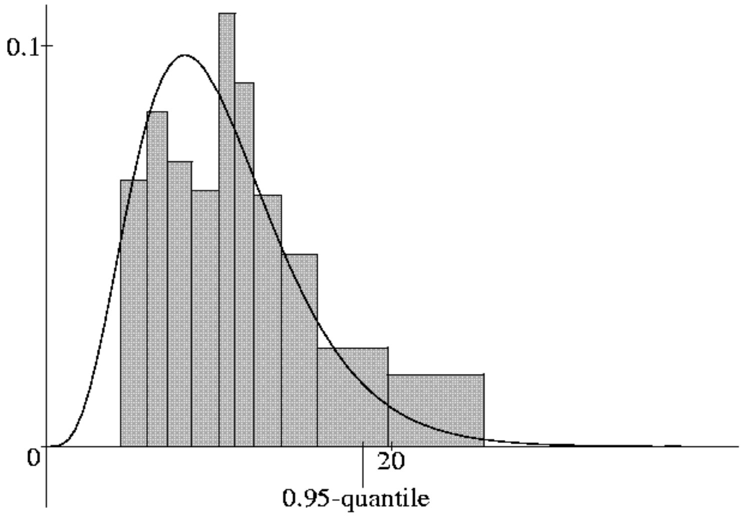
UNIVERSITÄT POTSDAM
Institut für Mathematik

Testing the Hazard Rate, Part I

Hannelore Liero



Mathematische Statistik und
Wahrscheinlichkeitstheorie

Universität Potsdam – Institut für Mathematik

Mathematische Statistik und Wahrscheinlichkeitstheorie

Testing the Hazard Rate, Part I

Hannelore Liero

Department of Mathematics, University of Potsdam, 14469 Potsdam, Germany

e-mail: liero@rz.uni-potsdam.de

Preprint 2003/17

August 2003

Impressum

Herausgeber: Mathematische Statistik und Wahrscheinlichkeitstheorie
am Institut für Mathematik

Adresse: Universität Potsdam
Am Neuen Palais 10
14469 Potsdam
Telefon:
Fax: +49-331-977 1500
E-mail: +49-331-977 1578
neisse@math.uni-potsdam.de

ISSN 1437-739X

Testing the Hazard Rate Part I

Hannelore Liero
Institute of Mathematics, University of Potsdam
e-mail: liero@rz.uni-potsdam.de

Abstract

We consider a nonparametric survival model with random censoring. To test whether the hazard rate has a parametric form the unknown hazard rate is estimated by a kernel estimator. Based on a limit theorem stating the asymptotic normality of the quadratic distance of this estimator from the smoothed hypothesis an asymptotic α-test is proposed. Since the test statistic depends on the maximum likelihood estimator for the unknown parameter in the hypothetical model properties of this parameter estimator are investigated. Power considerations complete the approach.

Keywords and phrases: kernel estimator of the hazard rate, goodness of fit, maximum likelihood estimator, limit theorem for integrated squared difference, censoring,

AMS subject classification: Primary 62N03, 62G10

1 Introduction and Main Result

This is the first part of a paper on testing the hazard rate in survival models with censored data. In this part we consider the model without covariates, that is we assume that the survival times are independent and identically distributed (i.i.d.) random variables. In the second part we extend our approach to the case that the survival times depend on covariates.

We start with some notation: Let $Y_1, \ldots, Y_n$ be a sequence of i.i.d. survival times with absolutely continuous distribution function F. As often occurs in applications the Y_i's are subject to random right censoring, i.e. the observations are

$$T_i \;=\; \min(Y_i, C_i) \quad \text{and} \quad \delta_i \;=\; 1(Y_i \leq C_i)$$

where $C_1, \ldots, C_n$ are i.i.d. random censoring times which are independent of the Y-sequence. The δ_i indicates whether Y_i has been censored or not. The function of interest is the hazard rate λ which is defined by

$$\lambda(t) \;=\; \lim_{s \downarrow 0} \frac{1}{s} \mathsf{P}(t < Y_i \leq t + s | Y_i \geq t).$$

We wish to test whether λ lies in a paramatric class of functions, i.e.

$$\mathcal{H}: \quad \lambda \in \mathcal{L} \;=\; \{\lambda(\cdot, \theta) \,|\, \theta \in \Theta \subseteq \mathbb{R}^k\} \quad \text{versus} \quad \mathcal{K}: \quad \lambda \notin \mathcal{L}.$$

Since no parametric form of the alternative is assumed we will use a nonparametric estimator of λ for testing $\mathcal{H}$ against $\mathcal{K}$. The idea for the construction of such a nonparametric estimator goes back to the paper of Watson and Leadbetter (1964), who considered the case without censoring. To describe the estimation procedure we introduce the distribution function of the observations T_i and the subdistribution function of the uncensored observations:

$$H(t) := \mathsf{P}(T_i \leq t) \quad \text{and} \quad H^U(t) := \mathsf{P}(T_i \leq t, \delta_i = 1).$$

Since

$$1 \,-\, H(t) \;=\; (1 \,-\, G(t))\,(1 \,-\, F(t))$$

and

$$H^U(t) \;=\; \int_0^t (1 - G(s))\, \mathrm{d}F(s),$$

where G is the distribution function of the censoring times C_i, the cumulative hazard function

$$\Lambda(t) := \int_0^t \lambda(s)\, \mathrm{d}s$$

can be written as

$$\Lambda(t) \;=\; \int_0^t \frac{\mathrm{d}F(s)}{1 - F(s_-)} \;=\; \int_0^t \frac{\mathrm{d}H^U(s)}{1 - H(s_-)}.$$

Now, for estimating Λ we replace H^U and H by their empirical versions, that is by

$$\hat{H}_n^U(t) \;=\; \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(T_i \le t, \delta_i = 1) \quad \text{and} \quad \hat{H}_n(t) \;=\; \frac{1}{n}\sum_{i=1}^{n} \mathbb{1}(T_i \le t).$$

The resulting estimator

$$\hat{\Lambda}_n(t) := \int_0^t \frac{\mathrm{d}\hat{H}_n^U(s)}{1-\hat{H}_n(s_-)} \;=\; \sum_{i=1}^{n} \frac{\mathbb{1}(T_{(i)} \le t)\,\delta_{[i]}}{n-i+1}$$

is the Nelson-Aalen estimator of Λ. Here $T_{(1)} \le \cdots \le T_{(n)}$ are the ordered observations and $\delta_{[i]} = \delta_j$ iff $T_j = T_{(i)}$.
As estimator of the derivative of Λ we define the kernel smoothed Nelson-Aalen estimator

$$\hat{\lambda}_n(t) := \frac{1}{b_n}\int K\left(\frac{t-s}{b_n}\right)\mathrm{d}\hat{\Lambda}_n(s) \;=\; \frac{1}{b_n}\sum_{i=1}^{n} \frac{K\left(\frac{t-T_{(i)}}{b_n}\right)\delta_{[i]}}{n-i+1}.$$

Here K is a kernel function and $\{b_n\}$ is a sequence of bandwidths tending to zero with an appropriate rate.

