Discussiones Mathematicae
Probability and Statistics 25 (2005) 5–37

ESTIMATION OF THE HAZARD RATE FUNCTION WITH A REDUCTION OF BIAS AND VARIANCE AT THE BOUNDARY

Bożena Janiszewska and Roman Różański

Institute of Mathematics and Informatics
Wrocław University of Technology
Wybrzeże Wyspiańskiego 27, PL 50–370 Wrocław, Poland

e-mail: Bozena.Janiszewska@pwr.wroc.pl
e-mail: Roman.Rozanski@pwr.wroc.pl

Abstract

In the article, we propose a new estimator of the hazard rate function in the framework of the multiplicative point process intensity model. The technique combines the reflection method and the method of transformation. The new method eliminates the boundary effect for suitably selected transformations reducing the bias at the boundary and keeping the asymptotics of the variance. The transformation depends on a pre-estimate of the logarithmic derivative of the hazard function at the boundary.

Keywords: hazard rate function, multiplicative intensity point process model, Ramlau-Hansen kernel estimator, reduction of the bias, reflection, transformation.

2000 Mathematics Subject Classification: 62G05 , 62N02 , 62M99.

1. Introduction

An investigation into the intensity of the occurrence of phenomena observed as some processes is the most common subject of interest in the theory of point processes. In the most simple cases, which occur in the reliability theory or survival analysis, the object observed can take only one of two states, defined as "working" and "broken" (alive, dead). The transition between these states can be thought of as an event of some point process.

In such cases the intensity of the events "death" is given by the hazard function for the distribution of the survival time. Let us note that the model described above is a special case of the so called multiplicative intensity model [1] which has played a key role in the theory of point processes. The model assumes that the intensity of a given point process $N_n(t)$ is the product of a deterministic function $\alpha(t)$ and an observable factor $Y_n(t)$:

$$\lambda(t) = \alpha(t)Y_n(t).$$

The deterministic part of this model (for a two state model), $\alpha(t)$ is interpreted as the hazard function, whereas $Y_n(t)$ denotes the total number of objects among n, which are at risk at time t. When a collection of individuals is observed it is often impossible to wait for the event to happen for all the objects observed - it is only known that the event had not yet happened at some specified time and in this case the observation of the time to the occurrence of the event is censored. A well known estimator of the function α in this multiplicative model is the Ramlau-Hansen (R-H) estimator [4]. However, this estimator gives poor estimates at the end points of the domain of the hazard function. This phenomenon is analogous to the boundary effect observed when estimating a density function.

One of the methods proposed in order to eliminate this undesirable effect, occurring in both the estimation of a density function, as well as of a hazard rate function, is the method of transformation, see, e.g., [5], [6], [9], [10], [3]. The transformed estimator of the hazard function (the kernel-diffeomorphic estimator) proposed in [3] leads to a bias of order $O(h^2)$, whereas for the R-H estimator it is of order $O(h)$, where h is the bandwidth parameter. This result is true for any diffeomorphic transformation φ. The problem of estimation of the hazard rate function with a uniform accuracy in the whole domain seems to be important due to possible applications in biostatistic, demographic, epidemiologic and survival analysis.

In this article, we construct a new estimator of the hazard rate function, which combines the reflection method (used in the estimation of the density function [7]) and the method of transformation. Under the following conditions on a monotonically increasing transformation φ

- $\varphi(0) = 0$
- $\varphi^{(1)}(0) = 1$

we obtain estimates for the bias and asymptotic variance of order $O(h^2)$ and $O(\frac{1}{nh})$ respectively. The choice of the transformation φ is an important element of the new estimator. Since the conditions on the transformation φ depend on the unknown function α we have presented a method of estimating such a transformation function φ. Due to the proposed method of estimation of the transformation function φ, it is important for the hazard function we estimate to satisfy the condition $\alpha > 0$.

The plan of the paper is as follows. Section 2 contains preliminary results concerned with the multiplicative point process model. In Section 3, we introduce a new estimator of the hazard rate function. Moreover, we describe conditions on the transformation and present a method for its estimation. We also formulate a lemma where the formula defining the bias and variance of the introduced estimator is given. Furthermore, we formulate theorems on asymptotic properties of the estimator. Section 4 is devoted to the presentation of some simulation results involving the estimator considered. Section 5 contains proofs of all theorems and statements from previous sections.

2. The description of the model

In this section, we describe the multiplicative intensity point process model introduced by Aalen (1978). Let $(\Omega, P, \mathcal{F})$ be a probability space on which a sequence of point processes $\{N_n(t), n \in N\}$ is defined. We assume that the processes are adapted to their filtrations $\{\mathcal{F}_{t,n}; n \in N, t \geq 0\}$. We consider such models for which a stochastic intensity $\lambda(t)$ of the process $N_n(t)$ exists and can be defined in the following way

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

The point process $N_n(t)$, belongs to the class of multiplicative intensity model if the intensity function $\lambda(t)$ has the following form

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

where $Y_n(t)$ is an observable nonnegative process, left continuous with right-hand limits (or more generally predictable), and $\alpha(t)$ is the unknown nonnegative deterministic function to be estimated.

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

Suppose, we observe 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 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

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

where

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

D_i − is the indicator of time of death being observed $D_i = 1\{\tilde{T}_i = T_i\}$

and

$$\mathcal{F}_{i,t} = \{\tilde{T}_i < s, s \leq t, D_i\}.$$

Further, we have

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

with

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

Assuming the independence of the risk of a failure (death) and censoring process, (see Anderson, Gill [2]) and that the subjects observed represent a homogeneous population, we can define a sequence of one-dimensional point processes $N_n(t)$ whose value at time t represents the number of failures (deaths) in the interval $[0, t]$. The intensity of the process is given by $\lambda_n(t) = \alpha(t)Y_n(t)$, where $Y_n(t)$ represents the number of elements still functioning and being under observation up to time t. We can write

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

where $M_n(t)$ is a martingale with respect to the filtration $\{\mathcal{F}_{n,t}\} = \sigma(\mathcal{F}_{i,t}; i = 1, 2, \ldots, n)$. To estimate the hazard function $\alpha(t)$, $t \in [0, T]$, one can use the following estimator

$$\hat{\alpha}_n(x) = \frac{1}{h}\int_0^T K\left(\frac{x-s}{h}\right)\frac{J_n(s)}{Y_n(s)}dN_n(s), \tag{2.1}$$

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

In addition:

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

(2.3) h – bandwidth (a positive parameter), $h \to 0$ as $n \to \infty$.

The form of this estimator was derived by Ramlau-Hansen (see [4]). It is an asymptotically unbiased, consistent and asymptotically normal estimator of the hazard rate function (see [4]). However, the results of simulations carried out using this estimator are unsatisfactory at the end points of the time interval on which the process $N_n(t)$ is observed. To correct this so called boundary effect a kernel-diffeomorphic estimator has been defined in [3] which gives improved results at the boundary. It is of the following form

$$\begin{aligned}&\hat{\alpha}_{n,\varphi}(x) = \\ &= \frac{1}{h}\int_a^b K\left(\frac{\varphi(x)-s}{h}\right)\frac{\varphi^{(1)}(\varphi^{-1}(s))}{Y_n(\varphi^{-1}(s))}J_n(\varphi^{-1}(s))dN_n(\varphi^{-1}(s)),\end{aligned} \tag{2.4}$$

where φ is a diffeomorphism from $(0, c)$ on (a, b), and a, b, c are allowed to be infinite. The estimator has been proved to be asymptotically unbiased and asymptotically normal. Moreover, for any diffeomorphic transformation φ the order of bias is $O(h^2)$ for all $x \in [0, T]$ and also for $x = ch$, $1 \geq c \geq 0$. Despite of satisfactory properties of the bias, this estimator has a somewhat larger variance than that of Ramlau-Hansen.

