Youngnam Statistical Letters
June 1992, Vol. 3, No. 1, 79-90

A Kernel Estimator of Hazard Ratio

Myong–Hui Choi, In–Suk Lee and Jae–Kee Song *

ABSTRACT

We consider hazard ratio as a descriptive measure to compare the hazard experience of a treatment group with that of a control group with censored survival data. In this paper, we propose a kernel estimator of hazard ratio. The uniform consistency and asymptotic normality of a kernel estimator are proved by using counting process approach via martingale theory and stochastic integrals.

1. Introduction

Being compared survival across treatment groups in a clinical trial, it is useful to have a descriptive measure of the difference in survival between groups. If the hazard functions in two groups are roughly proportional, then the ratio of hazard functions has the interpretation of relative risk. Otherwise the ratio of hazard functions depends on t and puts a construction on a hazard ratio and has intuitive appeal as a descriptive statistic.

The proportional hazard model (here after will be omitten by PHM) proposed by Cox (1972) specified that the hazard rate for the survival time T of an individual with covariate vector z has the form

$$\alpha(t|z) = \alpha_0(t) exp(\beta' z), \quad t \geq 0$$

where β is a p-vector of unknown regression coefficients and $\alpha_0(t)$ is an arbitrary and unspecified baseline hazard function.

* Department of Statistics, Kyungpook National University, Taegu, Korea.

Consider the special case in which $p = 1$. Let z be the indicator function that the obsevation is from treatment group. Then the Cox PHM is reduced to

$$\alpha_2(t) = e^{\beta}\alpha_1(t)$$

where $\theta = e^{\beta}$ is the relative risk, the proportionality constant of the hazard functions in two groups, and α_i is the corresponding hazard function in each group. In this case, $S_2(t) = (S_1(t))^{\theta}$,where $S_i(t)$ is the survival function of the ith group. If $\theta = 1$, then there is no difference between the two survival functions.

Aalen (1975, 1978) demonstrated that most of the two sample rank tests were special cases of tests based on his multiplicative intensity model for counting processes and that the asymptotic properties of these tests could be derived by martingale central limit theory. Also, Gill (1980) used martingale method to treat the two sample problem for censored data.

On the other hand, Kalbfleish and Prentice (1981) studied an estimator of the average hazard ratio depending on weight functions which are power transformations of the product-limit estimators of the survival distributions in the two groups Andersen (1983) introduced the generalized rank estimator of the hazard ratio as a new interpretation of the linear nonparametric two sample tests for censored data, and established asymptotic normality by using counting process and martingale theory. O'Sullivan (1986) studied nonparametric estimation of the relative risk in the Cox model as alternative to the local scoring method of Hastie and Tibshirani (1986). This method involves penalized partial likelihood. Dabrowska and Doksum and Song (1989) discussed graphs, confidence procedures and tests that could be used with censored survival data to compare the hazard experience of a treatment group with that of a control group. In particular, they considered the relative change $\Delta(t)$ in a cumulative hazard function, which equals to $\theta(t) - 1$.

In this paper, we shall propose a kernel estimator of hazard ratio , where we use the kernel function $K(t)$ instead of a predictable random weight function in the generalized rank estimator (See Andersen, 1983). Particularly we extend their estimator in Dabrowska et al (1989)'s paper by the kernel method.

In the next section we review various estimators of the hazard ratio and introduce the proposed kernel estimator of the hazard ratio. In Section 3, we derive uniform consistency and asymptotic normality of the proposed kernel estimator of hazard ratio by the Lenglart's inequality and the martingale central limit theorem. In Section 4, we summarize the results of this paper.

