

Bożena Janiszewska, Roman Róžański

TRANSFORMED DIFFEOMORPHIC KERNEL ESTIMATION OF HAZARD RATE FUNCTION

Dedicated to Professor Kazimierz Urbanik

Abstract. In the article, a transformed diffeomorphic kernel estimator of the hazard rate function in the presence of censoring is constructed. The estimator is defined in the framework of multiplicative intensity point process model. It is shown that the proposed estimator is asymptotically unbiased, consistent and asymptotically normal. Analysis in the reduction of the bias of the diffeomorphic estimator is carried out. Some simulation results comparing the obtained estimator with the Ramlau-Hansen estimator are also presented.

1. Introduction

It is known that the Parzen-Rosenblatt kernel density estimator of a probability density function performs poorly near the boundary of the support of the estimated density. To remedy this so called boundary effect different methods have been proposed:

- the reflection method (Schuster ([11]), Cline and Hart [5])
- the boundary kernel method (Cheng, Fan and others [4])
- the transformation method (Ruppert and Marron [10], Saoudi, Ghorbel and others [9], Zhang, Karunamuni and others [13]).

The same problem one meets when estimates a hazard rate function in survival model. Since Aalen (1978) [1] it is known that the survival model with censored observations can be described as a multiplicative model of a point process $N_n(t)$ with the intensity function $\alpha(t)Y_n(t)$, where $\alpha(t)$ is the hazard rate function to be estimated and $Y_n(t)$ denotes the total number

Key words: hazard rate function, censoring, multiplicative intensity point process model, Ramlau-Hansen kernel estimator, kernel-diffeomorphic estimator, asymptotical unbiasedness, reduction of the bias, consistency, asymptotical normality.

2000 Mathematics Subject Classification: 62G05, 62N02, 62M.

of objects among n being at risk either of failure or censoring up to time t . The process $N_n(t)$ is observed on the interval $[0, T]$. Let $T=1$, for simplicity. In [7] the kernel estimator (the Ramlau-Hansen estimator) of the function α in a multiplicative point process model has been defined. In the mentioned paper, asymptotic unbiasedness, consistency and asymptotic normality of the Ramlau-Hansen estimator have been proved.

Estimating the unknown function α (hazard rate function) at the points near the boundary of the time interval on which the point process $N(t)$ is observed, one can see that the analogous boundary effect appears as in the problem of kernel density estimation. Taking into account applications in biostatistical and insurance models the problem is rather essential. To correct the severe bias of Ramlau-Hansen estimator near the boundary of the time interval we propose a transformed estimator defining a diffeomorphic kernel estimator of the function α . We have proved that the diffeomorphic estimator is asymptotically unbiased, consistent and asymptotically normal. It has been also showed that the bias of the estimator is reduced to the order b_n^2 , where b_n is a window bandwidth, the same order as at the points in the interior of the time interval. The method proposed gives better order of bias reducing than the reflection method in density estimators, and the same order as in transformed density estimators introduced in [12].

The presented paper is organized as follows. Section 2 contains preliminary results concerned with multiplicative point process model and asymptotic properties of the Ramlau-Hansen estimator. In Section 3 we introduce a kernel diffeomorphic estimator of the hazard rate function. Section 4 contains theorems and their proofs on asymptotic properties of the estimator introduced. In Section 5, approximation describing reducing of the bias for the diffeomorphic estimator is given. Section 6 is devoted to presentation of some simulation results concerned with the estimator considered.

2. Multiplicative model of a point process. Nonparametric hazard rate estimation

Let us consider a point process $\mathcal{N} = \{N(t)\}$, $t \in [0, T]$, on a given probability space $(\Omega, P, \mathcal{F}, \{\mathcal{F}_t\})$, where $\{\mathcal{F}_t\}_{t \geq 0}$ is a non-decreasing right-continuous family of σ -algebras. Under some appropriate assumptions the following Doob-Meyer decomposition holds,

$$N(t) = \Lambda(t) + M(t), \quad t \in [0, T],$$

where $\Lambda(t)$ is a predictable nonnegative non-decreasing process and $M(t)$ is a martingale (see Andersen, Gill [2], for all details). We will refer to the process $\Lambda(t)$ as the compensator of the point process $N(t)$. As long as appropriate regularity conditions hold there exists a predictable process $\lambda(t)$ such that

$$\Lambda(t) = \int_0^t \lambda(s) ds, \quad t \in [0, T],$$

where $\lambda(t)$ is called an intensity of the point process $N(t)$ and

$$\lambda(t) = \lim_{h \rightarrow 0} \frac{1}{h} P\{N(t+h) - N(t) \geq 1 \mid \mathcal{F}_{t-}\}.$$

In the multiplicative model, the intensity function is of the form

$$\lambda(t) = \alpha(t)Y(t),$$

where

$\alpha(t)$ — is a non-negative deterministic function,

$Y(t)$ — is a non-negative left-continuous and observable (predictable) process adapted to the filtration \mathcal{F}_t .

An important example of application of the multiplicative intensity point process model is the operation-failure model under censoring.

