5-2011

local likelihood in regression analysis of proportional mean residual life model with censored survival data

Yang Ni
Clemson University, yangn@clemson.edu

Local Likelihood in Regression Analysis of Proportional Mean Residual Life Model with Censored Survival Data

A Thesis
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Master of Science
Mathematics

by
Yang Ni
May 2011

Accepted by:
Dr. Karunarathna Kulasekera, Committee Chair
Dr. Colin Gallagher
Dr. Chanseok Park

Abstract

As a function of time t, mean residual life (MRL) is the remaining life expectancy of a subject given its survival to t. In survival analysis, the relationship between a survival time and a covariate can be conveniently modeled with the proportionality mean residual life (MRL) model proposed by Oakes and Dasu(1990) and provides an alternative to the Cox proportionality hazards model, Cox (1972). In this paper we consider the proportional MRL regression model with a nonparametric covariate effect. We discuss estimation of the proportional function when the baseline MRL function is not specified. We develop the asymptotic properties of the proportionality function. Simulation studies are presented to assess the finite sample behavior.

Dedication

The thesis is dedicated to my dear parents: Yang, Yong AND Ni, Jianming who introduce me to the joy of exploring from birth , enabling such of a study to take place!

Table of Contents

Chapter 1

Introduction

In the field of reliability and survival analysis, numerous research have been done on mean residual life function and hazard rate function due to their wide applications. It is often of interest to analyze the mean residual life as a function of time to characterize the survival over time. For a nonnegative survival time T with finite expectation, the mean residual life at time $t \geq 0$ is defined as

$$m(t) = E(T - t|T > t)$$

For a comprehensive review of previous research on this function, readers are refereed to Guess & Proschan (1988) and Csörgő & Zitikis (1996).

In many instances, scientists are interested in exploring the possible relationship between a survival time T and a covariate Z. To assess the effect of covariates on the mean residual life, we consider the proportional mean residual life model proposed by Oakes and Dasu (1990):

$$m(t|Z) = m_0(t) \exp(\beta^T Z) \tag{1.1}$$

where $m(t|Z)$ is the mean residual life corresponding to the p-vector covariate Z, $m_0(t)$ is some unknown baseline mean residual life function, and β is the regression parameter. Here β^T denotes the transpose of β. When one considers a nonparametric covariate effect, it is appropriate to use the model:

$$m(t|Z) = m_0(t)\Psi(Z) \tag{1.2}$$

where Z is the corresponding covariate, which could be one dimensional or a p-vector. In this article we only consider Z as a one dimensional covariate, and assume $\Psi(Z) = \exp(\psi(Z))$ to ensure the positivity of the MRL function.

Previous work on the mean residual life has focused on single-sample and two-sample cases with model (1.1). Oakes and Dasu (2003) has outlined the methods for these cases in their recent work, while Magulari and Zhang (1994) provide a way to estimate the parameters in model (1.1), but mainly for uncensored survival data. In most applications, the survival times of some subject are censored instead of being fully observed. To accommodate censoring, Chen and Jewell (2002) proposed an inference procedure for model (1.1) in the presence of censoring based on Magulari and Zhang's work (1994). Another straightforward approach is to apply the inverse-probability-of-censoring-weighted paradigm of Robins and Rotnitzky (1992) to the estimating equations from complete event times; however, this would require estimating or modeling the censoring distribution.

When it comes to a general form $\psi(Z)$ and a nonparametric baseline MRL function $m_0(t)$ in model (1.2), we employ localized likelihood equation and counting process theory to develop the estimation procedure for baseline MRL function $m_0(t)$, and $\psi(Z)$. Our inference procedure mimics the Semi-parametric Regression Analysis of Mean Residual Life with Censored Survival Data for model (1.1) by Chen and Cheng (2005), and the Local Likelihood and Local Partial Likelihood in Hazard Regression by Fan and Gijbels (1997).

Chapter 2

Estimation Procedure

Let T and C be the failure time and potential censoring time, respectively, and let $X = \min(T, C)$. Here, conditional on the covariate Z, T and C are assumed to be independent. In order to avoid lengthy technical discussion of the tail behavior of limiting distributions, we further assume that $0 < \tau = \inf\{t : P(X > t) = 0\} < \infty$. If necessary, Ying's (1993) treatment of the asymptotic properties beyond τ can be adapted.

Suppose the observed data set consists of n independent triplets (X_i, δ_i, Z_i), where $i = 1, \cdots, n$, $X_i = \min(T_i, C_i)$ and $\delta_i = I(T_i \leq C_i)$. Here $I(\cdot)$ is the indicator function. In addition, let $N_i(t) = I(X_i \leq t)\delta_i$, $Y_i(t) = I(X_i \geq t)$. Let $\Lambda_i(t)$ be the cumulative hazard rate function of T_i.

2.1 Survival Function

The conditional survival function of time T with conditional pdf $f(t|Z = z)$ is defined by $S(t|Z = z) = \int_t^\infty f(u|Z = z)\mathrm{d}u$. It is well known that the survival function of T given Z is related to the MRL function of time T given Z (Lawless 1998) in a proportional MRL model as

$$S(t|Z) = P(T \geq t|Z) = \frac{m(0|Z)}{m(t|Z)} \exp\left\{-\int_0^t \frac{1}{m(u|Z)}\mathrm{d}u\right\}. \tag{2.1}$$

From (2.1) and under model (1.2) we develop the relationship between $S(t|Z)$ and the hazard rate function $\lambda(t|Z)$. Taking log on both sides of (2.1) we get

$$\log S(t|Z) = \log \frac{m_0(0)\Psi(Z)}{m_0(t)\Psi(Z)} - \frac{1}{\Psi(Z)} \int_0^t \frac{1}{m_0(u)} \mathrm{d}u$$

