JOURNAL OF MULTIVARIATE ANALYSIS **62**, 273–309 (1997)
ARTICLE NO. MV971692

Multivariate Hazard Rates under Random Censorship

Jean-David Fermanian

INSEE–CREST, Malakoff Cedex, France

The data consists of multivariate failure times under right random censorship. By the kernel smoothing technique, convolutions of cumulative multivariate hazard functions suggest estimators of the so-called multivariate hazard functions. We establish strong i.i.d. representations and uniform bounds of the remainder terms on some compact sets of the underlying space. Thus asymptotic normality and uniform consistency on such sets are obtained. The asymptotic mean squared error gives an optimal bandwidth by the plug-in method. Simulations assess the performance of our estimators.

1. INTRODUCTION

In survival analysis, hazard functions characterize distributions of lifetimes and are precious tools for describing the instantaneous probability of occurrence of some conditional events. The non-parametric estimation of hazard functions was initiated by the works of Watson and Leadbetter (1964a, 1964b) and Rice and Rosenblatt (1976). Two types of estimators of the univariate hazard function were proposed and their generalizations to right censored data generated a large amount of literature: for the most recent papers, see Lo *et al.* (1989), Gijbels and Wang (1993), or Xiang (1994).

However, when we work with vectors of lifetimes independently right censored by vectors of failures, no estimators of hazard functions have ever been studied in the literature (to our knowledge). We propose a kernel estimator for the multivariate hazard function by smoothing the cumulative hazard function. Using this estimator, we will be able to evaluate instantaneous probabilities of occurrence of some durations at different dates, given that other events did not occur yet.

Formally, define $\mathbf{T} = (T_1, ..., T_d)$ to be a d-vector of positive variables (called failure times or lifetimes) with a continuous survival function

Received April 24, 1996; revised March 25, 1997.

AMS 1980 subject classifications: 62G05, 62E20.

Key words and phrases: Hazard functions, multidimensional lifetimes, right-censoring, kernel estimators, simulations.

0047-259X/97 $25.00

$F(\mathbf{x}) = P(\mathbf{T} > \mathbf{x}) = P(T_1 > x_1, ..., T_d > x_d)$ and a vector of censoring times $\mathbf{C} = (C_1, ..., C_d)$ independent from $\mathbf{T}$ whose survival function is G.

The data is thus an i.i.d. sample $((\mathbf{X}_1, \vec{\delta}_1, ..., (\mathbf{X}_n, \vec{\delta}_n))$ where $\mathbf{X} = (X_1, ..., X_d)$ with $X_i = T_i \wedge C_i$ and $\vec{\delta} = (\delta_1, ..., \delta_d)$ with $\delta_i = \mathbf{1}\{T_i \leqslant C_i\}$. Here, $\mathbf{1}\{\ \}$ will be the indicator function of an event. Note that the survival function of $\mathbf{X}$ is $H = F.G$ and that the single censoring case is included in our model.

For example, when $d = 2$, T_1 could be the duration of the spell in unemployment, T_2 the duration between the entry in the unemployed state and the starting of a serious disease, and C_1 (resp. C_2) could be the data of attrition in the files of the Employment Agency (resp. the hospital).

In the univariate case ($d = 1$) define the cumulative hazard function

$$\Lambda(x) = -\int \mathbf{1}\{t \leqslant x\}\ F(dt)/F(t-)$$

Let us assume that T has a density f. Then $\Lambda(x) = \int \mathbf{1}\{t \leqslant x\}\ f(t)/F(t)\ dt = -\ln(F(x))$ and the hazard function λ is defined for all x such that $F(x) > 0$ by

$$\lambda(x) = f(x)/F(x).$$

With right censored data, there are two nonparametric main approaches to estimate the hazard function: First, we can smooth the "natural" estimator of the survival function F, i.e., the Kaplan–Meier estimator (1958). We obtain

$$\lambda_n^{(1)}(x) = -\int K_h(x-u)\ \Gamma_n(du)/\Gamma_n(x)$$

with Γ_n the (slightly modified) Kaplan–Meier estimator. Lo *et al.* (1989) have obtained asymptotic properties via strong representations of $\lambda_n^{(1)}$. More recently, Xiang (1994) has established strong uniform consistency.