Suppose we are observing the course of n life insurance policies. Let T_i be the random variable representing the i -th person insured, where the deterministic function in the multiplicative model is interpreted as a hazard rate function of the form

$$P(T_i \in [t, t+dt) \mid T_i \geq t) = \alpha_i(t)dt.$$

Obviously, in practice we do not have an access to the complete set of observations regarding the lifetime of the people insured. Data censoring appears in many cases, i.e. loss of contact with the client before death. We define the point process of such a model to be

$$N^{(i)}(t) = 1\{\tilde{T}_i \leq t, D_i = 1\}, \quad i = 1, 2, \dots, n,$$

where

\tilde{T}_i — is either the time of death or censoring,

D_i — is the indicator of loss of contact $D_i = 1\{\tilde{T}_i = T_i\}$.

In order to guarantee the conditional independence of the risk of a failure (death) and censoring process we assume that (see Anderson, Gill [2])

$$P(\tilde{T}_i \in [t, t+dt), D_i = 1 \mid \mathcal{F}_{t-}) = \begin{cases} \alpha_i(t)dt & \text{for } \tilde{T}_i \geq t, \\ 0 & \text{for } \tilde{T}_i < t. \end{cases}$$

Defining the process

$$Y^{(i)}(t) = 1\{\tilde{T}_i \geq t\}$$

with the filtration

$$\mathcal{F}_t = \sigma(\{\tilde{T}_i < s, s \leq t, D_i\}),$$

we obtain

$$P(dN^{(i)}(t) = 1 \mid \mathcal{F}_{t-}) = \alpha_i(t)Y^{(i)}(t)dt.$$

As a result we obtain an n -dimensional point process $N = (N^{(1)}, \dots, N^{(n)})$ with intensity function

$$\lambda^{(i)}(t) = \alpha_i(t)Y^{(i)}(t), \quad i = 1, 2, \dots, n.$$

In certain case, it is reasonable to assume that the subjects observed represent a homogeneous population, which in our interpretation is equivalent to the equality of the hazard functions, $\alpha_i(t) = \alpha(t)$ for $i = 1, 2, \dots, n$. In such cases we define a sequence of one-dimensional point processes

$$N_n(t) = \sum_{i=1}^n N^{(i)}(t),$$

whose value at time t represents the number of failures (deaths) in the interval $[0, t]$. The intensity of the process $N_n(t)$ is given by

$$\lambda_n(t) = \sum_{i=1}^n \lambda^{(i)}(t) = \alpha(t)Y_n(t),$$

where

$$Y_n(t) = \sum_{i=1}^n Y^{(i)}(t) = \#\{i : \tilde{T}_i \geq t\}$$

represents the number of elements still functioning and being under observation up to time t . We can write

$$dN_n(t) = \alpha(t)Y_n(t)dt + dM_n(t).$$

In the multiplicative model of intensity

$$\lambda_n(t) = \alpha(t)Y_n(t), \quad t \in [0, T]$$

we use the following estimator of the deterministic hazard function $\alpha(t)$,

$$\hat{\alpha}_n(t) = \frac{1}{b_n} \int_0^T K\left(\frac{t-s}{b_n}\right) \frac{J_n(s)}{Y_n(s)} dN_n(s),$$

where $J_n(s) = 1\{Y_n(s) > 0\}$ and $J_n(s)/Y_n(s) = 0$, when $Y_n(s) = 0$.

In addition:

$$(2.1) \quad K - \text{the kernel function, with support } [-1, 1], \quad \int_{-1}^1 K(u)du = 1,$$

$$(2.2) \quad b_n - \text{bandwidth (a positive parameter), } b_n \rightarrow 0 \text{ as } n \rightarrow \infty.$$

The form of this estimator was derived by Ramlau-Hansen (see [7]). It is an asymptotically unbiased, consistent and asymptotically normal estimator of the hazard rate function (see [7]). However, the results of simulations

carried out using this estimator are unsatisfactory at the end points of the time interval on which the process $N(t)$ is observed. In order to eliminate this undesired effect, we will introduce a modified estimator of the hazard rate function. In the sequel, we assume that (2.1) and (2.2) are in force.

3. The kernel-diffeomorphic estimator of the hazard rate

Let $\{N_n(t)\}$ be a sequence of observed point processes and let $\varphi : (0, c) \rightarrow (a, b)$ be a diffeomorphic transformation, that means φ is continuously differentiable bijection from $(0, c)$ on (a, b) , where a, b and c are allowed to be infinite. Let us define

$$N_{\varphi,n}(t) = N_n(\varphi^{-1}(t))$$

and

$$\begin{aligned}\tilde{\alpha}(t) &= \alpha(\varphi^{-1}(t)) , \\ Y_{\varphi,n}(t) &= Y_n(\varphi^{-1}(t)) ,\end{aligned}$$

thus we obtain

$$dN_{\varphi,n}(t) = \tilde{\alpha}(t)Y_{\varphi,n}(t)\frac{1}{\varphi'(\varphi^{-1}(t))}dt + M_n(d\varphi^{-1}(t)) .$$

If we treat $M_n(\varphi^{-1}(dt))$ as a perturbation, we may estimate $\tilde{\alpha}(t)dt$ with

$$\frac{N_{\varphi,n}(dt)\varphi'(\varphi^{-1}(t))}{Y_{\varphi,n}(t)} .$$

Assume that $t = \varphi(u)$. This leads to

$$\tilde{\alpha}(\varphi(u)) = \alpha(\varphi^{-1}(\varphi(u))) = \alpha(u) .$$

Thus the function $\alpha(u)$ can be estimated by

$$\frac{1}{b_n} \int_a^b K\left(\frac{\varphi(u) - s}{b_n}\right) \frac{\varphi'(\varphi^{-1}(s))}{Y_n(\varphi^{-1}(s))} dN_{\varphi,n}(s) .$$

Hence, with the aid of the diffeomorphic transformation φ we define the transformed estimator of the hazard function as follows.

