International Journal of Statistical Sciences ISSN 1683–5603
Vol. 3, 2004, pp 21–27

Estimation of Survivability of Clinically Diagnosed Aids Cases Under Competing Risks Using Cox's Regression Model

Gurprit Grover and Dulumoni Das
Department of Statistics
Faculty of Mathematical Sciences
University of Delhi
Delhi, India

[Received April 19, 2003; Revised July 29, 2003; Accepted November 3, 2003]

Abstract

This paper presents the Cox's regression model[1] to estimate the survivability based on an exposure of 10 years of clinically diagnosed AIDS cases. The hazard rate is taken to be a function of the explanatory variables and a constant. The parameters are estimated using Newton-Raphson iteration method[2]. An application of the developed model is shown and the survivabilities of AIDS cases corresponding to different covariates are compared with the help of graphs.

Keywords and Phrases: HIV mortality rate, Incubation period, T-4 cells, Double decrement life table, Sero-positivity, Simple birth and death process.

AMS Classification: 62Q05; 60G$_{XX}$.

1 Introduction

In this paper an effort is made to estimate the survivability of clinically diagnosed AIDS patients who may be exposed to the many other diseases or risks operating in the population. Moreover, the hazard rate at any point of time may not depend only on time but also on a host of explanatory variables or covariates, some of which may not be expressed in quantitative form. For analyzing censored survival data allowing for covariates, the suitable consideration is the Cox's regression model[1], which can be elegantly adapted in the present context. The model is based on estimation technique and the estimates of the parameters concerning the individual covariates, independent of the parameters concerning the hazard rate with respect to time are obtained using Newton-Raphson iteration method[2].

2 Development of the Model

Suppose that we have a sample of n clinically diagnosed AIDS cases and let corresponding to every AIDS case $j = 1, 2, \cdots, n$ the random variable T_j be the survival time, each of whom is observed for a fixed time C. Let the hazard rate of the j^{th} individual at time 't', be given by

$$\lambda_j(t) = \lim_{dt \to 0} \frac{P[t < T_j \leq t + dt/T_j > t]}{dt}$$

Using the simplest form of the Cox's regression model[1] to fit the present situation, the hazard rate can be taken as

$$\lambda_j(t) = \lambda\ exp[\underset{\sim}{\beta}' \underset{\sim}{Z}_j(t)] \tag{1}$$

Where λ= the overall hazard rate irrespective of other covariates. $exp[\underset{\sim}{\beta}' \underset{\sim}{Z}_j(t)]$= hazard rate with respect to p factors having corresponding time independent intensities $\beta_1, \beta_2, \cdots, \beta_p$ with which j^{th} individual will be affected at time t. $\underset{\sim}{\beta}' = (\beta_1, \beta_2, \cdots, \beta_p)$ and $\underset{\sim}{Z}_j(t) = (Z_{j1}(t), Z_{j2}(t), \cdots, Z_{jp}(t))_{1 \times p}$. Let $R(t) = \{j : T_j, 0 \leq t \leq C\}$ be the risk set i.e. the set of individuals exposed to the risk under observations.

Since this analysis is restricted to the effect of covariates like Tuberculosis, Diarrhoea, Measles and Others in the mortality of AIDS patients, therefore, we have chosen Cox's proportional hazard model with four covariates. Assuming Chiang's proportionality assumption holds in respect of the competing risks due to Tuberculosis, Diarrhoea, Measles and Others the selection of Cox's proportional hazard model is justified.