3. Estimation of the hazard function with transformation and reflection of data. Properties of the estimator

Let T_i be the points of discontinuity of the point process N_n, which is observed on the interval $[0, \infty)$. Moreover, let $\varphi : [0, \infty) \to [0, \infty)$ be a nonnegative, continuous, and monotonically increasing function. We transform the points of discontinuity of the process, T_i, to $\varphi(T_i)$, and then reflect them around the point $t = 0$. Hence, we obtain $-\varphi(T_i)$. Using these new pseudo-data we can write

$$\hat{\alpha}_{ref}(x) = \frac{1}{h}\sum_{T_i} K\left(\frac{x-(-\varphi(T_i))}{h}\right)\frac{1}{Y_n(T_i)}$$

$$= \frac{1}{h}\int_0^\infty K\left(\frac{x+\varphi(s)}{h}\right)\frac{1}{Y_n(s)}dN_n(s). \tag{3.5}$$

Combining the formula for the R-H (2.1) estimator and making use of the pseudo-data (3.5) we can define a new estimator of the hazard rate function as follows.

Definition 3.1 . The reflected after transforming estimator $\tilde{\alpha}_n$ of the hazard rate function α is defined by the following formula

$$\tilde{\alpha}_n(x) = \frac{1}{h}\int_0^\infty \left[K\left(\frac{x-s}{h}\right) + K\left(\frac{x+\varphi(s)}{h}\right)\right]\frac{J_n(s)}{Y_n(s)}dN_n(s), \tag{3.6}$$

where $J_n(s) = 1\{Y_n(s) > 0\}$. If $Y_n(s) = 0$ we define $J_n(s)/Y_n(s) = 0$. The kernel K and parameter h satisfy conditions (2.2), (2.3), respectively.

It can be easily shown that the form of this estimator reduces to the R-H estimator for $x > h$. This fact follows from the boundedness of the support of the kernel K. For $x > h$ the above estimator has properties analogous to those of the R-H estimator. Thus, the bias is of order $O(h^2)$. In order to describe the behaviour of the new estimator at the boundary we need the explicit form of the bias expression of the estimator (3.6).

Throughout the paper, we assume that

$$\lim_{n\to\infty} \sup_{s\in[0,T]} E\left|\frac{nJ_n(s)}{Y_n(s)} - \frac{1}{y(s)}\right| = 0 \tag{3.7}$$

and

$$\lim_{n\to\infty} \sup_{s\in[0,T]} E\left|J_n(s) - 1\right| = \lim_{n\to\infty} \sup_{s\in[0,T]} P\left(Y_n(s) = 0\right) = o\left(\frac{1}{n}\right), \tag{3.8}$$

where $y(s)$ is a positive and continuous function and T is any positive number.

Remark 3.1 .The assumption 3.8 holds in a model of observing n objects in the presence of right censoring, see [2] for details.

3.1. Reduction of the bias

The transformation φ plays an important role in the definition of the estimator (3.6). The choice of the function determines the form and the order of the bias.

Lemma 3.1 . *Assume that $\alpha^{(2)}(\cdot)$, $\varphi^{(3)}(\cdot)$ exist and they are continuous. Furthermore, assume that $\varphi^{-1}(0) = 0$, and $\varphi^{(1)}(0) = 1$, where φ^{-1} is the inverse function of φ, and $\alpha^{(i)}$ and $\varphi^{(i)}$ are the* i*th derivatives of α and φ, $i \geq 0$ ($\alpha^{(0)} = \alpha, \varphi^{(0)} = \varphi$). Moreover, assume that condition* (3.8) *holds. Then for $x = ch$, $0 \leq c \leq 1$, we have*

$$
\begin{aligned}
E\tilde{\alpha}_n(x) - \alpha(x) &= h\left(2\alpha^{(1)}(0) - \varphi^{(2)}(0)\alpha(0)\right)\int_c^1 K(t)(t-c)dt \\
&+ \frac{h^2}{2}\alpha^{(2)}(0)\int_{-1}^1 K(t)t^2 dt \\
&- \frac{h^2}{2}\left[\varphi^{(3)}(0)\alpha(0) + 3\varphi^{(2)}(0)\left(\alpha^{(1)}(0) - \varphi^{(2)}(0)\alpha(0)\right)\right] \\
&\times \int_c^1 K(t)(t-c)^2 dt + o(h^2). \qquad (3.9)
\end{aligned}
$$

The proof of the lemma is given in Section 5. The primary goal of our transformation φ is to eliminate the first-order term in the bias expression (3.9). Assume that $\alpha(0) > 0$, then it suffices that

$$\varphi^{(2)}(0) = \frac{2\alpha^{(1)}(0)}{\alpha(0)}, \qquad (3.10)$$

to eliminate the first order term. Consequently, let us choose a transformation φ satisfying the following conditions:

(1) $\varphi(0) = 0$

(2) $\varphi^{(1)}(0) = 1$

(3) $\varphi^{(2)}(0) = \frac{2\alpha^{(1)}(0)}{\alpha(0)}$

(4) φ is monotonically increasing.

It is easy to find such transformations fulfilling the conditions above. For instance, let us choose the transformation φ of the form:

$$\varphi(x) = x + Dx^2 + AD^2x^3, \qquad (3.11)$$

where

$$D = \frac{\alpha^{(1)}(0)}{\alpha(0)}, \tag{3.12}$$

and $3A > 1$. This function obviously satisfies conditions (1)–(5), and the bias (3.9) takes the following form:

$$\begin{aligned} E\tilde{\alpha}_n(x) - \alpha(x) &= \frac{h^2}{2}\alpha^{(2)}(0)\int_{-1}^{1} K(t)t^2 dt \\ &- \frac{h^2}{2}6D(A-1)\alpha^{(1)}(0)\int_{c}^{1} K(t)(t-c)^2 dt + o(h^2). \end{aligned}$$

3.2. Asymptotic distribution of the estimator $\tilde{\alpha}_n$

To study the statistical properties of $\tilde{\alpha}_n$, it is convenient to introduce the quantities

$$\tilde{\alpha}_n^*(x) = \frac{1}{h}\int_0^\infty \left[K\left(\frac{x-s}{h}\right) + K\left(\frac{x+\varphi(s)}{h}\right)\right] J_n(s)\alpha(s)ds, \tag{3.13}$$

$$\alpha^0(x) = \frac{1}{h}\int_0^\infty \left[K\left(\frac{x-s}{h}\right) + K\left(\frac{x+\varphi(s)}{h}\right)\right] \alpha(s)ds. \tag{3.14}$$

The following theorem gives the form of the asymptotic distribution of $\tilde{\alpha}_n$.

Theorem 3.1. *Assume that φ and h satisfy conditions* (1)–(4), (2.3) *respectively, and $nh \to \infty$ when $n \to \infty$. Further, assume that the function α is twice continuously differentiable and conditions* (3.7), (3.8) *hold. Then for $x = ch$, $0 \le c \le 1$ we have*

$$\sqrt{nh}\,(\tilde{\alpha}_n(x) - \tilde{\alpha}_n^*(x)) \xrightarrow{D} N(0,V) \quad \text{as } n \to \infty, \tag{3.15}$$

where V is of the form

$$V = \frac{\alpha(0)}{y(0)} \int_{-1}^{1} K^2(t)dt + 2\frac{\alpha(0)}{y(0)} \int_{-1}^{c} K(t)K(2c-t)dt. \tag{3.16}$$

In addition to the theorem one may be interested in extra conditions which ensure that $\sqrt{nh}(\alpha^0(x) - \alpha(x))$ is asymptotically negligible. Applying the mean value theorem, this occurs if $nh^5 \to 0$ as $n \to \infty$. Moreover, we can write

Theorem 3.2 . *Assume that φ and h satisfy conditions* (1)–(4), (2.3) *respectively, and $nh \to \infty$, $nh^5 \to 0$ when $n \to \infty$. Further, assume that the function α is twice continuously differentiable and conditions* (3.7), (3.8) *hold. Then for $x = ch$, $0 \le c \le 1$ we have*

$$\sqrt{nh}\,(\tilde{\alpha}_n(x) - E\tilde{\alpha}_n(x)) \xrightarrow{D} N(0, V) \quad as \quad n \to \infty \tag{3.17}$$

and

$$\sqrt{nh}\,(\tilde{\alpha}_n(x) - \alpha(x)) \xrightarrow{D} N(0, V) \quad as \quad n \to \infty, \tag{3.18}$$

where V is of the form 3.16.