DEFINITION 3.1. *Diffeomorphic estimator $\hat{\alpha}_{\varphi,n}$ of the hazard rate function α is defined by the following formula*

$$(3.1) \quad \hat{\alpha}_{\varphi,n}(u) = \frac{1}{b_n} \int_a^b K\left(\frac{\varphi(u) - s}{b_n}\right) \frac{\varphi'(\varphi^{-1}(s))}{Y_n(\varphi^{-1}(s))} J_n(\varphi^{-1}(s)) dN_{\varphi,n}(s) ,$$

where $J_n(t) = 1\{Y_n(t) > 0\}$. If $Y_n(t) = 0$ we define $J_n(t)/Y_n(t) = 0$. The kernel K being a symmetric function and parameter b_n satisfy conditions (2.1), (2.2) respectively.

4. Asymptotic properties of the diffeomorphic estimator of the hazard rate

Let us now consider the properties of the estimator (3.1). We, will show that the estimator constructed above is asymptotically unbiased. First, let us calculate the expected value of the estimator (3.1). Using the fact that the expected value of a stochastic integral of a predictable process with respect to a martingale is equal to zero and taking into account that the kernel K is equal to 0 outside $[-1, 1]$ we obtain

$$\begin{aligned}
 (4.1) \quad E[\hat{\alpha}_{\varphi,n}(u)] &= E\left[\frac{1}{b_n} \int_a^b K\left(\frac{\varphi(u)-s}{b_n}\right) \frac{\varphi'(\varphi^{-1}(s))}{Y_n(\varphi^{-1}(s))} J_n(\varphi^{-1}(s)) dN_{\varphi,n}(s)\right] \\
 &= \frac{1}{b_n} \int_a^b K\left(\frac{\varphi(u)-s}{b_n}\right) E J_n(\varphi^{-1}(s)) \frac{\varphi'(\varphi^{-1}(s))}{Y_n(\varphi^{-1}(s))} \times \\
 &\quad \times \alpha(\varphi^{-1}(s)) Y_n(\varphi^{-1}(s)) \frac{1}{\varphi'(\varphi^{-1}(s))} ds \\
 &= \int_{-1}^1 K(v) \alpha(\varphi^{-1}(vb_n + \varphi(u))) E J_n(\varphi^{-1}(vb_n + \varphi(u))) dv,
 \end{aligned}$$

for sufficiently large n such that $\min\left(\frac{\varphi(u)-a}{b_n}, \frac{\varphi(u)-b}{b_n}\right) < -1$ and $\max\left(\frac{\varphi(u)-a}{b_n}, \frac{\varphi(u)-b}{b_n}\right) > 1$. Thus, we get

THEOREM 4.1. *Suppose that α is a continuous function. If $\forall u \quad E J_n(u) \rightarrow 1$ uniformly in some neighbourhood of u then*

$$(4.2) \quad E[\hat{\alpha}_{\varphi,n}(u)] \xrightarrow{n \rightarrow \infty} \alpha(u).$$

Proof. We have

$$\begin{aligned}
 E[\hat{\alpha}_{\varphi,n}(u)] &\stackrel{4.1}{=} \int_{-1}^1 K(v) \alpha(\varphi^{-1}(vb_n + \varphi(u))) E J_n(\varphi^{-1}(vb_n + \varphi(u))) dv \\
 &\stackrel{n \rightarrow \infty}{\cong} \int_{-1}^1 K(v) E J_n(u) \alpha(u) dv \\
 &\cong \alpha(u) \int_{-1}^1 K(v) dv = \alpha(u).
 \end{aligned}$$

Moreover, the following results concerned with the convergence to zero of the variance and the mean square error (MSE) of the estimator can be also obtained. Let us first calculate the variance