2. Formulation of the problem

Let $(\Omega, \mathcal{F}, \mathcal{P})$ be a probability space and let $\{\mathcal{F}_t, t \in [0,1]\}$ be an increasing, right-continuous family of sub-sigma algebras of $\mathcal{F}$. We take $\mathcal{F}_t$ to represent the information collected during the period $[0,t]$. The counting processes $N_i(t), t \in [0,1]$, $i = 1,2$, are stochastic processes counting the number of events of interest in each of the two groups and adapted to $\{\mathcal{F}_t\}$, where each sample path is a right-continuous step function with $N(0) = 0$ and a finite number of jumps, each of size $+1$. No two components jump at the same time. In survival study $N_i(t)$ will be the number of deaths in group i in the interval $[0,1]$. We also assume that $EN_i(1) < \infty$, $i = 1,2$. Since N_i is increasing and hence a submartingale it follows from the Doob - Meyer decomposition that $N_i = A_i + M_i$, where A_i is a predictable increasing process and M_i is a martingale. We shall assume that there exists a nonnegative left-continuous process Λ_i adapted to $\{\mathcal{F}_t\}$, with right-hand limits such that $A_i(t) = \int_0^t \Lambda_i(s)ds$. Then ,by Aalen (1978),

$$M_i(t) = N_i(t) - \int_0^t \Lambda_i(s)\, ds, \quad i = 1,2, \tag{2.1}$$

are square integrable martingales with variance processes

$$< M_i > (t) = \int_0^t \Lambda_i(s)\, ds. \tag{2.2}$$

The process Λ_i is called the intensity process of N_i.

This paper contributes to the study of the multiplicative intensity model (Aalen ,1978), where it is assumed that Λ_i can be written in the form

$$\Lambda_i(t) = \alpha_i(t) Y_i(t), \quad t \in [0,1], \tag{2.3}$$

where α_i is an unknown function and Y_i is an observable ,adapted left stochastic process. In a survival study $Y_i(t)$ will be the number of individuals at risk at time $t-$ and α_i is the hazard function.

Also, in order to introduce the stochastic integrals for counting processes, we assume that M_i is a martingale and H_i is a predictable process. Define a process $U_i = \{U_i(t) : t \geq 0\}$ by $U_i(t) = \int_0^t H_i(s)dM_i(s)$ or equivalently, $dU_i(t) = H_i(t)dM_i(t)$. Then U_i is a stochastic integral and also a martingale. Also, $< U_i > (t) = \int_0^t H_i(s)^2 d < M_i > (s)$ by the relation,

$$\begin{aligned} Var[dU_i(t)|\mathcal{F}_{t-}] &= Var[H_i(t)dM_i(t)|\mathcal{F}_{t-}] \\ &= H_i(t)^2 Var[dM_i(t)|\mathcal{F}_{t-}] \\ &= H_i(t)^2 d < M_i > (t). \end{aligned} \tag{2.4}$$

In this section, we wish to estimate $\theta(t)$ with randomly censored data. In the case of the proportional hazard model, the estimation of the relative risk has been considered by many authors.

First, the Cox (1972, 1975) maximum partial likelihood estimator is given by

$$\hat{\theta}_{Cox} = e^{\hat{\beta}}, \tag{2.5}$$

where $\hat{\beta}$ is the Cox's estimator, the solution to $\frac{d}{d\beta}C(\beta, 1)$ and the log partial likelihood is

$$C(\beta, t) = \sum_{i=1}^{n} \int_0^t \beta' z_i(s) dN_i(s) - \int_0^t \log\left\{\sum_{i=1}^{n} Y_i(s) e^{\beta' z_i(s)}\right\} d\overline{N}(s)$$

where $\overline{N} = \sum_{i=1}^{n} N_i$.

Andersen and Gill (1982) proved the asymptotic properties of $\hat{\beta}$ in framework of counting process. Andersen (1983) presented the asymptotic variance of $\hat{\theta}_{Cox}$ and showed that the Cox-estimator $\hat{\theta}_{Cox}$ of θ always has a smaller asymptotic variance than any estimator of the form $\hat{\theta}_W(t)$.

Secondly, Andersen (1983) and Gill and Schumacher (1987) consider the generalized rank estimator defined by

$$\hat{\theta}_W(t) = \frac{\int_0^t W(s)d\hat{\beta}_2(s)}{\int_0^t W(s)d\hat{\beta}_1(s)}, \quad t \in (0, 1], \tag{2.6}$$

where $W(t)$ is a predictable process and $\hat{\beta}_i(t)$ is a estimator introduced by Nelson (1972) and generalized by Aalen (1978) such that

$$\hat{\beta}_i(t) = \int_0^t \frac{dN_i(s)}{Y_i(s)}, \quad i = 1, 2.$$

As noted by Aalen the choice $W = Y_1 Y_2$ corresponds to the generalized Wilcoxon test of Gehan (1965) and $W = Y_1 Y_2/(Y_1 + Y_2)$ corresponds to the log rank test.

