

Type-I censoring occurs if the event is observed only if it occurs prior to some pre-specified time. As a result, it is impossible for us to observe such an individual that the life time of it is infinity. However, we could observe some cures in an independent sample. Hence if there are $m \geq 1$ cures, the likelihood function from Eq.(2.3.5) becomes:

$$L(\pi, F_0) = \pi^m \cdot \left[\prod_{i=1}^{k} (z_i - z_{i-1}) \right] \cdot (1 - z_{k+1})^{n-k}, \tag{4.1.1}$$

where $z_i = (1-\pi) \cdot F_0(t_i)$ for $i = 1, \cdots k$, and $z_{k+1} = \cdots = z_n = (1-\pi) \cdot F_0(C_r)$, $0 \leq z_1 \leq \cdots \leq z_{k+1} \leq 1 - \pi$. From Eq.(2.3.7), we have

$$L(\pi, \tilde{z}) = \pi^m \cdot \frac{1}{k^k} \cdot z_k^{\,k} \cdot (1 - z_{k+1})^{n-k}. \tag{4.1.2}$$

Hence depending on the value of π, we get:

$$\underset{0 \leq \pi \leq 1}{Max} \left\{ \underset{0 \leq z_1 \leq \cdots \leq z_n \leq 1-\pi}{Max} \left\{ z_k^{\,k} \cdot (1 - z_k)^{n-k} \right\} \cdot \pi^m \right\}$$

$$= Max \begin{cases} \underset{0 \leq \pi \leq 1}{Max} \left\{ \left(\frac{k}{n}\right)^k \cdot \left(1 - \frac{k}{n}\right)^{n-k} \cdot \pi^m \right\}, & \text{if } \pi \leq 1 - \frac{k}{n} \\ \underset{0 \leq \pi \leq 1}{Max} \left\{ (1-\pi)^k \cdot \pi^{n-k} \cdot \pi^m \right\}, & \text{if } \pi > 1 - \frac{k}{n} \end{cases}$$

$$= Max \begin{cases} \left(\frac{k}{n}\right)^k \cdot \left(1 - \frac{k}{n}\right)^{n-k+m}, & \text{if } \pi \leq 1 - \frac{k}{n} \\ \left(1 - \frac{m+n-k}{m+n}\right)^k \cdot \left(\frac{m+n-k}{m+n}\right)^{n-k+m}, & \text{since } \frac{m+n-k}{m+n} > 1 - \frac{k}{n} \text{ always true} \end{cases}$$

$$=\left(1-\frac{m+n-k}{m+n}\right)^{k}\cdot\left(\frac{m+n-k}{m+n}\right)^{n-k+m}, \text{ with } \hat{\pi}=\frac{m+n-k}{m+n}=1-\frac{k}{m+n}$$

Which gives
$$\hat{z}_i=\begin{cases}\dfrac{i}{m+n}, & i=1,\cdots,k-1\\ \dfrac{k}{m+n}, & i=k,\cdots,n\end{cases} \tag{4.1.3}$$

and
$$\hat{y}_i=F_0(t_i)=\frac{\hat{z}_i}{1-\hat{\pi}}=\begin{cases}\dfrac{i}{k}, & i=1,\cdots,k-1\\ 1, & i=k,\cdots,n\end{cases} \tag{4.1.4}$$

The MLE of the cure-rate is then
$$\hat{\pi}=1-\frac{k}{m+n} \tag{4.1.5}$$

4.2 NPMLE for Type-II censoring with observing some cures

Type-II Censoring is defined when the study continues until the failure of the first r individuals, where r is a predetermined integer ($r<n$). We get the MLEs of the underlying distribution function and cure-rate by replacing k by r in (4.1.4) and (4.1.5).

4.3 NPMLE for random censoring with observing some cures

4.3.1 Introduction

In our discussion, we adopt some of the notations and conventions of Miller (1981). The usual observations of survival times and censoring indicators are augmented by the possibility of observing that an individual in the mixed population is in fact a cure. The primary variable of interest is time to death from a disease for which the probability of recovery, i.e., cure, is π. An individual is known to recover if survival is greater than a known time, say, *C*.

Let X' be a random variable denoting the conditional survival time of the non-cured population taking values in the possibly infinite interval $\left[0\ ,\ C\right)$, and $X_i^{'}$s are independently and identically distributed (i.i.d.), with F_0.

Let *X* be a random variable with $P(X = X') = 1 - \pi$ and $P(X = C) = \pi$. $X_i^{'}$s are i.i.d. with F_π, where $F_\pi(x) = (1-\pi)\cdot F_0(x)$.

Let *Y* be a censoring random variable that is independent of *X*, and δ an indicator function which is defined by $\delta = 1\left\{X \le Y\right\}$.

An outcome that is a cure, regardless of when it is noted, is mathematically equivalent to observing $X = C < Y$. Let $Z = \min(X, Y)$, we observe the pair (Z, δ).

Let S_π be the survival distribution function of the observable survival random variable *X*, π the probability that a randomly chosen member of the population is cured, and S_0 the conditional survival distribution function of individuals not cured.

