Journal of Multivariate Analysis 128 (2014) 226–238

Contents lists available at ScienceDirect

Journal of Multivariate Analysis

journal homepage: www.elsevier.com/locate/jmva

Semiparametric varying-coefficient study of mean residual life models

Guangren Yang [a,b], Yong Zhou [b,c,*]

[a] Department of Statistics, School of Economics, Jinan University, Guangzhou 510632, China
[b] School of Statistics and Management, Shanghai University of Finance and Economics, Shanghai 200433, China
[c] Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China

ARTICLE INFO

Article history:
Received 11 August 2013
Available online 1 April 2014

AMS 2000 subject classifications:
62N01
62N02
62N99

Keywords:
Mean residual life
Varying-coefficient models
Counting process
Estimating equations
Right censoring

ABSTRACT

In this paper, we consider a flexible class of semiparametric varying-coefficient mean residual lifetime (MRL) models that depended on an exposure variable where some effects may be functions of the exposure variables and some may be constants. We develop three-step estimation procedures to estimate parametric and nonparametric parts in the semiparametric varying-coefficient MRL model under the right censoring. We first establish a local estimating equation with inverse probability of censoring weighting (IPCW) approach, and estimate parametric and nonparametric parts simultaneously. In the second step, substituting the nonparametric estimator into estimating equations, we can obtain the global parametric estimating equation to refine the estimators of the parametric part. The asymptotic normality of the parametric estimator is established, meanwhile it has been shown that the estimators achieve the $\sqrt{n}$ convergence rate under some smoothed conditions. In the third step, substituting the refined parametric estimator into the local estimating equations, we can obtain updated local nonparametric estimating equations to estimate the nonparametric part, and show that the asymptotic normality of nonparametric estimator is still true. Some numerical simulations are conducted to illustrate performance of the proposed methods.

1. Introduction

The MRL function of a subject is defined as the expected remaining (residual) lifetime of the subject, when the subject has survived up to a given time point. The MRL function can play a very useful role in many fields of applied sciences such as biomedical studies, reliability models, actuarial sciences and economics, where the goal is often to characterize the residual life expectancy.

In many fields of application, we are interested in the remaining life expectancy rather than the probability of immediate failure or the distribution of a failure time. For example, a cancer patient may care much more on how long he/she can survive from the time of diagnosis. For another example, an insurance company would be interested in the mean residual life of its customers. Sometimes, the mean residual life function can serve as a more useful tool than the hazard function in applied sciences, because the mean residual life function has a direct interpretation in terms of average behavior. For example, patients in a clinical trial might be more interested to know how many more years they are expected to survive

* Corresponding author at: Academy of Mathematics and Systems Science, Chinese Academy of Sciences, Beijing 100190, China.
E-mail addresses: yzhou@amss.ac.cn, yzhou@amss.com.cn (Y. Zhou).

http://dx.doi.org/10.1016/j.jmva.2014.03.011
0047-259X/

given that they began a treatment at a certain time ago as compared to their risk of instantaneous dying given that they have started the treatment at a given time.

There are many authors to study mean residual life model in the literature. Bickel et al. [1], Rojo and Ghebremichael [23], Oakes and Dasu [20] and Maguluri and Zhang [18] proposed some estimating methods of MRL function with complete data. Chen et al. [10] and Chen and Cheng [8] focused on proportional mean residual life models with censored data. A class of additive mean residual life models were studied in [9]. Recently [26] studied a family of general transformation models which includes many common models, such as additive model, proportional mean residual model as special cases. Sun et al. [24] consider a time varying coefficient model of the mean residual life under censored data. Chan et al. [7] study the proportional mean residual life model for right-censored length-bias data by modified partial likelihood method. However, no work has been done to extend these methods to the semiparametric varying-coefficient MRL models with an exposure variable for censored data.

The varying coefficient model has gained considerable interest because it is a simple but useful extension of classical generalized linear model since pioneering work by Hastie and Tibshirani [14]. It is easy to model dynamical systems that led to applications in areas including nonparametric generalized regression [3,4], functional data modeling [22], time series analysis [16,3,4,2], longitudinal data analysis [15,12,13], and survival analysis [11,5] and nonparametric quantile regression [6].

We propose a robust inverse probability weight to adjust the bias induced by the censored data that can be used to estimate parameters and varying coefficients in the semiparametric varying coefficient function model with a specific link function. Moreover, we allow the inverse probability weight dependent on covariates which is more realistic. To improve the efficiency of estimators, we develop a three-stage estimating procedure using local linear regression technique and establish the asymptotic normality of proposed estimators for both the parametric and non-parametric components. The first step, we establish local estimating equations with inverse probability of censoring weighting (IPCW) approach, and estimate parametric and nonparametric parts simultaneously. The second step, substituting the nonparametric estimator into estimating equations, we can obtain global parametric estimating equations to refine the estimator of the parametric part. The asymptotic normality of the parametric estimator is established, it is to show that the estimators achieve the $\sqrt{n}$ convergence rate under some smoothed conditions. The third step, substituting the refined parametric estimator into local estimating equations, we can obtain updated the local nonparametric estimating equations to estimate the nonparametric part. The three-stage procedure not only gives consistent estimates of the coefficient functions and parameters but also allows us to improve the efficiency of estimator of the parameters. It is easy to show that nonparametric estimators have the asymptotic normality with the nonparametric convergence rate $\sqrt{nh}$ and parameter estimators converge to normality with root of n as the usual convergence rates.

The rest of this paper is organized as follows. In the next section, by using method of estimating equation and IPCW approach, we derive estimation procedures both the nonparametric and parametric components of the model, and establish three estimating equations. Section 3 studies the asymptotic properties of the resulting estimators. Section 4 reports results from simulation studies conducted for evaluating the proposed methods.

2. Models

Let T be a positive valued random variable that represents the lifetime of a subject on a probability space $(\Omega, \mathcal{F}, P)$ and has a finite expectation. The mean residual life (MRL) function of a survival time T, denoted by $m(t)$, is the expected remaining lifetime given survival up to time t. That is,

$$m(t) \equiv E[T - t|T > t] = \int_t^\infty \frac{S(\mu)}{S(t)} d\mu,$$

where $S(t) = Pr(T > t)$ is the survival function and $0/0 \equiv 0$. The relation of $S(\cdot)$ and $m(\cdot)$ is given by the inversion formula as

$$S(t) = 1 - F(t) = \frac{m(0)}{m(t)} \exp\left\{-\int_0^t \frac{1}{m(\mu)} d\mu\right\}.$$

From $m(t)$ and $S(t)$ it follows that $m(\cdot)$ uniquely determines $S(\cdot)$ and vice versa.

To assess the effects of covariates on the mean residual life, Oakes and Dasu [20] proposed the proportional, and takes the form

$$m(t|Z) = m_0(t) \exp(\beta^T Z), \tag{1}$$

where $m(t|Z)$ denotes the MRL function corresponding to a p-dimensional covariate vector Z, $m_0(t)$ is some unknown baseline MRL function and β is a vector of unknown regression parameters. Maguluri and Zhang [18] derived estimation procedures without censoring time. An additive mean residual life model was proposed by Chen and Cheng [9]. More recently, a class of transformed MRL models were proposed [26] which take the form

$$m(t|Z) = g\left\{m_0(t) + \beta^T Z\right\}, \tag{2}$$