DOI: http://doi.org/10.32792/utq.jceps.10.01.019

Efficient Estimation Of Left-Truncated And Right Censored Data With Additive Hazard Model

Riyadh R. Al-Mosawi[1], Wissam Kamil Ghafil[2]

1.Department of Mathematics, Thi-Qar University, Iraq.
2.Department of Mathematics, University of the-Qar, Iraq.

Received 1/10/2016 Accepted 26/10/2016 Published 20/1/2020

Abstract:
In this paper, an additive hazard model is considered as a semiparametric model in which the unknown parameter consists of finite dimensional and infinite dimensional parts with left truncated and right censored data. The full likelihood function for the model is obtained for the parametric part and also for the nonparametric part using linear sieve procedure and then compute the maximum likelihood estimators for the two parts. The consistency of the maximum likelihood estimators is also proved for the two type of parameters. The score operators for the parametric and nonparametric parts are obtained and their adjoint score operators are computed. Finally, a simulation study using Monti-Carlo method and R language is implemented to compute the maximum likelihood estimators and compare the results of the proposed method with the true values. As a real life application, Stanford Heart transplant data is considered and the maximum likelihood estimators are computed.

1. Introduction:

he data collected from survival experiments is usually referred to as survival data, time-toevent data, or failure time data and represented by a nonnegative random variable T. The survival function of T is defined as $S(t) = P(T > t) = 1 - F(t)$, where $F(t)$ is the Cumulative Distribution Function (CDF). The Survival Function $S(t)$ represents the probability that the event occurred not before the time t. Another functions which are commonly used in modeling T are the hazard function and the cumulative hazard function of T. When T is continuous, the hazard function of T is defined as

$$\lambda(t) \lim_{h\to 0}\frac{1}{h} P(t \le T < t + h|T \ge t) = f(t)\, S(t) \quad (1.1)$$

where $f(t)$ is the probability density function (p.d.f.) of T. It is easy to prove that

$$S(t) = e^{-\Lambda(t)}, \quad (1.2)$$

where $\Lambda(t) = \int_0^t \lambda(t)dt$ (1-3)

is the cumulative hazard function of T. Assume the time of event random variables of subjects follow a specific type of semiparametric model which known as additive hazard model (see for example Aalen (1989) and Cox and Oakes (1984, page 147)). In this model, we assume that the covariates act on an unknown baseline hazard rate in an additive manner. Assume that the time of event $T0$ and covariates $Z0$ are modelled as additive hazard model given by $\lambda(t|z) = \lambda(t) + \beta^T z$
(1.4)
where $\lambda(t|z)$ is the conditional hazard function of T^0 given $Z^0 = z$, β is an unknown $k$dimension regression parameter and λ is an unknown baseline hazard function i.e. the conditional hazard function when $z = 0$. The cumulative baseline hazard function of this model is given by $\Lambda(t) = \int_0^t \lambda(s)ds$ the cumulative hazard function is given by

$$\Lambda(t|z) = \int_0^t \lambda(s)ds + \beta^T zt = \Lambda(t) + \beta^T zt \quad (1.5)$$

and the survival function is given by

$$F(t|z) = F(t) \exp\{-\beta^T zt\}, \quad (1.6)$$

where $F(t) = \exp\{-\Lambda(t)\}$ is the baseline survival function i.e. the conditional survival function when $z = 0$. For the above model, the parameters are $\theta = (\beta, \Lambda)$, where $\beta \in Rk$ is a finite dimensional parameter (parametric part) whereas Λ is in an infinite dimensional2 parameter (nonparametric part), hence our model is semiparametric. Usually, β is called the parameter of interest and Λ is the nuisance parameter. The true value of parameter $\theta = (\beta, \Lambda)$ is $\theta0 = (\beta0, \Lambda0)$. Assume further that time of event random variable $T0$ is subject to left-truncation time $X0$ so that the subject is observed whenever $T^0 \geq X^0$. Moreover, we assume that $T0$ is also subject to right-censorship mechanism at monitoring time C^0. So that our observation will be (Y^0, Δ^0, X^0), where $Y0$ is the smallest of T^0 and C^0 i.e $Y^0 = T^0 \wedge C^0$ and $\Delta0 = I(T^0 \leq C^0)$ is the right-censored indicator and $I(A)$ is the indicator function of the set A. Let $T0$ has distribution function F with density f, C^0 has distribution function G with density g, Z^0 has density q and $X0$ has distribution function H with density h and let F, G and H represent the survival functions of T^0, C^0 and X^0, respectively. We consider the triplet (Y, Δ, X) as the observed data which has the same joint distribution as (Y^0, Δ^0, X^0) when $T^0 \geq X^0$ where Y is the observed survival time, Δ is the right censoring indicator and X is the truncation time. Consider H as the class of all bounded, continuous and non-decreasing functions over $[\tau0, \tau1]$. Then the parameter space of the parameter $\theta = (\beta, \Lambda)$ is defines as $\Theta = B \times H$ i.e. Λ is considered as a bounded, continuous and non-decreasing function over $[\tau0, \tau1]$. Throughout the paper, in addition to the assumptions $A1$ and $A2$, we assume the following assumptions:

A1 : The parameter space of B of β is a bounded open subset of Rk.

