Journal of Applied Sciences

ISSN 1812-5654

science alert

ANSInet
an open access publisher
http://ansinet.com

Journal of Applied Sciences 6 (6): 1258-1264, 2006
ISSN 1812-5654

An Analytical Approach on Non-parametric Estimation of Cure Rate Based on Uncensored Data

[1]M. Taj Uddin, [2]A. Sen, [3]M.S. Noor, [1]M.N. Islam and [1]Z.I. Chowdhury
[1]Department of Statistics, Shahjalal University of Science and Technology, Sylhet, Bangladesh
[2]Department of Mathematics and Statistics, Concordia University, Canada
[3]Department of Mathematics, Shahjalal University of Science and Technology, Sylhet, Bangladesh

Abstract: This study deals with the analysis of non parametric estimation of Cure rate parameter. Frailty models have been proposed for cure rate estimation. In this research we have tried to estimate the cure parameter from the Frailty model using non-parametric maximum likelihood estimation (NPMLE) method. To perform the analysis, we consider two cases: i) cured and non cured group simultaneously and ii) non cured group only. The analysis showed that cute rate estimator can not be obtained from the analytical solution of the estimating equation but may be obtained from the numerical solution of the equation.

Key words: NPMLE method, cure rate parameter, Frailty model

INTRODUCTION

When analyzing survival data from clinical trials, it is sometimes clear that a non-zero proportion of patients can be considered as cured. Cantor and Shuster (1992) make a constructive discussion about parametric and non-parametric methods for estimating cure rate based on censored survival data. They use Kaplan-Meier (1958) method for non-parametric estimation of cure proportion. On the other hand for parametric estimation of cure rates, they assume a survival function S(t) for which $\lim_{t\to\infty} S(t) = S(\infty) > 0$ i.e., in a proportion of patients the event never occurs, using MLE one can estimate $S(\infty)$, which is considered as cure rate. Since, in many cases, distributional form clearly does not fit the survival data well, the use of non-parametric methods is attractive. However, the Kaplan-Meier method is not free of problems and limitations. Miller (1983) shows that the asymptotic efficiency of the Kaplan–Meier estimator tends to be low relative to the MLE for a given distribution.

Tsodikov (2001) studied the estimation of survival function based on proportional hazard model with cure. He proposed an algorithm to fit the proportional hazard model restricted by the fixed survival rates at the end of observation period. He used parametric cure model to estimate the proportion of long term survivors. To combine the stability of the parametric method, the survival function is estimated non-parametrically conditional on the cure rates provided by the parametric analysis. Peng and Carriere (2002) have proposed parametric and semi-parametric cured models for cure rate estimation. In their study, several parametric and semi-parametric models are compared and their estimation methods are discussed within the framework of EM algorithm. They showed that semi-parametric cure models can achieve efficiency levels similar to those of parametric cure models, provided that the failure time distribution is well specified and non-cured patients have an increasing hazard rate. They also recommended that the semi-parametric model is a viable alternative to parametric cure models. They have proposed mixture models for these analyses. The main objective of this analysis is to develop the non-parametric estimating equation to estimate the cure parameter of the model.

Cure model: Cure models were first proposed 50 years ago (Boag, 1949; Berkson *et al*., 1952; Haybittle, 1959) and have since received regular attention in the statistical literature (Haybittle, 1965; Mould *et al*., 1975; O'Neill, 1979; Farewell, 1982; Goldman, 1984; Sposto and Sather, 1985; Farewell, 1986; Halpern, 1987; Goldman, 1991; Sposto *et al*., 1992; Sposto, 2002; Kuk *et al*., 1992; Maller *et al*., 1992; Cantor *et al*., 1992; Ghitany *et al*., 1994; Yakovlev, 1994; Cantor *et al*., 1994; Ghitany *et al*., 1995; Lee *et al*., 1995; Zhou *et al*., 1995; Laska *et al*., 1992; Tsodikov *et al*., 1998; Peng *et al*., 1998; Gieser *et al*., 1998; Tsodikov, 1998; Hauck *et al*., 1997; De Angelis *et al*., 1999; Sy and Tayler, 2000). However, they have not attained wide use or acceptance in the medical research

Corresponding Author: Mohammed Taj Uddin, Department of Statistics, Shahjalal University of Science and Technology, Sylhet-3114, Bangladesh

literature, perhaps in part because of their reliance on particular parametric forms. However, parametric cure models provide a good empirical description of outcome in paediatric cancer data (Sposto and Sather, 1985; Sposto *et al*., 1992). Most importantly, they provide a single analytic method within which the effect of treatments and prognostic factors on the proportion cured can be assessed separately from their effect on the time to failure.