Second, we can smooth the "natural" estimator of the cumulative hazard function, i.e. the Nelson–Aalen estimator of Λ. So we obtain

$$\lambda_n^{(2)}(x) = \sum_{j=1}^{n} K_h(x - X_{(j)})\ \mathbf{1}\{\delta_{(j)} = \mathbf{1}\}/(n-j+1).$$

The properties of $\lambda_n^{(2)}$ can be found in Tanner and Wong (1983) (variance, asymptotic normality), Ramlau-Hansen (1983) (uniform consistency), and Yandell (1983) (strong approximations, confidence bounds).

In the bivariate (and more generally multivariate) case, we can try to adapt the univariate methods, but the best estimators of the survival function in these larger dimensions are more complicated than Γ_n, with the exception of the one proposed by Lin and Ying (1993) when $C_1 = C_2$; the most successful were probably proposed by Dabrowska (1988, 1989) and Prentice and Cai (1992) (see Gill, Van der Laan and Wellner (1993) for a recent analysis). See also Campbell (1981), Campbell and Földes (1982), Landberg and Shaked (1982), Hanley and Parnes (1983), Pruitt (1991) and Tsai, Leurgans and Crowley (1986).

This explains why we prefer the estimator $\lambda_n^{(2)}$ which smoothes the multivariate Nelson–Aalen estimator.

First, let us precise some notations. Orderings on $\mathbb{R}^d$ are defined coordinate by coordinate; for example, $\mathbf{u} \leqslant \mathbf{x}$ if for all $i \in \{1, ..., d\}$, $u_i \leqslant x_i$. For each subset A of $\{1, ..., d\}$, each vector $v \in \mathbb{R}^d$, the vector of $\mathbb{R}^{|A|}$ whose coordinates are $(v_k, k \in A)$ is denoted v_A. In the remainder of the paper, we fix a subset of $\{1, ..., d\}$ and we denote r its cardinal and J its complement.

If the domain of integration is not specified, it is $\mathbb{R}^r$. The bandwidth sequence $(h_n)_{n \geqslant 0}$ is such that $h_n \to 0$ and $nh_n^r \to +\infty$.

Let the following distribution functions and their empirical counterparts be

$$S_I(\mathbf{x}) = P(\mathbf{X} > \mathbf{x}, \vec{\delta}_I = \mathbf{1}_I) \qquad \hat{S}_I(\mathbf{x}) = n^{-1} \sum_{i=1}^{n} \mathbf{1}\{\mathbf{X}_i > \mathbf{x}, \vec{\delta}_{iI} = \mathbf{1}_I\}$$

Note that $H = S_\varnothing$. Consider the subset $\tau = \prod_{k=1}^{d} [0, \tau_k] \subset \mathbb{R}^d$ such that there exist $\varepsilon > 0$ and $H(\tau_1 + \varepsilon, ..., \tau_d + \varepsilon) > 0$. Define also the I-cumulative hazard function on $\mathbf{x}$ such that $H(\mathbf{x}) > 0$ by the r-dimensional Lebesgue-Stieltjes integral (see Fermanian (1996) for theoretical justifications)

$$\Lambda_I(x) = (-1)^r \int \mathbf{1}\{\mathbf{u}_I \leqslant \mathbf{x}_I\} \frac{S_I(d\mathbf{u}_I, \mathbf{x}_J)}{H(\mathbf{u}_I -, \mathbf{x}_J)} \tag{1.1}$$