A2 : Given the covariate Z^0, the time of event T^0 the observation time C^0 and the truncated time X^0 are independent random variables. Assume also that $X0$ is independent of T^0, C^0 and Z^0.

A3 : The marginal densities of C^0, $X0$ and Z^0 i.e. g, h and q do not involve the parameter
$\theta = (\beta, \Lambda)$.

A4 : (a) The covariate Z is bounded with probability one i.e. there exists a positive constant M such that
$= \sqrt{Z_1^2 + Z_2^2 + \cdots + Z_k^2} \leq M$ with probability one.$\|Z\|_{R^k}$
β_1^T $Z) > 0(b)$ (Identifiability assumption) For any $\beta1 = \beta2 \in B$ we have that $P(\beta_1^T$ Z)

$A5$: There exists a positive number η such that $P(C0 - X0 \geq \eta) = 1$.
$A6$: There exist $0 < \tau 0 < \tau 1$ and $0 < m0 < m1 < \infty$ such that
$P(\tau 0 \leq X0 < C0 \leq \tau 1) = 1$ and $m0 < \Lambda(\tau 0) < \Lambda(\tau 1) < m1$.
$A7$: For $t \in [\tau 0, \tau 1]$, assume that $\lambda(t|z) > 0$
The problem of estimating the Regresion parametric and a semiparametric additive hazard model was Lin and Ying (1994) developed simple procedures with high efficiencies to estimate the regression parameters for the additive risk model with an unspecified baseline hazard function. Huang (1999) studied the partly linear additive Cox model as an extension of the (linear) Cox model. He investigated the asymptotic properties of the maximum partial likelihood estimator of this model with right censored data using polynomial splines. Zeng and Shen (2006) considered a semiparametric additive hazards model with interval censored event and their interest focused on the estimation of the effect of risk factors. Wang et al. (2010) proposed an estimating equation based approach for regression analysis of interval-censored failure time data with the additive hazards model. Song et al. (2011) studied a semiparametric additive hazards regression model for right censored data that allows some censoring indicators to be missing at random. Huang et al. (2013) considered additive hazards model with left-truncated and right-censored data. They used a pairwise pseudo likelihood to eliminate nuisance parameters from the marginal likelihood. Lu and Song (2014) considered the partly linear additive hazards regression model with current status data. They used polynomial splines approach to estimate both cumulative baseline hazard function with monotonicity constraint and nonparametric regression functions with no such constraint. Shen (2014b) analyzed left truncated and right censored data using additive hazard model models. He used the integrated square error to select an optimal bandwidth of the weighted least-squared estimator. Wang et al. (2015) studied the regression analysis of the the additive hazards model with left-truncation and current status censored data. They derived the maximum likelihood estimators of the unknown parameters. For the infinite dimensional parameter, they used the sieve estimation approach that approximates the baseline cumulative hazard function by linear functions. Feng et al. (2015) discussed regression analysis of additive hazards model with current status failure time data. They focused on the case when some covariates could be missing but there may4 exist some auxiliary information about the missing covariates. Chen (2016) investigated the additive hazards model with left-truncated and right-censored data. He developed a pseudo likelihood estimation approach for the parameters of interests.

1.1 Maximum Likelihood Estimatio:

In this section, we compute the maximum likelihood estimators for the parameter of interest β and the nuisance parameter Λ using sieve procedure. The likelihood of β and Λ based on a sample of n independent observations $Vi = (Yi, \Delta i, Xi, Zi)$, $i = 1, 2, \cdots, n$, can be written

$$l\,n(\beta, \Lambda) = \Pi_{i=1}^{n} \frac{[G(yi|zi)]^{\delta i}\,[g(yi|zi)]^{1-\delta i}\,[\lambda(yi) + \beta^T zi]\delta i \exp(-\Lambda(yi) - \beta^T\ ziyi)h(xi)q(zi)}{R(zi;\beta,\Lambda)} \quad (1.7)$$

where

$$R(zi; \beta, \Lambda) = \int_{\tau 0}^{\tau 1} \exp\{-\Lambda(u) - \beta^T\ zu\}\, h(u)du. \quad (1.8)$$