Several asymptotic properties of this estimator are known. Let us mention here the papers of Singpurwalla and Wong (1983), Tanner and Wong (1983) and the results of Diehl and Stute (1988). Diehl and Stute gave an approximation for the difference between the estimator $\hat{\lambda}_n$ and the smoothed hazard rate by a sum of i.i.d. random variables. On the basis of this i.i.d. representation asymptotic normality at a fixed point t and a limit theorem for the maximal deviation were derived. Here, for our test problem we consider the quadratic deviation. As test statistic we choose the L_2-distance of $\hat{\lambda}_n$ from the hypothesis, that is from a function which characterizes $\mathcal{H}$. To avoid problems arising from the bias of $\hat{\lambda}_n$ we do not take the distance of $\hat{\lambda}_n$ from an element of the hypothetical class $\mathcal{L}$, but from a smoothed version of this, which is given by

$$\tilde{\lambda}_n(t,\theta) := \int K_{b_n}(t-s)\,\lambda(s,\theta)\,\mathrm{d}s \;=\; \int K_{b_n}(t-s)\,\mathrm{d}\Lambda(s,\theta),$$

where $\Lambda(t,\theta) = \int_0^t \lambda(s,\theta)\,\mathrm{d}s$ and $K_b(t) \;=\; \frac{1}{b}K(t/b)$. So, we define

$$Q_{0n} := \int \left(\hat{\lambda}_n(t) \;-\; \tilde{\lambda}_n(t,\theta)\right)^2 a(t)\,\mathrm{d}t. \tag{1.1}$$

Here the weight function a is introduced to control the region of integration and has to be chosen by the statistician. Since the parameter θ in (1.1) is unknown we have to replace it by a suitable estimator. We propose to take the maximum likelihood estimator, say $\hat{\theta}_n$. Thus, finally we obtain as test statistic

$$\hat{Q}_{0n} := \int \left(\hat{\lambda}_n(t) \;-\; \tilde{\lambda}_n(t,\hat{\theta}_n)\right)^2 a(t)\,\mathrm{d}t.$$

From Theorem 2.1 below it will follow that for all $\theta \in \Theta$ the standardized Q_{0n} is asymptotically normally distributed if $\mathcal{H}$ is true, that is the following limit statement holds:

$$n b_n^{1/2} \,(Q_{0n} \,-\, m_{0n}) \quad \xrightarrow{\mathcal{D}} \quad \mathsf{N}(0, \sigma_0^2) \tag{1.2}$$

with

$$m_{0n} \;=\; (n b_n)^{-1}\, \kappa_1 \int \frac{\lambda(t,\theta)}{1 - H(t)}\; a(t)\,\mathrm{d}t$$

and

$$\sigma_0^2 \;=\; 2\,\kappa_2 \int \left(\frac{\lambda(t,\theta)}{1 - H(t)} \right)^2 a^2(t)\,\mathrm{d}t,$$

where $\kappa_1 = \int K^2(x)\,\mathrm{d}x$ and $\kappa_2 = \int (K * K)^2(x)\,\mathrm{d}x$ and "$*$" denotes the convolution. Under certain regularity conditions the maximum likelihood estimator $\hat{\theta}_n$ is $\sqrt{n}$-consistent. Therefore the limit statement (1.2) remains true for $\hat{Q}_{0n}$. Furthermore, in the standardizing terms the unknown distribution function H can be replaced by $\hat{H}_n$ without changing the limit distribution (see Consequence 3.1). Thus, finally we obtain an asymptotic α-test by the rule: Reject $\mathcal{H}$, iff

$$\hat{Q}_{0n} \,\geq\, \frac{z_\alpha\, \hat{\sigma}_{0n}}{n b_n^{1/2}} \,+\, \hat{m}_{0n} \tag{1.3}$$

where

$$\hat{m}_{0n} \;=\; (n b_n)^{-1}\, \kappa_1 \int \frac{\lambda(t,\hat{\theta}_n)}{1 - \hat{H}_n(t)}\, a(t)\,\mathrm{d}t,$$

$$\hat{\sigma}_{0n}^2 \;=\; 2\,\kappa_2 \int \left(\frac{\lambda(t,\hat{\theta}_n)}{1 - \hat{H}_n(t)} \right)^2 a^2(t)\,\mathrm{d}t,$$

and $\Phi(z_\alpha) = 1 - \alpha$.

2 Asymptotic Normality of the Quadratic Functional

In this section we present a theorem stating that the quadratic functional

$$Q_n := \int \left(\hat{\lambda}_n(t) \,-\, \tilde{\lambda}_n(t) \right)^2 a(t)\,\mathrm{d}t.$$

is asymptotically normal. This theorem is formulated not only for the behavior under the null hypothesis, but for general hazard rate λ. We define

$$\tilde{\lambda}_n(t) := \int K_{b_n}(t - s)\,\lambda(s)\,\mathrm{d}s.$$

Further, let T_H be the right end point of the distribution H.

Theorem 2.1 *Suppose that*

(i) K *is a continuous density function vanishing outside the interval* $[-L, L]$ *for some* $L > 0$.

(ii) λ *and* H *are Lipschitz continuous.*

(iii) *The function* a *is continuous and* $a(t) \equiv 0$ *for all* $t > T_H$.

(iv) $b_n \to 0$ *and* $nb_n^2 \to \infty$.

Then for $n \to \infty$

$$nb_n^{1/2}\,(Q_n - m_n) \;\xrightarrow{\mathcal{D}}\; \mathsf{N}(0, \sigma^2) \tag{2.1}$$

with

$$m_n = (nb_n)^{-1}\,\kappa_1 \int \frac{\lambda(t)}{1 - H(t)}\; a(t)\,\mathrm{d}t$$

and

$$\sigma^2 = 2\,\kappa_2 \int \left(\frac{\lambda(t)}{1 - H(t)}\right)^2 a^2(t)\,\mathrm{d}t.$$

3 The Maximum Likelihood Estimator of the Parameter

Let us now investigate the maximum likelihood estimator of the unknown parameter θ. The likelihood function is given by

$$L_n(\theta, T_1, \delta_1, \ldots, T_n, \delta_n) = \prod_{i=1}^{n} L(\theta, T_i, \delta_i)$$

with

$$\begin{aligned} L(\theta, t, \delta)) &= (1 - G(t))^{\delta}\;(1 - F(t, \theta))^{1-\delta}\; f(t, \theta)^{\delta}\, g(t)^{1-\delta} \\ &= \lambda(t, \theta)^{\delta}\;\exp(-\Lambda(t, \theta))\;(1 - G(t))^{\delta}\; g(t)^{1-\delta}, \end{aligned} \tag{3.1}$$

where g is the density of the censoring times. Thus, the maximum likelihood estimator $\hat{\theta}_n$ is a (measurable) maximizer of

$$l_n(\theta) = \sum_{i=1}^{n} \left(\delta_i \log \lambda(T_i, \theta) - \Lambda(T_i, \theta)\right).$$

To conclude from Theorem 2.1 to the asymptotic normality of our test statistic we use the $\sqrt{n}$-consistency of the maximum likelihood estimator. For that purpose we formulate the following regularity conditions:

(i) For all $t \in [0,\infty)$ and all $i,j = 1,\ldots,k$ the second derivatives $\nabla_i\nabla_j\lambda(t,\theta)$ and $\nabla_i\nabla_j\Lambda(t,\theta)$ exist and are continuous on Θ^o, the open kernel of Θ.

(ii) For all $\theta \in \Theta^o$ and all $i,j = 1,\ldots,k$

$$\nabla_i\int \lambda(t,\theta)\,\mathrm{d}t = \int \nabla_i\lambda(t,\theta)\,\mathrm{d}t,$$
$$\nabla_i\nabla_j\int \lambda(t,\theta)\,\mathrm{d}t = \int \nabla_i\nabla_j\lambda(t,\theta)\,\mathrm{d}t.$$

(iii) For any $\theta \in \Theta^o$ there exist a ν-neighborhood $U(\theta,\nu) \subset \Theta^o$ of θ, and a measurable function $M(\cdot,\cdot,\theta)$ with $\mathsf{E}M(T_1,\delta_1,\theta) < \infty$ such that

$$\left|\,\nabla_i\nabla_j \log L(\theta',\cdot,\cdot)\right| \le M(\cdot,\cdot,\theta) \qquad \text{for all } \theta' \in U(\theta,\nu)$$

for all $i,j = 1,\ldots,k$.