Then

$$\begin{aligned} S_\pi(x) &= 1 - F_\pi(x) \\ &= 1 - (1-\pi)\cdot(1 - S_0(x)) \\ &= \pi + (1-\pi)\cdot S_0(x) \\ \text{and if } C < \infty,\ & S_0(x) = 0 \text{ for } X > C. \end{aligned} \tag{4.3.1}$$

4.3.2 NPMLE of the cure-rate π

This case was already considered by Laska and Meisner (1992). We present their result below. Given the *N* observations (Z_i, δ_i), $i = 1, \cdots, n$, and *m* cures, where *n*+*m*=*N*. Based on our result of Section 3.2, augmented by observing some cures, maximizing the likelihood of the observations for the cure model S_π is equivalent to maximizing L given by:

$$L(\pi,F_0){=}\pi^{\mathrm{m}}\cdot\prod_{i=1}^{n}\left[(1-\pi)\cdot F_0(z_i)-(1-\pi)\cdot F_0(z_{i-1})\right]^{\delta_i}\cdot\left[\pi+(1-\pi)\cdot S_0(z_i)\right]^{1-\delta_i} \quad (4.3.2)$$

Let

$$\begin{aligned} q_i &= F_0(z_i)-F_0(z_{i-1}) \\ &= S_0(z_{i-1})-S_0(z_i), \qquad (4.3.3) \\ & i=1,\cdots,n, \text{ and } Z_0=0 \end{aligned}$$

$$\Rightarrow S_0(z)=\sum_{j:Z_j>z} q_j \quad , \quad S_0(z_0)=\sum_{j=1}^{n} q_j=1$$

put

$$p_i=(1-\pi)\cdot q_i \qquad (4.3.4)$$

$$\begin{aligned} &=(1-\pi)\cdot\left[F_0(z_i)-F_0(z_{i-1})\right] \\ &=F_\pi(z_i)-F_\pi(z_{i-1}) \qquad (4.3.5) \\ &=\mathrm{S}_\pi(z_{i-1})-S_\pi(z_i), \quad i=1,\cdots,n \end{aligned}$$

Note that $\sum_{j=1}^{n} p_j=1-\pi$, and $\lim_{x\to C} S_\pi(x)=\pi$

$$\Rightarrow L(\pi,\tilde{p})=\pi^m\cdot\prod_{i=1}^{n} p_i^{\delta_i}\cdot\left[\pi+\sum_{j:Z_j>Z_i} p_j\right]^{1-\delta_i}$$

$$=\pi^m\cdot\prod_{i=1}^{n} p_i^{\delta_i}\cdot\left[1-\sum_{j:Z_j\le Z_i} p_j\right]^{1-\delta_i} \qquad (4.3.6)$$

Using again the conditional probability

$$\lambda_i=\frac{p_i}{1-\sum_{j=1}^{i-1} p_j} \qquad (4.3.7)$$

then

$$1-\lambda_i=\frac{1-\sum_{j=1}^{i} p_j}{1-\sum_{j=1}^{i-1} p_j}$$

$$\prod_{j=1}^{i}(1-\lambda_j)=1-\sum_{j=1}^{i}p_j \qquad (4.3.8)$$

Hence $$\pi^m=(1-\sum_{i=1}^{n}p_i)^m=\left[\prod_{i=1}^{n}(1-\lambda_i)\right]^m=\prod_{i=1}^{n}(1-\lambda_i)^m \qquad (4.3.9)$$

and $$p_i^{\delta_i}\cdot\left[1-\sum_{j:Z_j\le Z_i}p_j\right]^{1-\delta_i}=\lambda_i^{\delta_i}\cdot\left[1-\sum_{j:Z_j\le Z_i}p_j\right]$$

$$=\lambda_i^{\delta_i}\cdot\prod_{j=1}^{i-1}(1-\lambda_j) \qquad (4.3.10)$$

Put (4.3.8), (4.3.9) and (4.3.10) into (4.3.6)

$$\Rightarrow L(\pi,\tilde{\lambda})=\left[\prod_{i=1}^{n}(1-\lambda_i)^m\right]\cdot\prod_{i=1}^{n}\left[\lambda_i^{\delta_i}\cdot\prod_{j=1}^{i-1}(1-\lambda_j)\right] \qquad (4.3.11)$$

Notice that $$\prod_{i=1}^{n}\left[\prod_{j=1}^{i-1}(1-\lambda_j)\right]=\prod_{i=1}^{n}(1-\lambda_i)^{n-i} \qquad (4.3.12)$$

$$\Rightarrow L(\pi,\tilde{\lambda})=\prod_{i=1}^{n}\lambda_i^{\delta_i}\cdot(1-\lambda_i)^{n+m-i} \qquad (4.3.13)$$

Hence the MLEs for λ_i's are given by

$$\hat{\lambda}_i=\frac{\delta_i}{n+m-i+\delta_i} \qquad (4.3.14)$$

Since $$\prod_{i=1}^{n}(1-\lambda_i)=1-\sum_{i=1}^{n}p_i$$

And $$\sum_{i=1}^{n}p_i=1-\pi$$

$$\Rightarrow \hat{\pi}=\prod_{i=1}^{n}(1-\hat{\lambda}_i)$$

$$=\prod_{i=1}^{n}(1-\frac{\delta_i}{n+m-i+\delta_i}) \qquad (4.3.15)$$

is the NPMLE of the cure-rate.

4.3.3 NPMLE of the survival function

We get the NPMLE of the survival function motivatied by of the Kaplan-Meier product limit (PL) estimator of 1958.