$$\begin{aligned}
\sigma_{\varphi}^2(u) &= \text{Var } \hat{\alpha}_{n,\varphi}(u) = E[\hat{\alpha}_{n,\varphi}(u) - E\hat{\alpha}_{n,\varphi}(u)]^2 \\
&= E\left[\frac{1}{b_n} \int_a^b K\left(\frac{\varphi(u) - s}{b_n}\right) \frac{\varphi'(\varphi^{-1}(s))}{Y_n(\varphi^{-1}(s))} J_n(\varphi^{-1}(s)) dN_{\varphi,n}(s) \right. \\
&\quad \left. - \frac{1}{b_n} \int_a^b K\left(\frac{\varphi(u) - s}{b_n}\right) E J_n(\varphi^{-1}(s)) \alpha(\varphi^{-1}(s)) ds\right]^2 \\
&\stackrel{n \rightarrow \infty}{\cong} E\left[\frac{1}{b_n} \int_a^b K\left(\frac{\varphi(u) - s}{b_n}\right) \frac{\varphi'(\varphi^{-1}(s))}{Y_n(\varphi^{-1}(s))} J_n(\varphi^{-1}(s)) dM_n(\varphi^{-1}(s))\right]^2 \\
&= \frac{1}{b_n^2} E\left[\int_a^b K^2\left(\frac{\varphi(u) - s}{b_n}\right) \left(\frac{\varphi'(\varphi^{-1}(s))}{Y_n(\varphi^{-1}(s))}\right)^2 J_n(\varphi^{-1}(s)) d\langle M_n(\varphi^{-1}(s)) \rangle\right] \\
&= \frac{1}{b_n^2} E\left[\int_a^b K^2\left(\frac{\varphi(u) - s}{b_n}\right) \frac{\varphi'(\varphi^{-1}(s))}{Y_n(\varphi^{-1}(s))} J_n(\varphi^{-1}(s)) \alpha(\varphi^{-1}(s)) ds\right] \\
&\stackrel{n \rightarrow \infty}{\cong} \frac{1}{b_n^2} \int_a^b K^2\left(\frac{\varphi(u) - s}{b_n}\right) \varphi'(\varphi^{-1}(s)) \alpha(\varphi^{-1}(s)) E\left\{\frac{J_n(\varphi^{-1}(s))}{Y_n(\varphi^{-1}(s))}\right\} ds,
\end{aligned}$$

where $\langle M_n(\cdot) \rangle$ denotes the predictable variation process of M_n .

PROPOSITION 4.1. *Suppose the following conditions are satisfied*

- $nb_n \rightarrow \infty$ as $n \rightarrow \infty$,
- $\int K^2(t) dt < \infty$,
- $nE\left\{\frac{J_n(s)}{Y_n(s)}\right\} \rightarrow \frac{1}{y(s)}$ uniformly in a neighbourhood of u ,
- $\alpha(s)$ and $y(s)$ are continuous in a neighbourhood of u .

Then

$$(4.3) \quad \sigma_{\varphi}^2(u) \xrightarrow{n \rightarrow \infty} 0.$$

Proof. We have

$$\begin{aligned}
\sigma_{\varphi}^2(u) &\cong \frac{1}{b_n^2} \int_a^b K^2\left(\frac{\varphi(u) - s}{b_n}\right) \varphi'(\varphi^{-1}(s)) \alpha(\varphi^{-1}(s)) E\left\{\frac{J_n(\varphi^{-1}(s))}{Y_n(\varphi^{-1}(s))}\right\} ds \\
&\cong \frac{1}{b_n} \int_{-1}^1 K^2(t) \varphi'(u) \alpha(u) E\left\{\frac{J_n(u)}{Y_n(u)}\right\} dt \\
&= \frac{1}{b_n} \varphi'(u) \alpha(u) E\left\{\frac{J_n(u)}{Y_n(u)}\right\} \int_{-1}^1 K^2(t) dt \\
&\cong \frac{1}{nb_n} \frac{\alpha(u) \varphi'(u)}{y(u)} \int_{-1}^1 K^2(t) dt \xrightarrow{n \rightarrow \infty} 0.
\end{aligned}$$

Furthermore, from (4.2) and (4.3) it follows that

$$E(\hat{\alpha}_{\varphi,n}(t) - \alpha(t))^2 \xrightarrow{n \rightarrow \infty} 0$$

and thus the estimator $\hat{\alpha}_{\varphi,n}$ is consistent.

4.1. Asymptotic normality

Using Rebolledo's theorem (see [8]), we show that the kernel diffeomorphic estimator is asymptotically normal.

THEOREM 4.2. *Let us assume that the following conditions hold:*

- $nb_n \rightarrow \infty$ as $n \rightarrow \infty$,
- $\int K^2(t)dt < \infty$,
- α is a continuous function,
- there exists a continuous function y such that

$$\forall \varepsilon \sup_{s \in [t-\varepsilon; t+\varepsilon]} \left| \frac{1}{n} Y_n(s) - y(s) \right| \xrightarrow{P} 0.$$

Then

$$\sqrt{nb_n}(\hat{\alpha}_{\varphi,n}(t) - E\hat{\alpha}_{\varphi,n}(t)) \xrightarrow{D} N(0, \tau^2(t)),$$

where

$$\tau^2(t) = \frac{\varphi'(t)\alpha(t)}{y(t)} \int_{-1}^1 K^2(u)du.$$

Proof. The following result can be obtained by similar calculations to those carried out earlier

$$\begin{aligned} & \sqrt{nb_n}(\hat{\alpha}_{\varphi,n}(t) - E\hat{\alpha}_{\varphi,n}(t)) \\ &= \sqrt{n/b_n} \int_a^b K\left(\frac{\varphi(t)-s}{b_n}\right) \frac{\varphi'(\varphi^{-1}(s))}{Y_n(\varphi^{-1}(s))} J_n(\varphi^{-1}(s)) dM_n(\varphi^{-1}(s)). \end{aligned}$$