The above theorems give us some description of the asymptotic variance of the estimator $\tilde{\alpha}_n$ which is to be of the rate $\frac{V}{nh}$.

Remark 3.2 .Let us observe that for $x > h$, the estimator $\tilde{\alpha}(x)$ has the same asymptotic normal distribution as the Ramlau-Hansen estimator.

3.3. Estimation of the transformation φ

The transformation φ given by (3.11) is difficult to use in practice, because D defined by (3.12) is unknown (it depends on the unknown function α). Taking into account that

$$D = (d/dx)\, \log\, \alpha(x)|_{x=0} \tag{3.19}$$

one can estimate D by

$$D_n = \frac{\log \alpha_n(h) - \log \alpha_n(0)}{h}, \tag{3.20}$$

where $h \to 0$ as $n \to \infty$, and

$$\alpha_n(h) = \hat{\alpha}_n(h) + \frac{1}{n^2}, \tag{3.21}$$

where $\hat{\alpha}_n(h)$ is defined by (2.1), and

$$\alpha_n(0) = max\left(\hat{\alpha}_n(0), \frac{1}{n^2}\right) \tag{3.22}$$

with

$$\hat{\alpha}_n(0) = \frac{1}{h}\int_0^\infty K_{(0)}\left(\frac{-s}{h}\right)\frac{J_n(s)}{Y_n(s)}dN_n, \tag{3.23}$$

for $K_{(0)}$ satisfying

$$\int_{-1}^0 K_{(0)}(t)dt = 1, \qquad \int_{-1}^0 tK_{(0)}(t)dt = 0$$

and

$$\int_{-1}^0 t^2 K_{(0)}(t)dt \neq 0.$$

The factor $\frac{1}{n^2}$ in 3.21 and 3.22 is used to keep $\alpha_n(h)$ and $\alpha_n(0)$ bounded away from 0. Assuming that $h = O(n^{-1/5})$ we obtain the rate of convergence of $\hat{\alpha}_n$ to α and D_n to D for $x = 0, h$. Namely, we have

Lemma 3.2. *Let* $\alpha_n(h)$ *and* $\alpha_n(0)$ *be defined by* (3.21) *and* (3.22), *respectively. Suppose that* $\alpha^{(2)}(\cdot)$ *is continuous near* 0. *Then*

(3.24) $$|\alpha_n(x) - \alpha(x)| \overset{P}{\cong} O(h^2)$$

and

(3.25) $$E\left(\alpha_n(x) - \alpha(x)\right)^2 \cong O(h^4)$$

for $x = 0, h$.

Lemma 3.3. *Let* D_n *be defined by* (3.20). *Assume that* $\alpha(x) > 0$ *for* $x = 0, h$ *and that* $\alpha^{(2)}(\cdot)$ *is continuous near* $x = 0$. *Then*

(3.26) $$|D_n - D| \overset{P}{\cong} O(h)$$

and

(3.27) $$E\left(D_n - D\right)^2 \cong O(h^2).$$

Thus we define

(3.28) $$\varphi_n(x) = x + D_n x^2 + A D_n^2 x^3$$

as our estimator of $\varphi(x)$.

3.4. The proposed estimator

Based on the estimator φ_n defined in (3.28), we propose a new estimator of the form presented in (3.6). It is defined as follows

(3.29) $$\hat{\alpha}_{new}(x) = \frac{1}{h}\int_0^\infty \left[K\left(\frac{x-s}{h}\right) + K\left(\frac{x+\varphi_n(s)}{h}\right)\right]\frac{J_n(s)}{Y_n(s)}dN_n(s).$$

Since the estimator has the same form as the estimator (3.6), it can be shown that for $x > h$ it reduces to the R-H (2.1) estimator. Thus (3.29) is a natural extension of the R-H estimator. The properties of the bias and asymptotic distribution of the estimator (3.29) are given in the following theorems.

Theorem 3.3. *Assume that* $\alpha(x) > 0$ *for* $x = 0, h$, *and* $\alpha^{(2)}(x)$ *is continuous in a neighbourhood of* 0. *Moreover, assume that conditions* (3.7) *and* (3.8) *hold. Then for* $x = ch$, $0 \le c \le 1$, *we have*

$$E\hat{\alpha}_{new}(x) - \alpha(x) = E\tilde{\alpha}_n(x) - \alpha(x) + O(h^2), \tag{3.30}$$

where $\hat{\alpha}_{new}$ *and* $\tilde{\alpha}_n$ *are given by* (3.29) *and* (3.6)*, respectively.*

Let $\hat{\alpha}^*_{new}$ be defined as follows

$$\hat{\alpha}^*_{new}(x) = \frac{1}{h}\int_0^\infty \left[K\left(\frac{x-s}{h}\right) + K\left(\frac{x+\varphi_n(s)}{h}\right)\right]\alpha(s)J_n(s)ds$$

and α^0 be defined by 3.14. We now turn to a study of the asymptotic distribution of the proposed estimator $\hat{\alpha}_{new}$.

Theorem 3.4. *Assume that* φ *and* h *satisfy conditions* (1)–(4) *and* 2.3 *respectively, and* $nh^5 \to 0$ *when* $n \to \infty$*. Further, assume that the function* α *is twice continuously differentiable and conditions* 3.7, 3.8 *hold. Then for* $x = ch$, $0 \le c \le 1$ *we have*

$$\sqrt{nh}\left(\hat{\alpha}_{new}(x) - \hat{\alpha}^*_{new}(x)\right) \xrightarrow{D} N(0,V) \quad as \quad n \to \infty, \tag{3.31}$$

$$\sqrt{nh}\left(\hat{\alpha}_{new}(x) - \alpha^0(x)\right) \xrightarrow{D} N(0,V) \quad as \quad n \to \infty. \tag{3.32}$$

Moreover

$$\sqrt{nh}\left(\hat{\alpha}_{new}(x) - \alpha(x)\right) \xrightarrow{D} N(0,V) \quad as \quad n \to \infty \tag{3.33}$$

and

$$\sqrt{nh}\left(\hat{\alpha}_{new}(x) - E\hat{\alpha}_{new}(x)\right) \xrightarrow{D} N(0,V) \quad as \quad n \to \infty, \tag{3.34}$$

where V *is of the form* 3.16.

Proofs of these theorems are given in Section 5.

4. Simulation results

We assume that failure times $T_1, \ldots, T_n$, where $n = 1000$ are observed on the interval $(0, \infty)$ and data are generated from

(i) the exponential distribution with $\lambda = 5$, here the hazard rate is fixed $\alpha(x) = 5$,

(ii) the Gompertz distribution with the hazard function of the form $\alpha(x) = \theta \exp(\gamma x)$, where $\theta = 1$ and $\gamma = 1$.

In simulations, we used the following kernels

$$K(x) = \frac{3}{4}(1 - x^2)1\{[-1, 1]\}$$

and

$$K_0(x) = 12(1 + t)\left(t + \frac{1}{2}\right)1\{[-1, 0]\}.$$

The value of A used in the simulations was 0.55, which gave relatively good results. In all cases, we set $h = 0,24$, which is of order $n^{-1/5}$. We have to keep in mind that the estmator $\hat{\alpha}_{nev}$ differs from the R-H estimator at points $x = ch$, where $c \in (0, 1)$; whereas for points $x > h$ these esimators are equal. We estimate the hazard rate function on the interval $(0, 1)$ and we do not have complete observations of the lifetime. We assume that $T_1, \ldots, T_n$ are censored by n independent censoring time $U_1, \ldots, U_n$ generated from the exponential distribution with the mean λ. The parameter $\lambda = 3$ was chosen for the life time T_i from the exponential distribution and $\lambda = 1.5$ for the life time T_i from the Gompertz distribution. 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 $N_n(t) = \sum_{i=1}^{n} 1\{\tilde{T}_i < t\}$. This process counts the total number of failures in the interval $(0, t)$, and the intensity function is given by $\lambda(t) = \alpha(t)Y_n(t)$, where $Y_n = \sum_{i=1}^{n} 1\{\tilde{T} \geq t\}$. The process $Y_n(t)$ counts the number of elements at risk of failure just to time t. Figures 1 and 2 present realizations of the estimators $\hat{\alpha}_{new}$ and $\hat{\alpha}_n$ obtained from the censored data. We observe a significantly better behavior of the estimator $\hat{\alpha}_{new}$ at the boundary (in the neighborhood of zero). We draw analogous conclusions from Figure 3 where the bias of both estimators is presented.