(iv) The determinant of the Fisher information $I(\theta) = (I_{ij}(\theta))_{i,j=1,\ldots,k}$ with

$$I_{ij}(\theta) = \int \nabla_i\lambda(t,\theta)\,\nabla_j\lambda(t,\theta)\,\frac{(1-H(t,\theta))}{\lambda(t,\theta)}\,\mathrm{d}t$$

is nonzero for all $\theta \in \Theta^o$.

Under these conditions we have:

Theorem 3.1 *Suppose that conditions* (i) - (iv) *are satisfied and that $\hat{\theta}_n$ is consistent. Then under $\mathcal{H}$*

$$\sqrt{n}\left(\hat{\theta}_n - \theta\right) \xrightarrow{\mathcal{D}} \mathsf{N}\left(0, I(\theta)^{-1}\right)$$

for any $\theta \in \Theta^o$.

Remark: Very often the consistency of the maximum likelihood estimator can be verified directly.
Generally one can show that under the following conditions a consistent maximum likelihood estimator exists: For all $\theta,\theta' \in \Theta^o$

$$\mathsf{E}\log\frac{L(\theta',T_1,\delta_1)}{L(\theta,T_1,\delta_1)} < \infty,$$

where the expectation is taken with respect to the distribution depending on θ. Further, assume that the set

$$\mathcal{A}_n(t_1,\delta_1,\ldots,t_n,\delta_n) = \{\theta \,|\, \nabla_\theta \log L_n(\theta,t_1,\delta_1,\ldots,t_n,\delta_n) = 0\}$$

is not empty and the function $\nabla_\theta \log L(\theta,\cdot,\cdot)$ is continuous in θ.

From Theorem 3.1 it follows that the maximum likelihood estimator $\hat{\theta}_n$ is $\sqrt{n}$-consistent. This, together with the $\sqrt{n}$- consistency of the empirical distribution function $\hat{H}_n$ is sufficient to show that the limit statement (2.1) remains valid for $\hat{Q}_{0n}$ with estimated standardizing terms, so we have

Consequence 3.1 *Suppose that for each θ the vector of the partial derivatives $\nabla_\theta \lambda(\cdot, \theta)$ is uniformly continuous with respect to the first argument and that the conditions of Theorem 2.1 and Theorem 3.1 are satisfied. Then*

$$\frac{n b_n^{1/2}}{\hat{\sigma}_{0n}} \left(\hat{Q}_{0n} - \hat{m}_{0n} \right) \quad \xrightarrow{\mathcal{D}} \quad \mathsf{N}(0,1). \tag{3.2}$$

4 Power Considerations

We consider local alternatives of the form

$$\mathcal{K}_n : \quad \lambda_n^*(t) \;=\; \lambda(t, \theta) \;+\; \varepsilon_n \, d(t), \tag{4.1}$$

where θ is arbitrarily fixed, d is a function satisfying some regularity conditions, and $\{\varepsilon_n\}$ is a sequence of positive numbers. Let $\mathsf{P}_{\mathcal{K}_n}$ be the probability under the alternative, that is, when the distribution of the survival times Y_i is F_n^* with

$$1 \,-\, F_n^*(t) \;=\; (1 \,-\, F(t,\theta)) \exp(-\varepsilon_n D(t)), \qquad D(t) = \int_0^t d(s)\,\mathrm{d}s$$

and the censoring times are distributed according G. The power of the test proposed by (1.3) is given by

$$\beta_n \;=\; \mathsf{P}_{\mathcal{K}_n} \left(\frac{n b_n^{1/2}}{\hat{\sigma}_{0n}} \left(\hat{Q}_{0n} \,-\, \hat{m}_{0n} \right) \geq z_\alpha \right).$$

Suppose that for each n the function λ_n^* satisfies the conditions of Theorem 2.1 in Section 2. Then we have for all z_n

$$\left| \, \mathsf{P}_{\mathcal{K}_n} \left(\frac{n b_n^{1/2}}{\sigma_n^*} \left(Q_n^* \,-\, m_n^* \right) \leq z_n \right) \,-\, \Phi(z_n) \right| \to 0, \tag{4.2}$$

where

$$Q_n^* \;=\; \int \left(\hat{\lambda}_n(t) \,-\, \tilde{\lambda}_n^*(t) \right)^2 a(t)\,\mathrm{d}t \quad \text{with} \quad \tilde{\lambda}_n^*(t) = \int K_{b_n}(t-s) \lambda_n^*(s)\,\mathrm{d}s,$$

$$m_n^* \;=\; (n b_n)^{-1}\, \kappa_1 \int \frac{\lambda_n^*(t)}{1 - H_n^*(t)}\; a(t)\,\mathrm{d}t$$

and

$$\sigma_n^{*2} = 2\,\kappa_2 \int \left(\frac{\lambda_n^*(t)}{1 - H_n^*(t)} \right)^2 a^2(t)\,\mathrm{d}t$$

with

$$1 - H_n^*(t) = (1 - F_n^*(t))(1 - G(t)).$$

To formulate our theorem about the power β_n we need conditions on the behavior of the parameter estimator under $\mathcal{K}_n$. We assume that there exists a sequence of asymptotically normally distributed r. v.'s Z_n with expectation zero and a nonrandom vector S such that under $\mathcal{K}_n$ the following expansion holds for $n \to \infty$

$$\hat{\theta}_n - \theta = n^{-1/2} Z_n + \varepsilon_n S\,(1 + o(1)). \tag{4.3}$$

Theorem 4.1 *Suppose that the assumptions of Theorem* 2.1 *are satisfied, that* λ_n^* *and* H_n^* *are Lipschitz continuous, that the partial derivatives* $\nabla_j \lambda(\cdot, \theta)$, $j = 1, \ldots, k$ *are uniformly continuous with respect to the first argument and that* (4.3) *holds. If*

$$b_n^{-1/2} \varepsilon_n \to 0 \quad \textit{or} \quad n b_n \varepsilon_n \to \infty \quad \textit{or} \quad \int \frac{d(t) - S^t \nabla_\theta \lambda(t, \theta)}{1 - H(t)}\, a(t)\,\mathrm{d}t = 0$$

then the test is asymptotically unbiased, that is, $\lim_n \beta_n \geq \alpha$. *Furthermore,*

$$\lim_n \beta_n = \begin{cases} \alpha & \textit{if} \quad n b_n^{1/2}\, \varepsilon_n^2 \to 0, \\ \beta & \textit{if} \quad n b_n^{1/2}\, \varepsilon_n^2 \to c^* > 0, \\ 1 & \textit{if} \quad n b_n^{1/2}\, \varepsilon_n^2 \to \infty, \end{cases}$$

where the number β *lies between* α *and* 1 *and depends on* $\sigma_0^2, \nabla_\theta \lambda(\cdot, \theta), H, d, S, a$ *and* c^*.

From Theorem 4.1 it follows that the power of the test tends to a nontrivial limit if $n b_n^{1/2} \varepsilon_n^2$ converges to a positive number, or in other words, if the squared weighted L_2-norm of the deviation from the hypothesis $||\varepsilon_n\, d\, a^{1/2}||_2^2$ converges to $c > 0$. The local alternative (4.1) is a very simple one, more general alternatives are for example

$$\lambda_n^*(t) = \lambda(t, \theta) + d_n(t) \tag{4.4}$$

where d_n tends to zero in a certain sense. The investigation of the behavior of the power under such alternatives requires more technical conditions (Lipschitz condition for d_n with constant depending on n, behavior of $\hat{\theta}_n$ under (4.4)). But the main result, namely that the convergence of the squared weighted L_2-norm of the disturbing function is essential for the distinguishability of the test, remains valid also under (4.4). So, we do not consider the technically more complicated alternatives.