To allow for possible ties in the data, suppose that the events occur at N distinct times $t_1 < t_2 < \cdots < t_N$, and that at time t_i there are d_i events. Let Y_i be the number of individuals who are at risk at time t_i, i.e., Y_i is the number of individuals who are alive at t_i or experience the event of interest at t_i. The quantity $\frac{d_i}{Y_i}$ provides an estimate of the conditional probability that an individual who survives to just prior to time t_i experiences the event at time t_i. The Kaplan-Meier product limit estimator of the survival function is hence given by:

$$\hat{S}(t) = \begin{cases} 1, & \text{if } t < t_1 \\ \prod_{t_i \le t} (1 - \frac{d_i}{Y_i}), & \text{if } t_1 \le t \end{cases}$$

In our approach, since the MLEs for the conditional probability are given by (4.3.14) (c.f., Laska and Meisner, 1992):

$$\hat{\lambda}_i = \frac{\delta_i}{n + m - i + \delta_i},$$

the MLEs of the survival distributions, $\hat{\mathrm{S}}_\pi$ and $\hat{\mathrm{S}}_0$, are given by

$$\hat{S}_\pi(t) = \prod_{i: Z_i \le t} (1 - \frac{\delta_i}{n + m - i + \delta_i}) \tag{4.3.16}$$

and
$$\hat{S}_0(t) = \frac{\hat{S}_\pi(t) - \hat{\pi}}{1 - \hat{\pi}}. \tag{4.3.17}$$

Moreover, it is proven that the KM-PL estimator is the generalized maximum likelihood estimator (GMLE) (Kaplan-Meier, 1958). Hence, the above estimators are the GMLEs of $\hat{S}_\pi(t)$ and $\hat{S}_0(t)$, respectively.

4.4 NPMLE for interval censoring-case 1 with observing some cures

4.4.1 Background

We adopt the notations illustrated in Section 3.3. When applying the cure-mixture model with the assumption that the cure-rate is π, we have:

$$\begin{cases} F_\pi(t) = (1-\pi)F_0(t) \\ S_\pi(t) = 1 - F_\pi(t) = \pi + (1-\pi)\cdot S_0(t) \end{cases} \tag{4.4.1}$$

On assuming some cures are observed, the likelihood function is given by:

$$\begin{aligned} L(\pi, F_0) &= \pi^m \cdot \prod_{i=1}^{n} [F_\pi(t_i)]^{\delta_i} \cdot [1 - F_\pi(t_i)]^{1-\delta_i} \\ &= \pi^m \cdot \prod_{i=1}^{n} [(1-\pi)\cdot F_0(t_i)]^{\delta_i} \cdot [\pi + (1-\pi)\cdot (1 - F_0(t_i))]^{1-\delta_i} \end{aligned} \tag{4.4.2}$$

where $m = \sum_{i=1}^{n} 1\{x_i = \infty\}$ is the number of cures, and N=m+n is the total number of individuals.

Put $x_i = (1-\pi)\cdot F_0(t_i)$, obviously $0 \le x_1 \le \cdots \le x_n \le 1-\pi$.

$$\Rightarrow L(\pi, \tilde{x}) = \pi^m \cdot \prod_{i=1}^{n} x_i^{\delta_i} \cdot (1-x_i)^{1-\delta_i}$$

The log-likelihood function is thus given by:

$$\ln L(\pi, \tilde{x}) = m \cdot \ln \pi + \sum_{i=1}^{n} \left\{ \delta_i \ln x_i + (1-\delta_i) \cdot \ln(1-x_i) \right\} \tag{4.4.3}$$

To maximize the log-likelihood function is equivalent to first maximizing the function:

$$\phi(\tilde{x}) \overset{def}{=} \sum_{i=1}^{n} \{ \delta_i \cdot \ln x_i + (1-\delta_i) \cdot \ln(1-x_i) \} \tag{4.4.4}$$

under the condition that $$0 \le x_1 \le \cdots \le x_n \le c \tag{4.4.5}$$

where c is denoted as $c = 1 - \pi$, and then solving

$$\max_{0\le\pi\le 1}\left\{m\ln\pi+\max_{0\le x_1\le\cdots\le x_n\le 1-\pi}\phi(\tilde{x})\right\} \tag{4.4.6}$$

4.4.2 Optimal solution to the maximization problem of Eq. (4.4.4)

(i) Let $(y_1,\cdots,y_n)$ be the optimal solution to the general interval censoring case-1 problem, which is without applying the cure- mixture model or observing any cures,

i.e., $\tilde{y}=\arg\max_{0\le x_1\le\cdots\le x_n\le 1}\left\{\sum_{i=1}^{n}\{\delta_i\cdot\ln x_i+(1-\delta_i)\cdot\ln(1-x_i)\}\right\}$

Based on the property of concave function, since $\phi(\tilde{x})$ is concave, we have

$$\phi(\tilde{x})-\phi(\tilde{y})\le\langle\nabla\phi(\tilde{x}),\tilde{x}-\tilde{y}\rangle \quad \text{for } \forall\tilde{x},\tilde{y}\in R^n$$

$$\tilde{y}=\arg\max\phi(\tilde{x})\Leftrightarrow(\nabla\phi(\tilde{y}))^T\cdot(\tilde{x}-\tilde{y})\le 0 \quad \text{for } \forall\tilde{x}\in R^n \tag{4.4.7}$$

Equivalently, $\sum_{i=1}^{n}\left\{\frac{\delta_i}{y_i}-\frac{1-\delta_i}{1-y_i}\right\}\cdot(x_i-y_i)\le 0$ for all $0\le x_1\le\cdots\le x_n\le 1$ (4.4.8)

More specifically, Groeneboom and Wellner (1992) have given a so-called "max-min formula" as the solution to this problem:

$$y_m=\max_{i\le m}\min_{k\ge m}\frac{\sum_{i\le j\le k}\delta_j}{k-i+1} \tag{4.4.9}$$

(ii) We will first give the solution to our problem, then follows the proof. To show that the optimal solution to maximizing (4.4.4) under condition (4.4.5) is

$$\begin{aligned} y_m^* &= \min(c,\max_{i\le m}\min_{k\ge m}\frac{\sum_{i\le j\le k}\delta_j}{k-i+1}) \\ &= c\wedge y_m, \end{aligned} \tag{4.4.10}$$

it is enough to show that

$$\sum_{i=1}^{k}\left\{\frac{\delta_i}{y_i}-\frac{1-\delta_i}{1-y_i}\right\}\cdot(x_i-y_i)+\sum_{i=k+1}^{n}\left\{\frac{\delta_i}{c}-\frac{1-\delta_i}{1-c}\right\}\cdot(x_i-c)\le 0 \tag{4.4.11}$$

for all $0\le x_1\le\cdots\le x_n\le c$, where $1\le k\le n$ is such that

$0 \le y_1 \le \cdots \le y_k \le c \le y_{k+1} \le \cdots \le y_n$.