Obviously, the kth coordinate of $(\mathbf{u}_I, \mathbf{x}_J)$ is u_k (resp. x_k) if $k \in I$ (resp. $k \in J$). Its natural estimator is then

$$\hat{\Lambda}_I(\mathbf{x}) = (-1)^r \int \mathbf{1}\{\mathbf{u}_I \leqslant \mathbf{x}_I\} \frac{\hat{S}_I(d\mathbf{u}_I, \mathbf{x}_J)}{\hat{H}(\mathbf{u}_I -, \mathbf{x}_J)} \tag{1.2}$$

$$= \sum_{i=1}^{n} \frac{\mathbf{1}\{\mathbf{X}_{iI} \leqslant \mathbf{x}_I, \mathbf{X}_{iJ} > \mathbf{x}_J, \vec{\delta}_{iI} = \mathbf{1}_I\}}{\sum_{j=1}^{n} \mathbf{1}\{\mathbf{X}_{jI} \geqslant \mathbf{X}_{iI}; \mathbf{X}_{jJ} > \mathbf{x}_J\}} \tag{1.3}$$

Suppose that the density f of $\mathbf{T}$ exists. For each I subset of $\{1, ..., d\}$ and each $\mathbf{x} \in \tau$, define the hazard function (or hazard rate) at point $\mathbf{x}$ such that $F(\mathbf{x}) > 0$ by

$$\lambda_I(\mathbf{x}) = \lim_{h_i \to 0\, \forall i \in I} \left(\prod_{i \in I} h_i \right)^{-1} P\left(\bigcap_{i \in I} T_i \in [x_i, x_i + h_i] \,\Big|\, T_I \geqslant \mathbf{x}_I, T_J > \mathbf{x}_J \right)$$

$$= \frac{1}{F(\mathbf{x}_I -, \mathbf{x}_J)} \int f(\mathbf{x}_I, \mathbf{u}_J)\, \mathbf{1}\{\mathbf{u}_J > \mathbf{x}_J\}\, d\mathbf{u}_J$$

Notice that

$$\Lambda_I(d\mathbf{x}_I, \mathbf{x}_J) = (-1)^r \frac{S_I(d\mathbf{x}_I, \mathbf{x}_J)}{H(\mathbf{x}_I -, \mathbf{x}_J)} = (-1)^r \frac{F(d\mathbf{x}_I, \mathbf{x}_J)}{F(\mathbf{x}_I -, \mathbf{x}_J)} = \lambda_I(\mathbf{x})\, d\mathbf{x}_I$$

Remark 1.1. Some authors consider a so-called "bivariate cumulative hazard function" defined by $\Lambda = -\ln(F)$ (see Campbell and Földes (1982) or Ruymgaart (1989)). In our opinion, this function is not well-adapted to our problem, because its derivatives (i.e. $\partial^2_{1,2}$) cannot be interpreted as instantaneous probabilities.

Consider a kernel K: $\mathbb{R}^r \to \mathbb{R}$ i.e. a bounded Lebesgue-integrable function with integral 1. Possible assumptions on the kernel K are:

$(K1)$ K is compactly supported with support $[-\mathbf{A}, \mathbf{A}] = \prod_{i=1}^r [-A_i, A_i]$;

$(K2)$ K is of bounded variation;

$(K3)$ K is symmetrical;

$(K4)$ K is lipschitzian.

For each kernel K on $\mathbb{R}^r$, denote K_h the function:

$$K_h: \mathbb{R}^r \to \mathbb{R}$$

$$(x_1, ..., x_r) \to K(x_1/h, ..., x_r/h)/h^r$$

$\lambda_I(\mathbf{x})$ is estimated by

$$\hat{\lambda}_I(\mathbf{x}) = \int K_h(\mathbf{x}_I - \mathbf{u}_I)\, \hat{\Lambda}_I(d\mathbf{u}_I, \mathbf{x}_J) \tag{1.4}$$

$$= \sum_{i=1}^n K_h(\mathbf{x}_I - \mathbf{X}_{iI}) \frac{\mathbf{1}\{\vec{\delta}_{iI} = \mathbf{1}_I, \mathbf{X}_{iJ} > \mathbf{x}_J\}}{\sum_{j=1}^n \mathbf{1}\{\mathbf{X}_{jI} \geqslant \mathbf{X}_{iI}, \mathbf{X}_{jJ} > \mathbf{x}_J\}} \tag{1.5}$$