5 Proofs

Proof of Theorem 2.1: The proof consists of two parts. In the first part we show that the difference

$$V_n(t) \;=\; \hat{\lambda}_n(t) \;-\; \tilde{\lambda}_n(t)$$

can be approximated by the sum of i.i.d. random variables

$$V_{n1}(t) \;=\; n^{-1} \sum_{i=1}^{n} Z_i(t, b_n)$$

with

$$Z_i(t, b_n) \;=\; \frac{K_{b_n}(t - T_i)\delta_i}{1 - H(T_i)} \;-\; \int \frac{\mathbf{1}(T_i \geq s)\, K_{b_n}(t-s)}{(1 - H(s))^2} \, \mathrm{d}H^U(s)$$

In the second step the asymptotic normality of the approximating integral

$$\int V_{n1}^2(t)\, a(t)\, \mathrm{d}t$$

is proved.

Lemma 5.1 *Suppose that* (i) *and* (ii) *are satisfied. Then for all* $T < T_H$

$$\sup_{0 \leq t \leq T} |V_n(t) \;-\; V_{n1}(t)| \;=\; O\left(\frac{\log\log n}{n} \;+\; \left(\frac{b_n \log n}{n}\right)^{1/2}\right) \quad a.s. \tag{5.1}$$

Proof of Lemma 5.1: Standard computations lead to

$$\begin{aligned} V_n(t) \;-\; V_{n1}(t) = & \int \frac{K_{b_n}(t-s)(\hat{H}_n(s) - H(s))^2}{(1 - \hat{H}_n(s))(1 - H(s))} \, \mathrm{d}H^U(s) \\ & + \int \frac{K_{b_n}(t-s)(\hat{H}_n(s) - H(s))}{(1 - \hat{H}_n(s))(1 - H(s))} \, \mathrm{d}(\hat{H}_n^U(s) - H^U(s)) \end{aligned} \tag{5.2}$$

To estimate this difference we use that for all $\tilde{T} < T_H$

$$\sup_{0 \leq t \leq \tilde{T}} |\hat{H}_n(t) - H(t)| \overset{\text{a.s.}}{=} O\left(n^{-1/2} (\log\log n)^{1/2}\right).$$

Consider the first summand in (5.2). Since K has a bounded support we have for sufficiently large n and for some $T' < T_H$

$$\begin{aligned} & \sup_{0 \leq t \leq T} \Big| \int \frac{K_{b_n}(t-s)(\hat{H}_n(s) - H(s))^2}{(1 - \hat{H}_n(s))(1 - H(s))} \, \mathrm{d}H^U(s)) \Big| \\ \leq & \sup_{0 \leq t \leq T} \sup_{0 \leq s \leq T'} \frac{(\hat{H}_n(s) - H(s))^2}{(1 - \hat{H}_n(s))(1 - H(s))} \int K_{b_n}(t-s)\, \mathrm{d}H^U(s) \\ \leq & \frac{1}{(1 - \hat{H}_n(T'))(1 - H(T'))} \sup_{0 \leq t \leq T} \int K_{b_n}(t-s)\, \mathrm{d}H^U(s) \;\cdot\; O\left(n^{-1} \log\log n\right) \\ = & \; O\left(n^{-1} \log\log n\right) \quad a.s. \quad . \end{aligned}$$

Now, let us investigate the second summand in (5.2). We take a partition of the interval $[0,T]$ into intervals $[t_i, t_{i+1})$ of length Lb_n, $i = 1, \dots, k_n$, $k_n = [\frac{T}{Lb_n}]$. Then with

$$\psi_n(t) \;=\; \frac{(\hat{H}_n(t) - H(t))}{(1-\hat{H}_n(t))(1-H(t))}$$

we have

$$\begin{aligned}
&\sup_{0\le t\le T} \Big| \int \frac{K_{b_n}(t-s)(\hat{H}_n(s)-H(s))}{(1-\hat{H}_n(s))(1-H(s))}\, \mathrm{d}(\hat{H}_n^U(s) - H^U(s))\Big| \\
&\le \max_{1\le i\le k_n} \sup_{t_i\le t<t_{i+1}} \Big|\psi_n(t_i) \int K_{b_n}(t-s)\, \mathrm{d}(\hat{H}_n^U(s) - H^U(s))\Big| && (5.3)\\
&+ \max_{1\le i\le k_n} \sup_{t_i\le t<t_{i+1}} \Big| \int (\psi_n(s) - \psi_n(t_i))\; K_{b_n}(t-s)\, \mathrm{d}\hat{H}_n^U(s)\Big| && (5.4)\\
&+ \max_{1\le i\le k_n} \sup_{t_i\le t<t_{i+1}} \Big| \int (\psi_n(s) - \psi_n(t_i))\; K_{b_n}(t-s)\, \mathrm{d}H^U(s)\Big|. && (5.5)
\end{aligned}$$

Since $\sup_{0\le t\le T} |\int K_{b_n}(t-s)\, \mathrm{d}(\hat{H}_n^U(s) - H^U(s))| \overset{\text{a.s.}}{=} O\left((nb_n)^{-1/2}(\log n)^{1/2}\right)$ we obtain that the summand (5.3) is almost surely of order $O\left(n^{-1}b_n^{-1/2}(\log n \log\log n)^{1/2}\right)$.
Note that $\sup_{0\le t\le T} \int K_{b_n}(t-s)\, \mathrm{d}\hat{H}^U(s)$ and $\sup_{0\le t\le T} \int K_{b_n}(t-s)\, \mathrm{d}\hat{H}_n^U(s)$ are (almost surely) bounded. Thus the summands (5.4) and (5.5) are bounded by

$$\max_{1\le i\le k_n} \sup_{\substack{s:|t-s|\le Lb_n \\ t_i\le t<t_{i+1}}} |\psi_n(s) \;-\; \psi_n(t_i)| \cdot O(1). \qquad (5.6)$$

Furthermore set

$$\tilde{\psi}_n(t) \;=\; \frac{(\hat{H}_n(t) - H(t))}{(1-H(t))^2}.$$

Since

$$\sup_{0\le s\le T} |\psi_n(s) - \tilde{\psi}_n(s)| \overset{\text{a.s.}}{=} O\left(n^{-1}\log\log n\right)$$

it is enough to consider $\max_{1\le i\le k_n} \sup_{\substack{s:|t-s|\le Lb_n \\ t_i\le t<t_{i+1}}} |\tilde{\psi}_n(s) - \tilde{\psi}_n(t_i)|$. For some constant C we have

$$\begin{aligned}
&\max_{1\le i\le k_n} \sup_{\substack{s:|t-s|\le Lb_n \\ t_i\le t<t_{i+1}}} |\tilde{\psi}_n(s) \;-\; \tilde{\psi}_n(t_i)| \\
&\le C \max_{1\le i\le k_n} \sup_{\substack{s:|t-s|\le Lb_n \\ t_i\le t<t_{i+1}}} |\hat{H}_n(s) - H(s) - \hat{H}_n(t_i) + H(t_i)| \\
&+ O\left(n^{-1/2}(\log\log n)^{1/2}\right) \cdot O(b_n).
\end{aligned}$$