Now take the derivative with respect to t on both sides to get

$$-\mathrm{d} \log S(t|Z) = \frac{\mathrm{d}m_0(t)}{m_0(t)} + \frac{1}{\Psi(Z)} \frac{\mathrm{d}t}{m_0(t)} \tag{2.2}$$

By the connection of survival function and hazard rate function, we also have

$$-\mathrm{d} \log S(t|Z) = \frac{f(t|Z)\mathrm{d}t}{S(t|Z)} = \lambda(t|Z)\mathrm{d}t = \mathrm{d}\Lambda(t|Z) \tag{2.3}$$

Now, (2.2) and (2.3) would give us

$$m_0(t)\mathrm{d}\Lambda(t|Z) = \frac{1}{\Psi(Z)}\mathrm{d}t + \mathrm{d}m_0(t) \tag{2.4}$$

2.2 Likelihood Function

Under model (1.2), suppose temporarily, that the baseline MRL function $m_0(\cdot)$ is known, and that $\psi(z)$ has been parameterized as $\psi(z) = \beta^T \mathbf{Z}$. Under independent and noninformative censoring, which means that the distribution of C does not involve the parameters β, it can be shown that the conditional likelihood function is given by

$$L = \prod_{i=1}^{n} \left\{ [f(X_i|Z_i)]^{\delta_i} \prod_{i=1}^{n} [S(X_i|Z_i)]^{1-\delta_i} \right\}.$$

This kind of likelihood appears often in the literature [c.f. Aitkin and Clayton(1980)]. Then the log-likelihood function is

$$\log L = \sum_{i=1}^{n} [\delta_i \log \lambda(X_i|Z_i) + \log S(X_i|Z_i)] \tag{2.5}$$

Now, using the proportional MRL model (1.1), the associated score function becomes

$$\frac{\partial \log L}{\partial \beta} = \sum_{i=1}^{n} \exp(-\beta^T \mathbf{Z}_i)\mathbf{Z}_i \left[-\frac{\delta_i}{\exp\{-\beta^T \mathbf{Z}_i\} + m_0'(X_i)} + \int_0^{X_i} \frac{1}{m_0(u)} \mathrm{d}u \right]. \tag{2.6}$$

The details of this derivation are given in the Appendix A.

2.3 Local Likelihood

Suppose now that the form of $\psi(z)$ is not specified, and that the p^{th} order derivative of $\psi(z)$ exists. Then, by Taylor's expansion,

$$\psi(Z) \approx \psi(z) + \psi'(z)(Z - z) + \cdots + \frac{\psi^{(p)}(z)}{p!}(Z - z)^p,$$

for Z in neighborhood of a single value z. Let h be the bandwidth parameter that controls the size of the local neighborhood and let K be the kernel function that smoothly weighs down the contribution of remote data points. Set

$$\mathbf{Z} = \{1, Z - z, \cdots, (Z - z)^p\}^T$$

and

$$\mathbf{Z}_i = \{1, Z_i - z, \cdots, (Z_i - z)^p\}^T$$

Then, locally around z, $\psi(Z)$ can be modeled as

$$\psi(Z) \approx \beta^T \mathbf{Z}, \tag{2.7}$$

where $\beta = (\beta_0, \cdots, \beta_p)^T = \{\psi(z), \cdots, \frac{\psi^{(p)}(z)}{p!}\}^T$. Using the local model(2.7), and incorporating the localizing weights, we obtain the local log-likelihood by (2.4), (2.5) and (2.6) as

$$l_n(\beta) = n^{-1} \sum_{i=1}^{n} \left[\delta_i \frac{\exp\{-\beta^T \mathbf{Z}_i\} + m_0'(X_i)}{m_0(X_i)} + \log \frac{m_0(0)}{m_0(X_i)} + \exp\{-\beta^T \mathbf{Z}_i\} \int_0^{X_i} \frac{1}{m_0(u)} \mathrm{d}u \right] K_h(Z_i - z) \tag{2.8}$$

with associated score function

$$\frac{\partial l_n(\beta)}{\partial \beta} = n^{-1} \sum_{i=1}^{n} \exp(-\beta^T \mathbf{Z}_i)\mathbf{Z}_i \left[-\frac{\delta_i}{\exp\{-\beta^T \mathbf{Z}_i\} + m_0'(X_i)} + \int_0^{X_i} \frac{1}{m_0(u)} \mathrm{d}u \right] K_h(Z_i - z), \tag{2.9}$$

where $K_h(t) = h^{-1}K(t/h)$. It can be shown that $l_n(\beta)$ is strictly concave with respect to β(details are given in section 2.5). Thus, this local log-likelihood has a unique maximizer with respect to β. Let $\widehat{\beta}$ be the maximizers of (2.8). Then, according to our parametrization, a natural estimator of $\psi^{(v)}(z), v = 0, \cdots, p$, for a known baseline MRL function $m_0(t)$ is

$$\widehat{\psi}^{(v)}(z) = v!\widehat{\beta}_v. \tag{2.10}$$

2.4 Sampling Properties

Our goal in this work is to establish the asymptotic normality of the estimator $\hat{\psi}(z)$. Since $\hat{\psi}(z) = \hat{\beta}_0$, a solution to the local score function, we begin this section with proving Bartlett identities for the local likelihood (2.9).

Proposition 1. First-order Bartlett identity

$$E\left\{\Psi(z)^{-1}\mathbf{Z}\left[-\frac{\delta}{\Psi(z)^{-1}+m_0'(X)}+\int_0^X \frac{1}{m_0(u)}\mathrm{d}u\right]\big|Z=z\right\}=0 \tag{2.11}$$

where $\Psi(z) = \exp\{\psi(z)\}$.