Thus, when $d=2$, $\lambda_{1,2}(s,t)$ is the instantaneous rate of double failure at the point (s,t), given that the first failure time had not been observed until date s and the second one until t. Further, $\lambda_1(s,t)$ (resp. $\lambda_2(s,t)$) represents the rate of a single rate failure at time s (resp. t) given the same conditions.

Our first goal is to give an expansion of $\hat{\Lambda}_I$ (resp. $\hat{\lambda}_I$) as a sum of i.i.d. random variables and a remainder term which has a certain rate uniformly on τ. This will allow us to establish consistency and asymptotic properties of $\hat{\lambda}_I$. The performance of our estimators will be tested using simulations. This method provides a powerful way to study the rate of convergence of many estimators, because classical laws of large numbers, central limit theorems and sometimes the theory of empirical processes can be applied to the i.i.d. random variable in the expansion. Here, it can be used e.g. to study an asymptotically optimal bandwidth selector (see Fermanian (1996)).

We will use some classical techniques of discretization instead of the more modern framework of multivariate counting processes. Indeed, this last theory suffers from a lack of powerful and practical tools which are now commonly used in the univariate case; particularly, to our knowledge, there exists no functional central limit theorem for weak martingales that could play the part of Reboleddo's theorem. See Pons (1986) for some results in the bivariate case and Andersen *et al.* (1993) for a discussion.

2. RESULTS

Our proofs are given in the appendix. First, we state a generalization in $\mathbb{R}^d$ of some well-known strong expansions of the univariate cumulative hazard function where they give some expansions of the product-limit estimator, the mainly studied point: Lo and Singh (1986), Major and Rejtö (1988). See brief surveys of the topic in Gijbels and Wang (1993) and Stute (1994). To use some integrations by parts, suppose that the total variation of $1/H(\cdot,\mathbf{x}_J)$ on τ_I is finite and bounded uniformly in $\mathbf{x}_J\in\tau_J$. Note that this is true if H is C^d on τ or more generally if, for all $\mathbf{x}_J\in\tau_J$, the density (with respect to the Lebesgue measure) of X_I with respect to $\{X_J>\mathbf{x}_J\}$ exists on τ_I and is bounded uniformly on τ.

THEOREM 2.1. *For all $\mathbf{x}\in\tau$ we have the decomposition*

$$\hat{\Lambda}_I(\mathbf{x})-\Lambda_I(\mathbf{x})=\frac{1}{n}\sum_{i=1}^{n}\eta_{iI}(\mathbf{x})+r_n(\mathbf{x}) \tag{2.1}$$

with

$$\eta_{iI}(\mathbf{x}) = \frac{\mathbf{1}\{\vec{\delta}_{iI} = \mathbf{1}_I, \mathbf{X}_{iI} \leqslant \mathbf{x}_I, \mathbf{X}_{iJ} > \mathbf{x}_J\}}{H(\mathbf{X}_{iI}-, \mathbf{x}_J)} + (-1)^{r+1} \int \mathbf{1}\{\mathbf{X}_{iI} \wedge \mathbf{x}_I \geqslant \mathbf{u}_I, \mathbf{X}_{iJ} > \mathbf{x}_J\} \frac{S_I(d\mathbf{u}_I, \mathbf{x}_J)}{H^2(\mathbf{u}_I-, \mathbf{x}_J)} \tag{2.2}$$

and, if S_I and H are continuous on τ, then

$$\sup_{\mathbf{x} \in \tau} |r_n(\mathbf{x})| = O(n^{-1} \ln n) \qquad a.e. \tag{2.3}$$

Remark 2.1. (a) If $J = \varnothing$, the result has been stated under the weaker hypothesis: the r.v. T_i are continuous on τ_i for all $i = 1, ..., d$.