Using the Assumption $A3$, the above function can be written (up to terms do not involve (β, Λ)) as

$$l\,n(\beta, \Lambda) = \Pi_{i=1}^{n} \frac{[\lambda(yi) + \beta^T zi]^{\delta i} \exp(-\Lambda(yi) - \beta^T\, ziyi)}{R(zi;\beta,\Lambda)}$$

and the corresponding log likelihood has the form

$$\sum_{i=1}^{n}[\delta i \log(\lambda(yi) + \beta^{T} zi) - \Lambda(yi) - \beta^{T} ziyi - \log(R(zi;\ \beta,\Lambda))], \quad (1.9)$$

For a fixed Λ, the maximum likelihood estimators $\beta\hat{}$ for β is defined as

$\beta\hat{}n$ = $\text{argmax}_{\beta\in\Theta} l\, n(\beta,\Lambda)$

i.e. $\beta\hat{}n$ is the value that maximizes the log likelihood $ln(\beta, \Lambda)$ for all $\beta \in \Theta$. The maximum likelihood estimator for β is simply computed by differentiating ln with respect to β and equating to zero as given below. For βj, $j = 1, \cdots, k$, we have

$$= \ zj\sum_{i=1}^{n} \left\{\frac{\delta i}{\lambda(yi) + \beta^{T} zi} - yi + \frac{R_1(zi;\ \beta,\Lambda)}{R(zi;\ \beta,\Lambda)}\right\} \qquad \text{where}\frac{\partial ln(\beta,\Lambda)}{\partial \beta j}$$

$$R_1(zi;\ \beta,\Lambda) = \int_{\tau 0}^{\tau 1} \exp\{-\Lambda(u) - \beta^{T}\ zu\}u\, h(u)du.$$

Hence to find the maximum likelihood estimator for β we need to solve the following equation

$$Uj(\beta)=zj\sum_{i=1}^{n} \left\{\frac{\delta i}{\lambda(yi)+\beta^{T} zi} - yi + \frac{R_1(zi;\ \beta,\Lambda)}{R(zi;\ \beta,\Lambda)}\right\} = \quad for\ j\ = 0 \quad 1, 2,\cdots$$

$, k$ (1.10)

approach For fixed β, to find the maximum likelihood estimator of Λ, we apply the sieve estimation example Huang and using piecewise linear function with knots at pre-specified locations (see, for Rossini (1997) and Wang et al. (2015)). The basic idea of the sieve method is that a sequence of increasing subspaces (sieves) is used to approximate a large parameter space such that, asymptotically, the original parameter space. The main idea of sieve the closure of the limiting subspace contains Let $\tau 0 = t0 < t1 < t2 < \cdots < tq = \tau 1$ be a partition of the interval method is given in the following steps. $[\tau 0, \tau 1]$, where $\tau 0$ ad $\tau 1$denotes the smallest of truncation times and the largest observation times, respectively and the number q is called the sieve number. Following Wang et al. (2015), the dimension of the sieve number, q, is usually an increasing integer along with n, sample size, at rate $O(nr)$ with $0 < r < 1/2$. Define, for $i = 1, 2, \cdots, q$, $\Lambda(ti) = hi$ and let

$$\Lambda(t) =\sum_{i=1}^{q} I_i(t)\left\{h_{i-1)} + \frac{(t-t_{i-1})}{t_i-t_{i-1}} \quad (\mathrm{h} - h_{i-1})\right\} \qquad (1-11)$$

and

$$\lambda(t) = \frac{\partial\ \Lambda(t)}{\partial\ t}.=\sum_{i=1}^{q} I_i(t)\ \left\{\frac{\mathrm{h}_i-h_{i-1}}{t_i-t_{i-1}}\right\} \qquad (1.12)$$

where

$(t)= I(t_{i-1} \le t < t_i)I_i$

and $I(A)$ is the indica for function of the set A. Here and in the sequel we adopt the convention that $\sum_a^b f = 0$ whenever $b < a$. Following Wang et al. (2015), to remove the monotonicity condition of hi, we assume $h_i = \sum_{j=1}^{i} e^{\gamma j}$ and $\mathrm{h}_0 = 0$. Then Equations $(1-11)$ and (1.12) become

$$\Lambda(t) =\sum_{i=1}^{q} I_i(t)\ \left\{\sum_{j=1}^{i-1} e^{\gamma j} + \frac{(t-t_{i-1})}{t_i-t_{i-1}}\ e^{\gamma j}\right\} \qquad (1-13)$$

and

$$\lambda(t)= \sum_{j=1}^{q}\ \ I_i(t)\ \ \left\{\frac{e^{\gamma j}}{t_i-t_{i-1}}\right\} \qquad (1-14)$$