(iii) Take $0 \le x_1 \le \cdots \le x_n \le c$, apply the inequality (4.4.8) to the vector $(x_1, \cdots, x_k, y_{k+1}, \cdots y_n)$ to get

$$\sum_{i=1}^{k} \left\{ \frac{\delta_i}{y_i} - \frac{1-\delta_i}{1-y_i} \right\} \cdot (x_i - y_i) \le 0 \tag{4.4.12}$$

Hence, in view of (4.4.12), (4.4.11) will follow if we show that

$$\sum_{i=k+1}^{n} \left\{ \frac{\delta_i}{c} - \frac{1-\delta_i}{1-c} \right\} \cdot (x_i - c) \le 0 \tag{4.4.13}$$

$$\Leftrightarrow \frac{\sum\limits_{i=k+1}^{n} (\delta_i - c)(x_i - c)}{c(1-c)} \le 0$$

$$\Leftrightarrow \sum_{i=k+1}^{n} (\delta_i - c) a_i \ge 0, \text{ where } a_i = c - x_i, \text{ and } a_{k+1} \ge \cdots \ge a_n \ge 0$$

$$\Leftrightarrow c \le \frac{\sum\limits_{i=k+1}^{n} \delta_i \cdot a_i}{\sum\limits_{i=k+1}^{n} a_i}, \tag{4.4.14}$$

Where it is assumed that at least one $a_i > 0$, otherwise the result is trivial.

Put $$p_i = \frac{a_i}{\sum\limits_{i=k+1}^{n} a_i}, \text{ for } i = k+1, \cdots, n.$$

We have $p_{k+1} \ge \cdots \ge p_n \ge 0$, and $p_{k+1} + \cdots + p_n = 1$. The inequality (4.4.14) is hence equivalent to

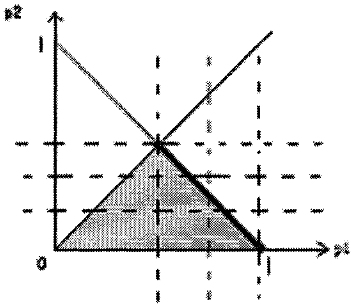
$$c \le \sum_{i=k+1}^{n} \delta_i \cdot p_i$$

(iv) Hence to prove (4.4.13), it is enough to prove that

$$c \le \min \left\{ \sum_{i=k+1}^{n} \delta_i \cdot p_i \,\middle|\, p_{k+1} \ge \cdots \ge p_n \ge 0, \text{ and } p_{k+1} + \cdots + p_n = 1 \right\}$$

$$= \min\left\{\delta_{k+1}, \frac{\delta_{k+1}+\delta_{k+2}}{2}, \frac{\delta_{k+1}+\delta_{k+2}+\delta_{k+3}}{3}, \cdots \frac{\delta_{k+1}+\cdots+\delta_n}{n-k}\right\} \tag{4.4.15}$$

The last equality holds because the minimum is attained at one of the extreme points. To get all these points, we start at a simple case: minimize $\{\delta_1 p_1 + \delta_2 p_2\}$ with respect to $p_1 \geq p_2 \geq 0$, and $p_1 + p_2 = 1$ (see bold segment in the following graph).



The extreme points are hence (1,0) and (1/2,1/2). In general, to minimize $\{\delta_1 p_1 + \delta_2 p_2 + \cdots + \delta_n p_n\}$ w.r.t. $p_1 \geq \cdots \geq p_n \geq 0$ and $p_1 + \cdots + p_n = 1$, the extreme points are: $(1,0,0,\cdots,0)$, $\left(\frac{1}{2},\frac{1}{2},0,\cdots,0\right)$, $\left(\frac{1}{3},\frac{1}{3},\frac{1}{3},0\cdots,0\right)$, ……, and $\left(\frac{1}{n},\frac{1}{n},\cdots,\frac{1}{n}\right)$.

(v) Now note that

$$c < y_{k+1} = \max_{i \leq k+1} \min_{j \geq k+1} \frac{\sum_{i \leq l \leq j} \delta_l}{j-i+1}$$

$$\Leftrightarrow c < \min\left\{\frac{\sum_{i_0 \leq l \leq j} \delta_l}{j-i_0+1}, k+1 \leq j \leq n\right\}, \text{ for some } 1 \leq i_0 \leq k+1. \tag{4.4.16}$$

Meanwhile, $$c > y_k = \max_{i \leq k} \min_{j \geq k} \frac{\sum_{i \leq l \leq j} \delta_l}{j-i+1}$$

$$\Leftrightarrow c > \min\left\{ \frac{\sum\limits_{l=i}^{j} \delta_l}{j-i+1}, k \le j \le n \right\}, \text{ for all } 1 \le i \le k. \qquad (4.4.17)$$

① In (4.4.16) if $i_0 = k+1$, $\Rightarrow c \le \frac{\sum\limits_{k+1 \le l \le j} \delta_l}{j-k}$ for all $k+1 \le j \le n$,

$$\Leftrightarrow c \le \min\left\{ \delta_{k+1}, \frac{\delta_{k+1}+\delta_{k+2}}{2}, \frac{\delta_{k+1}+\delta_{k+2}+\delta_{k+3}}{3}, \cdots \frac{\delta_{k+1}+\cdots+\delta_n}{n-k} \right\}$$

Hence (4.4.15) is proved.