Let us define the following predictable process

$$H_{\varphi,n}(s) = \sqrt{n/b_n} K\left(\frac{\varphi(t)-s}{b_n}\right) \frac{\varphi'(\varphi^{-1}(s))}{Y_n(\varphi^{-1}(s))} J_n(\varphi^{-1}(s)),$$

and let $\{\tilde{M}_n\}$ denote a sequence of martingales of the form

$$\tilde{M}_n(z) = \int_a^z H_{\varphi,n}(s) dM_n(\varphi^{-1}(s)).$$

We will show that $\langle \tilde{M}_n(z) \rangle \xrightarrow{P} V(z)$, where V is non-decreasing and non-negative function and $n \rightarrow \infty$. We have

$$\langle \tilde{M}_n(z) \rangle = \int_a^z H_{\varphi,n}^2(s) d\langle M_n(\varphi^{-1}(s)) \rangle$$

$$\begin{aligned}
& \stackrel{n \rightarrow \infty}{\cong} \int_{-1}^1 K^2(v) \frac{\varphi'(\varphi^{-1}(vb_n + \varphi(t)))}{y(\varphi^{-1}(vb_n + \varphi(t)))} \\
& \quad \times J_n(\varphi^{-1}(vb_n + \varphi(t))) \alpha(\varphi^{-1}(vb_n + \varphi(t))) dv \\
& \stackrel{n \rightarrow \infty}{\cong} \frac{\varphi'(t)}{y(t)} \alpha(t) \int_{-1}^1 K^2(v) dv = \tau^2(t).
\end{aligned}$$

Let $\{\tilde{M}_{n,\varepsilon}\}$, be the following sequence of martingales

$$\tilde{M}_{n,\varepsilon}(z) = \int_a^z H_{\varphi,n}(s) 1\{|H_{\varphi,n}(s)| > \varepsilon\} dM_n(\varphi^{-1}(s)).$$

It remains to show that

$$\langle \tilde{M}_{n,\varepsilon}(z) \rangle \xrightarrow{P} 0, \quad \text{when } n \rightarrow \infty.$$

Observe that

$$\begin{aligned}
\langle \tilde{M}_{n,\varepsilon}(z) \rangle &= \int_a^z H_{\varphi,n}^2(s) 1\{|H_{\varphi,n}(s)| > \varepsilon\} d\langle M_n(\varphi^{-1}(s)) \rangle \\
&= n/b_n \int_a^z K^2\left(\frac{\varphi(t) - s}{b_n}\right) \frac{\varphi'(\varphi^{-1}(s))}{Y_n(\varphi^{-1}(s))} 1\{|H_{\varphi,n}(s)| > \varepsilon\} \\
&\stackrel{P}{\cong} 1/b_n \int_a^z K^2\left(\frac{\varphi(t) - s}{b_n}\right) \frac{\varphi'(\varphi^{-1}(s))}{y(\varphi^{-1}(s))} 1\{|H_{\varphi,n}(s)| > \varepsilon\} \xrightarrow{P} 0
\end{aligned}$$

because

$$\begin{aligned}
1\{|H_{\varphi,n}(s)| > \varepsilon\} &= 1\left\{nK\left(\frac{\varphi(t) - s}{b_n}\right) \frac{\varphi'(\varphi^{-1}(s))}{Y_n(\varphi^{-1}(s))} > \varepsilon\sqrt{nb_n}\right\} \\
&\cong 1\left\{K\left(\frac{\varphi(t) - s}{b_n}\right) \frac{\varphi'(\varphi^{-1}(s))}{y(\varphi^{-1}(s))} > \varepsilon\sqrt{nb_n}\right\}
\end{aligned}$$

converges to zero uniformly in probability, when $nb_n \rightarrow \infty$ as $n \rightarrow \infty$. Finally, all the assumptions of Rebolledo's theorem ([2]) hold. Thus we have

$$\sqrt{nb_n}(\hat{\alpha}_{\varphi,n}(t) - E\hat{\alpha}_{\varphi,n}(t)) \xrightarrow{D} N(0, \tau^2), \quad \text{as } n \rightarrow \infty.$$

5. Reducing the bias

As mentioned before, the Ramlau-Hansen estimator does not give accurate results at the end points of the time interval. The estimator has a larger bias at the end points than in the interior. A diffeomorphic estimator might be a solution to this problem. We now consider the order of the bias for the Ramlau-Hansen estimator (R-H) as well as for the diffeomorphic estimator. We assume that the function to be estimated is defined on the interval (0,1). We consider points in the neighbourhood of zero, defined by