Now, the maximum likelihood estimators for $\gamma = (\gamma_1, \cdots, \gamma_q)$ can be obtained by differentiating the log likelihood function with respect to $\gamma = (\gamma_1 \cdots, \gamma_q)$ and solving

the resulting equations for $\gamma = (\gamma_1, \cdots, \gamma_q)$ as given in the following steps.
Differentiating Equation(**1-7**) with respect to γ_r , $r = 1, \cdots, q$, gives us

$$= \sum_{i=1}^{n} \left\{ \frac{\delta i}{\lambda(yi) + \beta^{T} zi} \frac{\partial \lambda(yi)}{\partial \gamma_r} - \frac{\partial \Lambda(yi)}{\partial \gamma_r} + \frac{R_2(zi; \beta, \Lambda)}{R(zi; \beta, \Lambda)} \right\} \qquad (1-15) \quad \frac{\partial ln(\beta, \gamma)}{\partial \gamma_r}$$

where

$$R_2(zi;\ \beta, \Lambda) = \int_{\tau 0}^{\tau 1} \frac{\partial \Lambda(u)}{\partial \gamma_r} \exp\{-\Lambda(u) - \beta^T zu\}\, h(u) du. \qquad (1-16)$$

and

$$\frac{\partial \lambda(t)}{\partial \gamma_r} = \frac{\partial}{\partial \gamma_r} = \sum_{i=1}^{q} I_i(t) \left\{ \frac{e^{\gamma_j)}}{t_i - t_{i-1}} \right\} = \left\{ \frac{e^{\gamma_j)}}{t_i - t_{i-1}} \right\} I_r(t) \quad (1-17)$$

$$\frac{\partial \Lambda(t)}{\partial \gamma_r} = \frac{\partial}{\partial \gamma_r} = \sum_{i=1}^{q} I_i(t) \left\{ \sum_{j=1}^{i-1} e^{\gamma_j} + \frac{(t - t_{i-1})}{t_i - t_{i-1}}\ e^{\gamma_j} \right\} \qquad (1-18)$$

Hence to find the maximum likelihood estimator for γ, we need to solve the following equations

$$Ur(\gamma) = \sum_{i=1}^{n} \left\{ \frac{\delta i}{\lambda(yi) + \beta^{T} zi} \frac{\partial \lambda(yi)}{\partial \gamma_r} - \frac{\partial \Lambda(yi)}{\partial \gamma_r} + \frac{R_2(zi; \beta, \Lambda)}{R_1(zi; \beta, \Lambda)} \right\} = 0$$

for i=1,2,…….,q (1-19)

To find the maximum likelihood estimators, we may use the profile likelihood approach as given in the following iteration procedure.

Step(1) Select initial values $\beta^{(0)} = (\beta_1^{(0)}, \cdots, \beta_k^{(0)})^T$ and $\gamma^{(0)} = (\gamma_1^{(0)}, \cdots, \gamma_k^{(0)})^T$.

Step(2) For the ith iteration, set $\beta = \beta^{(i-1)}$ and solve Equation (**1-19**) to find $\gamma^{(\mathrm{i})}$.

Step(3) Set $\gamma = \gamma^{(\mathrm{i})}$and solve Equation (**1-18**) to find $\beta^{(\mathrm{i})}$.

Step(4) For a given $\epsilon > 0$, if

$$\left\|\beta^{(\mathrm{i})} - \beta^{(\mathrm{i-1})}\right\|^2 + \left\|\gamma^{(\mathrm{i})} - \gamma^{(\mathrm{i-1})}\right\|^2 \le \epsilon$$

then set $\hat{\beta}_n = \beta^{(\mathrm{i})}$and $\hat{\gamma}_n = \gamma^{(\mathrm{i})}$ and go to **Step (5)**, otherwise, go to **Step (2)**. **Step (5)** Compute Λ and λ using Equations (**1-13**) and (**1-14**), respectively. To perform **Step (2)** and **Step (3)**, we may use, for example, Newton-Raphson method. As a special case, for additive hazard model with right censored data, the likelihood function can be simply obtained by substituting $R(z; \beta, \Lambda) = 1$ in Equation (**1-19**) and consequently, the maximum likelihood estimators can be obtained, by solving

$$Uj(\beta) = zj \sum_{i=1}^{n} \left\{ \frac{\delta i}{\lambda(yi) + \beta^{T} zi} - yi \right\} = 0 \qquad j = 1, \dots, k$$

$$Ur(\gamma) = \sum_{i=1}^{n} \left\{ \frac{\delta i}{\lambda(yi) + \beta^{T} zi} \frac{\partial \lambda(yi)}{\partial \gamma_r} - \frac{\partial \Lambda(yi)}{\partial \gamma_r} \right\} = 0 \qquad r = 1, \dots, q$$