② If $i_0 \le k$, in view of (4.4.16) and (4.4.17), we have

$$\left\{ \begin{array}{l} c < \min\left\{ \frac{\delta_{i_0}+\cdots+\delta_{k+1}}{k-i_0+2}, \frac{\delta_{i_0}+\cdots+\delta_{k+1}+\delta_{k+2}}{k-i_0+3}, \cdots \frac{\delta_{i_0}+\cdots+\delta_n}{n-i_0+1} \right\} = a_{i_0}(k,n), \text{ say,} \\ \quad \text{for some } 1 \le i_0 \le k \\ c > \min\left\{ \frac{\delta_i+\cdots+\delta_k}{k-i+1}, \frac{\delta_i+\cdots+\delta_{k+1}}{k-i+2}, \cdots, \frac{\delta_i+\cdots+\delta_n}{n-i+1} \right\} = \min\left\{ \frac{\delta_i+\cdots+\delta_k}{k-i+1}, a_i(k,n) \right\} \\ \quad \text{for all } 1 \le \text{i} \le \text{k} \end{array} \right.$$

Specializing the 2nd inequality to $i = i_0$, we have

$$c > \min\left\{ \frac{\delta_{i_0}+\cdots+\delta_k}{k-i_0+1}, a_{i_0}(k,n) \right\};$$

meanwhile, from the 1st inequality, $c < a_{i_0}(k,n)$.

Hence $\frac{\delta_{i_0}+\cdots+\delta_k}{k-i_0+1} < c < a_{i_0}(k,n)$.

$$\Rightarrow \frac{\delta_{i_0}+\cdots+\delta_k}{k-i_0+1} \left\{ \begin{array}{l} < \frac{\delta_{i_0}+\cdots+\delta_{k+1}}{k-i_0+2} \Leftrightarrow \frac{\delta_{i_0}+\cdots+\delta_{k+1}}{k-i_0+2} < \delta_{k+1} \\ < \frac{\delta_{i_0}+\cdots+\delta_{k+1}+\delta_{k+2}}{k-i_0+3} \Leftrightarrow \frac{\delta_{i_0}+\cdots+\delta_{k+1}+\delta_{k+2}}{k-i_0+3} < \frac{\delta_{k+1}+\delta_{k+2}}{2} \\ \quad \vdots \\ < \frac{\delta_{i_0}+\cdots+\delta_n}{n-i_0+1} \Rightarrow \frac{\delta_{i_0}+\cdots+\delta_n}{n-i_0+1} < \frac{\delta_{k+1}+\cdots+\delta_n}{n-k} \end{array} \right.$$

Now we have

$$c < \min\left\{\frac{\delta_{i_0}+\cdots+\delta_{k+1}}{k-i_0+2}, \frac{\delta_{i_0}+\cdots+\delta_{k+1}+\delta_{k+2}}{k-i_0+3}, \cdots \frac{\delta_{i_0}+\cdots+\delta_{n}}{n-i_0+1}\right\}$$

$$< \min\left\{\delta_{k+1}, \frac{\delta_{k+1}+\delta_{k+2}}{2}, \frac{\delta_{k+1}+\delta_{k+2}+\delta_{k+3}}{3}, \cdots \frac{\delta_{k+1}+\cdots+\delta_{n}}{n-k}\right\}$$

Hence, the inequalities (4.4.16) and (4.4.17) imply that (4.4.15) holds, consequently (4.4.13) is true, and then we get (4.4.11). By this approach, we have confirmed that the optimal solution to maximizing (4.4.4) under condition (4.4.5) is

$$y_m^* = \min(c, \max_{i\le m} \min_{k\ge m} \frac{\sum_{i\le j\le k} \delta_j}{k-i+1}) \tag{4.4.18}$$

for *m=1,...,n*, where $c = 1-\pi$, and π is the cure-rate.

4.4.3 NPMLE of the cure-rate

Now we turn our attention to finding the NPMLE of the cure-rate. The log-likelihood function is $\ln L(\tilde{x},c) = m\cdot\ln(1-c) + \sum_{i=1}^{n}\{\delta_i \cdot \ln x_i + (1-\delta_i)\cdot\ln(1-x_i)\}$, where $m = \sum_{i=1}^{N} 1\{X_i = \infty\}$ is the number of observed cures among all the *N* individuals, and *N=m+n*.

By following the approach of Groeneboom and Wellner (1992), we get the optimal solution to the maximization problem without involving a cure-rate, i.e., $y_m = \max_{i\le m}\min_{k\ge m}\frac{\sum_{i\le j\le k}\delta_j}{k-i+1}$ for *m=1,...,n*. Denote $y_{j_1}, y_{j_2}, \cdots, y_{j_r}$ as the distinct values of $y_1, y_2, \cdots, y_n$, which form j_r+1 intervals for *c* to pick up a value from. Since

$y_m^* = \min(c, y_m)$, we determine the value of each y_m^* depending on which interval c lies in. We then solve

$$\max_{1 \le k \le n} \left\{ \max_{c \in A_k} \left[m \cdot \ln(1-c) + \max_{0 \le x_1 \le \cdots \le x_n \le c} \phi(\tilde{x}) \right] \right\} \tag{4.4.19}$$

where $A_k = \left(y_{j_{k-1}}, y_{j_k} \right]$.

