Volume : 5 | Issue : 10 | October 2016 • ISSN No 2277 - 8179 | IF : 3.508 | IC Value : 69.48

Original Research Paper

Disease Diagnosis by Using Cox PH and Frailty Models

Statistic

KEYWORDS : Cox PH, Frailty, TB.

E.Kumar	DepartmentOf Statistics, Presidency College, Chennai-5
M. Muthuvijayalakshmi	Department Of Statistics,National Institute for Research in Tuberculosis, (ICMR), Chennai-31
R.Ravanan	DepartmentOf Statistics, Presidency College, Chennai-5

ABSTRACT

Disease diagnosis is one of the challenging tasks in the medical field. Several techniques are available for classification , In this paper, we have concentrated on the Cox PH model and Frailty model with applications to clinical trails. The broad objectives of this paper to study new approaches for treatment-patient interactions using frailty models. The output results are carried out by comparing the models which gives the better result.

Introduction

The early efforts in development of survival analysis were predominantly focused on estimation of the hazard function and the survival function Cox(1972) , Tsiatis(1975) and Peterson (1976) established that neither the hazard function nor the survival function is identifiable if there are censored observations. Chiang (1968) referred to the hazard function estimated with censored observations as crude hazard and denoted as h(t). The actual hazard function was called the net hazard. Therefore, in most survival analysis applications, as key assumption is made regarding the equality of the crude hazard and the net hazard. Survival models and their applications have been investigated and recommended in the last two decades. Semi parametric regression models contain components of parametric and non parametric components, and the advantages of parametric and non parametric regression models. In case of an optimal combination of the two approaches, the semi parametric regression model has the efficiency (low variance) of parametric models and the flexibility (small bias) of non parametric models. For the four decades the Cox PH model has been used comprehensively to examine the covariate effects on the hazard function for the failure time variable.

Cox proportional hazard (PH) model

Let $T_1, T_2, \ldots, T_n$; $C_1, C_2, \ldots, C_n$ be independent random variables. C_1 is the censoring time associated with the survival time T_i. We observe $(Y_1, \delta_1), \ldots, (Y_n, \delta_n)$.

Where $Y_i = T_i \wedge C_i$, $\delta_i = I(T_i \leq C_i)$

Also available are $x_1, x_2, \ldots, x_n$, where $x_i = (x_{i1}, x_{i2}, \ldots, x_{ip})'$

Is the vector of independent variables or covariates associated with the dependent variable T_i.

The Hazard function is $\lambda(t; x) = \dfrac{f(t,x)}{1-F(t,x)}$ ----------(1)

Where we have included the dependence of the distribution of T upon the covariates in x. The proportional hazard model assumes

$$\lambda(t; x) = e^{\beta' x} \lambda_0(t), \qquad \text{----------(2)}$$

Where = (β_1 , ..., β_p) , is the vector of regression coefficients. The hazard rate is the product of a scalar and the function $\lambda_o(t)$, where the scalar depends on the regression coefficients and the covariates. The theory could work if $e^{\beta' x}$ where replaced by any sensible $h(\beta, x)$ where h is positive. Both the regression coefficients β and the underlying hazard $\lambda_o(t)$ are unknown.

Frailty model

Ordinary survival models deal with the simplest case of independent and identically distributed data. This is based on the assumption that the study population is homogeneous. But it is a basic observation of medical statistics that individuals differ greatly. So do the effects of a drug, or the influence of various explanatory variables. This heterogeneity is often referred to as variability and it is generally recognized as one of the most important sources of variability in medical and biological applications. To address the problem of heterogeneity in a population resulting from unobserved co-variates, Vaupel et al. (1979) suggested a random effects model for durations.

Volume : 5 | Issue : 10 | October 2016 • ISSN No 2277 - 8179 | IF : 3.508 | IC Value : 69.48

They introduced the notion or frailty and applied it to population data. The classical and mostly applied frailty model assumes a proportional hazards model that is conditional on the random effect (frailty). To be more specific, the hazard of an individual depends additionally on an unobservable, age-independent random variable Z, which acts multiplicatively on the baseline hazard function λ_0

$$\lambda\,(t,z) \;=\; Z\,\lambda_0\,(t) \qquad \text{----------(3)}$$

Database and software

The sample data has been collected from the TRC (ICMR) at Chennai. The sample is about the pulmonary TB patients who were undergone diagnosis with different treatments. The data is recorded for 681 patients, for whom survival times are given in months. Time to sputum conversion during treatment period is the event of interest and all others are censored. A variable associated with the event of interest of an individual; the values 1 that is, the patient has cured from the disease pulmonary TB (free from pulmonary TB) and zero if the survival time is right censored. The original is condensed and presented in table and analysis is carried out using SPSS. The variable status has been arrived using the smear test. The status has been taken '1' if it is an event and '0' if it is censored. The time has been obtained in terms of month where the smear test of an individual turns to be zero (free from TB).The ten different types of regimen were compared with the base alone. The dependence of the hazard function on a treatment is to be modeled, where the Treatment has ten levels including control.