Now, for $\theta_i = (\beta_i, \Lambda_i) \in \Theta$, $i = 1, 2$, define the distance d on the parameter space Θ as

$$d(\theta_1, \theta_2) = d((\beta_1, \Lambda_1), (\beta_2, \Lambda_2))$$

$$= |\beta_1 - \beta_2| + \left[\int_{\tau 0}^{\tau 1} (\Lambda_1(u) - \Lambda_2(u))^2\, G(u)\right]^{1/2} \qquad (1.20)$$

In the following, we prove the consistency of maximum likelihood estimators for $(\beta.\Lambda)$.

Theorem 1.1 *Suppose that the conditions* (A1) − (A7) *hold. Then we have*

$d(\hat{\theta}_n, \theta_2) = d((\hat{\beta}_n, \widehat{\Lambda}_n), (\beta_0, \Lambda_0)) \to 0$ *a.s*

i.e. the maximum likelihood estimators $(\hat{\beta}_n, \widehat{\Lambda}_n)$, *are consistent estimators of* (β_0, Λ_0).

Proof : We borrow some rdeas of the proof from Wang et al.(2015).Let

$p_n(y) = p(y; \hat{\beta}_n, \widehat{\Lambda}_n)$ and $p_0(y) = p(y; \beta_0, \Lambda_0)$, where

$p(y, \beta, \Lambda) = \frac{(\lambda_0(y)+\beta^T z)^{\delta} \exp(-\Lambda_0(y) - \beta^T z\, y)}{R(z;\beta,\Lambda)}$

Since $(\hat{\beta}_n, \widehat{\Lambda}_n)$ maximizes the likelihood function over Θ and $(\beta_0, \Lambda_0) \in \Theta$ then

$\frac{1}{n}\sum_{i=1}^{n} \log\frac{p_n(Y_i)}{p_0(Y_0)} \geq 0$

By concavity of log function, for any $0 < \alpha < 1$, we have

$\frac{1}{n}\sum_{i=1}^{n} log\left\{1-\alpha+\alpha\frac{p_n(Y_i)}{p_0(Y_0)}\right\} \geq \frac{\alpha}{n}\log\frac{p_n(Y_i)}{p_0(Y_0)}$

n is the empirical measure of $(\delta_i, Y_i, X_i, Z_i)$, $i = 1, 2, ..., n$ and P is the joint probability measure of (δ, Y, X, Z). Now

$$\frac{1}{n}\sum_{i=1}^{n} log\left\{1-\alpha+\alpha\frac{p_n(Y_i)}{p_0(Y_0)}\right\} = \int log\left\{1-\alpha+\alpha\frac{p_n(Y_i)}{p_0(Y_0)}\right\} d\,P_n(y)$$
$$= \int log\left\{1-\alpha+\alpha\frac{p_n(Y_i)}{p_0(Y_0)}\right\} d\,(P_n - P)(y)$$
$$+\int log\left\{1-\alpha+\alpha\frac{p_n(Y_i)}{p_0(Y_0)}\right\} d\,P(y) \qquad (1\text{-}21)$$

Since B is bounded subset of R^k then $\hat{\beta}_n \in B$ is a sequence of bounded vectors. From Bolzano-Weierstrass theorem, for any subsequence of $\hat{\beta}_n$ we can find a subsequence converging to $\beta^* \in \bar{B}$, the closure of B. Since $\hat{\eta}_n$ is a sequence of bounded nondecreasing functions then using Helly's selection theorem, for any subsequence of $\hat{\eta}_n$ we can find a subsequence converging to some nondecreasing function $\eta*$. Without loss of generality, we assume that $(\hat{\beta}_n, \hat{\eta}_n)$ converges to (β^*, η^*). To prove the theorem, it is sufficient to prove that $\beta^* = \beta_0$ and $\eta^* = \eta_0$. Since $(\hat{\beta}_n, \hat{\eta}_n)$ are the maximum likelihood estimators of (β_0, η_0) we obtain $l_n(\hat{\beta}_n, \hat{\eta}_n) \geq l_n(\beta_0, \eta_0)$.

Let the sample space Ω be defined as $\Omega = \{(\delta_i, Y_i, X_i, Z_i), i = 1, 2, ...\}$ the space of all infinite sequences endowed with the σ-field generated by the product topology on $\Pi_{i=1}^{n}\{0,1\} \times R^d \times (R^+)^2$ and the product measure P.