Frailty model: Chen *et al.* (1999) have proposed a different type of cure rate model, which is considered as a frailty model. The frailty model can be derived as follows:

Suppose that for an individual in the population, N denotes the total number of carcinogenic cells (often called clonogens) for that individual left active after the initial treatment and then assume that N has a Poisson distribution with mean θ. Also let Z_i denote the random time for the ith clonogenic cell to produce a detectable cancer mass. That Z_i can be viewed as an incubation time for the ith clonogenic cell. The variable Z_i, i = 1, 2, ….., are assumed to be identically independently distributed (i.i.d.) with a common distribution function F(t)=1-S(t)and are independent of N. Where S(t)is the survival function. The time to relapse of cancer can be defined by the random variable $T=\min\{Z_i, 0\le i\ge N\}$, $P(Z_0=\infty)=1$ and N is independent of the sequence Z_i, Z_2, … The survival function for T and hence the survival function for the population is given by

$$
\begin{aligned}
S_P(t) &= P(T>t)P(\text{no cancer by time t})\\
&= P(N=0)+P(Z_1>t,.....Z_N>t, N\ge 1)\\
&= \exp(-\theta) + P(Z_1>t)P(Z_2).....P(Z_N)P(N\ge 1)\\
&= \exp(-\theta)+S(t)S(t).....S(t)\{P(N=1)+P(N=2)+...\}\\
&= \exp(-\theta) + S(t)^k\sum_{k=1}^{\infty} P(N=k)\\
&= \exp(-\theta) + \sum_{k=1}^{\infty} S(t)^k\frac{e^{-\theta}\theta^k}{k!}\\
&= \exp(-\theta) + \exp(-\theta)\sum_{k=1}^{\infty}\frac{\{S(t)\theta\}^k}{k!}\\
&= \exp(-\theta)(1+\sum_{k=1}^{\infty}\frac{\{S(t)\theta\}^k}{k!})\\
&= \exp(-\theta)\sum_{k=1}^{\infty}\frac{\{S(t)\theta\}^k}{k!}\\
&= \exp(-\theta)\exp(\theta S(t))\\
&= \exp(-\theta + \theta S(t))\\
&= \exp(-\theta + \theta(1-F(t))), \text{ since } S(t) = 1-F(t)\\
&= \exp(-\theta F(t)), \qquad (1)
\end{aligned}
$$

The model (1) is not a proper survival function. Because $S_P(\infty) = \exp(-\theta)$. Note that F(0) = 0 and $F(\infty) = 1$. We observe that model (1) shows explicitly the contribution to the failure time of two distinct characteristics of tumor growth: the initial number of carcinogenic cells and the rate of their progression. Thus the model incorporates parameters bearing clear biological meaning. The model (1) is suitable for any type of failure data that has a surviving fraction (cure rate). Thus the model can be useful for modeling various types of failure time data, including time to relapse, time to death, time to first infection and so forth.

We also observe that the cure rate π is given by

$$S_P(\infty) = P(N = 0) = \exp\text{-}\theta) \qquad (2)$$

As $\theta\to\infty$, the cure rate tends to zero, whereas as $\theta\to o$, the cure rate tends to 1. i.e., the cure rate lies between 0 and 1. Note that by taking first derivative of (1), we get,

$$
\begin{aligned}
S'_P(t) &= \exp(-\theta F(t))(-\theta F'(t))\\
&= -\theta f(t)\exp(-\theta F(t))
\end{aligned}
$$

Where, $S'_P(t)$ and F'(t) denote the first derivative of $S_P(t)$ and F(t), respectively and F'(t) = f(t)
Or, $-S'_P(t) = \theta f(t)\exp(-\theta F(t))$
Since, $-S'_P(t) = \theta f_P(t)$

The density function corresponding to model (1) is given by

$$f_P(t) = \theta f(t)\exp(-\theta F(\theta)), \qquad (3)$$

We observe that $S_P(t)$ is not a proper survival function, because $S_P(\infty)\ne 0$. Therefore $f_P(t)$ is not a proper probability density function. But f(t) in model (3) is a proper density function.