Proof. We know that

$$E\left\{m_0(t)\mathrm{d}N(t) - Y(t)\left[\Psi(z)^{-1}\mathrm{d}t + \mathrm{d}m_0(t)\right]\big|Z=z\right\}=0 \tag{2.12}$$

We also note that

$$\begin{aligned}
& \Psi(z)^{-1}\mathbf{Z}\left[-\frac{\delta}{\Psi(z)^{-1}+m_0'(X)}+\int_0^X \frac{1}{m_0(u)}\mathrm{d}u\right] \\
= \quad & -\int_0^\tau \frac{\Psi(z)^{-1}\mathbf{Z}}{m_0(t)\{\Psi(z)^{-1}+m_0'(t)\}}\left\{m_0(t)\mathrm{d}N(t)-Y(t)\left[\Psi(z)^{-1}\mathrm{d}t+\mathrm{d}m_0(t)\right]\right\}
\end{aligned}$$

Hence, (2.11) follows by taking the conditional expectation of the above equality with respect to X given $Z = z$. □

The next result shows the second order Bartlett identity for the score function $\frac{\partial l_n(\beta)}{\partial \beta}$. Namely, we will show that $E[-\frac{\partial^2 l}{\partial\beta^2}] = E[\frac{\partial l}{\partial\beta}(\frac{\partial l}{\partial\beta})^T]$.

Proposition 2. The following equation holds.

$$\begin{aligned} & E\left[\int_0^\tau -\frac{\Psi(z)^{-1}}{m_0(t)(\Psi(z)^{-1}+m_0'(t))} Y(t)\Psi(z)^{-1}dt \big| Z=z\right] \\ = & -E\left[\delta\left(\frac{\Psi(z)^{-1}}{\Psi(z)^{-1}+m_0'(X)}\right)^2 |Z=z\right]. \end{aligned} \tag{2.13}$$

Proof. Note that

$$\delta\left(\frac{\Psi(z)^{-1}}{\Psi(z)^{-1}+m_0'(X)}\right)^2 = \int_0^\tau \frac{\Psi(z)^{-2}\mathrm{d}N(t)}{(\Psi(z)^{-1}+m_0'(t))^2}.$$

Using this and (2.12) gives us (2.13). □

Before we state the main result of this section, we need to impose some conditions on $m_0(\cdot)$:

C1: For any $t \geq 0$,

$$|m_0'(t)| \leq 1. \tag{2.14}$$

This is not restrictive, because from Nanda, Bhattacharjee and Alam(2005), C1 is a necessary and sufficient condition for the existence of proportional mean residual life model.

C2: There exists an $\eta > 0$ such that

$$E\left\{\left|\int_0^X \frac{1}{m_0(u)}\mathrm{d}u\right|^{2+\eta} |Z\right\} \tag{2.15}$$

is finite and continuous at the point $Z = z$.

C3:

$$E\left\{\left|\int_0^X \frac{1}{m_0(u)}\mathrm{d}u\right| |Z\right\} \tag{2.16}$$

is finite and continuous at the point $Z = z$.

C4: As $n \to \infty$ and $h \to 0$, $nh \to \infty$.

Theorem 1. Under conditions above, there exists a solution $\hat{\beta}$ to the local likelihood equation (2.8) such that

$$H(\hat{\beta} - \beta_*) \xrightarrow{P} 0,$$

where $H = \text{diag}\{1, h, \cdots, h^p\}$, h is the bandwidth of the kernel function and β_* is the true value of β.

Theorem 2. Under conditions above, the solution given in Theorem 1 is asymptotically normal:

$$\sqrt{nh}H(\hat{\beta} - \beta_*) \xrightarrow{D} N\{0, f^{-1}(z)S^{-1}(z)G(z;\beta_*)S^{-1}(z)\},$$

where

$$\mathbf{S}(z) = \int \mathbf{u}\mathbf{u}^T K(u)\mathrm{d}uE\left[\delta\left(\frac{\Psi(z)^{-1}}{\Psi(z)^{-1} + m_0'(X)}\right)^2 \big| Z = z\right]$$

and

$$\mathbf{G}(z;\beta_*) = \int_{-\infty}^{\infty} K^2(v)E\left\{\left[\frac{\exp(-\mathbf{Z}^T\beta_*)U}{\exp(-\mathbf{Z}^T\beta_*) + m_0'(X)}\right]^{\bigotimes 2}\delta\big|Z = hv + z\right\}\mathrm{d}v$$

2.5 Asymptotic Concavity of the Local Likelihood

From the last section, we know that there exists a solution to the local likelihood equation that is consistent. But we don't know which solution is consistent if there are more than one solution. However, in this section we will establish the asymptotic strict concavity of the local likelihood which ensures the uniqueness of the solution with probability tending to 1. Therefore the unique solution must be consistent with probability one. Let $A^{\bigotimes 2}$ denote AA^T for a vector or matrix A. The Hessian matrix of $l_n(\beta)$ is given by

$$\begin{aligned}
& l_n''(\beta) \\
= \; & n^{-1}\sum_{i=1}^{n} \exp(-\beta^T\mathbf{Z}_i)\mathbf{Z}_i^{\bigotimes 2}\left[\frac{\delta_i m_0'(X_i)}{(\exp\{-\beta^T\mathbf{Z}_i\} + m_0'(X_i))^2} - \int_0^{X_i}\frac{1}{m_0(u)}\mathrm{d}u\right]K_h(Z_i - z), \\
\xrightarrow{P} \; & E\left\{\int_0^{\tau} -\frac{\exp(-\beta^T\mathbf{Z})\mathbf{Z}^{\bigotimes 2}}{m_0(t)[\exp(-\beta^T\mathbf{Z}) + m_0'(t)]}Y(t)\exp(-\beta^T\mathbf{Z})\mathrm{d}tK_h(Z_i - z)\right\} \\
= \; & E\left\{K_h(Z - z)\mathbf{Z}^{\bigotimes 2}E\left\{\int_0^{\tau} -\frac{\exp(-\beta^T\mathbf{Z})}{m_0(t)[\exp(-\beta^T\mathbf{Z}) + m_0'(t)]}Y(t)\exp(-\beta^T\mathbf{Z})\mathrm{d}t|Z\right\}\right\}
\end{aligned}$$

By Proposition 2,

$$l_n''(\beta) \xrightarrow{P} -E\left\{K_h(Z-z)\mathbf{Z}^{\bigotimes 2}E\left\{\delta\left[\frac{\exp(-\beta^T\mathbf{Z})}{\exp(-\beta^T\mathbf{Z})+m_0'(X)}\right]^2|Z\right\}\right\} < 0$$

This gives us the asymptotic strict concavity of the local likelihood.