Dabrowska et al (1989) considered the ratio of the Nelson - Aalen estimators of the each cumulative hazard function as an estimator of hazard ratio under the hypothesis that the PHM holds. This estimator can be extended by applying kernel method , which becomes a kernel-type estimator. Now we propose a kernel estimator of hazard ratio defined by

$$\hat{\theta}_{KER}(t) \quad = \quad \frac{\int_0^1 K((t-s)/b)d\hat{\beta}_2(s)}{\int_0^1 K((t-s)/b)d\hat{\beta}_1(s)} \ , \tag{2.7}$$

where

$$\hat{\beta}_i = \int_0^t \frac{dN_i(s)}{Y_i(s)}, \quad i = 1, 2,$$

are the Nelson-Aalen estimators, and $N_i(t)$ are the numbers of deaths in the group i in the interval $[0, t]$ and $Y_i(t)$ are the numbers of the individuals at risk at time $t-$ and K is a bounded function with integral 1, and b is a positive number. Since, for each $i = 1, 2$, $\hat{\beta}_i$ jumps at X_{ij} with jump size $\frac{1}{Y_i(X_{ij})}$, when $\delta_{ij} = 1, j = 1, 2, \cdots, n_i$, the kernel estimator $\hat{\theta}_{KER}(t)$ can be represented as follows:

$$\hat{\theta}_{KER}(t) = \frac{\displaystyle\sum_{\substack{j:X_{2j} \le t \\ \delta_{2j}=1}} \frac{K((t - X_{2j})/b)}{Y_2(X_{2j})}}{\displaystyle\sum_{\substack{l:X_{1l} \le t \\ \delta_{1l}=1}} \frac{K((t - X_{1l})/b)}{Y_1(X_{1l})}}. \tag{2.8}$$

Remark 1. The bandwidth b represents the amount of smoothing that tends to 0 as $n \longrightarrow \infty$, but $nb \longrightarrow \infty$. If a very small bandwidth b will reduce the bias, then the variance will become large. Also, choosing a large bandwidth will reduce the variance. Thus the appropriate selection of bandwidth is very important in kernel estimation of hazard ratio (See Silverman, 1985).

3. Asymptotic Results

Now we are concerned with the asymptotic properties of the kernel estimator $\hat{\theta}_{KER}(t)$ and we consider the sequences of counting processes $\{N_i^{(n)}\}$ on $[0,1]$ with the corresponding sequences of martingales given by

$$M_i^{(n)}(t) = N_i^{(n)}(t) - \int_0^t \Lambda_i^{(n)}(s)ds, \quad i = 1, 2, n = 1, 2, \cdots, \tag{3.1}$$

where$\{\Lambda_i^{(n)}\}$ is the sequences of intensity process. In the survival model with a total number $n = Y_1^{(n)}(0) + Y_2^{(n)}(0) = n_1 + n_2$ of individuals, $Y_i^{(n)}(t), i = 1, 2$, will have the same order of magnitude as n and $n_i = Y_i^{(n)}(0)$ is the i^{th} sample size ,$i = 1, 2$.

3.1 Consistency of the kernel estimator $\hat{\theta}_{KER}(t)$

To show the consistency, we refer the following lemma, so called the Lenglart's inequality.

Lemma 3.1. (Andersen & Gill) Let M be a local square integrable martingale. Then for all $\delta, \eta > 0$

$$P\left\{ \sup_{t\in[0,1]} |M(t)| > \eta \right\} \le \frac{\delta}{\eta} + P\{< M, M > (1) > \delta\}. \tag{3.2}$$

Now the uniform consistency of the kernel estimator is as follows:

Theorem 3.2. Assume that the following conditions hold :

(i) $\alpha(t)$ is continuous on $[0,1]$.

(ii) There exist funtions y_1, y_2 taking value in $(0,1)$ such that under the hypothesis that the true hazard ratio is θ $\sup_{t\in[0,1]} \left| \frac{Y_i^{(n)}(t)}{n} - y_i(t) \right| \xrightarrow{p} 0, \quad i = 1, 2;$ as $n \to \infty$.