MATERIALS AND METHODS

We try to estimate the cure parameter by using Non Parametric Maximum Likelihood Estimation (NPMLE) Method. The method is described as follows:

Non-parametric maximum likelihood method: Suppose that X is a random variable with probability density function $f(x;\theta)$, θ to be estimated and x_1, x_2, ... x_n is a random sample of size n. The joint probability density function of the random variable comprising the sample is called the likelihood function of the sample and is given by

$$L(x_1, x_2, \ldots x_n; \theta) = f(x_1; \theta)\, f(x_2; \theta) .. F(x_n; \theta) = \prod_{i=1}^{n} F(x_i; \theta), \quad (4)$$

If there exists a value $L(x_1,x_2,\ldots x_n;\theta^*) \geq L(x_1,x_2,\ldots x_n;\theta)$ such that θ for all possible choices of θ^*, then based upon the meaning of likelihood function, θ^* maximizes the Eq. (4) and this value θ^* is considered as the maximum likelihood estimate. Choosing the value of θ that makes it most likely that the data would be as obtained is certainly a reasonable approach. Therefore, if a value θ^* can be found such that θ^* maximizes the likelihood function (4) for a given set of sample values $x_1,x_2,\ldots x_n$, then θ^* is called the maximum likelihood estimate for the given set of sample values. Since the likelihood function is a function of parameter, θ under the sample information $(x_1,x_2,\ldots x_n)$, the maximum likelihood estimate will be a function of the sample values. If the sample functional relationship between the estimate and the sample information $(x_1,x_2,\ldots x_n)$ holds for all possible choices of the x_i, then that functional relationship can be taken as an estimation rule and the result will be an estimator $\hat{\theta}$ of θ, this estimator being known as maximum likelihood estimator or the maximum likelihood filter.

In Non-parametric maximum likelihood method, we write the non-parametric likelihood function as

$$L_n(F) = \prod_{i=1}^{n} (\Delta F(x_i)), \quad (5)$$

where, F(.) is the common distribution function of $X_i, 1 \leq i \leq n$ and $\Delta F(x_i)$-$F(x_i-)$ is the jump of F(.) at $X_i, 1 \leq i \leq n$. Putting $\Delta F(x_i) = p_i$ in (5), we can write

$$L_n(P_1,P_2,\ldots,P_n) = \prod_{i=1}^{n} Pi, \quad (6)$$

Now we maximize (6) subject to condition

$$\sum_{i=1}^{n} P_i = 1,\ P_i \geq 0, 1 \leq i \leq n, \quad (7)$$

The log-likelihood function becomes

$$\log Ln(P_i,p_2,\ldots P_n) = \sum_{i=1}^{n} \log Pi, \quad (8)$$

By using Lagrange multiplier method we can maximize (8). By adding a Lagrange multiplier λ, (8) becomes

$$\log L_n(P_1,P_2,\ldots P_n) = \sum_{i=1}^{n} \log Pi - \lambda\left(\sum_{i=1}^{n} Pi - 1\right) \quad (9)$$

Thus, the non-parametric maximum likelihood estimator of p_i is obtained by the solution of the following equations

$$\frac{\partial \log L}{\partial p_i} = 0, \frac{\partial \log L}{\partial \lambda} = 0 \quad I = 1,2,..,n, \quad (10)$$

Now, $\frac{\partial \log L}{\partial p_i} = 0$ gives

$$\frac{1}{p_i} - \lambda = 0$$

Therefore, $p_i = \frac{1}{\lambda}$, $i = 1,2,\ldots,n$, (11)