To estimate the difference in the first summand we introduce the intervals $J_i = [t_i - Lb_n, t_i + Lb_n]$ and take a subpartition of these intervals into intervals $[t_{ij}, t_{ij+1})$ of length $n^{-1/2}(b_n \log n)^{1/2}$, $j = 1, \ldots, r_n$. We have:

$$\begin{aligned}
&\max_{1\le i\le k_n} \sup_{\substack{s:|t-s|\le Lb_n \\ t_i\le t<t_{i+1}}} |\hat{H}_n(s) - H(s) - \hat{H}_n(t_i) + H(t_i)| \\
\le\ &\max_{1\le i\le k_n} \sup_{s\in J_i} |\hat{H}_n(s) - H(s) - \hat{H}_n(t_i) + H(t_i)| \\
\le\ &\max_{1\le i\le k_n} \max_{1\le j\le r_n} \sup_{t_{ij}\le s<t_{ij+1}} |\hat{H}_n(s) - \hat{H}_n(t_{ij})| && (5.7)\\
&+ \max_{1\le i\le k_n} \max_{1\le j\le r_n} |\hat{H}_n(t_{ij}) - H(t_{ij}) - \hat{H}_n(t_i) + H(t_i)| && (5.8)\\
&+ \max_{1\le i\le k_n} \max_{1\le j\le r_n} \sup_{t_{ij}\le s<t_{ij+1}} |H(s) - H(t_{ij})|. && (5.9)
\end{aligned}$$

The summands (5.7) and (5.9) are of order $O\left(\left(\frac{b_n \log n}{n}\right)^{1/2}\right)$. It remains to investigate (5.8). The term $\hat{H}_n(t_{ij}) - H(t_{ij}) - \hat{H}_n(t_i) + H(t_i)$ is a sum of i.i.d. (bounded) random variables with expectation zero and variance

$$\mathsf{Var}(\hat{H}_n(t_{ij}) - \hat{H}_n(t_i)) \le Cn^{-2} \max_{1\le j,k\le r_n} |t_{ij} - t_{ik}| \;=\; O\left(n^{-2}b_n\right).$$

From a lemma about strong uniform consistency (see Liero (1999)) it follows that for all constants $\rho > 0$ there exists a constant C_ρ such that for all $C \ge C_\rho$

$$\mathsf{P}\left(\left(\frac{n}{b_n \log n}\right)^{1/2} |\hat{H}_n(t_{ij}) - H(t_{ij}) - \hat{H}_n(t_i) + H(t_i)| > C\right) \le n^{-\rho}. \tag{5.10}$$

By a suitable choice of ρ it follows from (5.10) that

$$\max_{1\le i\le k_n} \max_{1\le j\le r_n} |\hat{H}_n(t_{ij}) - H(t_{ij}) - \hat{H}_n(t_i) + H(t_i)| \overset{\text{a.s.}}{=} O\left(\left(\frac{b_n \log n}{n}\right)^{1/2}\right)$$

and the proof is complete.

□

Lemma 5.2 *Suppose that the conditions of Theorem* 2.1 *are satisfied. Then*

$$nb_n^{1/2}\left(\int V_{n1}^2(t)\, a(t)\, \mathrm{d}t \,-\, m_n\right) \xrightarrow{\mathcal{D}} \mathsf{N}(0, \sigma^2). \tag{5.11}$$

Proof of Lemma 5.2: To investigate the distributional behavior of $\int V_{n1}^2(t)\, a(t)\, \mathrm{d}t$ it is useful to introduce the covariance function $C_n(t,s) = \mathsf{Cov}(V_{n1}(t), V_{n1}(s))$. Straightforward computations yield

$$C_n(t,s) \;=\; (nb_n)^{-1} \int \frac{K(z)K\left(\frac{s-t}{b_n} + z\right)\, \lambda(t - zb_n)}{1 - H(t - zb_n)}\, \mathrm{d}z. \tag{5.12}$$

The integral of interest can be decomposed into

$$\int V_{n1}^2(t)\, a(t)\, \mathrm{d}t \;=\; \frac{n-1}{n}\, U_n \;+\; S_n, \tag{5.13}$$

where

$$U_n \;=\; \frac{2}{n(n-1)} \sum\sum_{1\le i< j\le n} \eta_n(T_i,\delta_i,T_j,\delta_j)$$

is a U-statistic of degree 2 with kernel

$$\eta_n(T_i,\delta_i,T_j,\delta_j) \;=\; \int Z_i(t,b_n)\, Z_j(t,b_n)\, a(t)\, \mathrm{d}t \tag{5.14}$$

and

$$S_n \;=\; \frac{1}{n^2}\sum_{i=1}^{n} \eta_n(T_i,\delta_i,T_i,\delta_i).$$

At first we investigate the term S_n. With the help of the covariance function we get

$$\begin{aligned}\mathsf{E}S_n &= \int C_n(t,t)\, a(t)\, \mathrm{d}t \\ &= (nb_n)^{-1} \int\int \frac{K^2(z)\, \lambda(t-zb_n)}{1-H(t-zb_n)}\, a(t)\, \mathrm{d}z\, \mathrm{d}t.\end{aligned}$$

Since H and λ are Lipschitzian this leads to

$$\begin{aligned}\mathsf{E}S_n &= (nb_n)^{-1} \int \frac{\lambda(t)}{1-H(t)}\, a(t)\, \mathrm{d}t \int K^2(z)\, \mathrm{d}z \;+\; O(n^{-1}) \\ &= m_n \;+\; O(n^{-1}).\end{aligned}$$

Similarly one verifies that the variance of S_n is of order $O(n^{-3}b_n^{-2})$, therefore by the Chebyshev inequality we have

$$nb_n^{1/2}(S_n \;-\; m_n) \;=\; o_{\mathsf{P}}(1)$$

and in (5.13) S_n can be replaced by m_n, or in other words, it is sufficient to show that the distribution of $nb_n^{1/2}U_n$ tends to $\mathsf{N}(0,\sigma^2)$. To do that, we apply the method proposed by Hall (1984) and recall his Theorem 1. (We formulate it in the notation used here.)

Theorem (P. Hall) *Assume η_n is symmetric, $\mathsf{E}(\eta_n(T_1,\delta_1,T_2,\delta_2)|T_1,\delta_1) \;=\; 0$ almost surely and $\mathsf{E}(\eta_n^2(T_1,\delta_1,T_2,\delta_2)) < \infty$ for each n.*
Set $\mathbb{G}_n(t_1,\nu_1,t_2,\nu_2) \;=\; \mathsf{E}(\eta_n(T_1,\delta_1,t_1,\nu_1)\, \eta_n(T_1,\delta_1,t_2,\nu_2))$.
If

$$\frac{\mathsf{E}\mathbb{G}_n^2(T_1,\delta_1,T_2,\delta_2) \;+\; n^{-1}\mathsf{E}\eta_n^4(T_1,\delta_1,T_2,\delta_2)}{(\mathsf{E}\eta_n^2(T_1,\delta_1,T_2,\delta_2))^2} \to 0 \tag{5.15}$$

for $n \to \infty$, then U_n is asymptotically normally distributed with zero mean and variance given by $2\, n^{-2} \mathsf{E}\eta_n^2(T_1, \delta_1, T_2, \delta_2)$.

It is easy to see, that the first three assumptions of this theorem are satisfied by the U-kernel defined in (5.14). Let us now compute $\mathsf{E}\eta_n^2(T_1, \delta_1, T_2, \delta_2)$. We have

$$\begin{aligned}
&\mathsf{E}\eta_n^2(T_1, \delta_1, T_2, \delta_2) \\
&= n^2 \iint C_n^2(t,s)\, a(t)\, a(s)\, \mathrm{d}s\, \mathrm{d}t \\
&= b_n^{-2} \iiiint \frac{K(z)\, K\left(\frac{s-t}{b_n} + z\right)\, K(w)\, K\left(\frac{s-t}{b_n} + w\right)\, \lambda(t - zb_n)\lambda(t - wb_n)}{(1 - H(t - zb_n))\,(1 - H(t - wb_n))} \\
&\qquad\qquad \times a(t)\, a(s)\, \mathrm{d}z\, \mathrm{d}w\, \mathrm{d}s\, \mathrm{d}t \\
&= b_n^{-1} \iiiint \frac{K(z)\, K(s+z)\, K(w)\, K(s+w)\ \lambda(t - zb_n)\lambda(t - wb_n)}{(1 - H(t - zb_n))\,(1 - H(t - wb_n))} \\
&\qquad\qquad \times a(t)\, a(t + sb_n)\, \mathrm{d}z\, \mathrm{d}w\, \mathrm{d}s\, \mathrm{d}t \\
&= b_n^{-1} \int \left(\frac{\lambda(t)}{1 - H(t)}\right)^2 a^2(t)\, \mathrm{d}t \int (K * K)^2\,(s)\, \mathrm{d}s\, +\, O(1) \\
&= b_n^{-1}\, \sigma^2/2\, +\, O(1).
\end{aligned}$$