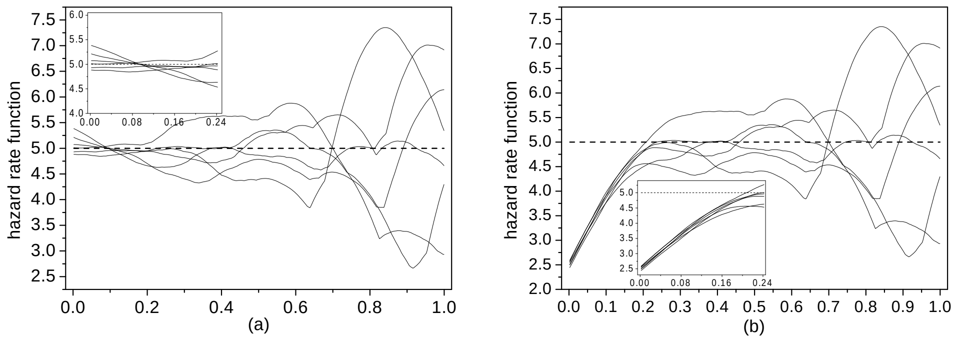

Figure 1. Data from exponential distribution:

(a) The estimator $\hat{\alpha}_{new}$ (solid line) and true hazard rate function (dashed line).

(b) The Ramlau-Hansen estimator (solid line) and true hazard rate function (dashed line).

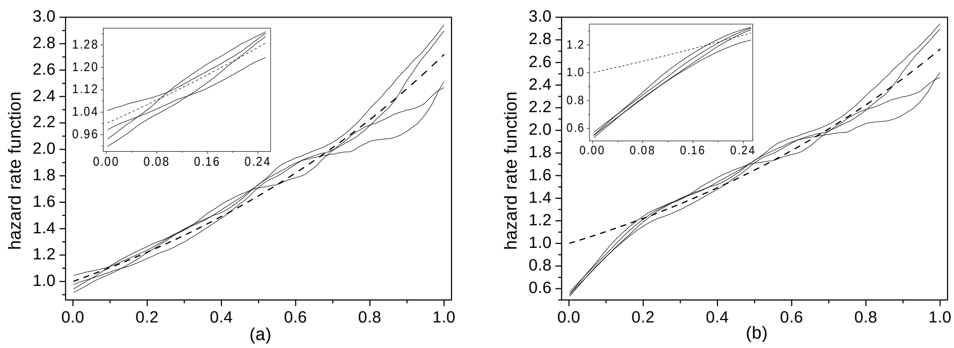

Figure 2. Data from Gompertz distribution

(a) The estimator $\hat{\alpha}_{new}$ (solid line) and true hazard rate function (dashed line).

(b) The Ramlau-Hansen estimator (solid line) and true hazard rate function (dashed line).

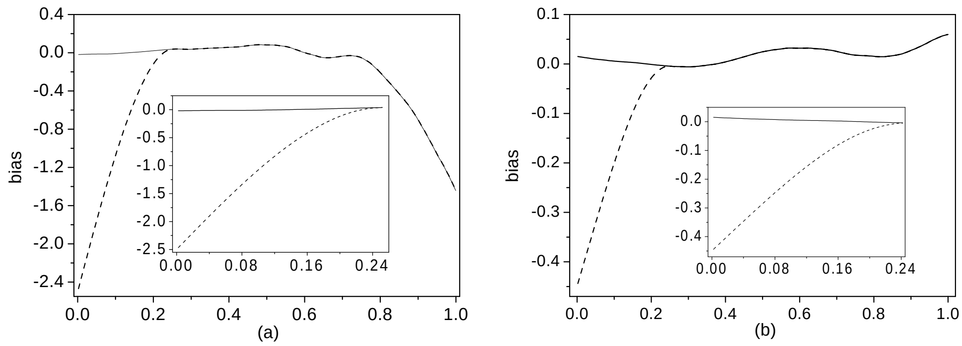

Figure 3. The bias of the estimator $\hat{\alpha}_{new}$ (solid line) and the Ramlau-Hansen estimator (dashed line):

(a) data from exponential distribution

(b) data from Gompertz distribution.

5. Proofs

***Proof of Lemma* 3.1.**
We assume that $x = ch$, $0 \leq c \leq 1$ and the kernel function K has a bounded support (i.e., [-1,1])

$$
\begin{aligned}
E\tilde{\alpha}_n(x) &= E\left[\frac{1}{h}\int_0^\infty \left[K\left(\frac{x-s}{h}\right) + K\left(\frac{x+\varphi(s)}{h}\right)\right]\frac{J_n(s)}{Y_n(s)}dN_n(s)\right] \\
&= \frac{1}{h}\int_0^\infty \left[K\left(\frac{x-s}{h}\right) + K\left(\frac{x+\varphi(s)}{h}\right)\right]\alpha(s)EJ_n(s)ds \\
&\cong \frac{1}{h}\int_0^\infty K\left(\frac{x-s}{h}\right)\alpha(s)ds + \frac{1}{h}\int_0^\infty K\left(\frac{x+\varphi(s)}{h}\right)\alpha(s)ds \\
&= I_1 + I_2,
\end{aligned}
$$

where

$$I_1 = \frac{1}{h}\int_0^{x+h} K\left(\frac{x-s}{h}\right)\alpha(s)ds = \int_{-1}^{x/h} K(t)\alpha(x-ht)$$

$$= \int_{-1}^{c} K(t)\left[\alpha(x) - \alpha^{(1)}(x)ht + \alpha^{(2)}(x)\frac{h^2t^2}{2} + o(h^2)\right]dt \quad (5.1)$$

and

$$I_2 = \frac{1}{h}\int_0^{\varphi^{-1}(h-x)} K\left(\frac{x+\varphi(s)}{h}\right)\alpha(s)ds$$

$$= \int_{x/h}^{1} K(t)\frac{\alpha(\varphi^{-1}(ht-x))}{\varphi^{(1)}(\varphi^{-1}(ht-x))}dt = \int_c^1 K(t)\frac{\alpha(\varphi^{-1}(h(t-c)))}{\varphi^{(1)}(\varphi^{-1}(h(t-c)))}dt$$

$$= \int_c^1 K(t)\left[\frac{\alpha(\varphi^{-1}(0))}{\varphi^{(1)}(\varphi^{-1}(0))} + (t-c)h\left(\frac{\alpha(\varphi^{-1}(\cdot))}{\varphi^{(1)}(\varphi^{-1}(\cdot))}\right)'_{(\cdot)=0}\right.$$

$$\left. + \frac{(t-c)^2h^2}{2}\left(\frac{\alpha(\varphi^{-1}(\cdot))}{\varphi^{(1)}(\varphi^{-1}(\cdot))}\right)''_{(\cdot)=0}\right]dt + o(h^2)$$

$$= \int_c^1 K(t)\left[\alpha(0) + (t-c)h\left(\alpha^{(1)}(0) - \varphi^{(2)}(0)\alpha(0)\right) + \frac{(t-c)^2h^2}{2}\right. \quad (5.2)$$