Then, under the proportional hazard model, we have

$$\sup_{t\in[0,1]} |\hat{\theta}^{(n)}_{KER}(t) - \theta| \xrightarrow{p} 0, \quad \text{as} \quad n \to \infty. \tag{3.3}$$

Proof. From the definition (2.7) of the kernel estimator $\hat{\theta}^{(n)}_{KER}(t)$, we have

$$\hat{\theta}^{(n)}_{KER}(t) - \theta = \frac{\int_0^1 K^{(n)}\left(\frac{t-s}{b_n}\right)\frac{dN_2^{(n)}(s)}{Y_2^{(n)}(s)}}{\int_0^1 K^{(n)}\left(\frac{t-s}{b_n}\right)\frac{dN_1^{(n)}(s)}{Y_1^{(n)}(s)}} - \theta$$

$$= \frac{\int_0^1 K^{(n)}\left(\frac{t-s}{b_n}\right)\left(\frac{dN_2^{(n)}(s)}{Y_2^{(n)}(s)} - \theta\frac{dN_1^{(n)}(s)}{Y_1^{(n)}(s)}\right)}{\int_0^1 K^{(n)}\left(\frac{t-s}{b_n}\right)\frac{dN_1^{(n)}(s)}{Y_1^{(n)}(s)}}$$

$$= \frac{\int_0^1 K^{(n)}(\frac{t-s}{b_n})(\frac{dM_2^{(n)}(s)}{Y_2^{(n)}(s)} - \theta\frac{dM_1^{(n)}(s)}{Y_1^{(n)}(s)})}{\int_0^1 K^{(n)}(\frac{t-s}{b_n})\frac{dN_1^{(n)}(s)}{Y_1^{(n)}(s)}},$$

here the last equality follows from the facts that $dN_i^{(n)}(s) = dM_i^{(n)}(s) + \alpha_i(s)Y_i^{(n)}(s)$ and $\alpha_2(s) = \theta\alpha_1(s)$ under the PHM. Then $\hat{\theta}^{(n)}_{KER}(t) - \theta$ can be represented by a stochastic integral as follows :

$$\hat{\theta}^{(n)}_{KER}(t) - \theta = \int_0^1 B_2^{(n)}(s)dM_2^{(n)}(s) - \int_0^1 B_1^{(n)}(s)dM_1^{(n)}(s)$$

where

$$B_1^{(n)}(s) = \left(\int_0^1 K^{(n)}\left(\frac{t-s}{b_n}\right)\frac{dN_1^{(n)}(s)}{Y_1^{(n)}(s)}\right)^{-1}\frac{K^{(n)}\left(\frac{t-s}{b_n}\right)}{Y_1^{(n)}(s)}$$

and

$$B_2^{(n)}(s) = \left(\int_0^1 K^{(n)}\left(\frac{t-s}{b_n}\right)\frac{dN_1^{(n)}(s)}{Y_1^{(n)}(s)}\right)^{-1}\frac{K^{(n)}\left(\frac{t-s}{b_n}\right)}{Y_2^{(n)}(s)},$$

are the stochastic integrals. Since $B_1^{(n)}$ and $B_2^{(n)}$ are the predictable processes and $M_1^{(n)}$ and $M_2^{(n)}$ are the martingales.

By combining Lemma 3.1 with (2.2) and (2.4), we have, for all $\delta, \eta > 0$,

$$P\left(\sup_{t\in[0,1]} |\hat{\theta}^{(n)}_{KER}(t) - \theta| > \eta\right)$$

$$\leq \frac{\delta}{\eta^2} + P(< \hat{\theta}^{(n)}_{KER} - \theta > (1) > \delta)$$

$$= \frac{\delta}{\eta^2} + P\Big(\Big(\int_0^1 [B_2^{(n)}(s)]^2 d < M_2^{(n)} > (s) + \int_0^1 [B_1^{(n)}(s)]^2 d < M_1^{(n)} > (s)\Big) > \delta\Big)$$

$$= \frac{\delta}{\eta^2} + P\left(\frac{\theta^2 \int_0^1 \frac{K^{(n)^2}((t-s)/b_n)}{n}\left(\frac{n}{\theta Y_2^{(n)}(s)} + \frac{n}{Y_1^{(n)}(s)}\right)\alpha_1(s)\, ds}{\left(\int_0^1 K^{(n)}\left(\frac{t-s}{b_n}\right)\frac{dN_1^{(n)}(s)}{Y_1^{(n)}(s)}\right)^2} > \delta\right).$$

Then since $\frac{Y_i^{(n)}(t)}{n} \xrightarrow{P} y_i(t) > 0$, for each t, the right-hand side converges to 0. So we get the desired result.