With

$$\begin{aligned}
\mathsf{E}\mathbb{G}_n^2(T_1, \delta_1, T_2, \delta_2) = n^4 &\iiiint C_n(t_1, t_2) C_n(t_2, t_3) \\
&\times C_n(t_3, t_4) C_n(t_4, t_1) a(t_1) a(t_2) a(t_3) a(t_4)\, \mathrm{d}t_1\, \mathrm{d}t_2\, \mathrm{d}t_3\, \mathrm{d}t_4
\end{aligned}$$

the term $\mathsf{E}\mathbb{G}_n^2$ can be handled similarly, and we get

$$\mathsf{E}\mathbb{G}_n^2(T_1, \delta_1, T_2, \delta_2)\; =\; O(b_n^{-1}).$$

Furthermore, we have

$$\mathsf{E}\eta_n^4(T_1, \delta_1, T_2, \delta_2)\; =\; O(b_n^{-3}).$$

Thus, condition (5.15) is satisfied, and the desired limit statement follows by the theorem of Hall.

□

To complete the proof of Theorem 2.1 we have to show that $\int V_n^2 a$ and $\int V_{n1}^2 a$ have the same asymptotic behavior. By the Cauchy-Schwarz inequality we get

$$\begin{aligned}
&n b_n^{1/2} \left| \int (V_n^2(t)\, -\, V_{n1}^2(t))\, a(t)\, \mathrm{d}t \right| \\
&= n b_n^{1/2} \left| \int (V_n(t)\, -\, V_{n1}(t))^2\, a(t)\, \mathrm{d}t\, +\, 2 \int V_{n1}(t)\, (V_n(t)\, -\, V_{n1}(t))\, a(t)\, \mathrm{d}t \right|
\end{aligned}$$

$$\leq n b_n^{1/2} \int (V_n(t) - V_{n1}(t))^2\, a(t)\, \mathrm{d}t$$

$$+ 2\, n b_n^{1/2} \sqrt{\int (V_n(t) - V_{n1}(t))^2\, a(t)\, \mathrm{d}t} \sqrt{\int V_{n1}^2(t)\, a(t)\, \mathrm{d}t}.$$

By Lemma 5.1 we have that the first summand is of order $O(b_n^{3/2} \log n)$. Lemma 5.2 implies

$$\sqrt{\int V_{n1}^2(t)\, a(t)\, \mathrm{d}t} = O_{\mathsf{P}}\left((n b_n)^{-1/2}\right).$$

Therefore the desired difference is of order $o_{\mathsf{P}}(1)$.

□

Proof of Theorem 3.1: The proof of the asymptotic normality is based on well-known results about efficiency of maximum likelihood estimators. The conditions stated here correspond Theorem 6.35 in Witting and Müller-Funk (1995). Using partial integration one can verify that assumptions (A1) to (A4) formulated there for a density f are satisfied by the likelihood function L defined in (3.1). So the proof is omitted.

□

Proof of Consequence 3.1: Let θ be arbitrarily fixed. By definition we have

$$\frac{n b_n^{1/2}\left(\hat{Q}_{0n} - \hat{m}_{0n}\right)}{\hat{\sigma}_{0n}} = \frac{\sigma_0}{\hat{\sigma}_{0n}} \left(\frac{n b_n^{1/2}\left(Q_{0n} - m_{0n}\right)}{\sigma_0} + \sum_{j=1}^{3} \xi_{nj} \right)$$

where

$$\xi_{n1} = \frac{n b_n^{1/2}}{\sigma_0} \int \left(\tilde{\lambda}_n(t,\theta) - \tilde{\lambda}_n(t,\hat{\theta}_n)\right)^2 a(t)\, \mathrm{d}t$$

$$\xi_{n2} = \frac{2 n b_n^{1/2}}{\sigma_0} \int \left(\hat{\lambda}_n(t) - \tilde{\lambda}_n(t,\theta)\right) \left(\tilde{\lambda}_n(t,\theta)) - \tilde{\lambda}_n(t,\hat{\theta}_n)\right) a(t)\, \mathrm{d}t$$

and

$$\xi_{n3} = \frac{n b_n^{1/2}}{\sigma_0} (m_{0n} - \hat{m}_{0n})$$

To prove the statement it is enough to show that $\sigma_0/\hat{\sigma}_{0n} \xrightarrow{\mathrm{P}} 1$ and $\xi_{nj} \xrightarrow{\mathrm{P}} 0$ under $\mathcal{H}$. Replacing $\lambda(\cdot,\hat{\theta}_n)$ by $\lambda(\cdot,\theta) + \nabla_\theta \lambda(\cdot,\tilde{\theta})^T(\hat{\theta}_n - \theta)$, where $\tilde{\theta}$ is a point between $\hat{\theta}_n$ and θ, and applying Theorem 3.1 we obtain:

$$\sigma_0/\hat{\sigma}_{0n} \xrightarrow{\mathrm{P}} 1\ ,\ \xi_{n1} = O_{\mathsf{P}}\left(b_n^{1/2}\right) \text{ and } \xi_{n3} = O_{\mathsf{P}}\left((n b_n)^{-1/2}\right).$$

Consider now term ξ_{n2}: We have for all fixed $\varepsilon > 0$

$$\begin{aligned}\mathsf{P}\left(|\xi_{n2}| \geq \varepsilon\right) \leq \mathsf{P}\left(n^{1/2}\nu_n \mid \hat{\theta}_n - \theta \mid \geq \tau\right) \\ + \mathsf{P}\Bigg(\frac{2nb_n^{1/2}}{\sigma_0} \sup_{\theta' \in U_n(\theta)} \Big| \int \Big(\hat{\lambda}_n(t) - \tilde{\lambda}_n(t,\theta)\Big) \\ \times \Big(\tilde{\lambda}_n(t,\theta) - \tilde{\lambda}_n(t,\theta')\Big)\, a(t)\,\mathrm{d}t\Big| \geq \varepsilon\Bigg) \end{aligned} \tag{5.16}$$

with $U_n(\theta) = \{\theta' : |\theta - \theta'| \leq n^{-1/2}\, \nu_n^{-1}\, \tau\}$, where $\{\nu_n\}$ is a sequence of positive numbers tending to zero specified later, and $\tau > 0$ is fixed.