Example 4.1. Let $n=5, \delta_1 = \delta_3 = \delta_4 = 1,$and $\delta_2 = \delta_5 = 0$. Then the vector maximizing the log-likelihood without observing any cures is given by

$$y_1 = y_2 = \frac{1}{2}, y_3 = y_4 = y_5 = \frac{2}{3}$$

(see Groeneboom & Wellner, 1992)

If we observe some cures in this example, the log-likelihood function with plugging in the optimal solution $y_m^* = \min(c, y_m)$ will be given by

$$f(\tilde{x}, c) \overset{def}{=} \begin{cases} m \cdot \ln(1-c) + \sum_{i=1}^{5} \{\delta_i \cdot \ln y_i + (1-\delta_i) \cdot \ln(1-y_i)\}, \text{ if } c \ge \frac{2}{3} \\ m \cdot \ln(1-c) + \ln y_1 + \ln(1-y_2) + \ln c + \ln c + \ln(1-c), \text{ if } \frac{1}{2} < c < \frac{2}{3} \\ m \cdot \ln(1-c) + \ln c + \ln(1-c) + \ln c + \ln c + \ln(1-c), \text{ if } 0 \le c \le \frac{1}{2} \end{cases}$$

The maximum of $f(\tilde{x}, c)$ is attained at the maximum of the following extreme values:

$$\max_{0\le c\le 1} f(c)=\begin{cases}(m+3)\cdot\ln\frac{1}{3}, \text{ corresponding to } c=\frac{2}{3}\\ (m+5)\cdot\ln\frac{1}{2}-\xi, \text{ corresponding to } c=\frac{1}{2}+\varepsilon\\ 3\cdot\ln\left(\frac{3}{\mathrm{m+5}}\right)+(m+2)\cdot\ln\left(\frac{m+2}{m+5}\right), \text{ corresponding to } c=\frac{3}{m+5}\end{cases}$$

where $\varepsilon > 0, \xi > 0$, ε and ξ are arbitrary.

Note that if *m=0*, $\max_{0\le c\le 1} f(c)=3\cdot\ln\frac{1}{3}\Rightarrow\hat{c}=\frac{2}{3}\Leftrightarrow\hat{\pi}=1-\hat{c}=\frac{1}{3}$.

If $m\ge 1$, it is always true that

$$(m+3)\cdot\ln\frac{1}{3}\le(m+5)\cdot\ln\frac{1}{2}-\xi \quad \text{and}$$

$$(m+5)\cdot\ln\frac{1}{2}-\xi\le 3\cdot\ln\left(\frac{3}{\mathrm{m+5}}\right)+(m+2)\cdot\ln\left(\frac{m+2}{m+5}\right)$$

with arbitrary ξ>0. Hence,

$$\max_{0\le c\le 1} f(c)=3\cdot\ln\left(\frac{3}{\mathrm{m+5}}\right)+(m+2)\cdot\ln\left(\frac{m+2}{m+5}\right)$$

$$\Rightarrow\hat{c}=\frac{3}{m+5}\Leftrightarrow\hat{\pi}=1-\hat{c}=\frac{m+2}{m+5}$$

A simulation study of the above approach is also presented in the appendix.

4.5 Summary and further research

In this thesis, we first have introduced the non-parametric maximum likelihood method and the mixture model as well as the “hidden” model for cure-rate. These two models have been shown to be equivalent under the nonparametric maximum likelihood method. As preliminary results, we have found the NPMLE of uncensored data as well as Type-I censoring and Type-II censoring. Afterwards, we have shown that, in the uncensored model and also, Random censoring and Interval censoring (case-1), the cure-rate is not estimable under the non-parametric maximum likelihood method when no cures are actually observed. However, we have been able to give explicit solutions for the event time distribution for all the above censoring models. A proof has been given of the almost sure convergence of $\sup_x F(x)$ to $(1-\pi)$, where $\sup_x F(x)$ is the supremum of the MLE of the underlying distribution function, and π is the true underlying cure-rate, for random censoring and interval censoring (case-1), even when no cures are observed.

Since in most applications, the data are interval censored, there has been a need for some theoretical and numerical results on MLE for complex interval censored event data. In this work, we have described and illustrated the "max-min formula", for the estimation of the distribution function, proposed by Groeneboom and Wellner (1992), and modifed it to get the optimal solution to the maximization problem under a cure-mixture model, when some cures are observed. Finally, we have performed a simulation study to give some numerical results as well.

For further research, one may seek the MLE for interval censoring case-2 (IC-2). Recent studies of interval censoring have focused on IC-2 data, which involves a time-to-event variable X whose value is never observed but is known to lie in the time interval between two consecutive inspections times T and U. Consider a random sample $X_1, \cdots, X_n$ from unknown distribution function $F_\pi(x) = (1-\pi) \cdot F_0(x)$ on the real line, where π represents for the cure-rate. Instead of observing this sample directly, for each i a quadruple with $(T_i, U_i, \Delta_i, \Gamma_i)$ is observed. Here T_i and U_i are random time points, independent of X_i, with $T_i < U_i$ a.s. and the indicators $\Delta_i = 1_{(-\infty, T_i]}(X_i)$ and $\Gamma_i = 1_{(T_i, U_i]}(X_i)$ give information on the position of X_i with respect to these time points. Given a realization $(t_i, u_i, \delta_i, \gamma_i)$ of the n quadruples, the log-likelihood of the distribution function can be defined as

$$\ln(\pi, F_0) = \sum_{i=1}^{n} \left\{ \delta_i \cdot \ln\left(c \cdot F_0(t_i)\right) + \gamma_i \cdot \ln\left(c \cdot \left(F_0(u_i) - F_0(t_i)\right)\right) + (1 - \delta_i - \gamma_i) \cdot \ln\left(1 - c \cdot F_0(u_i)\right) \right\}$$

where $c = 1 - \pi$.