$$\left. \times \left(\alpha^{(2)}(0) - \varphi^{(3)}(0)\alpha(0) - 3\varphi^{(2)}(0)\left[\alpha^{(1)}(0) - \varphi^{(2)}(0)\alpha(0)\right]\right)\right]$$

$$\times\ dt + o(h^2).$$

Summing (5.1) and (5.2) we obtain

$$\begin{aligned} E\tilde{\alpha}_n(x) &= \alpha(x)\int_{-1}^{c} K(t)dt - \alpha^{(1)}(x)h\int_{-1}^{c} K(t)tdt + \alpha^{(2)}(x)\frac{h^2}{2}\int_{-1}^{c} K(t)t^2dt \\ &+ \alpha(0)\int_{c}^{1} K(t)dt + \left(\alpha^{(1)}(0) - \varphi^{(2)}(0)\alpha(0)\right) h\int_{c}^{1} K(t)(t-c)dt \\ &+ \left[\alpha^{(2)}(0) - \varphi^{(3)}(0)\alpha(0) - 3\varphi^{(2)}(0)\left(\alpha^{(1)}(0) - \varphi^{(2)}(0)\alpha(0)\right)\right] \\ &\times \frac{h^2}{2}\int_{c}^{1} K(t)(t-c)^2dt + o(h^2). \qquad (5.3) \end{aligned}$$

By the assumed continuity of $\alpha^{(2)}(\cdot)$ near 0, we have that for $x = ch$

(a) $\alpha(0) = \alpha(x) - \alpha^{(1)}(x)ch + \frac{1}{2}\alpha^{(2)}(x)c^2h^2 + o(h^2)$,

(b) $\alpha^{(1)}(x) = \alpha^{(1)}(0) + ch\alpha^{(2)}(0) + o(h)$,

(c) $\alpha^{(2)}(x) = \alpha^{(2)}(0) + o(1)$.

Substituting (a)–(c) into (5.3), we obtain

$$\begin{aligned} E\tilde{\alpha}_n(x) &= \alpha(x) + h\left(2\alpha^{(1)}(0) - \varphi^{(2)}(0)\alpha(0)\right)\int_{c}^{1} K(t)(t-c)dt \\ &+ \frac{h^2}{2}\alpha^{(2)}(0)\int_{-1}^{1} K(t)t^2dt - \frac{h^2}{2} \\ &\times \left(\varphi^{(3)}(0)\alpha(0) + 3\varphi^{(2)}(0)\left[\alpha^{(1)}(0) - \varphi^{(2)}(0)\alpha(0)\right]\right) \\ &\times \int_{c}^{1} K(t)(t-c)^2dt + o(h^2) \end{aligned}$$

which completes the proof of Lemma 3.1.

***Proof of Theorem* 3.1.**

$$\sqrt{nh}\left(\tilde{\alpha}_n(x) - \tilde{\alpha}_n^*(x)\right) =$$

$$= \sqrt{nh}\left(\frac{1}{h}\int_0^\infty \left[K\left(\frac{x-s}{h}\right) + K\left(\frac{x+\varphi(s)}{h}\right)\right]\frac{J_n(s)}{Y_n(s)}dN_n(s)\right.$$

$$\left. - \frac{1}{h}\int_0^\infty \left[K\left(\frac{x-s}{h}\right) + K\left(\frac{x+\varphi(s)}{h}\right)\right] J_n(s)\alpha(s)ds\right)$$

$$= \sqrt{\frac{n}{h}}\int_0^\infty \left[K\left(\frac{x-s}{h}\right) + K\left(\frac{x+\varphi(s)}{h}\right)\right]\frac{J_n(s)}{Y_n(s)}dM_n(s)ds.$$

Let us consider a sequence of predictable processes

$$H_n(s) = \sqrt{\frac{n}{h}}\left[K\left(\frac{x-s}{h}\right) + K\left(\frac{x+\varphi(s)}{h}\right)\right]\frac{J_n(s)}{Y_n(s)}$$

with a fixed x and h at the moment, and let $\{\tilde{M}_n\}$ be a sequence of martingales as follows

$$\tilde{M}_n(z) = \int_0^z H_n(s)dM_n(s) \qquad z \in (0, \infty).$$

We show that

$$\langle \tilde{M}_n(z)\rangle \xrightarrow{P} V(z),$$

where V is a non-decreasing and a non-negative function and $\langle\cdot\rangle$ denotes the quadratic variation.

$$
\begin{aligned}
\langle \tilde{M}_n(z) \rangle &= \int_0^z H_n^2(s) d\langle M_n(s) \rangle \\
&= \frac{n}{h} \int_0^z \left[K\left(\frac{x-s}{h}\right) + K\left(\frac{x+\varphi(s)}{h}\right) \right]^2 \frac{J_n(s)}{Y_n(s)} \alpha(s) ds \\
&\overset{P}{\cong} \frac{1}{h} \int_0^z \left[K^2\left(\frac{x-s}{h}\right) + K^2\left(\frac{x+\varphi(s)}{h}\right) \right] \frac{\alpha(s)}{y(s)} ds \\
&+ \frac{2}{h} \int_0^z K\left(\frac{x-s}{h}\right) K\left(\frac{x+\varphi(s)}{h}\right) \frac{\alpha(s)}{y(s)} ds \\
&= I_{11} + I_{22}.
\end{aligned}
$$

Observe that

$$
\begin{aligned}
I_{11} &= \frac{1}{h} \int_0^{min(x+h,z)} K^2\left(\frac{x-s}{h}\right) \frac{\alpha(s)}{y(s)} ds \\
&+ \frac{1}{h} \int_0^{min(\varphi^{-1}(h-x),z)} K^2\left(\frac{x+\varphi(s)}{h}\right) \frac{\alpha(s)}{y(s)} ds \\
&= \int_{max(-1,c-\frac{z}{h})}^{c} K^2(t) \frac{\alpha(h(c-t))}{y(h(c-t))} dt \\
&+ \int_c^{min\left(1,c+\frac{\varphi(z)}{h}\right)} K^2(t) \frac{\alpha(\varphi^{-1}(h(t-c)))}{y(\varphi^{-1}(h(t-c)))} \frac{1}{\varphi^{(1)}(\varphi^{-1}(h(t-c)))} dt \\
&\cong \int_{max(-1,c-\frac{z}{h})}^{min\left(1,c+\frac{\varphi(z)}{h}\right)} K^2(t) \frac{\alpha(0)+o(1)}{y(0)} dt \qquad (5.4)
\end{aligned}
$$

and

$$
\begin{aligned}
I_{22} &= \frac{2}{h}\int_0^{min(x+h,z)} K\left(\frac{x-s}{h}\right) K\left(\frac{x+\varphi(s)}{h}\right)\frac{\alpha(s)}{y(s)}ds \\
&= 2\int_{max(-1,c-\frac{z}{h})}^{c} K(t)K\left(\frac{x+\varphi(h(c-t))}{h}\right)\frac{\alpha(h(c-t))}{y(h(c-t))}dt \\
&= 2\int_{max(-1,c-\frac{z}{h})}^{c} K(t)K\left(\frac{x+\varphi(0)+\varphi^{(1)}(0)h(c-t)+o(h^2)}{h}\right) \\
&\times \frac{\alpha(0)+o(1)}{y(h(c-t))}dt \cong 2\int_{max(-1,c-\frac{z}{h})}^{c} K(t)K(2c-t+o(h)) \\
&\times \frac{\alpha(0)+o(1)}{y(0)}dt, \qquad (5.5)
\end{aligned}
$$

where we make use of facts

(a) $\frac{\alpha(\varphi^{-1}(h(t-c)))}{\varphi^{(1)}(\varphi^{-1}(h(t-c)))} = \frac{\alpha(0)}{\varphi^{(1)}(0)} + [\alpha^{(1)}(0) - \varphi^{(2)}(0)\alpha(0)]h(t-c) + o(h) = \frac{\alpha(0)}{\varphi^{(1)}(0)} + o(1)$,