Chapter 3

Nonparametric Estimation of The Baseline MRL Function

Since we do not know the baseline MRL function $m_0(t)$, we now modify our procedure to get an estimator of the baseline and $\widehat{\beta}_v$ simultaneously. Accommodating the local likelihood, suppose we rewrite the model (1.2) as

$$m(t|Z) = \widetilde{m_0} \exp(\widetilde{\beta^T} \widetilde{\mathbf{Z}}), \tag{3.1}$$

where $\widetilde{m_0} = m_0(t)\exp(\beta_0)$, $\widetilde{\beta} = \{\beta_1, \cdots, \beta_p\}^T = \{\psi'(z), \cdots, \frac{\psi^{(p)}(z)}{p!}\}^T$ and $\widetilde{\mathbf{Z}} = \{Z - z, \cdots, (Z - z)^p\}^T$. It is shown in the Corollary 1.4.1 of Fleming and Harrington (1991) that

$$E\{\mathrm{d}N_i(t)|\Im_{t-}; \beta_*, m_*(\cdot)\} = Y_i(t)\mathrm{d}\Lambda_i(t; \beta_*, m_*(\cdot)) \tag{3.2}$$

where $\Im_t$ is the right-continuous filtration $\{\Im_t : t \geq 0\}$ defined by $\Im_t = \sigma\{N_i(u), Y_i(u+), Z_i : 0 \leq u \leq t, i = 1, \cdots, n\}$ and β_*, $m_*(\cdot)$ are the true values of the parameter β and $m_0(\cdot)$. Let, for $i = 1, \cdots, n$,

$$M_i(t; \beta, m_0) = N_i(t) - \int_0^t Y_i(s)\mathrm{d}\Lambda_i(s; \beta, m_0)$$

Then $\{M_i(t; \beta_*, m_*(t))\}$ are zero-mean $\Im_t$ martingales(Fleming and Harrington, 1991). Therefore it is natural to estimate $m_*(\cdot)$ and β_* from local estimating equations parallel to the partial score equations:

$$n^{-1}\sum_{i=1}^{n}\{\mathrm{d}N_i(t)-Y_i(t)\mathrm{d}\Lambda_i(t;\beta,m_0(t))\}K_h(Z_i-z)=0 \quad (0\le t\le\tau), \tag{3.3}$$

$$n^{-1}\sum_{i=1}^{n}\int_0^{\tau} Z_i\{\mathrm{d}N_i(t)-Y_i(t)\mathrm{d}\Lambda_i(t;\beta,m_0(t))\}K_h(Z_i-z)=0. \tag{3.4}$$

Under model (3.1), by (2.4), (3.3) and (3.4), it is equivalent to estimate $\widetilde{m_0}(t)$ and $\widetilde{\beta}$ from the following equations:

$$n^{-1}\sum_{i=1}^{n}\{\widetilde{m_0}(t)\mathrm{d}N_i(t)-Y_i(t)[\exp(-\widetilde{\beta}^T\widetilde{\mathbf{Z}_i})+\mathrm{d}\widetilde{m_0}(t)]\}K_h(Z_i-z)=0 \quad (0\le t\le\tau), \tag{3.5}$$

$$n^{-1}\sum_{i=1}^{n}\int_0^{\tau}\widetilde{\mathbf{Z}}_i\{\widetilde{m_0}(t)\mathrm{d}N_i(t)-Y_i(t)[\exp(-\widetilde{\beta}^T\widetilde{\mathbf{Z}_i})+\mathrm{d}\widetilde{m_0}(t)]\}K_h(Z_i-z)=0. \tag{3.6}$$

In fact (3.5) is a first-order linear ordinary differential equation in $\widetilde{m_0}(t)$, so we can solve for $\widetilde{m_0}(t)$ in terms of $\widetilde{\beta}$ as

$$\widehat{\widetilde{m_0}}(t;\widetilde{\beta})=\widehat{S_{NA}}^{-1}(t)\int_t^{\tau}\widehat{S_{NA}}(u)Q(u;\widetilde{\beta})\mathrm{d}u \tag{3.7}$$

where

$$\widehat{S_{NA}}(t)=\exp\left\{-\int_0^t\frac{\sum_{i=1}^n K_h(Z_i-z)\mathrm{d}N_i(u)}{\sum_{i=1}^n Y_i(u)K_h(Z_i-z)}\right\}$$

$$Q(t;\widetilde{\beta})=\frac{\sum_{i=1}^n Y_i(t)\exp(-\widetilde{\beta}^T\widetilde{\mathbf{Z}_i})K_h(Z_i-z)}{\sum_{i=1}^n Y_i(t)K_h(Z_i-z)}$$

Details of this derivation are shown in section 5.