Similarly, $\frac{\partial \log L}{\partial \lambda} = 0$ gives

$$\sum_{i=1}^{n} p_i = 1 \quad (12)$$

From (11), we can write

$$\sum_{i=1}^{n} p_i = \frac{n}{\lambda} \quad (13)$$

Using (12) in (13) we obtain

$$\hat{\lambda} = n \quad (14)$$

Therefore, the non-parametric maximum likelihood estimator of p_i is

$$\hat{p}_i = \frac{1}{n}, \quad i = 1,2,\ldots\ldots,n$$

RESULTS AND DISCUSSION

Suppose that T is the life time of a patient. Then $P(X=\infty) = \lim_{t\to\infty} P(X > t) = e^{-\theta}$, which is considered as cure rate. On the other hand, $P(X<\infty) = 1 - e^{-\theta}$, $0 \leq t \leq \infty$ which is the probability of non-cured.

By using Non-Parametric Maximum Likelihood Estimation (NPMLE) method, we can estimate the cure parameter. For uncensored data we consider the following cases:

Case-(a): $F_0(.)$, $f_0(.)$ and θ are unknown. Here we observe both cured and non-cured group.

Suppose that we have the data in the form (x_i,ϵ_i), $i = 1,2,..,n,$, where x_i denotes the survival time for the ith patient, ϵ_i is the cured indicator with 1 if x_i is not cured and 0 otherwise i.e., $\epsilon_i = 1_{\{xi<\infty\}}$.

$$\text{Let } F(x) = \int_0^x \theta f_0(t) e^{-\theta F_0(t)} dt$$

$$= \int_0^{F_0(x)} \theta e^{-\theta z} \quad dz \text{ [by putting } F_o(t) = z]$$

$$= \theta[\frac{e^{-\theta z}}{-\theta}]_0^{F(x)}$$

$$= 1\text{-}e^{-\theta F_0(x)}, \tag{15}$$

Therefore, the non-parametric likelihood function is given by

$$L_n(\theta, F) = \prod_{i=1}^{n} (\Delta F(x_i))^{\in_i} (e^{-\theta})^{1-\in_i} \tag{16}$$

Where, $\Delta F(x_i)$ = jump of F(.) at x_i

$$= 1 - e^{-\theta F_0(x_i)} - (1 - e^{-\theta F_0(x_{i-})})$$

$$= e^{-\theta F_0(x_{i-})} - e^{-\theta F_0(x_i)}$$

$$= e^{-\theta F_0(x_{i-})}(1 - \frac{e^{-\theta F_0(x_i)}}{e^{-\theta F_0(x_{i-})}})$$

$$= e^{-\theta F_0(x_{i-})}(1 - e^{-\theta[F_0(x_i) - F_0(x_{i-})]})$$

$$= e^{-\theta(p_1+p_2+......+p_{i-1})}(1 - e^{-\theta p_i}) \tag{17}$$

Where, $p_i = F_0(x_i) - F_0(x_{i-})$

Therefore, the above likelihood function (16) can be written as

$$L_n(\theta, p_1, p_2, ...p_n) = \prod_{i=1}^{n}$$

$$(e^{-\theta(p_1+p_2+......+p_{i-1})}(1 - e^{-\theta p_i}))^{\in_i} (e^{-\theta})^{1-\in_i} \tag{18}$$

We want to maximize (18) subject to condition

$$P_1 + P_2 + ... P_n = 1, p_i \geq 0, \tag{19}$$

The log-likelihood function becomes

$$\log L_n(\theta, p_1, p_2, ..., p_n) =$$

$$-\theta \sum_{i=1}^{n} \in_i (p_1 + p_2 + ... + p_{i-1}) +$$

$$\sum_{i=1}^{n} \in_i \log(1 - e^{-\theta p_i}) - \theta \sum_{i=1}^{n} (1 - \in_i)$$

$$\text{Or,} \log L_n(\theta, p_1, p_2, ..., p_n) =$$

$$-\theta \sum_{i=1}^{n-1} p_i (\sum_{j=i+1}^{n} \in_i) + \sum_{i=1}^{n} \in_i \log(1 - e^{-\theta p_i}) - \theta \sum_{i=1}^{n} (1 - \in_i) \tag{20}$$