(b) $\alpha(h(c-t)) = \alpha(0) + o(1)$

and conditions (1)–(4). Thus we get

$$
\begin{aligned}
\langle \tilde{M}_n(z)\rangle &\overset{P}{\cong} \int_{max(-1,c-\frac{z}{h})}^{min\left(1,c+\frac{\varphi(z)}{h}\right)} K^2(t)\frac{\alpha(0)+o(1)}{y(0)}dt \\
&+ \ 2\int_{max(-1,c-\frac{z}{h})}^{c} K(t)K(2c-t+o(h))\frac{\alpha(0)+o(1)}{y(0)}.
\end{aligned}
$$

Further, let $\tilde{M}_{n,\epsilon}$ be of the form

$$\tilde{M}_{n,\epsilon}(z) = \int_0^z H_n(s)1\{|H_n(s)| > \epsilon\}dM_n(s).$$

Then

$$\begin{aligned}
&\langle \tilde{M}_{n,\epsilon}(z)\rangle = \\
&= \int_0^z H_n^2(s)1\{|H_n(s)| > \epsilon\}d\langle M_n(s)\rangle \\
&= \frac{n}{h}\int_0^z \left[K\left(\frac{x-s}{h}\right) + K\left(\frac{x+\varphi(s)}{h}\right)\right]^2 \frac{J_n(s)}{Y_n(s)}\alpha(s)1\{|H_n(s)| > \epsilon\}ds \\
&\stackrel{P}{\cong} \frac{1}{h}\int_0^{min(x+h,z)} \left[K\left(\frac{x-s}{h}\right) + K\left(\frac{x+\varphi(s)}{h}\right)\right]^2 \frac{\alpha(s)}{y(s)}1\{|H_n(s)| > \epsilon\}ds \\
&= \int_{max(-1,c-\frac{z}{h})}^c \left[K(t) + K\left(\frac{ch+\varphi(h(c-t))}{h}\right)\right]^2 \\
&\times \frac{\alpha(h(c-t))}{y(h(c-t))}1\{|H_n(x-th)| > \epsilon\}dt \stackrel{P}{\longrightarrow} 0
\end{aligned}$$

because

$$\begin{aligned}
&1\{|H_n(x-ht)| > \epsilon\} = \\
&= 1\left\{\left|n\left[K(t) + K\left(\frac{x+\varphi(x-ht)}{h}\right)\right]\frac{J_n(x-ht)}{Y_n(x-ht)}\right| > \epsilon\sqrt{nh}\right\} \\
&\stackrel{P}{\cong} 1\left\{\frac{K(t)}{y(0)} > \epsilon\sqrt{nh} - \frac{K\left(c+\frac{\varphi(h(c-t))}{h}\right)}{y(0)}\right\} \longrightarrow 0
\end{aligned}$$

uniformly in probability. Finally, all the assumptions of Rebolledo's theorem [2] are satisfied and thus

$$\tilde{M}_n(z) \xrightarrow{D} G(z),$$

where G is a Gaussian martingale with variance V. Taking into account that the kernel has a support in $[-1, 1]$ and $h = O(n^{-1/5})$ we obtain that

$$\begin{aligned} \langle \tilde{M}_n(z) \rangle \quad & \xrightarrow{n\to\infty} \quad \frac{\alpha(0)}{y(0)} \int_{-1}^{1} K^2(t)dt + 2\frac{\alpha(0)}{y(0)} \int_{-1}^{c} K(t)K(2c-t)dt \\ & = \quad V, \end{aligned} \tag{5.6}$$

uniformly in probability. This completes the proof of Theorem 3.1.

***Proof of Theorem* 3.2.**
First, we observe that $\sqrt{nh}\left(\alpha^0(x) - E\tilde{\alpha}_n(x)\right) \to 0$ by the assumption (3.8). Moreover,

$$\sqrt{nh}(\alpha^0(x) - \alpha(x)) \cong \sqrt{nh}\,(E\tilde{\alpha}_n(x) - \alpha(x)) = \sqrt{nh}O(h^2) = O(\sqrt{nh^5}) \to 0.$$

and

$$\begin{aligned} \sqrt{nh}\,(\tilde{\alpha}_n(x) - E\tilde{\alpha}_n(x)) &= \sqrt{nh}\,(\tilde{\alpha}_n(x) - \tilde{\alpha}_n^*(x)) + \sqrt{nh}\,(\tilde{\alpha}_n^*(x) - E\tilde{\alpha}_n(x)) \\ &= A_1 + A_2, \end{aligned}$$

and

$$\begin{aligned} \sqrt{nh}\,(\tilde{\alpha}_n(x) - \alpha(x)) &= \sqrt{nh}\,(\tilde{\alpha}_n(x) - \tilde{\alpha}_n^*(x)) + \sqrt{nh}\left(\tilde{\alpha}_n^*(x) - \alpha^0(x)\right) \\ &+ \sqrt{nh}\left(\alpha^0(x) - \alpha(x)\right) = A_1 + B_2 + B_3, \end{aligned}$$

where A_1 converges to $N(0, V)$ from Theorem 3.1, and it is easy to show that A_2, B_2 and B_3 converge to 0. This finishes the proof of 3.17 and 3.18.

***Proof of Lemma* 3.2.**
Without loss of generality we prove this lemma only for $x = h$. First, we show that $|\hat{\alpha}_n(x) - \alpha(x)| \overset{P}{\cong} O(h^2)$ and $E(\hat{\alpha}_n(x) - \alpha(x))^2 \cong O(h^4)$, where $\hat{\alpha}_n$ is defined by (2.1). Formulas (3.24) and (3.25) are a direct consequence of the facts mentioned above (respectively) and the assumption that $h = O(n^{-1/5})$. Namely, we have

$$\begin{aligned}
|\hat{\alpha}_n(x) - \alpha(x)| &= \left|\frac{1}{h}\int_0^\infty K\left(\frac{x-s}{h}\right)\frac{J_n(s)}{Y_n(s)}dN_n(s) - \alpha(x)\right| \\
&\leq \left|\frac{1}{h}\int_0^\infty K\left(\frac{x-s}{h}\right)J_n(s)\alpha(s)ds - \alpha(x)\right| \\
&+ \left|\frac{1}{h}\int_0^\infty K\left(\frac{x-s}{h}\right)\frac{J_n(s)}{Y_n(s)}dM_n(s)\right| \\
&= J_{11}(x) + J_{22}(x). \qquad (5.7)
\end{aligned}$$

It can be shown that

$$\begin{aligned}
J_{11} &\overset{P}{\cong} \left|\frac{1}{h}\int_0^{x+h} K\left(\frac{x-s}{h}\right)\alpha(s)ds - \alpha(x)\right| \\
&= \left|\int_{-1}^1 K(t)(\alpha(x-ht) - \alpha(x))dt\right| \\
&= \left|\int_{-1}^1 K(t)\left(\alpha(x) - \alpha^{(1)}(x)ht + \frac{1}{2}\alpha^{(2)}(x)h^2t^2 - \alpha(x) + o(h^2)\right)\right| \\
&= \left|-\alpha^{(1)}(x)h\int_{-1}^1 K(t)tdt + \frac{1}{2}\alpha^{(2)}(x)h^2\int_{-1}^1 K(t)t^2dt\right| + o(h^2) \\
&= \left|\frac{1}{2}\alpha^{(2)}(x)h^2\int_{-1}^1 K(t)t^2\right| + o(h^2) \\
&\overset{P}{=} O(h^2). \qquad (5.8)
\end{aligned}$$

Note that by Chebyshev inequality

$$J_{22} \overset{P}{\cong} O(h^2). \tag{5.9}$$

Moreover, by Vallee-Poussin theorem ([8], p.17) J_{22} is uniformly integrable. Substituting (5.8) and (5.9) into (5.7) we obtain

$$|\hat{\alpha}_n(x) - \alpha(x)| \overset{P}{\cong} O(h^2) \quad \text{for } x = h. \tag{5.10}$$