Jongbloed (1998) has suggested a modified iterative convex minorant (ICM) algorithm and the expectation maximization (EM) algorithm for IC-2 data. However, so far, very few theoretical results on the IC-2 have been addressed in scientific publications, and this particular situation cannot be handled straightforwardly by statistical software packages, even in the case of no cure-rate (i.e., $\pi = 0$). For this reason, we shall keep attempting to find an optimal solution.

APPENDIX

For the following simulation work, we implement the approach illustrated in Section 4.4 in S-plus.

First, generate a sample of size *N*+*n*. The first data set of size *N* is generated using $(V_1,V_2,\cdots V_N)$~$Uniform(0,1)$, and *m*, the number of cures, is defined as $m=\sum_{i=1}^{N}1\{V_i \le \pi_0\}$, where π_0 is an arbitrary initial value of the cure-rate. The second data set of size n , independent of the first sample, is generated using $(C_1,C_2,\cdots C_N)$~$Exponential(\lambda)$, $(T_1,T_2,\cdots T_N)$~$k\cdot Gamma(\alpha,\beta)$, and $(U_1,U_2,\cdots U_N)$~$Uniform(0,1)$, where *k* is some appropriate integer to make T_i and C_i comparable. The indicator, δ_i , is found by $\delta_i=\begin{cases}0 & \text{if } U_i \le \pi_0, \text{or } U_i>\pi_0 \text{ and } \mathrm{T}_i>\mathrm{C}_i \\ 1 & \text{if } U_i>\pi_0 \text{ and } \mathrm{T}_i \le C_i\end{cases}$. Then make the pairs of observations (C_i,δ_i) ordered statistics corresponding to C_i. The log-likelihood function based on the entire sample is given by

$$\ln L(\tilde{x},c)=m\cdot\ln(1-c)+(N-m)\cdot\ln c+\sum_{i=1}^{n}\{\delta_i\cdot\ln x_i+(1-\delta_i)\cdot\ln(1-x_i)\}$$

where $x_i=(1-\pi)\cdot F_0(C_i)=c\cdot F_0(C_i)$.

(i) Take $N=30,\ n=50,\ \pi_0=0.35,\ \lambda=0.005,\ k=800,\ \alpha=0.5,\ \beta=2$, it turns out that $m=12,\ N-m=18,$ and $\sum_{i=1}^{50}\delta_i=22$.

Applying the "max-min formula" (4.4.9), we have

$$y_i = \begin{cases} 1/6 & \text{for } i = 1,\cdots,6 \\ 1/4 & \text{for } i = 7,\cdots,10 \\ 1/3 & \text{for } i = 11,12,13 \\ 2/5 & \text{for } i = 14,\cdots,28 \\ 4/9 & \text{for } i = 29,\cdots,37 \\ 1/2 & \text{for } i = 38,39 \\ 2/3 & \text{for } i = 40,\cdots,48 \\ 1 & \text{for } i = 49,50 \end{cases}$$

And the likelihood function becomes

$$\ln L(\tilde{x},c) = 12\cdot\ln(1-c) + 18\cdot\ln c + \sum_{i=1}^{50}\{\delta_i\cdot\ln x_i + (1-\delta_i)\cdot\ln(1-x_i)\}$$

Take $y_i^* = y_i \wedge c$, discuss the extreme values of the log-likelihood function $\ln L(\tilde{x},c)$, with respect to different intervals which c lies in.

$$\ln L(\tilde{x},c) = \begin{cases} 40\ln(1-c)+40\ln c \overset{def}{=} f_1(c),\ \text{if } 0\le c\le 0.16667 \\ 35\ln(1-c)+39\ln c+3\ln(0.17\times 0.83) \overset{def}{=} f_2(c),\ \text{if } 0.16667 < c \le 0.25 \\ 32\ln(1-c)+38\ln c+3\ln(0.17\times 0.83)+2\ln(0.25\times 0.75) \overset{def}{=} f_3(c),\ \text{if } 0.25 < c \le 0.33333 \\ 30\ln(1-c)+37\ln c+3\ln(0.17\times 0.83)+2\ln(0.25\times 0.75)+\ln(0.33\times 0.67^2) \overset{def}{=} f_4(c), \\ \qquad \text{if } 0.33333 < c \le 0.4 \\ 21\ln(1-c)+31\ln c+3\ln(0.17\times 0.83)+2\ln(0.25\times 0.75)+\ln(0.33\times 0.67^2)+\ln(0.4^5\times 0.6^{10}) \\ \qquad \overset{def}{=} f_5(c),\ \text{if } 0.4 < c \le 0.44444 \\ 16\ln(1-c)+27\ln c+3\ln(0.17\times 0.83)+2\ln(0.25\times 0.75)+\ln(0.33\times 0.67^2)+\ln(0.4^5\times 0.6^{10}) \\ \qquad +\ln(0.44^4\times 0.56^5) \overset{def}{=} f_6(c),\ \text{if } 0.44444 < c \le 0.5 \\ 15\ln(1-c)+26\ln c+3\ln(0.17\times 0.83)+2\ln(0.25\times 0.75)+\ln(0.33\times 0.67^2)+\ln(0.4^5\times 0.6^{10}) \\ \qquad +\ln(0.44^4\times 0.56^5)+2\ln 0.5 \overset{def}{=} f_7(c),\ \text{if } 0.5 < c \le 0.66667 \\ 12\ln(1-c)+20\ln c+3\ln(0.17\times 0.83)+2\ln(0.25\times 0.75)+\ln(0.33\times 0.67^2)+\ln(0.4^5\times 0.6^{10}) \\ \qquad +\ln(0.44^4\times 0.56^5)+2\ln 0.5+\ln(0.67^6\times 0.33^3) \overset{def}{=} f_8(c),\ \text{if } 0.66667 < c \le 1 \end{cases}$$