Now we substitute $\widehat{\widetilde{m_0}}(t;\widetilde{\beta})$ into (3.6) to get an estimator for β_*. And it is easy to see that the resulting equations (3.6) are algebraically equivalent to

$$U(\beta)=n^{-1}\sum_{i=1}^{n}\int_0^{\tau}\{\widetilde{\mathbf{Z}_i}-\bar{Z}(t)\}\{\widehat{\widetilde{m_0}}(t;\widetilde{\beta})\mathrm{d}N_i(t)-Y_i(t)\exp(-\widetilde{\beta}^T\widetilde{\mathbf{Z}_i})\mathrm{d}t\}K_h(Z_i-z)=0,$$

where

$$\bar{Z}(t)=\frac{\sum_{i=1}^n\widetilde{\mathbf{Z}}_iY_i(t)K_h(Z_i-z)}{\sum_{i=1}^n Y_i(t)K_h(Z_i-z)}.$$

We are not able to solve for β_0 directly from the above equations because β_0 is cancelled out. However, once we have the estimator of β_1, $\widehat{\beta_1}=\widehat{\psi}'(z)$, then integrating $\widehat{\psi}'(z)$ from 0 to z would give us $\widehat{\psi}(z)$, while assuming $\psi(0) = 0$, which is a reasonable condition. As a consequence, we can recover the estimator of $m_0(t)$, $\widehat{m_0}(t)$ by

$$\widehat{m_0}(t) = \frac{\widehat{\widetilde{m_0}}(t; \widetilde{\beta})}{\exp\{\widehat{\psi}(z)\}}. \tag{3.8}$$

Chapter 4

Simulation study

Numerical studies were conducted to assess the finite-sample properties of the proposed estimation procedure. We consider the sample size n to be 50 with a one dimensional covariate Z for each of n subjects. Suppose $\psi(Z) = -Z^2$ and $\Psi(Z) = \exp(-Z^2) \approx \beta^T \mathbf{Z} = (\beta_0, \beta_1)^T (1, Z - z)$ where Z is a uniform random variable on $[-1, 1]$. We estimate $\psi(Z)$ at $z_1, z_2, \cdots, z_{50}$ which are fixed values uniformly selected from $[-1, 1]$. We select C uniformly from the interval $[1, 3]$ in order for the total censoring rate to be about $5\% - 15\%$. The censoring variable C is assumed to be independent of Z and T. The normal kernel function is adopted throughout the simulations.

4.1 Model 1

In this model, we suppose $m_0(t) = 1$. By (2.1), we have

$$S(t|Z) = \exp\left\{-\int_0^t \frac{1}{m(u|Z)} \mathrm{d}u\right\}.$$

Therefore,

$$t = -\frac{\log S(t|Z)}{\exp(Z^2)}. \tag{4.1}$$

Now we can use (4.1) to generate the lifetime T_i associated with the covariate Z_i.

4.2 Model 2

In this model, we assume $m_0(t) = t + 1$, and by (2.1)

$$\begin{aligned} S(t|Z) &= \frac{\exp(-Z^2)}{(t+1)\exp(-Z^2)} \exp\left\{-\int_0^t \frac{1}{(u+1)\exp(-Z^2)} \mathrm{d}u\right\} \\ &= (t+1)^{-1-\exp(Z^2)}. \end{aligned}$$

Therefore,

$$t = S(t|Z)^{\frac{1}{-1-\exp(Z^2)}} - 1. \tag{4.2}$$

Now we can use (4.2) to generate the lifetime T associated with the covariate Z_i.

Chapter 5

Derivations and Proofs

5.1 Derivation of (2.6)

By (2.2) and (2.3), we have the following form for the hazard rate function

$$\lambda(t|Z) = \frac{\exp(-\beta^T \mathbf{Z}) + m_0'(t)}{m_0(t)}$$

Then we have

$$\frac{\partial \delta \log \lambda(X|Z)}{\partial \beta} = \exp(-\beta^T \mathbf{Z})\mathbf{Z}\frac{\delta}{\exp(-\beta^T \mathbf{Z}) + m_0'(t)},$$

and

$$\frac{\partial \log S(X|Z)}{\partial \beta} = \exp(-\beta^T \mathbf{Z})\mathbf{Z}\int_0^X \frac{1}{m_0(u)}\mathrm{d}u.$$

Thus, by (2.5), we can easily get (2.6)

$$\frac{\partial \log L}{\partial \beta} = \sum_{i=1}^{n} \exp(-\beta^T \mathbf{Z}_i)\mathbf{Z}_i \left[-\frac{\delta_i}{\exp\{-\beta^T \mathbf{Z}_i\} + m_0'(X_i)} + \int_0^{X_i} \frac{1}{m_0(u)}\mathrm{d}u\right].$$

5.2 Proof of Theorem 1

We follow the ideas from Fan, Gijbels, and King (1996) in this proof. Let $\gamma = H(\beta - \beta_*)$, $\hat{\gamma} = H(\hat{\beta} - \beta_*)$, $\gamma_* = 0$ and $U_i = -H^{-1}\mathbf{Z}_i$. Put

$$\begin{aligned} l_n(\gamma) &= n^{-1}\sum_{i=1}^{n}[\delta_i \log \frac{\exp(U_i^T\gamma - \mathbf{Z}_i^T\beta_*) + m_0'(X_i)}{m_0(X_i)} + \log \frac{m_0(0)}{m_0(X_i)} \\ &\quad - \exp(U_i^T\gamma - \mathbf{Z}_i^T\beta_*) \int_0^{X_i} \frac{1}{m_0(u)} \mathrm{d}u] K_h(Z_i - z). \end{aligned}$$

Then, the problem is equivalent to showing that there exists a solution $\hat{\gamma}$ to the associated score function

$$\begin{aligned} 0 &= l_n'(\gamma) \\ &= n^{-1}\sum_{i=1}^{n}\left[\delta_i \frac{\exp(U_i^T\gamma - \mathbf{Z}_i^T\beta_*)}{\exp(U_i^T\gamma - \mathbf{Z}_i^T\beta_*) + m_0'(X_i)} - \exp(U_i^T\gamma - \mathbf{Z}_i^T\beta_*)\int_0^{X_i}\frac{1}{m_0(u)}\mathrm{d}u\right] U_i K_h(Z_i - z) \end{aligned} \tag{5.1}$$

such that

$$\hat{\gamma} \xrightarrow{P} 0.$$

Let S_ϵ denote the sphere centered at γ_* with radius ϵ. We will show that for any $\epsilon > 0$, the probability that

$$\sup_{\gamma \in S_\epsilon} l_n(\gamma) < l_n(\gamma_*) \tag{5.2}$$

tends to one.