Given that we have a set of n AIDS cases in the sample the probability that the j^{th} person dies at time t, $0 \leq t \leq C$ (assuming that death occurs independently) is given by

$$P_j = \frac{\lambda_j(t)}{\sum_{j \in R(t)=[0,C]} \lambda_j(t)} = \frac{\lambda\ exp[\underset{\sim}{\beta}' \underset{\sim}{Z}_j(t)]}{\sum_{j \in R(t)} \lambda \exp[\underset{\sim}{\beta}' \underset{\sim}{Z}_j(t)]} = \frac{exp[\underset{\sim}{\beta}' \underset{\sim}{Z}_j(t)]}{\sum_{j \in R(t)} \exp[\underset{\sim}{\beta}' \underset{\sim}{Z}_j(t)]}$$

Then the conditional likelihood[2] is given by

$$L(\underset{\sim}{\beta}') = \prod_{j=1}^{n} \frac{\exp[\underset{\sim}{\beta}' \underset{\sim}{Z}_j(t)]}{\sum_{j \in R(t)} \exp[\underset{\sim}{\beta}' \underset{\sim}{Z}_j(t)]} \tag{2}$$

3 Estimation of Parameters

Taking log on both sides of equation (2) we get,

$$LL(\underset{\sim}{\beta}) = \sum_{j=1}^{n}[\{\underset{\sim}{\beta}'\underset{\sim}{Z_j}(t)\} - log\{\sum_{j\in R(t)} exp[\underset{\sim}{\beta}'\underset{\sim}{Z_j}(t)\}]$$

The p normal equations are

$$\frac{\delta LL(\underset{\sim}{\beta}')}{\delta\beta_i} = 0; \quad i = 1, 2, \cdots, p$$

$$\Rightarrow A_i(\beta_1, \beta_2, \cdots, \beta_p) = \frac{\delta LL(\underset{\sim}{\beta}')}{\delta\beta_i} = 0$$

$$\Rightarrow A_i(\beta_1, \beta_2, \cdots, \beta_p) = \sum_{j=1}^{n}\left[Z_{ji}(t) - \frac{\sum_j Z_{ji}(t)\exp[\underset{\sim}{\beta}'\underset{\sim}{Z_j}(t)]}{\sum_{j\in R(t)}\exp[\underset{\sim}{\beta}'\underset{\sim}{Z_j}(t)]}\right]; \quad i = 1, 2, \cdots p \qquad (3)$$

Differentiating the set of equations in (3) with respect to $\beta_1, \beta_2, \cdots, \beta_p$ respectively we get

$$B_i(\beta_1, \beta_2, \cdots, \beta_p) = \frac{\delta^2 LL(\beta)}{\delta^2\beta_i} = 0; \quad i = 1, 2, \cdots, p$$

$$\begin{aligned}\Rightarrow B_i(\beta_1, \beta_2, \cdots, \beta_p) &= -\sum_{j=1}^{n}\left[\frac{\sum Z_{ji}(t)\exp\{\beta' Z_j(t)\}}{\sum_{j\in R(t)}\exp\{\beta' Z_j(t)\}} + \left\{\frac{\sum Z_{ji}(t)\exp\{\beta' Z_j(t)\}}{\sum_{j\in R(t)}\exp\{\beta' Z_j(t)\}}\right\}^2\right] \\ &= 0 \qquad (4)\end{aligned}$$

Using Newton-Raphson iteration method[2], the parameters $\beta_1, \beta_2, \cdots, \beta_p$ can be easily estimated.

4 Distribution of Survival Time

Since T_j is a random variable representing the life time of the j^{th} AIDS case, who is observed for a fixed censoring time C, therefore, for every t, $0 \le t \le C$, the hazard rate at time t, is given by,

$$\begin{aligned} S_j(t) &= P[\text{of survival of the } j^{th} \text{ individual at least time } t] \\ &= 1 - F_j(t); \quad 0 \le t \le C \end{aligned}$$

$$\begin{aligned} \text{where } F_j(t) &= \text{Failure function of the } j^{th} \text{ individual at time } t \\ &= \frac{1-\exp[-\int_0^t \lambda_j(x)dx]}{1-\exp[-\int_0^C \lambda_j(x)dx]}; \quad 0 \le t \le C \end{aligned} \tag{5}$$

$$\begin{aligned} \therefore\ S_j(t) &= F_j(t) \\ &= 1-\frac{1-\exp[-\int_0^t \lambda_j(x)dx]}{1-\exp[-\int_0^C \lambda_j(x)dx]} = 1-\frac{1-\exp[-\int_0^t \lambda \exp\{\underset{\sim}{\beta}' \underset{\sim}{Z}_j(x)\}dx]}{1-\exp[-\int_0^C \lambda \exp\{\underset{\sim}{\beta}' \underset{\sim}{Z}_j(x)\}dx]} \end{aligned} \tag{6}$$