By using Lagrange multiplier method we can maximize (20). Adding Lagrange multiplier λ, we can write (20) as follows

$$\log L_n(\theta, p_1, p_2, .., p_n) = -\theta \sum_{i=1}^{n-1} p_i (\sum_{j=i+1}^{n} \in_j) +$$

$$\sum_{i=1}^{n} \in_i \log(1 - e^{-\theta p_i}) -$$

$$\theta \sum_{i=1}^{n} (1 - \in_i) \; - \lambda(\sum_{i=1}^{n} p_i - 1) \tag{21}$$

Therefore, the non-parametric maximum likelihood estimators of θ and P_i are obtained by the solution of the following equations

$$\frac{\partial \log L}{\partial \theta} = 0, \frac{\partial \log L}{\partial p_i} = 0 \tag{22}$$

Now $\frac{\partial \log L}{\partial \theta} = 0$ gives,

$$-\sum_{i=1}^{n-1} p_i (\sum_{j=i+1}^{n} \in_j) + \sum_{i=1}^{n} \in_i \frac{p_i e^{-\theta p_i}}{(1 - e^{-\theta p_i})} - \sum_{i=1}^{n} (1 - \in_i) = 0 \tag{23}$$

Similarly, $\frac{\partial \log L}{\partial p_i} = 0$ gives, I = 1,2,...,n-1

$$-\theta \sum_{j=i+1}^{n} \in_j + \in_i \frac{\theta e^{-\theta p_i}}{(1 - e^{-\theta p_i})} - \lambda = 0 \tag{24}$$

and $\frac{\partial \log L}{\partial p_n} = 0$ gives,

$$\in_n \frac{\theta e^{-\theta p_n}}{(1 - e^{-\theta p_n})} - \lambda = 0 \tag{25}$$

Multiplying (24) by p_i and summing over i from 1 to n, we obtain the following equation

$$-\theta \sum_{i=1}^{n} p_i (\sum_{j=i+1}^{n} \in_j) + \sum_{i=1}^{n} p_i \in_i \frac{\theta e^{-\theta p_i}}{(1 - e^{-\theta p_i})} - \lambda \sum_{i=}^{n} p_i = 0 \tag{26}$$

Since, $\sum_{i=1}^{n} p_i = 1$, so the equation (26) becomes

$$-\theta \sum_{i=1}^{n} p_i \sum_{j=i+1}^{n} \in_j + \sum_{i=1}^{n} p_i \in_i \frac{\theta e^{-\theta p_i}}{(1 - e^{-\theta p_i})} - \lambda = 0 \tag{27}$$

From (23) and (27), we obtain

$$\hat{\lambda} = \theta \sum_{i=1}^{n} (1 - \in_i) \tag{28}$$

Now using the estimate of λ in (24), we get

$$-\theta\sum_{j=i+1}^{n}\epsilon_j+\epsilon_i\frac{\theta e^{-\theta p_i}}{(1-e^{-\theta p_i})}-\theta\sum_{i=1}^{n}(1-\epsilon_i)=0 \quad , i=1,2, \dots,n, \qquad (29)$$

Therefore, $$\sum_{j=i+1}^{n}\epsilon_j-\epsilon_i\frac{e^{-\theta p_i}}{(1-e^{-\theta p_i})}+\sum_{i=1}^{n}(1-\epsilon_i)=0 \qquad (30)$$

Comment: This Eq. 30 can be considered as an estimating equation of P_i which can not be solved analytically but may be solved numerically. So the solution of this equation is our desired estimate of P_i.

Again using (28) in (25), we may obtain the estimate of P_n

$$\epsilon_n\frac{\theta e^{-\theta p_n}}{(1-e^{-\theta p_n})}-\theta\sum_{i=1}^{n}(1-\epsilon_i)=0$$

Comment: The above equation also can not be solved analytically but may be solved numerically.

Finally the estimate of θ may be obtained from the numerical solution of Eq. (23).