Now, for the first term of (**??**), since the class of functions

$$\left\{\log\left(1-\alpha+\alpha\frac{p_n}{p_0}, (\beta,\Lambda)\right) \in \Theta\right\}$$

is uniformly bounded and uniformly Lipschitz of order 1 (see Wang et al. (2015)), so there exists a set $\Omega_0 \subset \Omega$ with $P(\Omega_0) = 1$ such that

$\int log\left\{1-\alpha+\alpha\frac{p_n}{p_0}\right\} d\,(P_n - P)(y) \to 0$ for every $\omega \in \Omega_0$

Let $p_*(Y) = p(Y, \beta_*, \Lambda_*)$. By the bounded convergence theorem

$$\lim_{n\to\infty} E\int log\left\{1-\alpha+\alpha\frac{p_n(Y)}{p_0(Y)}\right\} = E\int log\left\{(1-\alpha)+\alpha\frac{p_*(Y)}{p_0(Y)}\right\} \geq 0 \quad (1-22)$$

since $\quad E\int log\left\{1-\alpha+\alpha\frac{p_n(Y)}{p_0(Y)}\right\} \geq 0$

Now, by the concavity of log function and using Jensen's inequality, we get

$$\int log\left\{(1-\alpha)+\alpha\frac{p_*(Y)}{p_0(Y)}\right\} dP(y) = Elog\left\{(1-\alpha)+\alpha\frac{p_*(Y)}{p_0(Y)}\right\}$$
$$\leq log\left\{(1-\alpha)+\alpha E\left(\frac{p_*(Y)}{p_0(Y)}\right)\right\}$$
$$= log\{(1-\alpha)+\alpha\int p_*(Y)dy\}$$
$$=0 \qquad (1.23)$$

Therefore Equations (**??**) and (**??**) gives us

$$\int log\left\{1-\alpha+\alpha\frac{p_*(Y)}{p_0(Y)}\right\}dP(y)=0\Rightarrow p_*(Y)=p_0(Y)$$

which implies that

$\Lambda_0(y)+\beta_0^T zy=\Lambda_*(y)+\beta_*^T zy \Leftrightarrow \Lambda_0(y)-\Lambda_*(y)+(\beta_0^T-\beta_*^T)zy=0$ (1.24)

Let $z_1=z_2$. with z_1 and z_2 we obtain

$\Lambda_0(y)-\Lambda_*(y)+(\beta_0^T-\beta_*^T)\,z_1y=0$ (1.25)

$\Lambda_0(y)-\Lambda_*(y)+(\beta_0^T-\beta_*^T)\,z_2y=0$ (1.26)

By subtracting (**??**) from (**??**) we get

$(\beta_0^T-\beta_*^T)\,(z_1-z_2)y=0$

Since $z_1=z_2$ and $y=0$ and from Assumption A_4 (b) we have $\beta_0=\beta_*$ and we get $\Lambda_0=\Lambda_*$ a.s. Finally, using the bounded convergence theorem, we obtain

$$\lim_{n\to\infty} d((\beta_0,\Lambda_0),(\hat\beta_n,\hat\Lambda_n))=\lim_{n\to\infty}\|\beta_0-\hat\beta_n\|+\lim_{n\to\infty}\left[\int\left(\Lambda_0(u)-\hat\Lambda_n(u)\right)^2 G(u)\right]^{1/2}$$
$$=\|\beta_0-\hat\beta_n\|+\left[\int\left(\Lambda_0(u)-\hat\Lambda_n(u)\right)^2 G(u)\right]^{1/2}=\mathbf{0}$$

2 Simulation:

In this section, a Monte-Carlo simulation was carried out to assess the finite sample performance of the estimation approach proposed in Section 2 using R language 3.3.0. In the study, we assume that $Z\sim$ *Bernoulli* (0.5). The failure time T was generated from model (**??**) with $\lambda 0(t)=1$ as shown in the following steps. Since $F(t|z)=1-e^{-\Lambda(t|z)}=1-e^{-\Lambda(t)-\beta zt}$ then for fixed baseline hazard model, $\lambda(t)=1$, we have $F(t|z)=1-e^{-kt-\beta zt}\Rightarrow t=\frac{-\log(1-F(t|z))}{1+\beta z}$

Hence, to generate random numbers from additive hazard model, we use the following algorithm.

Step.1 Generate $u\sim U(0,1)$.

Step.2 Compute $t=\frac{-\log(1-u)}{1+\beta z}$.