First, we know that

$$\begin{aligned} l_n'(\gamma_*) &= n^{-1}\sum_{i=1}^{n}\left[\delta_i \frac{\exp(-\mathbf{Z}_i^T\beta_*)}{\exp(-\mathbf{Z}_i^T\beta_*) + m_0'(X_i)} - \exp(-\mathbf{Z}_i^T\beta_*)\int_0^{X_i}\frac{1}{m_0(u)}\mathrm{d}u\right] U_i K_h(Z_i - z) \\ &\xrightarrow{P} f(z)\int \mathbf{u}K(u)\mathrm{d}u E\left\{\Psi(z)^{-1}\mathbf{Z}\left[-\frac{\delta}{\Psi(z)^{-1} + m_0'(X)} + \int_0^X \frac{1}{m_0(u)}\mathrm{d}u\right] \big| Z = z\right\}. \end{aligned}$$

By (2.11), we get that

$$l_n'(\gamma_*) \xrightarrow{P} 0.$$

And, therefore, with probability tending to one,

$$\left|l_n'(\gamma_*)^T(\gamma-\gamma_*)\right| \le \epsilon^3. \tag{5.3}$$

Secondly,

$$\begin{aligned} l_n''(\gamma_*) &= n^{-1}\sum_{i=1}^n \left[\delta_i \frac{\exp(-\mathbf{Z}_i^T\beta_*)m_0'(X_i)}{(\exp(-\mathbf{Z}_i^T\beta_*)+m_0'(X_i))^2} - \exp(-\mathbf{Z}_i^T\beta_*)\int_0^{X_i}\frac{1}{m_0(u)}\mathrm{d}u\right] U_i^{\bigotimes 2} K_h(Z_i-z) \\ &\xrightarrow{P} f(z)\int \mathbf{u}\mathbf{u}^T K(u)\mathrm{d}u \times \\ &\quad E\left[\int_0^\tau -\frac{\Psi(z)^{-1}}{m_0(t)(\Psi(z)^{-1}+m_0'(t))} Y(t)\Psi(z)^{-1}\mathrm{d}t | Z=z\right]. \end{aligned}$$

By (2.13),

$$\begin{aligned} l_n''(\gamma_*) &= -f(z)\int \mathbf{u}\mathbf{u}^T K(u)\mathrm{d}u E\left[\delta\left(\frac{\Psi(z)^{-1}}{\Psi(z)^{-1}+m_0'(X)}\right)^2 |Z=z\right] + o_P(1). \\ &= -f(z)\mathbf{S}(z) + o_P(1). \end{aligned}$$

Hence, with probability tending to one,

$$(\gamma-\gamma_*)^T l_n''(\gamma_*)(\gamma-\gamma_*) < -af(z)\epsilon^2, \quad \text{for any } \gamma \in S_\epsilon, \tag{5.4}$$

where a is the smallest eigenvalue of $S(z)$. Let γ_j and γ_{*j} be the j^{th} elements of γ and γ_*, respectively. By Talyor expansion around the point γ_*,

$$l_n(\gamma) - l_n(\gamma_*) = l_n'(\gamma_*)^T(\gamma-\gamma_*) + \frac{1}{2}(\gamma-\gamma_*)^T l_n''(\gamma_*)(\gamma-\gamma_*) + R_n(\tilde{\gamma}), \tag{5.5}$$

where $\tilde{\gamma}$ is a point between γ_* and γ, and

$$R_n(\gamma) = \frac{1}{6}\sum_{j,k,l}(\gamma_j-\gamma_{*j})(\gamma_k-\gamma_{*k})(\gamma_l-\gamma_{*l})\frac{\partial^3}{\partial\gamma_j\partial\gamma_k\partial\gamma_l} l_n(\gamma).$$

By condition (2.16),

$$|R_n(\gamma)| \le \epsilon^3\{C + o_P(1)\} \tag{5.6}$$

for some constant $C > 0$. Substituting (5.3), (5.4), (5.6) into (5.5), we conclude that with probability tending to one that when ϵ is small enough,

$$l_n(\gamma) - l_n(\gamma_*) \leq 0 \quad \text{for all } \gamma \in S_\epsilon.$$

Hence, $l_n(\gamma)$ has a local maximum in the interior of S_ϵ. Since at a local maximum the score equation (5.1) must be satisfied, it follows that for any $\epsilon > 0$, with probability tending to one, the score equation has a solution $\hat{\gamma}(\epsilon)$ within S_ϵ. Let $\hat{\gamma}$ be the closest root to γ_*. Then

$$P\{||\hat{\gamma}||^2 \leq \epsilon\} \longrightarrow 1.$$

This in turn completes the proof of Theorem 1. □

5.3 Proof of Theorem 2

We still use the notation introduced in the proof of Theorem 1. By Taylor's expansion, we have

$$0 = l_n'(\hat{\gamma}) = l_n'(\gamma_*) + l_n''(\gamma_*)(\hat{\gamma} - \gamma_*) + O_P(||\hat{\gamma} - \gamma_*||^2).$$

Hence,

$$\begin{aligned} \hat{\gamma} - \gamma_* &= [l_n''(\gamma_*) + o_P(1)]^{-1} l_n'(\gamma_*) \\ &= [-f(z)\mathbf{S}(z) + o_P(1)]^{-1} l_n'(\gamma_*) \end{aligned} \tag{5.7}$$

Thus we only need to establish the asymptotic normality of $l_n'(\gamma_*)$. We first compute the mean and the variance of $l_n'(\gamma_*)$.