5 Application

The model developed in this paper is applied to a practical problem i.e. to Delhi's AIDS data[3] shown in Table 1 and Table 2. Here we are considering the four covariates viz: Tuberculosis, Diarrhoea, Measles and Others to which the AIDS patients are exposed. The clinically diagnosed AIDS patients may also be suffering from any of these four diseases. All those patients, who, either don't die or lost to follow up or withdrew from the study for some reason during 10 year observation period are supposed to have censored survival time $C = 10$ years.

Table 1: AIDS CASES IN DELHI

Year	AIDS cases	AIDS cases also suffering from			
		Tuberculosis	Diarrhoea	Measles	Others
Dec. 1994	19	16	1	1	1
Dec. 1995	25	21	3	0	1
Dec. 1996	64	53	2	3	6
Dec. 1997	66	57	3	3	3
Dec. 1998	68	57	3	5	3
Dec. 1999	72	55	7	6	4
Dec. 2000	139	117	9	4	9
Dec. 2001	158	132	9	9	8
Dec. 2002	106	89	5	5	7

Some adjustment is made to the above data for the smoothing of the survival curves. Using the adjusted data is used to the set of equations (3) and (4) and applying Newton-Raphson iteration method[2] the estimates of the parameters of our model are obtained and then we estimate the hazard rate as well as the survivability of each AIDS case, which are shown in Table 3.

Table 2: DEATH CASES IN DELHI

Year	Death cases	Death cases with			
		Tuberculosis	Diarrhoea	Measles	Others
Dec.1994	12	9	1	1	1
Dec.1995	14	11	2	0	1
Dec.1996	12	9	0	1	2
Dec.1997	16	13	1	1	1
Dec.1998	18	14	1	2	1
Dec.1999	6	4	0	1	1
Dec.2000	24	19	2	1	2
Dec.2001	27	21	2	2	2
Dec.2002	32	25	2	2	3

(Here the number of AIDS cases and the number of death cases in Delhi up to Dec. 1993 is 45 and 40 respectively)

Table 3: Estimates of Hazard rate and Survivability of AIDS case

	Tuberculosis		Diarrhoea		Measles		Others	
t	$\hat{\lambda_j}(t)$	$\hat{S_j}(t)$	$\hat{\lambda_j}(t)$	$\hat{S_j}(t)$	$\hat{\lambda_j}(t)$	$\hat{S_j}(t)$	$\hat{\lambda_j}(t)$	$\hat{S_j}(t)$
1.0	1.29099	0.27500	0.95932	0.38311	0.89850	0.40711	0.89741	0.40755
1.2	1.28841	0.23552	0.95250	0.33791	0.89703	0.35932	0.89257	0.36089
1.4	1.28583	0.19604	0.94567	0.29270	0.89555	0.31153	0.88773	0.31423
1.6	1.28325	0.15656	0.93885	0.24750	0.89408	0.26378	0.88290	0.26758
1.8	1.28067	0.11708	0.93202	0.20229	0.89260	0.21594	0.67806	0.22092
2.0	1.27809	0.07760	0.92520	0.15709	0.89113	0.16815	0.87322	0.17426
2.2	1.26468	0.06661	0.91083	0.14110	0.88333	0.15000	0.85672	0.15800
2.4	1.25127	0.05562	0.89645	0.12510	0.87553	0..13185	0.84022	0.14174
2.6	1.23786	0.04463	0.88208	0.10911	0.86772	0.11370	0.82371	0.12547
2.8	1.22445	0.03363	0.86770	0.09311	0.85992	0.09555	0.80721	0.10921
3.0	1.21104	0.02643	0.85333	0.07712	0.85212	0.07740	0.79071	0.09295