3.2 Asymptotic Normality

We discuss the asymptotic normality of estimators of hazard ratio. To do this, we need the martingale central limit theorem.

Lemma 3.3. (Andersen & Gill) Let $p \geq 1$ be fixed , and consider a sequence $N^{(n)}$ of k_n-variate counting processes with intensity processes $\Lambda^{(n)}$,and a sequence $H^{(n)}$ of $p \times k_n$-matrices of predictable processes, such that the stochastic integrals

$$U_j^{(n)}(t) = \int_0^t \sum_{h=1}^{k_n} H_{jh}^{(n)}(s)\{dN_h^{(n)}(s) - \Lambda_h^{(n)}(s)\, ds\};\ j = 1, \cdots, p;$$

are well defined. If, as $n \to \infty$,

$$< U_j^{(n)}, U_l^{(n)} > (t) \longrightarrow C_{jl}(t);\ j, l = 1, \cdots, p,\ t \in [0,1], \tag{3.4}$$

where C is $p \times p$ matrix of continuous functions on [0,1] forming the covariance function of a p-variate Gaussian martingale $U^{(\infty)}$ with $U^{(\infty)}(0) = 0$, and if for all

$\epsilon > 0$, as $n \to \infty$,

$$\int_0^1 \sum_{h=1}^{k_n} [H_{jh}^{(n)}(t)]^2 \Lambda_h^{(n)}(t) I\{|H_{jh}^{(n)}(t)|\} \, dt \xrightarrow{p} 0; \; j = 1, \cdots, p; \tag{3.5}$$

then

$$U^{(n)} \xrightarrow{\mathcal{D}} U^{(\infty)}, \quad \text{as } n \to \infty, \quad in \;\; D([0,1]^p).$$

Now we prove the asymptotic normality of the kernel estimator $\hat{\theta}_{KER}^{(n)}(t)$.

Theorem 3.4. Assume that the conditions of Theorem 3.2 are satisfied and that the following conditions hold :

(iii) $b_n \in (0, 1/2)$ and $b_n \to 0$ as $n \to \infty$.

(iv) The kernel has support within $[-1, 1]$ and is symmetric about zero.

(v) $\frac{1}{b_n} \int_0^1 K^{(n)}(\frac{t-s}{b_n}) \alpha_1(s) ds \xrightarrow{p} \alpha_1(t) \int_{-1}^1 K(u) du$, as $n \to \infty$. Furthermore, by Lemma 3.1, we have that

$$\frac{1}{b_n} \int_0^1 K^{(n)} \left(\frac{t-s}{b_n} \right) \frac{dN_1^{(n)}(s)}{Y_1^{(n)}(s)} \xrightarrow{p} \alpha_1(t) \int_{-1}^1 K(u) du, \quad \text{as} \quad n \to \infty.$$

Then

$$\sqrt{n}(\hat{\theta}_{KER}^{(n)}(t) - \theta) \xrightarrow{d} N(0, \sigma_{KER}^2(t)), \tag{3.6}$$

where

$$\sigma_{KER}^2(t) = \frac{\theta^2 \alpha_1(t)}{\left(\alpha_1(t) \int_{-1}^1 K(u) du\right)^2} \left(\frac{y_1(t) + \theta y_2(t)}{\theta y_1(t) y_2(t)} \right) \int_{-1}^1 K^2(u) du.$$