Case-(b): $F_0(.)$, $f_0(.)$ and θ are unknown and only non-cured group are observed. Suppose that we have the data in the form (x_i,ϵ_i), I = 1,2,..,n,where, x_i denotes the survival time for the ith patient, ϵ_i is the cured indicator with 1 if x_i is not cured and 0 otherwise. i.e., $\epsilon_i=1_{\{xi<\infty\}}$. The non-parametric likelihood function can be written as

$$L_n^{nc}(\theta,p_1,p_2,\dots,p_n)=\prod_{i=1}^{n}\frac{e^{-\theta(p_1+p_2+\dots p_{i-1})}(1-e^{-\theta p_i})}{1-e^{-\theta}}$$

$$=\frac{e^{-\theta\sum_{i=1}^{n}(p_1+p_2+\dots p_{i-1})}\prod_{i=1}^{n}(1-e^{-\theta p_i})}{(1-e^{-\theta})^n}$$

The log-likelihood function is given by

$$\log L_n^{nc}(\theta,p_1,p_2,\dots p_n)=-\theta\sum_{i=1}^{n}(p_1+p_2+\dots\dots p_{i-1})+\sum_{i=1}^{n}\log(1-e^{-\theta p_i})-n\log(1-e^{-\theta})$$

$$=-\theta\sum_{i=1}^{n-1}(n-i)p_i)+\sum_{i=1}^{n}\log(1-e^{-\theta p_i})-n\log(1-e^{-\theta}) \qquad (31)$$

We want to maximize (31) subject to condition $\sum_{i=1}^{n}p_i=1$, By using Lagrange multiplier method we can maximize (31). Adding Lagrange multiplier λ in (31), we can write

$$\log L_n^{nc}(\theta,p_1,p_2,\dots p_n)=-\theta\sum_{i=1}^{n-1}(n-i)p_i+\sum_{i=1}^{n}\log(1-e^{-\theta p_i})-n\log(1-e^{-\theta})-\lambda(\sum_{i=1}^{n}p_i-1) \qquad (32)$$

Therefore, the non-parametric maximum likelihood estimators of θ and p_i are obtained by the solution of following equations

$$\frac{\partial\ln L}{\partial\theta}=0,\frac{\partial\ln L}{\partial p_i}=0 \qquad (33)$$

Now $\frac{\partial\ln L}{\partial\theta}=0$ gives

$$-\sum_{i=1}^{n-1}(n-i)p_i+\sum_{i=1}^{n}\frac{p_ie^{-\theta p_i}}{1-e^{-\theta p_i}}-\frac{ne^{-\theta}}{1-e^{-\theta}}=0 \qquad (34)$$

and $\frac{\partial\ln L}{\partial p_i}=0$ gives, i =1,2,…,n-1

$$-(n-i)\theta+\frac{\theta e^{-\theta p_i}}{1-e^{-\theta p_i}}-\lambda=0 \qquad (35)$$

$\frac{\partial\log L}{\partial p_n}=0$ gives

$$\frac{\theta e^{-\theta p_n}}{1-e^{-\theta p_n}}-\lambda=0 \qquad (36)$$

Multiplying (35) by P_i and summing over i from1 to n we obtain

$$-\theta\sum_{i=1}^{n}p_i(n-i)+\sum_{i=1}^{n}p_i\frac{\theta e^{-\theta p_i}}{1-e^{-\theta p_i}}-\lambda\sum_{i=1}^{n}p_i=0 \qquad (37)$$

Since $\sum_{i=1}^{n}p_i=1$, so the above equation becomes

$$-\theta\sum_{i=1}^{n}p_i(n-i)+\sum_{i=1}^{n}p_i\frac{\theta e^{-\theta p_i}}{1-e^{-\theta p_i}}-\lambda=0 \qquad (38)$$

Multiplying Eq. (34) by θ and subtracting from (38), we obtain

$$-\lambda+\frac{n\theta e^{-\theta}}{1-e^{-\theta}}=0$$

Therefore, $$\hat{\lambda}=\frac{n\theta e^{-\theta}}{1-e^{-\theta}} \qquad (39)$$