For the left-truncation time X and the observation time C, they were assumed to follow the unform distributions $U(0,a)$ and $U(0,b)$, respectively where the constants a and b are chosen to give appropriate percentages of left-truncated and right-censored observations. Table 1 presents the simulation results of estimation the parameter β based on 500 replications with $\beta 0=0.5,1,2$, sieve number $q=5,9$ and sample size $n=200,500$. For each setup, the results include the bias (BIAS) given by the average of 500 point estimates minus the true value, the sample standard error (SSD) of the 500 point estimates, the average of 500 estimated standard errors (SEE) using the observed information matrix, and the 95% empirical coverage probability. Two methods, A and B, are used to perform simulation. In method A, we use Equation (**1-17**) and (**1-18**) and in method B, we use Equation (**1-19**) It can be seen from the simulation study that the two methods (using score operator's functions (A) and using the full likelihood function (B)) yield very close results. It can be seen that the estimates seem to be unbiased and the sample standard deviation is quite close to the estimated standard error, suggesting that the proposed variance estimate is reasonable. The empirical coverage probabilities seemed quite close to 95% in all setups. We ao investigated the proposed estimation procedure by considering different sieve numbers. It is clearly that different sieve numbers seem to give similar results although a large sieve number tends to result in slightly better estimate

n	q	LT	RC	β_0	A				B			
					Bias	SSD	ESD	CP	Bias	SSD	ESD	CP
200	5	0.352	0.324	2	0.069	0.511	0.479	0.927	0.059	0.505	0.479	0.927
		0.294	0.338	1	0.028	0.355	0.336	0.926	0.022	0.346	0.336	0.926
		0.257	0.345	0.5	0.015	0.274	0.276	0.929	0.004	0.272	0.269	0.929
500	5	0.337	0.320	2	0.086	0.315	0.315	0.937	0.079	0.314	0.315	0.937
		0.235	0.331	1	0.044	0.220	0.217	0.944	0.037	0.218	0.217	0.944
		0.251	0.339	0.5	0.030	0.179	0.170	0.930	0.021	0.176	0.170	0.930
n	q	LT	RC	β_0	A				B			
					Bias	SSD	ESD	CP	Bias	SSD	ESD	CP
200	9	0.352	0.324	2	0.028	0.497	0.480	0.932	0.070	0.508	0.480	0.932
		0.284	0.336	1	-0.012	0.617	0.345	0.977	0.028	0.354	0.345	0.977
		0.253	0.348	0.5	-0.043	0.732	0.270	0.903	0.012	0.285	0.270	0.903
500	9	0.352	0.324	2	0.028	0.497	0.480	0.932	0.070	0.508	0.480	0.932
		0.285	0.331	1	0.043	0.219	0.218	0.958	0.041	0.219	0.218	0.958
		0.251	0.339	0.5	0.012	0.377	0.171	0.944	0.025	0.179	0.171	0.944

Table 1: Simulation results for $Z \sim Bin(1, 0.5)$

3 Application:

In this section, we apply the results of Section 2 on Stanford Heart transplant data of Crowley and Hu (1977)(see for example, Kalbfleisch and Prentice (2002), Apendix I, page 230) as a real-life example. From 1967 to 1974, 103 patients were recruited in this study. Among them, 69(67%) patients received a transplant and 75(73%) patients died (45 transplanted and 30 not transplanted). Patients who received a transplant are treated as two cases, before and after the operation, so cases in the transplant group are in general both right-censored and left-truncated. For each individual, we have survival (time since acceptance into the transplantation program to transplant and to death), event (patient's status; censored or not), age (age at acceptance), year (year of acceptance), surgery (previous surgery;1 = yes; 0 = no) and transplant (1 for received transplant, 0 for not). For the analysis here, we interested on the 188 patients who were received the transplant (transplant=1) and study the effects of age, year and previous surgery on the
patients's survival
12 time. We apply the proposed procedure given in Section 2 by considering several sieve numbers (q = 10 or 15). With q = 10, we obtained $\hat{\beta} = -0.000657$ with the estimated standard error of 0.00003, $\hat{\beta} = -0.011654$ with the estimated standard error of 0.00075, and $\hat{\beta} = -0.05024$ with the estimated standard error of 0.00777, for age, year and surgery, respectively. The p-value is less than 0.01 for testing $\beta 0 = 0$ for all of the covariates. For q = 15, the method yielded $\hat{\beta} = -0.00043$ with the estimated standard error of 0.0003, $\hat{\beta} = -0.014745$ with the estimated standard error of 0.00089 and $\hat{\beta} = -0.088059$ with the estimated standard error of 0.00348. Similarly, the p-value is less than 0.01 for testing $\beta 0 = 0$ for all of the covariates. It can be seen that for all the cases, there are significantly

effects of the all covariates on the survival function. Figures **??** and **??** display the estimated cumulative hazard and survival functions for the two groups of surgery covariate
($Z = 0$ for no previous surgery and $Z = 1$, otherwise) using sieve number $q = 10$ and
$q = 15$, respectively.
q11_sur_HApp.jpg
Figure 1: Estimated survival and cumulative hazard functions using surgery covariate and $q = 10$
q15_sur_HApp.jpg
Figure 2: Estimated survival and cumulative hazard functions using surgery covariate and
$q =13, 15$