The extreme values are

$$\begin{cases} \max f_1(c) = -34.2933, \text{ corresponds to } c = \dfrac{1}{6} \\ \max f_2(c) = -29.0273, \text{ corresponds to } c = 0.25 \\ \max f_3(c) = -25.9164, \text{ corresponds to } c = \dfrac{1}{3} \\ \max f_4(c) = -24.3594, \text{ corresponds to } c = 0.4 \\ \max f_5(c) = -23.6429, \text{ corresponds to } c = \dfrac{4}{9} \\ \max f_6(c) = -22.9939, \text{ corresponds to } c = 0.5 \\ \max f_7(c) = -22.3452, \text{ corresponds to } c = \dfrac{26}{41} \\ \max f_8(c) = -22.3869 - \xi, \text{ corresponds to } c = 0.67 + \varepsilon \text{ (with arbitrary } \varepsilon, \xi > 0) \end{cases}$$

The maximum of $\ln L(\tilde{x}, c)$ is attained at the maximum of these extreme values, which is $\max f_7(c)$, the corresponding value of *c* is hence the NPMLE:

i.e.,

$$\hat{c} = \frac{26}{41} \Leftrightarrow \hat{\pi} = 1 - \hat{c} = \frac{15}{41}$$

which is close to the initial value (0.35) of the cure-rate.

(ii) In an attempt to find how the parameters, such as sample size and initial value of π, effect of the NPMLE of cure-rate, we fix the above generated data set, and start with different arbitrary values. It is obvious that the expression of the log-likelihood function varies according to different sample sizes, *N* and *n*, and hence yield different estimator of cure-rate. For our interest, we perform the above approach all over again for various initial values of π, and get the following table of results.

π_0	$\hat{\pi}$	$\lvert \hat{\pi} - \pi_0 \rvert$	Number of cures	$\sum_{i=1}^{50} \delta_i$
0.15	1/16	0.0875	2	26
0.25	3/16	0.0625	6	23
0.35	15/41	0.1659	12	22
0.45	7/16	0.0125	14	17
0.6	27/43	0.0279	21	15
0.7	30/43	0.0023	24	14
0.8	32/41	0.0195	26	10
0.9	44/51	0.0373	28	8

Procedure to get the NPMLE based on "max-min formula"

Input:

delta: indicator of whether an individual is censored or not, which is got from simulated data set

n: size of data set

```
begin
  for m=1 to n
  begin
     for i=1 to m
     begin
        for k=m to n
        begin
           sum:=0;
```

```
                for j=i to k
                begin
                    sum:=sum + delta[j];
                end;
                commin[k]:=sum / (k – i + 1);
            end;

            min:=commin[k];
            for k=m to 50;
            begin
                if min > commin[k] then min:=commin[k];
            end;
            commax[i]:=min;
        end;

        max:=commax[i];
        for i=1 to m
        begin
            if max < commax[i] then max:=commax[i];
        end;
        max;
    end;
end.
```

BIBLIOGRAPHY

[1] Aalen, O. O. (1988). Heterogeneity in survival analysis. Statistics in Medicine 7, 1121-1137.

[2] Aalen, O. O. (1992). Modelling heterogeneity in survival analysis by the compound Poisson distribution. Annals of Applied Probability 4, 951-972.

[3] Chen,M-H., Ibrahim, J.G., and Sinha, D. (1999). A new bayesian model for survival data with a surviving fraction, Journal of the American Statistical Association, September 1999, 94, 447, Thoery and Methods.

[4] David, H. A. (1981). Order Statistics. John Wiley & Sons, New York.

[5] Farewell, V. T. (1977). A model for a binary variable with time-censored observations. Biometrika 64, 43-46.

[6] Farewell, V. T. (1982). The use of mixture models for the analysis of survival data with long-term surviviors. Biometrika 38, 1041-1046.

[7] Groeneboom, P. and Wellner, J.A. (1992). Information bounds and nonparametric maximum likelihood estimation, Birkhauser Verlag, Basel.

[8] Jongbloed, G. (1998). The iterative convex minorant algorithm for nonparametric estimation, Journal of Computational and Graphical Statistics 7, 3, 310-321.

[9] Kaplan, E.L. and Meier, Paul (1958). "Nonparametric estimation from incomplete observations." Journal of the American Statistical Association 53, 457-481.

[10] Klein, J.P. and Moeschberger, M.L. (1997). Survival analysis—Techniques for

censored and truncated data, Springer, New York.

[11] Kuk, A. Y. C. and Chen, C. -H. (1992). A mixture model combining logistic regression with proportional hazards regression. Biometrika 79, 531-541.

[12] Laska, E.M. and Meisner, M.J. (1992). Nonparametric estimation and testing in a cure model, Biometrics 48, 1223-1234.

[13] Longini, I. M. and Halloran, M. E. (1996). A frailty mixture model for estimating vaccine efficacy. Applied Statistics 45. 165-173.

[14] Miller, R. G.(1981), Survival analysis, Wiley series in probability and mathematical statistics: applied probability and statistics.