Using the estimate of λin (35), we obtain

$$-(n-i)\theta+\frac{\theta e^{-\theta p_i}}{1-e^{-\theta p_i}}-\frac{n\theta e^{-\theta}}{1-e^{-\theta}}=0$$

$$\text{or,}\quad -(n-i)+\frac{e^{-\theta p_i}}{1-e^{-\theta p_i}}-\frac{ne^{-\theta}}{1-e^{-\theta}}=0$$

$$\text{or,}\quad \frac{e^{-\theta p_i}}{1-e^{-\theta p_i}}=(n-i)+\frac{ne^{-\theta}}{1-e^{-\theta}}$$

$$\text{Thus,}\quad p_i=-\frac{1}{\theta}[\log\{(n-i)+\frac{ne^{-\theta}}{1-e^{-\theta}})\}+\log(1-e^{-\theta p_i})] \quad (40)$$

Comment: This is an estimating equation of P_i, i =1,2,..,n-1. and it can not be solved analytically but it may be solved numerically and the solution of this equation is the desired estimate of P_i.

Again, using (39) in (36), we obtain

$$\frac{\theta e^{-\theta p_n}}{1-e^{-\theta p_n}}-\frac{n\theta e^{-\theta}}{1-e^{-\theta}}=0$$

$$\frac{e^{-\theta p_n}}{1-e^{-\theta p_n}}=\frac{ne^{-\theta}}{1-e^{-\theta}}$$

$$p_n=-\frac{1}{\theta}\{\log(\frac{ne^{-\theta}}{1-e^{-\theta}})+\log(1-e^{-\theta p_n})\} \quad (41)$$

We observe that the estimate of P_n may be obtained from the numerical solution of the Eq. (41)

Finally the estimate of θ can be obtained from the numerical solution of the Eq. (34).

CONCLUSIONS

Considering both non-cured and cured group, when we assume $f_0(.)$ and $F_0(.)$ are unknown, we found a non-parametric estimating equation of θ. Unfortunately we could not find an explicit solution for θ. But hopefully, this non-parametric estimating equation may be solved numerically by choosing an appropriate numerical method. Also we have found the same result when we consider non-cured group only. That is, in both the cases we have found a non-parametric estimating equation for θ.

REFERENCES

Berkson, J. and R.P. Gage, 1952. Survival curves for cancer patients following treatments. J. Am. Stat. Associ., 47: 501-515.

Boag, J.W., 1949. Maximum likelihood estimates of the proportion of patients cured by cancer therapy. J. R. Stat. Soc., Series B, 11: 15-44.

Cantor, A.B. and J.J. Shuster, 1992. Parametric versus non-parametric methods for estimating cure rates based on censored survival data. Statistics in Med., 11: 931-937.

Cantor, A.B. and J. Shuster, 1994. Parametric versus non-parametric methods for estimating cure rates based on censored survival data-reply. Statistics in Med., 13: 986-986.

Chen, M.H., J.G. Ibrahim and D. Sinha, 1999. A new Bayesian model for surviving data with a surviving fraction. J. Am. Stat. Assoc., 94: 909-919.

De Angelis, R., R. Capocaccia, T. Hakulinen, B. Soderman and A. Verdecchia, 1999. Mixture models for cancer survival analysis: Application to population based data with covariates. Statistics in Med., 18: 441-454.

Farewell, V.T., 1982. The use of mixture models for the analysis of survival data with long term survivors. Biometrics, 38: 1041-1046.

Farewell, V.T., 1986. Mixture model in survival analysis: are they worth the risk? Can. J. Stat., 14: 257-262.

Geiser, P.W., M.N. Chang, P.V. Rao, J.J. Shuster and J. Pullen, 1998. Modeling cure rates using the Gompertz model with covariate information. Statistics in Med., 17: 831-839.

Ghitany, M.E., R.A. Maller and S. Zhou, 1994. Exponential mixture models with long term survivors and covariates. J. Multivariate Analy., 49: 218-241.

Ghitany ME., Maller R. and S. Zhou, 1995. Estimating the proportion of immunes in censored samples- A simulation study. Statistics in Medicine, 14: 39-49.