References:

[1] Aalen, O. O. (1989). A Linear Regression Model for the Analysis of Life Times. Statisticin Medicine, 8, 907-925.
[2] Chen, L. and Sun, J. (2009). A multiple imputation approach to the analysis of current status data with the additive hazards model. Communications in Statistics-Theory and Methods, 38, 1009-1018.
[3] Feng, Y., Ma, L . and Sun, J. (2014). Regression analysis of current status data under the additive hazard model with auxiliary covariates. Jornal of Statistics, 42, 118-136.
[4] Wang, L. Sun, J. and Tong, X. (2010). Regression analysis of case II interval-censored failure time data with the additive hazards model. Statistical Sinica, 20, 1709-1723.
[5] Lu, X. and Peter, X.K. (2014). Efficient estimation of the partly linear additive hazards model with current status data. Jornal of Statistics, doi:10.1111/sjos.12108.
[6] Shen, S.P. (2014). A Generalization of turnbull's estimator for interval-censored and truncated data. Communications in Statistics, 43, 2958-2972.
[7] Huang, J. (1999). Efficient estimation of the partly linear additive cox model. The annals of Statistics, 27(5), 1536-1563.
[8] Pan, W. and Chappell, R. (2002). Estimation in the Cox proportional hazards model with left truncated and interval censored data. Jornal of the Inter, 58(1), 64-70.
[9] Kim, J.S. (2003). Efficient estimation for the proportional hazards model with lefttruncated and case 1 interval-censored data. Statistica Sinica, 13, 519-537.
[10] Zeng, D. Cai, J. and Shen, Y. (2006). Semiparametric additive risks model for intervalcensored data. Statistical Sinica, 16, 287-302.
[11] Wang, L. Sun, J. and Tong, X. (2010). Regression analysis of case II interval-censored failure time data with the additive hazards model. Statistica Sinica, Vol, 20, 1709-1723.
[12] Su, Y.R. and Wang, J.L. (2012). Modeling left-truncated and right-censored survival data with longitudinal covariates. The Annals of Statistics, 40(3), 1465-1488.14
[13] Lin, D. Y. Oakesm D. and Ying, Z. (1998). Additive hazarde regression with current status data. Biometrika, 85(2), 289-298.
[14] Huang, J and Wellner, J. A. (1996). Interval censored survival data. Biostatistics, 123, 123-169.

[15] Hang, J. (1995). Maximum likelihood esttimation for Proportional Odds Regression Model with Current Status Data. Monograph series, 27.
[16] Lin, D. Y. and Ying, Z. (1994). Semiparametric analysis of the additive risk model. Biometrika, 81(1), 61-71.
[17] Song, X., Sun, L., Mu, X. and Dinse, E. G. (2011). Additive hazareds regression with censoring indictors missing at random. Can. J. Stat.,38(3), 333-351.
[18] Qin, J. (2013). Semiparametric esttimation for the additive hazareds model with lefttruncated and right censored data. Biometrika, 100(4), 877-888.

[19] Shen, P. S. (2015). Semiparametric Analysis of Transformation Models with Dependently Left-truncated and Right-censored Data. Communications in StatisticsComputation and Simulation, 1532-4141.
[20] Shen, P. S. (2014). Proportional hazards regression with interval censored lefttruncated data. Journal of Statistical Computation and Simulation.
[21] Shen, P. S. (2014). Aalens Additive Risk Model for Left-Truncated and Right-Censored Data, Journal of Statistical Computation and Simulation.
[22] Chen, L. P. (2016). Pseudo Likelihood Estimation for the Additive Hazards Model with Data subject to LeftTruncation and Right-Censoring.
[23] Ma, L., Hu, T. and Sun, J. (2015). Sieve maximum likelihood regression analysis of dependent current status data. Biometrika, 18 doi: 10.1093/biomet/asv020.
[24] Feng, Y. and Sun, J. (2014). Regession Analysis of current status data under thr additive hazards model with auxiliary covariates. Scandinavian Journal of Statistics, doi: 10.1111/sjos.12098.15