Table1: Cox PH Regression

Case Processing Summary

		N	Percent
Cases available in analysis	Event(a)	639	93.7%
	Censored	42	6.2%
	Total	681	99.9%
Cases dropped	Cases with missing values	1	.1%
	Cases with negative time	0	.0%
	Censored cases before the earliest event in a stratum	0	.0%
	Total	1	.1%
Total		682	100.0%

a Dependent Variable: time

Block 1: Method = Enter

Omnibus Tests of Model Coefficients(a,b)

-2 Log Likelihood	Overall (score)			Change From Previous Step			Change From Previous Block		
	Chi-square	df	Sig.	Chi-square	df	Sig.	Chi-square	df	Sig.
7421.012	22.442	11	.021	22.296	11	.022	22.296	11	.022

a Beginning Block Number 0, initial Log Likelihood function: -2 Log likelihood: 7443.308
b Beginning Block Number 1. Method = Enter

Variables in the Equation

	B	SE	Wald	df	Sig.	Exp(B)	95.0% CI for Exp(B)	
							Lower	Upper
age	.006	.004	2.664	1	.103	1.006	.999	1.013
treatment			15.785	9	.071			
treatment(1)	.154	.181	.723	1	.395	1.166	.818	1.663
treatment(2)	.002	.186	.000	1	.993	1.002	.696	1.442
treatment(3)	.322	.195	2.726	1	.099	1.380	.941	2.023
treatment(4)	-.044	.233	.036	1	.849	.957	.607	1.509
treatment(5)	.427	.204	4.387	1	.036	1.532	1.028	2.283
treatment(6)	.190	.232	.675	1	.411	1.210	.768	1.905
treatment(7)	.306	.203	2.270	1	.132	1.358	.912	2.021
treatment(8)	.184	.223	.679	1	.410	1.201	.776	1.859
treatment(9)	.449	.217	4.279	1	.039	1.567	1.024	2.399
sexcode	-.244	.091	7.189	1	.007	.784	.656	.937

Table2. Cox Frailty Models

Cox regression --
Breslow method for ties Number of obs = 681
Gamma shared frailty Number of groups = 49
Group variable: age

No. of subjects = 681 Obs per group: min = 1
No. of failures = 639 avg = 13.89796
Time at risk = 3128 max = 67

Volume : 5 | Issue : 10 | October 2016 • ISSN No 2277 - 8179 | IF : 3.508 | IC Value : 69.48

```
                                Wald chi2(2)      =     10.11
Log likelihood =  -3716.7526                Prob > chi2       =    0.0064
------------------------------------------------------------------------------
          _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   treatment |   1.034194   .0148497     2.34   0.019     1.005495    1.063712
     sexcode |   .8211252    .072401    -2.24   0.025     .6908066    .9760281
-------------+----------------------------------------------------------------
       theta |   .0031815   .0117777
------------------------------------------------------------------------------
Likelihood-ratio test of theta=0: chibar2(01) =      0.08 Prob>=chibar2 = 0.387
```

Table3 Cox Frailty Models

```
stcox treatment sexcode age

        failure _d:  status == 1
  analysis time _t:  time

Iteration 0:   log likelihood = -3721.6539
Iteration 1:   log likelihood = -3715.5538
Iteration 2:   log likelihood = -3715.5474
Refining estimates:
Iteration 0:   log likelihood = -3715.5474

Cox regression -- Breslow method for ties

No. of subjects =          681                 Number of obs   =       681
No. of failures =          639
Time at risk    =         3128
                                LR chi2(3)      =     12.21
Log likelihood =  -3715.5474                Prob > chi2     =    0.0067
------------------------------------------------------------------------------
          _t | Haz. Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
   treatment |   1.034878   .0148534     2.39   0.017     1.006172    1.064404
     sexcode |   .7997046   .0717489    -2.49   0.013     .6707496    .9534518
         age |   1.005881   .0037156     1.59   0.012     1.006244    1.053189
```

Conclusion

The Event of interest (time to sputum conversion during the treatment period) is 93.7%.

Using Cox PH regression treatment and gender influence the event of interest (time until sputum conversion). Moreover the treatment groups and gender defined as categorical variable. The control treatment is defined as reference category. The rest of the test groups are compared with the control and it test groups influence the event of interest. In Gender female is the reference category and time to sputum conversion for male is slightly higher than the female based on their hazards rate value. Age is not influence the event of interest in both the form. Using Cox Frailty univariate approach, age also influence the event of interest. Moreover it accounts heterogeneity and the amount of heterogeneity is also presented in the table2&3.In the frailty model we conclude that the treatment, age, and sexcode are significant. It is concluded that this kind of time to event clinical trial data these techniques may be considered for better insight.©

References

1. Collect.D (2003), Modeling Survival Data in Medical Research, 2nd ed.,Chapman and Hall, New York
2. Cox D.R and Oakes.D (1984), Analysis of Survival Data,Chapman and Hall, New York.
3. David G.Kleinbaum Mitchel Klein(1996), Survival Analysis, 2nd ed., Springer, New York
4. Gupta,S.C. and Kapoor,V.K(2006), Fundamentals of Mathematical Statistics. Ed. 3, Sultan Chand and Sons, New Delhi
5. Rupert G. Miller, Jr. (1975), Survival Analysis, John Wiley and Son, New York.