Proof. Let $Z^{(n)}(t) \equiv \sqrt{n}(\hat{\theta}_{KER}^{(n)}(t) - \theta)$. Then, the process $Z^{(n)}(t)$ is simply

$$Z^{(n)}(t) = \frac{\sqrt{n} \int_0^1 K^{(n)} \left(\frac{t-s}{b_n} \right) \left(\frac{dM_2^{(n)}(s)}{Y_2^{(n)}(s)} - \theta \frac{dM_1^{(n)}(s)}{Y_1^{(n)}(s)} \right)}{\int_0^1 K^{(n)} \left(\frac{t-s}{b_n} \right) \frac{dN_1^{(n)}(s)}{Y_1^{(n)}(s)}}$$

$$= \int_0^1 H_2^{(n)}(s) dM_2^{(n)}(s) - \int_0^1 H_1^{(n)} dM_1^{(n)}(s)$$

where

$$H_i^{(n)}(s) = \sqrt{n}B_i^{(n)}(s), i = 1,2,$$

and $B_1^{(n)}(s)$, $B_2^{(n)}(s)$ are defined in the proof of Theorem 3.2. Since $H_1^{(n)}(s)$ and $H_2^{(n)}(s)$ are the predictable processes and $M_1^{(n)}(s)$ and $M_2^{(n)}(s)$ are the martingales, $Z^{(n)}(t)$ is the stochastic integral with respect to $M_1^{(n)}(t)$ and $M_2^{(n)}(t)$.

Now, we need check two conditions (3.4) and (3.5) to apply Lemma 3.3. Since

$$\{|H_1^{(n)}(s)| > \epsilon\} = \left\{ \left| \frac{1}{\sqrt{n}} \left(\int_0^1 K^{(n)} \left(\frac{t-s}{b_n} \right) \frac{dN_1^{(n)}(s)}{Y_1^{(n)}(s)} \right)^{-1} \frac{K^{(n)}(\frac{t-s}{b_n})}{Y_1^{(n)}(s)/n} \right| > \epsilon \right\}$$

and the condition (ii) of Theorem 3.2, we have

$$I\{|H_1^{(n)}(s)| > \epsilon\} \xrightarrow{p} 0 \quad \text{uniformly on} \quad [0,1].$$

Therefore we see that for all $\epsilon > 0$,

$$\int_0^1 \{H_1^{(n)}(s)\}^2 \alpha_1(s) Y_1(s) I(|H_1^{(n)}(s)| > \epsilon) ds \xrightarrow{p} 0.$$

And similarly

$$\int_0^1 \{H_2^{(n)}(s)\}^2 \alpha_2(s) Y_2(s) I(|H_2^{(n)}(s)| > \epsilon) ds \xrightarrow{p} 0.$$

That is, the condition (3.5) of Lemma 3.3 is satisfied.

Next, to verify the condition (3.4) of Lemma 3.3, we consider the variance process of the stochastic integral such that

$$
\begin{aligned}
&< Z^{(n)} > (t) \\
&= \int_0^1 (H_2^{(n)}(s))^2 d < M_2^{(n)} > (s) + \int_0^1 (H_1^{(n)}(s))^2 d < M_1^{(n)} > (s) \\
&= \frac{\theta^2}{\left(\int_{-1}^1 K^{(n)}(u) \frac{dN_1^{(n)}(t-b_n u)}{Y_1^{(n)}(t-b_n u)} \right)^2} \\
&\times \int_{-1}^1 K^{(n)}(u)^2 \left(\frac{n}{Y_1^{(n)}(t-b_n u)} + \frac{n}{\theta Y_2^{(n)}(t-b_n u)} \right) \alpha_1(t-b_n u) du \\
&\xrightarrow{p} \frac{\theta^2}{\left(\alpha_1(t) \int_{-1}^1 K(u) du\right)^2} \left(\frac{1}{\theta y_2(t)} + \frac{1}{y_1(t)} \right) \alpha_1(t) \int_{-1}^1 K(u)^2 du.
\end{aligned}
$$

because of the conditions (iii), (iv) and (v) and (i), (ii) of Theorem 3.2. Hence, $< Z^{(n)} > (t) \xrightarrow{p} \sigma^2_{KER}(t)$. Therefore, by Lemma 3.3, $Z^{(n)}(t)$ converges in distribution to $N(0, \sigma^2_{KER}(t))$. So, we complete the proof.

Remark 2.