$u = Cb_n$, where $C \in (0, 1)$, thus the points $u \in (0, b_n)$. Assume also that $J_n(s) \xrightarrow{P} 1$, as $n \rightarrow \infty$ and the functions α and φ are twice differentiable. First we estimate the bias of the Ramlau-Hansen estimator $\hat{\alpha}_n$

$$\begin{aligned} E\hat{\alpha}_n(u) &= E\left[\frac{1}{b_n} \int_0^1 K\left(\frac{u-t}{b_n}\right) \frac{J_n(t)}{Y_n(t)} dN_n(t)\right] \\ &= \frac{1}{b_n} \int_0^1 K\left(\frac{u-t}{b_n}\right) \alpha(t) E[J_n(t)] dt \\ &\stackrel{n \rightarrow \infty}{\cong} \frac{1}{b_n} \int_0^{b_n(1+C)} K(C-t/b_n) \alpha(t) dt \\ &= \int_{-1}^C K(v) \alpha(u-vb_n) dv \\ &= \int_{-1}^C K(v) \left\{ \alpha(u) - \alpha'(u)vb_n + \frac{\alpha''(u)}{2}(vb_n)^2 \right\} dv + o(b_n^2) \\ &= \alpha(u)\mu_0(C) - \alpha'(u)b_n\mu_1(C) + \frac{\alpha''(u)}{2}(b_n)^2\mu_2(C) + o(b_n^2), \end{aligned}$$

where $\mu_k(C) = \int_{-1}^C v^k K(v) dv$. In this case the bias at 0 is approximately proportional to $b_n \alpha'(u)$ as $b_n \rightarrow 0$.

Let us choose a diffeomorphism φ such that $\varphi(u) - b_n > a$ and $\varphi(u) + b_n < b$ for $n > n_0$. Thus in the case of diffeomorphic estimator $\hat{\alpha}_{\varphi,n}$ we obtain

$$\begin{aligned} E\hat{\alpha}_{\varphi,n}(u) &= \frac{1}{b_n} \int_a^b K\left(\frac{\varphi(u)-s}{b_n}\right) EJ_n(\varphi^{-1}(s)) \alpha(\varphi^{-1}(s)) ds \\ &= \frac{1}{b_n} \int_{\varphi(u)-b_n}^{\varphi(u)+b_n} K\left(\frac{\varphi(u)-s}{b_n}\right) EJ_n(\varphi^{-1}(s)) \alpha(\varphi^{-1}(s)) ds \\ &\stackrel{n \rightarrow \infty}{\cong} \int_{-1}^1 K(v) \alpha(\varphi^{-1}(\varphi(u) + vb_n)) dv = **, \end{aligned}$$

where in contrast to the previous case we integrate on the whole support of the kernel K (i.e. $[-1, 1]$). Further, applying Taylor's expansion we get.

$$\begin{aligned} (5.1) \quad ** &\stackrel{n \rightarrow \infty}{\cong} \int_{-1}^1 K(v) \alpha\left(\varphi^{-1}(\varphi(u)) + \frac{1}{\varphi'(\varphi^{-1}(\varphi(u)))} vb_n\right) dv \\ &= \int_{-1}^1 K(v) \left\{ \alpha(u) + \alpha'(u) \frac{vb_n}{\varphi'(u)} + \frac{\alpha''(u)}{2} \left(\frac{vb_n}{\varphi'(u)}\right)^2 \right\} dv + o(b_n^2) \end{aligned}$$

$$\begin{aligned}
&= \alpha(u)\mu_0(1) + \alpha'(u)\frac{b_n}{\varphi'(u)}\mu_1(1) + \frac{\alpha''(u)}{2}\left(\frac{b_n}{\varphi'(u)}\right)^2\mu_2(1) + o(b_n^2) \\
&= \alpha(u) + \frac{\alpha''(u)}{2}\left(\frac{b_n}{\varphi'(u)}\right)^2\mu_2(1) + o(b_n^2).
\end{aligned}$$

The results above give an approximation of the bias for diffeomorphic estimator. It is of the order $\alpha''(u)b_n^2/2$ for all points $u \in (0, 1)$, and thus also at the end points.

6. Simulation results

Computer simulations have been carried out, in order to illustrate the properties of the diffeomorphic estimator. We assume that failure times T_1, \dots, T_n , where $n = 1000$ are observed on the interval $(0, 1)$ and data are generated from

- (i) the Weibull distribution with hazard function of the form $\alpha(t) = \gamma\rho t^{\rho-1}$, where $\gamma = 1$ and $\rho = 0.5$,
- (ii) the distribution with U^2 density function $f(t) = 0.5t^{-0.5}$ for $t \in (0, 1)$, for which the hazard function is of the form $\frac{0.5}{\sqrt{t(1-\sqrt{t})}}$.

Additionally these lifetimes were censored by n independent censoring time U_1, \dots, U_n generated from exponential distribution with the mean 4. We define $\tilde{T}_i = \min(T_i, U_i)$ and the indicator of censoring $D_i = 1\{\tilde{T}_i = T_i\}$. In this case, the point process N is of the form $N(t) = \sum_{i=1}^n 1\{\tilde{T}_i \leq t\}$. This process counts the total number of failures in the interval $(0, t)$, where $t \in (0, 1)$ and the intensity function is given by $\lambda(t) = \alpha(t)Y_n(t)$, where $Y_n = \sum_{i=1}^n 1\{\tilde{T}_i \geq t\}$. The process $Y_n(t)$ counts the number of elements at risk of failure just to time t . In both cases, we set $b_n = 0.25$, which is of order $n^{-1/5}$ and the kernel $K(x) = \frac{3}{4}(1 - x^2)$ for $x \in [-1, 1]$. We take the diffeomorphism φ to be $\varphi(x) = \log(x/(1 - x))$ for $x \in (0, 1)$.