From Table 3, we have obtained the figures 1, 2, 3 and 4 by plotting survival functions of AIDS cases corresponding to different covariates. From the figures, we can have a fairly good idea about the survivability of the AIDS patient under already mentioned risks. Since the estimates of the parameters corresponding to covariates are the partial likelihood estimates, therefore, they have the asymptotic properties of maximum likelihood estimators, which provides most efficient estimators in large samples.

Figure 1: Survival function of AIDS cases corresponding to Tuberculosis

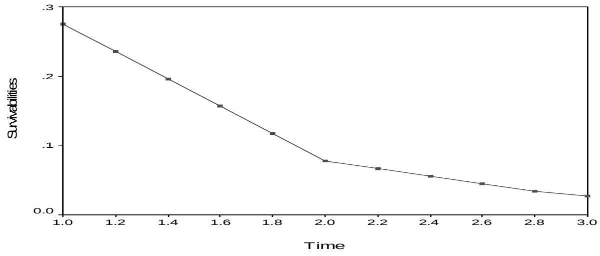


Figure 2: Survival function of AIDS cases corresponding to Diarrhoea

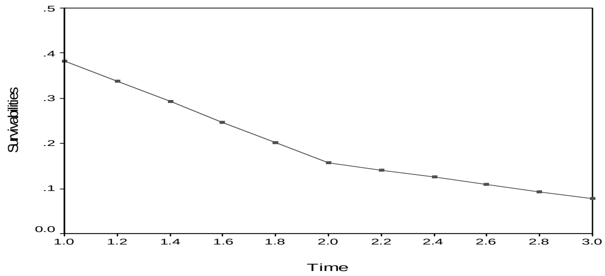


Figure 3: Graph of Survival function of AIDS cases corresponding to Measles

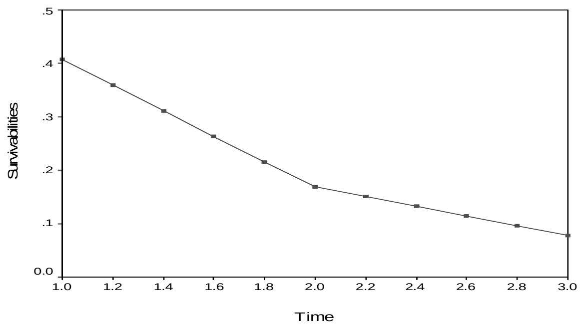
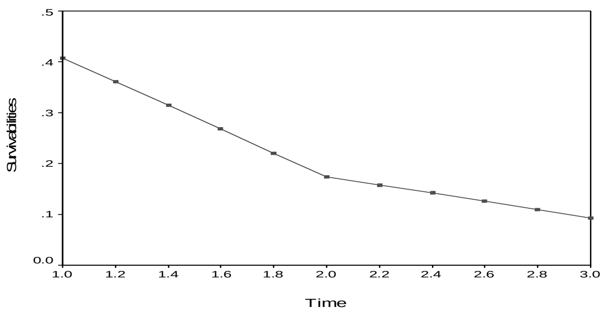


Figure 4: Survival function of AIDS cases corresponding to others



References

[1] Biswas, S. (1993) : Applied Stochastic Processes : A Biostatistical and Population Oriented Approach; New Age International Publishers Limited, Wiley Eastern Limited.

[2] Lee, Elisa. T. (1992) : Statistical methods for Survival Data Analysis. Second Edition, A. Wiley Interscience Publication, John Wiley & Sons, INC.

[3] National AIDS prevention and control policy, Delhi State AIDS control Society. Govt. of N.C.T. of Delhi.