$$\begin{aligned} El_n'(\gamma_*) &= E\left\{[\delta\frac{\exp(-\mathbf{Z}^T\beta_*)}{\exp(-\mathbf{Z}^T\beta_*)+m_0'(X)} - \exp(-\mathbf{Z}^T\beta_*)\int_0^X \frac{1}{m_0(u)}\mathrm{d}u]UK_h(Z-z)\right\} \\ &= E\left\{E\left\{[\delta\frac{\exp(-\mathbf{Z}^T\beta_*)}{\exp(-\mathbf{Z}^T\beta_*)+m_0'(X)} - \exp(-\mathbf{Z}^T\beta_*)\int_0^X \frac{1}{m_0(u)}\mathrm{d}u]UK_h(Z-z)\big|Z\right\}\right\}. \end{aligned}$$

But we know that, for any z'

$$E\left\{\left[\delta\frac{\exp(-\mathbf{Z}^T\beta_*)}{\exp(-\mathbf{Z}^T\beta_*)+m_0'(X)} - \exp(-\mathbf{Z}^T\beta_*)\int_0^X \frac{1}{m_0(u)}\mathrm{d}u\right]UK_h(Z-z)\big|Z=z'\right\} = 0.$$

Hence,

$$El_n'(\gamma_*) = 0$$

For variance,

$$\begin{aligned} Var(l_n'(\gamma_*)) &= n^{-1}E\left\{K_h^2(Z-z)\left\{\left[\delta\frac{\exp(-\mathbf{Z}^T\beta_*)}{\exp(-\mathbf{Z}^T\beta_*)+m_0'(X)} - \exp(-\mathbf{Z}^T\beta_*)\int_0^X \frac{1}{m_0(u)}\mathrm{d}u\right]U\right\}^{\otimes 2}\right\} \\ &= n^{-1}E\left\{K_h^2(Z-z)\left[\int_0^\tau \frac{\exp(-\mathbf{Z}^T\beta_*)U}{\exp(-\mathbf{Z}^T\beta_*)+m_0'(X)}\mathrm{d}M(t)\right]^{\otimes 2}\right\} \end{aligned}$$

By conditioning on Z and using

$$\mathrm{d} < M, M > (t) = Y(t)\frac{\exp(-\beta^T\mathbf{Z})+m_0'(t)}{m_0(t)}\mathrm{d}t,$$

(Fleming and Harrington,1991)we have

$$
\begin{aligned}
Var(l_n'(\gamma_*)) &= n^{-1}E\left\{K_h^2(Z-z)\left[\int_0^\tau \frac{\exp(-\mathbf{Z}^T\beta_*)U}{\exp(-\mathbf{Z}^T\beta_*)+m_0'(X)}\right]^{\bigotimes 2} Y(t)\frac{\exp(-\beta_*^T\mathbf{Z})+m_0'(t)}{m_0(t)}\mathrm{d}t\right\} \\
&= n^{-1}E\left\{K_h^2(Z-z)\left[\int_0^\tau \frac{\exp(-\mathbf{Z}^T\beta_*)U}{\exp(-\mathbf{Z}^T\beta_*)+m_0'(X)}\right]^{\bigotimes 2} \mathrm{d}N(t)\right\} \\
&= n^{-1}E\left\{K_h^2(Z-z)\left[\frac{\exp(-\mathbf{Z}^T\beta_*)U}{\exp(-\mathbf{Z}^T\beta_*)+m_0'(X)}\right]^{\bigotimes 2} \delta\right\} \\
&= n^{-1}\int_{-\infty}^{\infty} f_Z(t)E\left\{K_h^2(Z-z)\left[\frac{\exp(-\mathbf{Z}^T\beta_*)U}{\exp(-\mathbf{Z}^T\beta_*)+m_0'(X)}\right]^{\bigotimes 2} \delta\big|Z=t\right\}\mathrm{d}t \\
&= n^{-1}h^{-1}f(z)\int_{-\infty}^{\infty} K^2(v)E\left\{\left[\frac{\exp(-\mathbf{Z}^T\beta_*)U}{\exp(-\mathbf{Z}^T\beta_*)+m_0'(X)}\right]^{\bigotimes 2} \delta\big|Z=hv+z\right\}\mathrm{d}v + o(n^{-1}h^{-1}) \\
&= n^{-1}h^{-1}f(z)G(z;\beta_*)+o(n^{-1}h^{-1})
\end{aligned}
$$

To prove the asymptotic normality of $l_n'(\gamma_*)$, we use the Cramer-Wold device. For all $\theta\in\mathbb{R}^d$, we need to show

$$\sqrt{nh}\{\theta^T l_n'(\gamma_*) - \theta^T E l_n'(\gamma_*)\} \xrightarrow{D} N(0, f(z)\theta^T G(z;\beta_*)\theta).$$

To establish the asymptotic normality of $\theta^T l_n'(\gamma_*)$, we only need to verify the Lyapunov condition:

$$E\sum_{i=1}^n\left|\sqrt{nh}n^{-1}K_h(Z-z)\left[\delta\frac{\exp(-\mathbf{Z}^T\beta_*)}{\exp(-\mathbf{Z}^T\beta_*)+m_0'(X)} - \exp(-\mathbf{Z}^T\beta_*)\int_0^X \frac{1}{m_0(u)}\mathrm{d}u\right]\theta^T U\right|^{2+\eta} \longrightarrow 0,$$

for some $\eta>0$. By condition (2.14) and (2.15), the left-hand side of the above expression is bounded by $O\{(nh)^{-\eta/2}\}$ and therefore converges to zero. Hence,

$$\sqrt{nh}(\hat\gamma-\gamma_0) = \sqrt{nh}[-f(z)S(z)+o_P(1)]^{-1}l_n'(\gamma_*) \xrightarrow{D} N\{0, f^{-1}(z)S^{-1}(z)G(z;\beta_*)S^{-1}(z)\}$$

This completes the proof of Theorem 2.