Goldman, A.I., 1984. Survivorship analysis when cure is a possibility: A Monte- Corlo study. Statistics in Med., 3: 153-163.

Goldman, A.I., 1991. The cure model and time confounded risk in the analysis of survival and other timed events. J. Clin. Epidemiol., 44: 1327-1340.

Halpern, J. and B.W. Brown, 1987. Cure rate models: power of the log rank and generalized Wilcoxon tests. Statistics in Med., 6: 483-489.

Hauck, W.W., L.T. Mckee and B.J. Turner, 1997. Two-part survival models applied to administrative data for determining rate and predictors for maternal-child transmission of HIV. Statistics in Med., 16: 1683-1694.

Haybittle, J.L., 1965. A two parameters model for the survival curve of treated cancer patients. J. Am. Stat. Assoc., 60: 16-26.

Haybittle, J.L., 1959. The estimation of the proportion of patients cured after treatment for cancer of the breast. Br. J. Radiol., 32: 725-733.

Kaplan, E.L. and P. Meier, 1958. Non-parametric estimation from incomplete observations. J. Am. Stat. Assoc., 53: 457-481.

Kuk, A.Y.C. and C.H. Chen, 1992. A mixture model combining logistic regression with proportional hazard regression. Biometrika, 79: 531-541.

Laska, E.M. and M.J. Meisner, 1992. Non-parametric estimation and testing in a cure model. Biometrics, 48: 1223-1234.

Lee, J.W. and H.N. Sather, 1995. Group sequential methods for comparison of cure rates in clinical trials. Biometrics, 51: 756-763.

Maller, R.A. and S. Zhou, 1992. Estimating the proportion of immunes in a censored sample. Biometrika, 79: 731-739.

Maller, R.A. and S. Zhou, 1995. Testing for the presence of immune or cured individuals in censored survival data. Biometrics, 51: 1197-1205.

Miller, R.G., 1983. What price Kaplan-Meier? Biometrics, 39: 1077-1081.

Mould, R.F. and J.W. Boag, 1975. A test of several parametric models for estimating success rate in the treatment of Carcinoma cervix uteri. Br. J. Cancer, 32: 529-550.

O' Neill, T.J., 1979. Distribution free of cure time. Biometrika, 66: 184-187.

Peng, Y. and K.C. Carriere, 2002. An empirical comparison of parametric and semi parametric cure model. Biometrical. J., 44: 1002-1014.

Peng, Y., K.B. Dear and J.W. Denham, 1998. A generalized F mixture model for cure rate estimation. Statistics in Med., 17: 813-830.

Sposto, R. and H.N. Sather, 1985. Determining the duration of comparative clinical trials while allowing for cure. J. Chronic Dis., 38: 683-690.

Sposto, R., H.N. Sather and S.A. Baker, 1992. A comparison of tests of the difference in the proportion of patients who are cured. Biometrics, 48: 87-99.

Sposto, R., 2002. Cure model analysis in cancer: An application to data from the children's cancer group. Statistics in Med., 21: 293-312.

Sy, J.P. and J.M. Taylar, 2000. Estimation in a Cox proportional hazard cure model. Biometrics, 56: 227-236.

Tsodikov, A., 1998. A proportional hazard model taking account of long-term survivors. Biometrics, 54: 1508-1516.

Tsodikov, A., 2001. Estimation of Survival based on proportional hazards when cure is a possibility. Mathematical and Computer Modeling, 33: 1227-1236.

Tsodikov, A., M. Loeffler and A. Yakovlev, 1998. A cure model with time-changing risk factor: an application to the analysis of secondary Leukemia. A report from the International Database on Hodskins Disease. Statistics in Med., 17: 27-40.

Yakovlev, A.Y., 1994. Letter to the editor. Parametric versus non-parametric methods for estimating cure rates based on censored survival data. Statistics in Med., 13: 983-985.

Zhou, S. and R.A. Maller, 1995. The likelihood ratio test for the presence of immunes in a censored sample. Statistics in Med., 27: 181-201.