Now we show the second formula

$$\begin{aligned}
E\left(\hat{\alpha}_n(x) - \alpha(x)\right)^2 &= E\left(\frac{1}{h}\int_0^\infty K\left(\frac{x-s}{h}\right)\frac{J_n(s)}{Y_n(s)}dN_n(s) - \alpha(x)\right)^2 \\
&= E\left(\frac{1}{h}\int_0^{x+h} K\left(\frac{x-s}{h}\right)J_n(s)\alpha(s)ds - \alpha(x)\right)^2 \\
&+ E\left(\frac{1}{h}\int_0^{x+h} K\left(\frac{x-s}{h}\right)\frac{J_n(s)}{Y_n(s)}dM_n(s)\right)^2 \\
&+ 2\left(\frac{1}{h}\int_0^{x+h} K\left(\frac{x-s}{h}\right)J_n(s)\alpha(s)ds - \alpha(x)\right) \\
&\times \left(\frac{1}{h}\int_0^{x+h} K\left(\frac{x-s}{h}\right)\frac{J_n(s)}{Y_n(s)}dM_n(s)\right) \\
&= L_1 + L_2 + L_3.
\end{aligned} \tag{5.11}$$

It can be shown that

$$
\begin{aligned}
&L_1 = \\
&= \frac{1}{h^2} E\left(\int_0^{x+h}\int_0^{x+h} K\left(\frac{x-s}{h}\right) K\left(\frac{x-v}{h}\right)\alpha(s)\alpha(v)J_n(s)J_n(v)dsdv\right) \\
&+ \alpha^2(x) - \frac{2\alpha(x)}{h} E\int_0^{x+h} K\left(\frac{x-s}{h}\right) J_n(s)\alpha(s)ds \\
&\cong \frac{1}{h^2}\int_0^{x+h}\int_0^{x+h} K\left(\frac{x-s}{h}\right) K\left(\frac{x-v}{h}\right)\alpha(s)\alpha(v)dsdv \\
&+ \alpha^2(x) - \frac{2\alpha(x)}{h}\int_0^{x+h} K\left(\frac{x-s}{h}\right)\alpha(s)ds \\
&= \left(\int_{-1}^{1} K(t)\alpha(x-ht)dt\right)^2 + \alpha^2(x) - 2\alpha(x)\int_{-1}^{1} K(t)\alpha(x-ht)dt \\
&\cong \frac{\left[\alpha^{(2)}(x)\right]^2}{4} h^4 \left(\int_{-1}^{1} K(t)t^2 dt\right)^2 \cong O(h^4) \qquad (5.12)
\end{aligned}
$$

and

$$
\begin{aligned}
L_2 &= \frac{1}{h^2}\int_0^{x+h} K^2\left(\frac{x-s}{h}\right) E\frac{J_n(s)}{Y_n(s)}\alpha(s)ds \\
&\cong \frac{1}{h^2}\int_{-1}^{1} K^2\left(\frac{x-s}{h}\right)\frac{\alpha(s)}{ny(s)}ds \\
&\le \frac{C}{nh}\int_{-1}^{1} K^2(t)\alpha(x-ht)dt \cong O\left(\frac{1}{nh}\right) = O(h^4), \qquad (5.13)
\end{aligned}
$$

where we make use of the fact $h \cong n^{-1/5}$.

Moreover, by Schwartz's inequality we have

$$L_3 \ \leq \ 2\sqrt{L_1}\sqrt{L_2} \cong O(h^4). \tag{5.14}$$

Similarly, we can prove both above facts for $x = 0$. This finishes the proof of Lemma 3.2.

***Proof of Lemma* 3.3.**
Let D_n and D be defined by (3.20), (3.12), respectively, then

$$|D_n - D| = \left|\left(D_n - \frac{\log\alpha(h) - \log\alpha(0)}{h}\right) + \left(\frac{\log\alpha(h) - \log\alpha(0)}{h} - D\right)\right|$$

$$= |J_1 + J_2| \leq |J_1| + |J_2|. \tag{5.15}$$

By Taylor's expansion of $\log(\cdot)$, we have

$$|J_2| =$$

$$= \left|\frac{\log\alpha(0) + \frac{\alpha^{(1)}(0)}{\alpha(0)}h + \frac{1}{2}\frac{\alpha^{(2)}(0)\alpha(0) - (\alpha^{(1)}(0))^2}{\alpha^2(0)}h^2 - \log\alpha(0) + o(h^2)}{h} - \frac{\alpha^{(1)}(0)}{\alpha(0)}\right|$$

(5.16)

$$= \left|\frac{1}{2}h\frac{\alpha^{(2)}(0)\alpha(0) - (\alpha^{(1)}(0))^2}{\alpha^2(0)}\right| + o(h) = O(h).$$

Note that

$$\begin{aligned} |J_1| &= \left| \frac{\log \alpha_n(h) - \log \alpha_n(0)}{h} - \frac{\log \alpha(h) - \log \alpha(0)}{h} \right| \\ &= \left| \frac{\log \alpha_n(h) - \log \alpha(h)}{h} - \frac{\log \alpha_n(0) - \log \alpha(0)}{h} \right| \\ &\leq \frac{1}{h} \left(|\log \alpha_n(h) - \log \alpha(h)| + |\log \alpha_n(0) - \log \alpha(0)| \right). \end{aligned}$$

Applying Taylor's expansion of the function $\log(\cdot)$, for $x = 0, h$, we have

$$\begin{aligned} |\log \alpha_n(x) - \log \alpha(x)| &= \left| \frac{\alpha_n(x) - \alpha(x)}{\alpha(x) + \delta(\alpha_n(x) - \alpha(x))} \right| \\ &= \left| \frac{\alpha_n(x) - \alpha(x)}{\alpha(x)(1-\delta) + \delta \alpha_n(x)} \right| \\ &\leq \frac{1}{\alpha(x)(1-\delta)} |\alpha_n(x) - \alpha(x)| \\ &\stackrel{P}{=} O(h^2), \end{aligned}$$

where $0 \leq \delta \leq 1$ is a constant. The last equality follows from Lemma 3.2. Therefore,

$$J_1 \stackrel{P}{=} O(h). \tag{5.17}$$

Combining (5.15), (5.17) and (5.16) we conclude the proof of formula 3.26, 3.27 can be proved similarly.

***Proof of Theorem* 3.3.**
We have

$$E\hat{\alpha}_{new}(x)-\alpha(x) = (E\hat{\alpha}_{new}(x)-E\tilde{\alpha}_n(x))+(E\tilde{\alpha}_n(x)-\alpha(x))$$

$$= K+(E\tilde{\alpha}_n(x)-\alpha(x)),$$

where

$$|K| = \left|E\left(\frac{1}{h}\int_0^{\infty}\left[K\left(\frac{x-s}{h}\right)+K\left(\frac{x+\varphi_n(s)}{h}\right)\right]\frac{J_n(s)}{Y_n(s)}dN_n(s)\right.\right.$$

$$- \left.\left.\frac{1}{h}\int_0^{\infty}\left[K\left(\frac{x-s}{h}\right)+K\left(\frac{x+\varphi(s)}{h}\right)\right]\frac{J_n(s)}{Y_n(s)}dN_n(s)\right)\right|$$

$$= \left|E\left(\frac{1}{h}\int_0^{x+h}\left[K\left(\frac{x+\varphi_n(s)}{h}\right)-K\left(\frac{x+\varphi(s)}{h}\right)\right]\right.\right.$$

$$\times \left.\left.1\left\{s\in[0,ph],p>1\right\}\frac{J_n(s)}{Y_n(s)}dN_n(s)\right)\right|.$$