6.1. Data from the Weibull distribution

Figure 1(a) illustrates the diffeomorphic estimator together with the hazard function for the Weibull distribution and figure 1(b) illustrates the behaviour of the Ramlau-Hansen estimator for the same model. The oscillations in the diffeomorphic estimator in the neighbourhood of zero may be caused by the fact that the variance of the estimator is larger at the end points of the interval. It can be seen from these graphs that the diffeomorphic estimator approximates the hazard function well on the entire domain, in contrast to the R-H estimator, which gives a poor approximation, especially in the neighbourhood of zero. The results obtained by simulating the bias of both estimators (Figs. 2(a) and (b)) are a confirmation of this fact. It can be seen that the bias of the R-H estimator (Fig. 2(b)) is of much higher order,

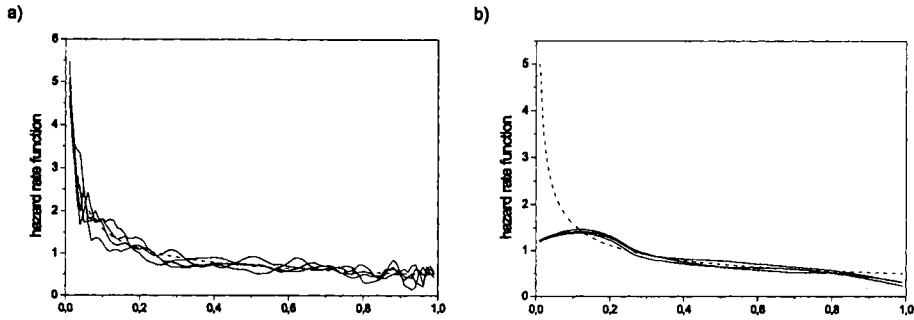


Figure 1. (a) The estimator $\hat{\alpha}_{n\varphi}$ (solid line) and true hazard rate function (dashed line), (b) The Ramlau-Hansen estimator (solid line) and true hazard rate function (dashed line).

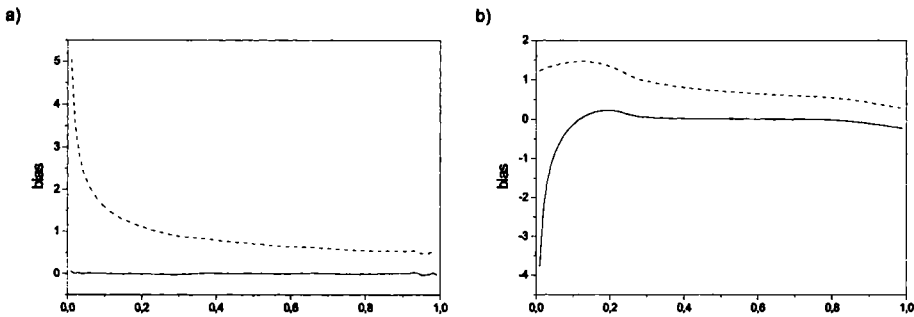


Figure 2. (a) the expected value of estimator $\hat{\alpha}_{\varphi,n}$ (dashed line), bias (solid line), (b) the expected value of estimator R-H (dashed line), bias (solid line).

especially in the neighbourhood of zero, than the bias of the diffeomorphic estimator (Fig. 2(a)). The difference in the bias is much less noticeable in the neighbourhood of 1. The bias of the diffeomorphic estimator is about 0.02, whilst the bias of the R-H estimator is about 0.3.

6.2. Data from the U^2 distribution

Figures 3(a) and 3(b) illustrate the diffeomorphic and R-H estimators, respectively, in the case of the U^2 -density. The hazard rate function in this case has two poles at 0 and at 1. Figures 4(a) and 4(b) give approximate values for the expected value and bias of these estimators. Since these graphs do not illustrate the behaviour of the estimators in the neighbourhood of zero well, figures 3(b) and 5 present the same estimates, but in this case restricted to the interval (0,0.1).

The quality of the simulation can be improved by choosing an appropriate value of the parameter b_n , depending upon the realisations of the failure times T_i . The method of cross-validation can be used to choose the optimal value of the parameter.

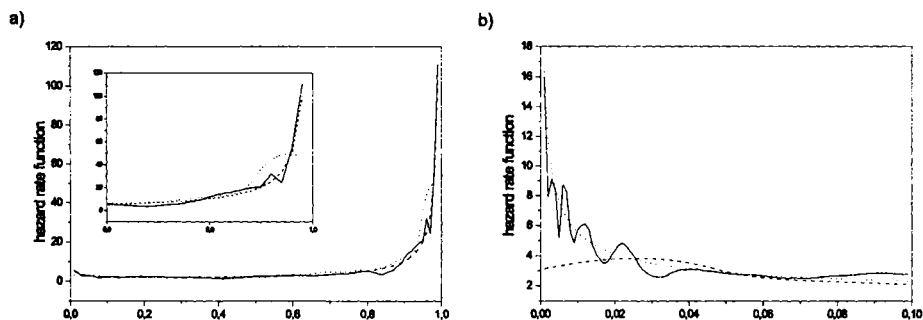


Figure 3. (a) The estimator $\hat{\alpha}_{n\varphi}$ (solid line), the estimator R-H (dotted line) and true hazard rate function (dashed line), (b) as above, but on $(0,0.1)$