(b) For simplicity as sake, we suppose the continuity of S_I and H. This hypothesis can be removed: see the discussion in Dabrowska (1989, p. 314) and Van Zuijlen (1978).

Proposition 2.1. *For all $(\mathbf{x}, \mathbf{y}) \in \mathbb{R}^{2d}$ such that $H(\mathbf{x})\, H(\mathbf{y}) > 0$, for all i in $\{1, ..., n\}$, the r.v. $\eta_{iI}(\mathbf{x})$ is centered and we have*

$$\begin{aligned}
&\operatorname{Cov}(\eta_{iI}(\mathbf{x}), \eta_{iI}(\mathbf{y})) \\
&\stackrel{\text{def}}{=} \mathscr{C}(\mathbf{x}, \mathbf{y}) \\
&= (-1)^r \int \frac{\mathbf{1}\{\mathbf{u}_I \leqslant \mathbf{x}_I \wedge \mathbf{y}_I\}\, S_I(d\mathbf{u}_I, \mathbf{x}_J \vee \mathbf{y}_J)}{H(\mathbf{u}_I-, \mathbf{x}_J)\, H(\mathbf{u}_I-, \mathbf{y}_J)} \\
&\quad - \int \frac{S_I(d\mathbf{u}_I, \mathbf{x}_J)\, S_I(d\mathbf{v}_I, \mathbf{x}_J \vee \mathbf{y}_J)}{H^2(\mathbf{u}_I-, \mathbf{x}_J)\, H(\mathbf{v}_I-, \mathbf{y}_J)} \mathbf{1}\{\mathbf{u}_I \leqslant \mathbf{x}_I, \mathbf{u}_I \leqslant \mathbf{v}_I \leqslant \mathbf{y}_I\} \\
&\quad - \int \frac{S_I(d\mathbf{u}_I, \mathbf{y}_J)\, S_I(d\mathbf{v}_I, \mathbf{x}_J \vee \mathbf{y}_J)}{H^2(\mathbf{u}_I-, \mathbf{y}_J)\, H(\mathbf{v}_I-, \mathbf{x}_J)} \mathbf{1}\{\mathbf{u}_I \leqslant \mathbf{y}_I, \mathbf{u}_I \leqslant \mathbf{v}_I \leqslant \mathbf{x}_I\} \\
&\quad + \int \frac{S_I(d\mathbf{u}_I, \mathbf{x}_J)\, S_I(d\mathbf{v}_I, \mathbf{y}_J)}{H^2(\mathbf{u}_I-, \mathbf{x}_J)\, H^2(\mathbf{v}_I-, \mathbf{y}_J)} H(\mathbf{u}_I \vee \mathbf{v}_I-, \mathbf{x}_J \vee \mathbf{y}_J)\, \mathbf{1}\{\mathbf{u}_I \leqslant \mathbf{x}_I, \mathbf{v}_I \leqslant \mathbf{y}_I\}
\end{aligned}$$

Using an integration by parts, we deduce from Theorem 2.1.

Corollary 2.1. *Under the conditions of Theorem* 2.1,

$$\left(\frac{n}{\ln_2 n}\right)^{1/2} \sup_{\mathbf{x} \in \tau} |\hat{\Lambda}_I(\mathbf{x}) - \Lambda_I(\mathbf{x})| \qquad \text{is bounded a.e.} \tag{2.4}$$

Let $\mathscr{D}_d(\tau)$ be the generalization in $\tau \subset \mathbb{R}^d$ of the well-known cadlag function space: see Neuhaus (1971) or Bickel and Wichura (1971).