Applying Taylor's expansion, we obtain

$$|K| \cong \left|E\left(\frac{1}{h}\int_0^{x+h}K^{(1)}\left(\frac{x+\varphi(s)}{h}+\frac{\varphi_n(s)-\varphi(s)}{h}\delta\right)\frac{\varphi_n(s)-\varphi(s)}{h}\right.\right.$$

$$\times \left.\left.1\left\{s\in[0,ph],p>1\right\}\frac{J_n(s)}{Y_n(s)}dN_n(s)\right)\right|.$$

$$\leq CE\left|\frac{1}{h^2}\int_0^{x+h}(\varphi_n(s)-\varphi(s))\,1\left\{s\in[0,ph],p>1\right\}\frac{J_n(s)}{Y_n(s)}dN_n(s)\right|$$

$$\leq \frac{C}{h^2}E\int_0^{x+h}|\varphi_n(s)-\varphi(s)|\,1\left\{s\in[0,ph],p>1\right\}\frac{J_n(s)}{Y_n(s)}dN_n(s),$$

where $0 \leq \delta \leq 1$ and C are constants. It is easy to see that $\varphi(x) \geq h$ for $x \geq ph$, where $p = \frac{4A}{4A-1}$ and by the assumption $A > 1/3$. Hence for $x \in (0, ph)$ we have

$$\begin{aligned}
|\varphi_n(x) - \varphi(x)| &= \left|x^2(D_n - D) + Ax^3(D_n^2 - D^2)\right| \\
&\leq \left|p^2h^2(D_n - D)\right| + \left|Ap^3h^3(D_n^2 - D^2)\right|.
\end{aligned}$$

So, we can write

$$\begin{aligned}
|K| &\leq \frac{C}{h^2} E\left[\int_0^{x+h} p^2h^2|D_n - D| 1\{s \in (0, ph)\} \frac{J_n(s)}{Y_n(s)} dN_n(s)\right. \\
&+ \left.\int_0^{x+h} Ap^3h^3|D_n^2 - D^2| 1\{s \in (0, ph)\} \frac{J_n(s)}{Y_n(s)} dN_n(s)\right] \\
&= \frac{C}{h^2} E \int_0^{x+h} p^2h^2|D_n - D| 1\{s \in (0, ph)\} \frac{J_n(s)}{Y_n(s)} dN_n(s) \\
&+ \frac{C}{h^2} E \int_0^{x+h} Ap^3h^3|D_n^2 - D^2| 1\{s \in (0, ph)\} \frac{J_n(s)}{Y_n(s)} dN_n(s) \\
&= K_1 + K_2. \qquad (5.18)
\end{aligned}$$

By Lemma 3.3 and uniform integrability of K_1 we have

$$K_1 = O(h^2). \qquad (5.19)$$

Also K_2 is uniformly integrable and

$$K_2 = O(h^4), \qquad (5.20)$$

which can be proved using the following

$$
\begin{aligned}
|D_n^2 - D^2| &= |D_n^2 - 2D_nD + D^2 + 2D_nD - 2D^2| \\
&= |(D_n - D)^2 + 2D(D_n - D)| \\
&\leq |D_n - D|^2 + 2D|D_n - D| \stackrel{P}{=} O(h^2) + O(h) \stackrel{P}{=} O(h).
\end{aligned}
$$

The last equality is a consequence of Lemma 3.3.

Combining (5.19) and (5.20) we obtain

$$
|K| = O(h^2), \tag{5.21}
$$

which completes the proof of (3.30).

***Proof of Theorem* 3.4.**
Note that

$$
\begin{aligned}
\sqrt{nh}\,(\hat{\alpha}_{new}(x) - \hat{\alpha}^*_{new}(x)) &= \sqrt{nh}\,(\hat{\alpha}_{new}(x) - \tilde{\alpha}_n(x)) \\
&+ \sqrt{nh}\,(\tilde{\alpha}_n(x) - \tilde{\alpha}^*_n(x)) \\
&+ \sqrt{nh}\,(\tilde{\alpha}_n(x) - \hat{\alpha}^*_{new}(x)) \\
&= J_1 + J_2 + J_3, \qquad (5.22)
\end{aligned}
$$

where

$$
\begin{aligned}
|J_1| &\leq \sqrt{\frac{n}{h^3}} C \int_0^{x+h} |\varphi_n(s) - \varphi(s)| 1\{s \in [0, ph]\} J_n(s)\alpha(s)ds \\
&+ \sqrt{\frac{n}{h^3}} C \int_0^{x+h} |\varphi_n(s) - \varphi(s)| 1\{s \in [0, ph]\} \frac{J_n(s)}{Y_n(s)} dM_n s \\
&= J_1^1 + J_1^2. \qquad (5.23)
\end{aligned}
$$

It can be shown by Lemma 3.3 that

$$
\begin{aligned}
J_1^1 &\le \sqrt{\frac{n}{h^3}} C \int_0^{x+h} \\
&\times \left(p^2 h^2 |D_n - D| + A p^3 h^3 |D_n - D|^2 + 2 A p^3 h^3 D |D_n - D|\right) \\
&\times 1\{s \in [0, ph]\} \alpha(s) ds \\
&\cong O(\sqrt{nh^5}) \longrightarrow 0. \qquad (5.24)
\end{aligned}
$$

Furthermore, by (3.7) we obtain

$$
\begin{aligned}
J_1^2 &\cong \sqrt{\frac{n}{h^3}} C \int_0^{x+h} |\varphi_n(s) - \varphi(s)| 1\{s \in [0, ph]\} \frac{1}{ny(s)} dM_n(s) \\
&\le \frac{C_1}{\sqrt{nh^3}} \int_0^{x+h} \left(ph^2 |D_n - D| + A p^3 h^3 |D_n - D|^2 + A p^3 h^3 2D |D_n - D|\right) \qquad (5.25) \\
&\quad \times 1\{s \in [0, ph]\} dM_n(s) \\
&\cong \frac{C_1}{\sqrt{nh^3}} O(h^4) \int_0^{x+h} dM_n(s) \overset{P}{\cong} C_2 h^{5/2} \int_0^{x+h} \frac{1}{\sqrt{n}} dM_n(s) \\
&\overset{P}{\longrightarrow} 0,
\end{aligned}
$$

where C and C_i are constants. So $J_1 \to 0$ in probability. The quantity $J_2 \overset{D}{\to} N(0, V)$, which follows from Theorem 3.1. Similarly as for J_1^1 we can prove that $J_3 \overset{P}{\cong} O(\sqrt{nh^5})$ and tends to 0 as $nh^5 \to 0$. This completes the proof of formula (3.31). Moreover, similarly we obtain formula (3.32) and by analogy to Theorem 3.2 we obtain formulas (3.33) and (3.34).

REFERENCES

[1] O. Aalen, *Nonparametric inference for a family of counting processes*, The Annals of Statistics **6** (4) (1978), 701–726.

[2] P. Andersen, O. Borgan, R. Gill and N. Keiding, *Statistical Models Based on Counting Processes*, Springer-Verlag, New York 1993.

[3] B. Janiszewska and R. Różański, *Transformed diffeomorphic kernel estimation of hazard rate function*, Demonstratio Mathematica **34** (2) (2001), 447–460.

[4] H. Ramlau-Hansen, *Smoothing counting process intensities by means of kernel funcions*, The Annals of Statistics **11** (2) (1983), 453–466.

[5] S. Saoudi, F. Ghorbel and A. Hillion, *Some statistical properties of the kernel-diffeomorphism estimator*, Applied Stochastic Models and Data Analysis **13** (1997), 39–58.

[6] D. Ruppert and J.S. Marron, *Transformations to reduce boundary bias in kernel density estimation*, J. Roy. Statist. Soc. Ser. B, **56** (4) (1994), 653–671.

[7] E.F. Schuster, *Incorporating support constraints into nonparametric estimators of densities*, Comm. Statist. A - Theory Methods. **14** (1985), 1125–1136.

[8] R.S. Liptser and A.N. Shiryayev, *Statistics of Random Processes I. General Theory*, Springer-Verlag, New York (1984) p. 17.

[9] M. Wand, J.S. Marron and D. Ruppert, *Transformations in density estimation (with discussion)*, J. Amer. Statist. Assoc. **86** (1991), 343–361.

[10] S. Zhang, R.J. Karunamuni and M.C. Jones, *An improved estimator of the density function at the boundary*, J. Amer. Statist. Assoc. **94** (1999), 1231–1240.

Received 20 July 2003