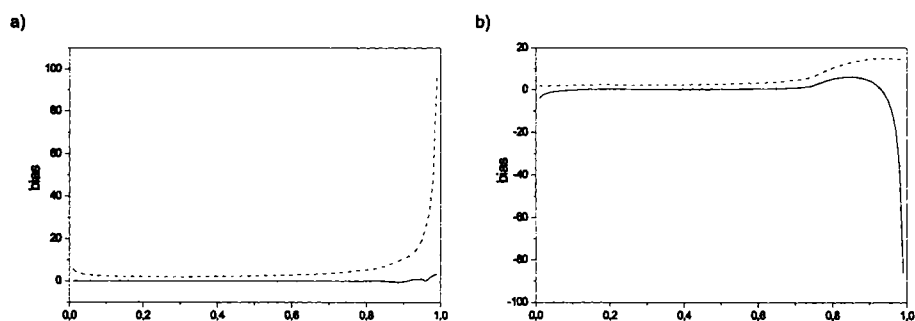


Figure 4. (a) the expected value of estimator $\hat{\alpha}_{\varphi,n}$ (dashed line), bias (solid line), (b) the expected value of estimator R-H (dashed line), bias (solid line).

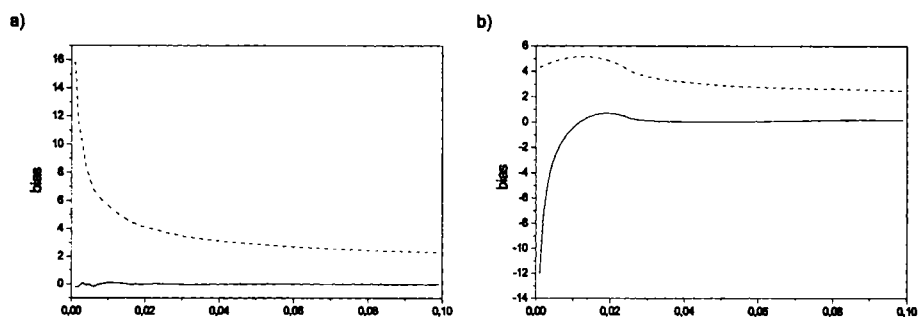


Figure 5. On $(0,0.1)$: (a) the expected value of estimator $\hat{\alpha}_{\varphi,n}$ (dashed line), bias (solid line), (b) the expected value of estimator R-H (dashed line), bias (solid line).

References

- [1] O. Aalen, (1978). *Nonparametric inference for a family of counting processes*, Ann. Statist. 6(4), 701–726.
- [2] P. Andersen, O. Borgan, R. Gill, and N. Keiding, (1993). *Statistical Models Based on Counting Processes*. Springer-Verlag, New York.
- [3] P. K. Andersen and O. Borgan, (1985). *Counting process models for life history*, Scand. J. Statist, 12, 97–158.
- [4] M. Y. Cheng, J. Fan, and J. S. Marron, (1997) *On automatic boundary corrections*, Ann. Statistics, 25, 1691–1.
- [5] D. B. H. Cline, and J. D. Hart, (1991) *Kernel estimation of densities with discontinuities or discontinuous derivatives*, Statistics, 22, 69–84.
- [6] T. R. Fleming, and D. P. Harrington, (1991) *Counting Processes and Survival Analysis*, John Wiley & Sons.
- [7] H. Ramlau-Hansen, (1983). *Smoothing counting process intensities by means of kernel functions*, Ann. Statist., 11(2), 453–466.
- [8] R. Rebolledo, (1980). *Central limit theorems for local martingales*, Z. Wahrsch. Verw. Gebiete, 51, 269–286.
- [9] S. Saoudi, F. Ghorbel, and A. Hillion, (1997) *Some statistical properties of the kernel-diffeomorphism estimator*, App. Stochastic Models Data Anal., Vol. 13, 39–58.
- [10] D. Ruppert, and J. S. Marron, (1994) *Transformations to reduce boundary bias in kernel density estimation*, J. Roy. Statist. Soc. Ser. B, 56(4), 653–671.
- [11] E. F. Schuster, (1985) *Incorporating support constraints into nonparametric estimators of densities*, Comm. Statist. A - Theory Methods., 14, 1125–1136.
- [12] M. Wand, J. S. Marron, D. Ruppert, (1991) *Transformations in density estimation (with discussion)*, J. Amer. Statist. Assoc., 86, 343–361.
- [13] S. Zhang, R. J. Karunamuni, and M. C. Jones, (1999) *An improved estimator of the density function at the boundary*, J. Amer. Statist. Assoc., 94, 1231–1240.

INSTITUTE OF MATHEMATICS
WROCLAW UNIVERSITY OF TECHNOLOGY
ul. Wybrzeże Wyspiańskiego 27
50-370 WROCLAW, POLAND

Received August 29, 2000; revised version December 22, 2000.