Since $\hat{\theta}_n$ is $\sqrt{n}$-consistent the first summand on the right hand side of (5.16) tends to zero. Divide the smallest cube containing $U_n(\theta)$ into l_n subcubes $\tilde{I}_r$ of equal volume and choose in each subcube $I_{nr}(\theta) = \tilde{I}_r \cap U_n(\theta)$ a point θ_r, $r = 1,\ldots,l_n$. Then the second summand on the right hand side of (5.16) is bounded by

$$\begin{aligned}&\mathsf{P}\Big(\frac{2nb_n^{1/2}}{\sigma_0} \max_{1\leq r\leq l_n} \sup_{\theta' \in I_{nr}(\theta)} \Big|\int (\hat{\lambda}_n(t) - \tilde{\lambda}_n(t,\theta))(\tilde{\lambda}_n(t,\theta_r) - \tilde{\lambda}_n(t,\theta'))\, a(t)\,\mathrm{d}t\Big| \geq \frac{\varepsilon}{2}\Big) \\ &+ \mathsf{P}\Big(\frac{2nb_n^{1/2}}{\sigma_0} \max_{1\leq r\leq l_n} \Big|\int (\hat{\lambda}_n(t) - \tilde{\lambda}_n(t,\theta))(\tilde{\lambda}_n(t,\theta) - \tilde{\lambda}_n(t,\theta_r))\, a(t)\,\mathrm{d}t\Big| \geq \frac{\varepsilon}{2}\Big). \end{aligned} \tag{5.17}$$

Using the Cauchy-Schwartz inequality and Theorem 2.1 we get for the first summand of (5.17)

$$\begin{aligned}&\Big(\int (\hat{\lambda}_n(t) - \tilde{\lambda}_n(t,\theta))(\tilde{\lambda}_n(t,\theta_r) - \tilde{\lambda}_n(t,\theta'))\, a(t)\,\mathrm{d}t\Big)^2 \\ &\leq \int (\hat{\lambda}_n(t) - \tilde{\lambda}_n(t,\theta))^2 a(t)\,\mathrm{d}t \int (\tilde{\lambda}_n(t,\theta_r) - \tilde{\lambda}_n(t,\theta'))^2\, a(t)\,\mathrm{d}t \\ &= O_{\mathsf{P}}\left((nb_n)^{-1}\right) \cdot O((|\theta' - \theta_r|)^2) = O_{\mathsf{P}}\left((nb_n)^{-1}\right) \cdot O\left(l_n^{-2/k}\nu_n^{-2} n^{-1}\right). \end{aligned}$$

Hence,

$$\begin{aligned}&nb_n^{1/2} \max_{1\leq r\leq l_n} \sup_{\theta' \in I_{nr}(\theta)} \Big| \int (\hat{\lambda}_n(t) - \tilde{\lambda}_n(t,\theta))(\tilde{\lambda}_n(t,\theta_r) - \tilde{\lambda}_n(t,\theta'))\, a(t)\,\mathrm{d}t \\ &= O_{\mathsf{P}}\left(l_n^{-1/k}\nu_n^{-1}\right). \end{aligned} \tag{5.18}$$

Consider the second summand of (5.17). With the notation introduced in the proof of Theorem 2.1 we have

$$\begin{aligned}&\int (\hat{\lambda}_n(t) - \tilde{\lambda}_n(t,\theta))(\tilde{\lambda}_n(t,\theta) - \tilde{\lambda}_n(t,\theta_r))\, a(t)\,\mathrm{d}t \\ &= \int V_{n1}(t)(\tilde{\lambda}_n(t,\theta) - \tilde{\lambda}_n(t,\theta_r))\, a(t)\,\mathrm{d}t + R_n(\theta,\theta_r),\end{aligned}$$

where the remainder term $R_n(\theta, \theta_r)$ can be estimated by the Cauchy-Schwartz inequality and Lemma 5.1 as follows:

$$\begin{aligned} nb_n^{1/2}|R_n(\theta,\theta_r)| &\le nb_n^{1/2} \left(\int (V_n(t) - V_{n1}(t))^2 a(t)\,\mathrm{d}t \right)^{1/2} \\ &\qquad\qquad \times \left(\int (\tilde{\lambda}_n(t,\theta) - \tilde{\lambda}_n(t,\theta_r))^2 a(t)\,\mathrm{d}t \right)^{1/2} \\ &= O_{\mathsf{P}}\left(b_n \nu_n^{-1} (\log n)^{1/2} \right). \end{aligned} \tag{5.19}$$

Using the covariance function $C_n(t,s)$ one shows that

$$\mathsf{Var} \int V_{n1}(t)(\tilde{\lambda}_n(t,\theta) - \tilde{\lambda}_n(t,\theta_r))\, a(t)\,\mathrm{d}t \;=\; O(n^{-2}\nu_n^{-2})$$

and get by the Chebyshev inequality

$$\begin{aligned} &\mathsf{P}\Big(nb_n^{1/2} \max_{1\le r\le l_n} |\int (\hat{\lambda}_n(t) - \tilde{\lambda}_n(t,\theta))(\tilde{\lambda}_n(t,\theta) - \tilde{\lambda}_n(t,\theta_r))\, a(t)\,\mathrm{d}t| \ge \varepsilon' \Big) \\ &\le \sum_r \mathsf{P}\Big(nb_n^{1/2} |\int (\hat{\lambda}_n(t) - \tilde{\lambda}_n(t,\theta))(\tilde{\lambda}_n(t,\theta) - \tilde{\lambda}_n(t,\theta_r))\, a(t)\,\mathrm{d}t| \ge \varepsilon' \Big) \\ &= O\left(l_n \cdot b_n \cdot \nu_n^{-2} \right), \end{aligned} \tag{5.20}$$

By a suitable choice of ν_n and l_n (for example $\nu_n = b_n^{1/(3(k+3))}$ and $l_n = [\nu_n^3 b_n^{-1}]$) equations (5.18), (5.19) and (5.20) imply $\xi_{n2} \xrightarrow{\mathrm{P}} 0$ and the proof of Consequence 3.1 is complete.

□

Proof of Theorem 4.1: Similarly as in the proof of Consequence 3.1 we have

$$\frac{nb_n^{1/2}\left(\hat{Q}_{0n} - \hat{m}_{0n}\right)}{\hat{\sigma}_{0n}} = \frac{\sigma_n^*}{\hat{\sigma}_{0n}} \left(\frac{nb_n^{1/2}\left(Q_n^* - m_n^*\right)}{\sigma_n^*} + \sum_{j=4}^{6} \xi_{nj} \right)$$

where

$$\xi_{n4} = \frac{nb_n^{1/2}}{\sigma_n^*} \int \left(\tilde{\lambda}_n^*(t) - \tilde{\lambda}_n(t,\hat{\theta}_n) \right)^2 a(t)\,\mathrm{d}t,$$

$$\xi_{n5} = \frac{2nb_n^{1/2}}{\sigma_n^*} \int \left(\hat{\lambda}_n(t) - \tilde{\lambda}_n^*(t) \right) \left(\tilde{\lambda}_n^*(t) - \tilde{\lambda}_n(t,\hat{\theta}_n) \right) a(t)\,\mathrm{d}t$$

and

$$\xi_{n6} = \frac{nb_n^{1/2}}{\sigma_n^*} \left(m_n^* - \hat{m}_{0n} \right).$$

Since Theorem 2.1 holds also under the alternative, that is $nb_n^{1/2}/\sigma_n^*(Q_n^* - m_n^*)$ is asymptotically normally distributed, it remains to examine the terms ξ_{nj}. First consider ξ_{n4}. By assumption (4.3) and the continuity of d and $\nabla_\theta \lambda$ one obtains

$$\begin{aligned}
\xi_{n4} &= \frac{nb^{1/2}}{\sigma_n^*} \int \Big(\varepsilon_n\, d(t) - \nabla_\theta \lambda(t,\theta)^t (n^{-1/2} Z_n \,+\, \varepsilon_n S)\Big)^2\, a(t)\,\mathrm{d}t\,(1 + o_{\mathsf{P}}\,(1)) \\
&= \frac{1}{\sigma_n^*} \int \Big(n^{1/2} b_n^{1/4} \varepsilon_n\,(d(t) \,-\, \nabla_\theta \lambda(t,\theta)^t S) \\
&\qquad\qquad - b_n^{1/4}\, \nabla_\theta \lambda(t,\theta)^t Z_n)\Big)^2\, a(t)\,\mathrm{d}t\,(1 + o_{\mathsf{P}}\,(1)).
\end{aligned}$$