(1) To simplify the mathematics, we assume that the kernel has support within $[-1, 1]$ and can extend to $[-a, a]$ in $\mathcal{R}$.

(2) The asymptotic variance was obtained by using counting process via martingale theory and stochastic integral. (For more details, see Ramlau - Hansen.)

minum .5pt By condition (ii) of Theorem 3.2 and condition (v) of Theorem 3.4, we estimate $\sigma^2_{KER}(t)$ consistently by

$$\hat{\sigma}^2_{KER}(t) = n\hat{\theta}^2 \times \frac{\int_{-1}^{1} K^2(u) \left(\frac{dN_2(t-bu) + dN_1(t-bu)}{Y_1(t-bu)\hat{\theta}Y_2(t-bu)} \right)}{\left(\int_{-1}^{1} K(u) \frac{dN_1(t-bu)}{Y_1(t-bu)} \right)^2}, \quad t \in [b, 1-b], \qquad (3.7)$$

where

$$\hat{\theta} = \hat{\theta}^{(n)}_{KER}(t).$$

4. Conclusion

In this paper we proposed a kernel estimator of the hazard ratio by using the counting processes. Next we proved a uniform consistency and a asymptotic normality of $\hat{\theta}_{KER}(t)$ by using the counting processes and the martingale theory. Finally, we proposed the estimator of the variance function $\sigma^2_{KER}(t)$.

References

1. Aalen, O. O. (1975), Statistical Inference for a Family of Counting Process, Ph.D. Dissertation, *University of California, Berkeley.*
2. Aalen, O. O. (1978), Nonparametric Inference for a Family of Counting Processes, *The Annals of Statistics*, 6, 701-726.

3. Andersen, P. K. (1983), Comparing Survival Distributions via Hazard Ratio Estimates, *Scandinavian Journal of Statistics*, 10, 77-85.
4. Andersen, P. K. and Borgan, O. (1985), Counting Process Models for Life History Data, a Review, *Scandinavian Journal of Statistics*, 12, 97-158.
5. Andersen, P. K. and Gill, R. D. (1982), Cox's Regression Model for Counting Processes : A Large Sample Study , *The Annals of Statistics*, 10, 1100 - 1120.
6. Begun, J. M. and Reid, N. (1983), Estimating the Relative Risk with Censored data, *Journal of the American Statistical Association*, 78, 337-341.
7. Cox, D. R. (1972), Regression Models and Life Tables (with Discussion), *Journal of The Royal Statistical Society*, Ser. B, 34, 187-220.
8. Cox, D. R. (1975), Partial Likelihood, *Biometrika*, 62, 269-279.
9. Dabrowska, D. M., Doksum, K. A. and Song, J. K. (1989), Graphical Comparison of Cumulative Hazard for two Populations, *Biometrika*, 76, 763-773.
10. Fleming, T, R. and Harrington, D. P. (1991), *Counting Processes and Survival Analysis*, Now York: John Wiley.
11. Gill, R. D. (1980), *Censoring and Stochastic Integrals*, Mathematical Centre Tracts 124, Mathematisch Centrum: Amsterdam.
12. Gill, R. D. and Schumacher, M. (1987), A Simple Test of the Proportional Hazards Assumption, *Biometrika*, 74, 289-300.
13. Harrington, D. P. and Fleming, T. R. (1982), A Class of Rank Test Procedures for Censored Survival Data, *Biometrika*, 69, 553-566.
14. Hastie, T. J. and Tibshirani, R. J. (1986), Generalized Additive Models (with Discussion), Statistical Science, 1, 297-318.
15. Kalbfleisch, J. D. and Prentice, R. L. (1980), *The Statistical Analysis of Failure Time data*, New York: John Wiley.
16. Kalbfleisch, J. D. and Prentice, R. L. (1981), Estimation of the Average Hazard Ratio, *Biometrika*, 68, 105 -12 .
17. Nelson, W. (1972), Theory and Applications of Hazard Plotting for Censored Failure Data, *Technometrics*, 14, 945-966.
18. O' Sullivan, F. (1986), Relative Risk Estimation, *Technical Report No. 76. University of California, Berkeley.*
19. Silverman, B. W. (1985), *Density Estimation for Statistics and Data Analysis*, London: Chapman and Hall.