5.4 Derivation of (3.7)

We mentioned that (3.5) is in fact a first-order linear ordinary differential equation in $\widetilde{m_0}(t)$, which is equivalent to

$$\begin{aligned} & \left\{\sum_{i=1}^{n} K_h(Z_i - z)\mathrm{d}N_i(t)\right\}\widetilde{m_0}(t) - \left\{\sum_{i=1}^{n} Y_i(t)K_h(Z_i - z)\right\}\mathrm{d}\widetilde{m_0}(t) \\ = \; & \left\{\sum_{i=1}^{n} Y_i(t)\exp(-\widetilde{\beta}^T\widetilde{\mathbf{Z}_i})K_h(Z_i - z)\right\}\mathrm{d}t \end{aligned} \tag{5.8}$$

Let

$$Q(t;\widetilde{\beta}) = \frac{\sum_{i=1}^{n} Y_i(t)\exp(-\widetilde{\beta}^T\widetilde{\mathbf{Z}_i})K_h(Z_i - z)}{\sum_{i=1}^{n} Y_i(t)K_h(Z_i - z)}$$

and

$$P(t;\widetilde{\beta}) = \frac{\sum_{i=1}^{n} K_h(Z_i - z)\mathrm{d}N_i(t)}{\sum_{i=1}^{n} Y_i(t)K_h(Z_i - z)},$$

and divide both sides of (5.8) by $\widetilde{m_0}(t)\sum_{i=1}^{n} Y_i(t)K_h(Z_i - z)$, we get

$$-P(t;\widetilde{\beta}) + \frac{\mathrm{d}\widetilde{m_0}(t)}{\widetilde{m_0}(t)} = \frac{Q(t;\widetilde{\beta})\mathrm{d}t}{-\widetilde{m_0}(t)}. \tag{5.9}$$

Let

$$-P(t;\widetilde{\beta}) = \frac{1}{\mu(t)}\mathrm{d}\mu(t)$$

which implies

$$\mu(t) = \widehat{S_{NA}}(t)$$

and (5.9) is equivalent to

$$Q(t;\widetilde{\beta})\mu(t) = -\widetilde{m_0}(t)\frac{\mathrm{d}\mu(t)}{\mathrm{d}t} - \mu(t)\frac{\mathrm{d}\widetilde{m_0}(t)}{\mathrm{d}t} = -\frac{\mathrm{d}\widetilde{m_0}(t)\mu(t)}{\mathrm{d}t}.$$

Hence,

$$\begin{aligned}
\widehat{\widetilde{m_0}}(t;\widetilde{\beta}) &= \widehat{S_{NA}}^{-1}(t)\int_t^{\tau}\widehat{S_{NA}}(u)Q(u;\widetilde{\beta})\mathrm{d}u + \widetilde{m_0}(\tau)\mu(\tau) \\
&= \widehat{S_{NA}}^{-1}(t)\int_t^{\tau}\widehat{S_{NA}}(u)Q(u;\widetilde{\beta})\mathrm{d}u
\end{aligned}$$

Bibliography

[1] Aitkin, M. and Clayton, D.G. (1980). "The fitting of exponential, Weibull and extreme value distributions to complex censored survival data using GLIM," *J. Roy. Statist. Soc. Ser. C*, **29**, 156-163.

[2] Asok K. Nanda, Subarna Bhattacharjee, S.S. Alam (2006). "Properties of Proportional Mean Residual Life Model," *Statistics and Probability Letters*, **76**, 880-890.

[3] Csörgő, M. Zitikis, R. (1996). "Mean residual life processes," *Ann. Statist*, **4**, 1717-39.

[4] Cox, D. T. (1972). "Regression models and life-tables(with discussion)," *J. R. Statist. Soc. B*, **34**, 187-220.

[5] Fleming, T. R. and Harrington, D. P. (1991), *Counting Process and Survival Analysis*, New York: John, Wiley.

[6] Gangaji Maguluri and Cun-Hui Zhang (1994). "Estimation in the Mean Residual Life Regression Model," *J.R. Statist. Soc. B*, **56**, No. 3, pp 477-489.

[7] Gentleman, R. and Crowley, J. (1991). "Local fully likelihood estimation for the proportional hazards model," *Biometrics*, **47**, 1283-1296.

[8] Guess, F. and Proschan, F. (1988). "Mean residual life: theory and applications," In *Krishnaiah*, Ed. P. R., Rao, C. R., Handbook of Statistics, Vol. 7. Elsevier, North-Holland, pp. 215-224.

[9] Jerald F. Lawless (1998), *Statistical Models and Methods for Lifetime Data*, New York: Wiley.

[10] Jianqing Fan, Irene Gijbels, Martin King (Aug., 1997). "Local Likelihood and Local Partial Likelihood in Hazard Regression," *The Annals of Statistics*, Vol. 24, No. 4, pp. 1661-1690.

[11] Maguluri, G. and Zhang, C.-H (1994). "Estimation in the mean residual life regression model," *J. R. Statist Soc. B*, **56**, 477-89.

[12] Oakes, D. and Dasu, T. (1990). "A note on residual life," *Biometrika*, **77**, 409-10.

[13] Oakes, D. and Dasu, T. (2003). "Inference for the proportional mean residual life model," *Crossing Boundaries: Statistical Essays in Honor of Jack Hall, Institute of Mathematical Statistics Lecture Notes Monograph Series*, Vol. **43**, Ed. J. E. Kolassa and D. Oakes, pp. 105-16.

[14] Robins, J. M. and Rotnitzky, A. (1992). "Recovery of information and adjustment for dependent censoring using surrogate markers," *AIDS Epidemiology, Methodological Issues*, Ed. N. P. Jewell, K. Dietz and V. Farewell, pp. 297-331. Boston: Birkhäuer.

[15] Ying Q. Chen and Nicholas P. Jewell (2002). "Inference for Proportional Mean Residual Life Model in the Presence of Censoring," *U.C. Berkeley Division of Biostatistics Working Paper Series*, Paper 108.

[16] Y. Q. Chen and S. Cheng (2005). "Semiparametric Regression Analysis of Mean Residual Life with Censored Survival Data," *Biometrika Trust*, **92**, 1, pp. 19-29.