For ξ_{n6} we get with the same expansions

$$\begin{aligned}
\xi_{n6} &= \frac{b_n^{-1/2}}{\sigma_n^*}\,\kappa_1 \int \frac{\varepsilon_n\, d(t) - \nabla_\theta \lambda(t,\theta)^t (n^{-1/2} Z_n \,+\, \varepsilon_n\, S)}{1 - H(t)}\, a(t)\,\mathrm{d}t\;(1 + o_{\mathsf{P}}\,(1)) \\
&= \frac{b_n^{-1/2} \varepsilon_n}{\sigma_n^*} \kappa_1 \int \frac{d(t) \,-\, \nabla_\theta \lambda(t,\theta)^t S}{1 - H(t)}\, a(t)\,\mathrm{d}t\;(1 + o_{\mathsf{P}}\,(1)).
\end{aligned}$$

And since $\sigma_n^*/\sigma_0 \to 1$ we have

$$\begin{aligned}
\xi_{n4} + \xi_{n6} = \sigma_0^{-1} \Big(&\varepsilon_n^2 n b_n^{1/2} \int (d(t) - \nabla_\theta \lambda(t,\theta)^t S)^2\, a(t)\,\mathrm{d}t \\
&+ \varepsilon_n\, b_n^{-1/2} \kappa_1 \int \frac{d(t) \,-\, \nabla_\theta \lambda(t,\theta)^t S}{1 - H(t)}\, a(t)\,\mathrm{d}t\Big)\,(1 + o_{\mathsf{P}}\,(1)).
\end{aligned}$$

It remains to investigate the term ξ_{n5}. With the notation used in the proof of Consequence 3.1 it can be written as

$$\xi_{n5} = \frac{nb_n^{1/2}}{\sigma_n^*} \int V_n^*(t)\, \Big(\tilde{\lambda}_n^*(t) \,-\, \tilde{\lambda}_n(t,\hat{\theta}_n)\Big)\, a(t)\,\mathrm{d}t$$

and decomposed into

$$\xi_{n5}^{(1)} = \frac{nb_n^{1/2}}{\sigma_n^*} \int V_{n1}^*(t)\, \Big(\tilde{\lambda}_n^*(t) \,-\, \tilde{\lambda}_n(t,\hat{\theta}_n)\Big)\, a(t)\,\mathrm{d}t \tag{5.21}$$

$$\xi_{n5}^{(2)} = \frac{nb_n^{1/2}}{\sigma_n^*} \int (V_n^*(t) - V_{n1}^*(t))\, \Big(\tilde{\lambda}_n^*(t) \,-\, \tilde{\lambda}_n(t,\hat{\theta}_n)\Big)\, a(t)\,\mathrm{d}t. \tag{5.22}$$

Further, the term (5.21) is the sum of

$$\xi_{n5}^{(11)} = \frac{nb_n^{1/2}\varepsilon_n}{\sigma_n^*} \int V_{n1}^*(t) \int K(s)\, \big(\nabla_\theta \lambda(t - sb_n,\theta)^t\, S \,+\, d(t - sb_n)\big)\, \mathrm{d}s\, a(t)\,\mathrm{d}t \\ \times (1 + o_{\mathsf{P}}\,(1))$$

and

$$\xi_{n5}^{(12)} = \frac{nb_n^{1/2}\varepsilon_n}{\sigma_n^*} \int V_{n1}^*(t) \int K(s)\, \Big(\nabla_\theta \lambda(t - sb_n,\theta)^t\, (\hat{\theta}_n - \theta - S_n)\Big)\, \mathrm{d}s\, a(t)\,\mathrm{d}t \\ \times (1 + o_{\mathsf{P}}\,(1))$$

where S_n is the deterministic part of the expansion given by (4.3), i.e. $S_n = \hat{\theta}_n - \theta - n^{-1/2} Z_n$.
The leading term of the expression $\xi_{n5}^{(11)}$ is a sum of i.i.d. r.v.'s. Using the formula for the covariance function given in (5.12) we get that this term has a variance of order $O\left(nb_n\varepsilon_n^2\right)$. Thus we have

$$\xi_{n5}^{(11)} \; = \; O_{\mathsf{P}}\left((nb_n)^{1/2}\varepsilon_n\right) \; = \; O_{\mathsf{P}}\left((nb_n^{1/2}\varepsilon_n^2)^{1/2} b_n^{1/4}\right). \tag{5.23}$$

Applying the same method of proof as in the investigation of term ξ_{n2} in the proof of Consequence 3.1 we get

$$\xi_{n5}^{(12)} \xrightarrow{\mathrm{P}} 0. \tag{5.24}$$

The difference in the approach here is that instead of the neighborhood $U_n(\theta)$ the neighborhood $\{\theta' \,:\, |\theta' - \theta - S_n| \leq n^{-1/2}\nu_n^{-1}\tau\}$ has to be chosen. For the leading term in (5.22) we get by Lemma 5.1

$$\xi_{n5}^{(1)} = O_{\mathsf{P}}\left((n\log n)^{1/2} b_n \varepsilon_n\right) \; = \; O_{\mathsf{P}}\left((nb_n^{1/2}\varepsilon_n^2)^{1/2} b_n^{3/4} \log n)^{1/2}\right). \tag{5.25}$$

From (5.23) to (5.25) it follows that the term ξ_{n5} is always asymptotically dominated by the term ξ_{n4} and we have

$$\lim_n \beta_n \; = \; \lim_n \left(\Phi(c_n - z_\alpha)\right)$$

with

$$c_n = \sigma_0^{-1}\left(nb_n^{1/2}\varepsilon_n^2 \int \left(d(t) - S^t \nabla_\theta \lambda(t,\theta)\right)^2 a(t)\,\mathrm{d}t \right.$$
$$\left. + \kappa_1 b_n^{-1/2}\varepsilon \int \left(d(t) - S^t \nabla_\theta \lambda(t,\theta)\right) a(t)\,\mathrm{d}t\right).$$

Under the conditions formulated in Theorem 4.1 the limit of c_n is nonnegative, thus $\lim_n \beta_n \geq \alpha$. If $nb_n^{1/2}\varepsilon_n^2 \to 0$, then $c_n \to 0$ and $\beta_n \to \alpha$. If $nb_n^{1/2}\varepsilon_n^2 \to \infty$, then $c_n \to \infty$ and $\beta_n \to 1$. And finally, if $nb_n^{1/2}\varepsilon_n^2 \to c^* > 0$, then

$$\lim_n c_n \; = \; \sigma_0^{-1}\, c^* \int \left(d(t) - S^t \nabla_\theta \lambda(t,\theta)\right)^2 a(t)\,\mathrm{d}t > 0$$

and $\beta_n \to \beta > \alpha$.

□

References

[1] S. Diehl and W. Stute. Kernel density and hazard function estimation in the presence of censoring. *J. Multivariate Anal.*, 25:299–310, 1988.

[2] P. Hall. Central limit theorem for integrated square error of multivariate nonparametric density estimators. *J. Multivariate Analysis*, 14:1–16, 1984.

[3] H. Liero. Global Measures of Deviation of Nonparametric Curve Estimators, 1999. Habilitationschrift Universität Potsdam.

[4] N. D. Singpurwalla and M. Y. Wong. Estimation of the failure rate - a survey of nonparametric methods, part I: Non- Baysian methods. *Commun. Statist.- Theory and Meth.*, 12:559–588, 1983.

[5] M. A. Tanner and W. H. Wong. The estimation of the hazard function from randomly censored data by the kernel method. *Ann. Statist.*, 11:989–993, 1983.

[6] G. S. Watson and M. R. Leadbetter. Hazard analysis I. *Biometrika*, 51:175–184, 1964.

[7] G. S. Watson and M. R. Leadbetter. Hazard analysis II. *Sankya Ser. A*, 26:110–116, 1964.

[8] H. Witting and U. Müller-Funk. *Mathematische Statistik II.* B. G. Teubner Stuttgart, 1995.