COROLLARY 2.2. *If the c.d.f. of X_k is continuous for all $k = 1,..., d$, then the sequence of $\mathscr{D}_d(\tau)$ valued processes $n^{1/2}(\hat{\Lambda}_I - \Lambda_I)$ converges weakly in the $\mathscr{S}$-topology (Skorohod topology) to a Gaussian process W such that, for all* $\mathbf{x}$ (resp. $\mathbf{y}$) *in τ, the r.v. $W(\mathbf{x})$ (resp. $W(\mathbf{y})$) is centered and*

$$\mathrm{Cov}(W(\mathbf{x}), W(\mathbf{y})) = \mathscr{C}(\mathbf{x}, \mathbf{y})$$

This last result was proved first by Pons (1986) when $d = 2$. Although the latter proposition could be proved using the functional delta-method (see e.g. Gill (1989) or Gill *et al.* (1993)), we prove this result through to more classical techniques, following Breslow and Crowley (1974) and the extensions of the weak convergence theory given by Bickel and Wichura (1971).

Suppose now the existence the density of $\mathbf{T}$ denoted f.

THEOREM 2.2. *For all $\mathbf{x} \in \tau$, we have the decomposition*

$$\hat{\lambda}_I(\mathbf{x}) - \lambda_I(\mathbf{x}) = (K_h * \lambda_I - \lambda_I)(\mathbf{x}) + n^{-1} \sum_{i=1}^{n} \xi_{iI}(\mathbf{x}) + R_n(\mathbf{x}) \qquad (2.5)$$

where, for all i,

$$\begin{aligned} \xi_{iI}(\mathbf{x}) &= \int K_h(\mathbf{x}_I - \mathbf{u}_I)\, \eta_{iI}(d\mathbf{u}_I, \mathbf{x}_J) \\ &= K_h(\mathbf{x}_I - \mathbf{X}_{iI}) \frac{\mathbf{1}\{\vec{\delta}_{iI} = \mathbf{1}_I, \mathbf{X}_{iJ} > \mathbf{x}_J\}}{H(\mathbf{X}_{iI}-, \mathbf{x}_J)} \\ &\quad + (-1)^{r+1} \int K_h(\mathbf{x}_I - \mathbf{u}_I) \frac{\mathbf{1}\{\mathbf{X}_{iI} \geqslant \mathbf{u}_I, \mathbf{X}_{iJ} > \mathbf{x}_J\}}{H^2(\mathbf{u}_I-, \mathbf{x}_J)} S_I(d\mathbf{u}_I, \mathbf{x}_J) \end{aligned}$$

and if

(i) *H is continuous on τ*

(ii) *for each k in J, the c.d.f. of C_k is γ-lipschizian on τ_k with $\gamma > 0$.*

(iii) *f_I, the density of T_I, is continuous on τ_I*

(iv) *K satisfies (K1)–(K2) and (K4)*

(v) *there exists $\varepsilon > 0$ such that $nh_n^{r(1+\varepsilon)}/\ln n \to \infty$,*

then

$$\sup_{\mathbf{x} \in \tau} |R_n(\mathbf{x})| = O(n^{-1} h_n^{-r/2} \ln n) \qquad a.e. \qquad (2.6)$$

Remark 2.2. (a) Restrict our assumptions to $(K1)$, $(K2)$, the continuity of H and S_I on τ; then the following upper bound is easily obtained using Theorem 2.1:

$$\sup_{\mathbf{x}\in\tau} |R_n(\mathbf{x})| = O(n^{-1}h_n^{-r}\ln n) \qquad \text{a.e.} \tag{2.7}$$

(b) Lemma 5.2 yields the pointwise upper bound of the remainder term: with the notations in the proof of Theorem 2.2, for all $\mathbf{x}\in\tau$, if f_I is continuous at $\mathbf{x}_I$ then $\tilde{R}_n(\mathbf{x}) = O(n^{-1}h_n^{-r/2}\ln n)$ and $\bar{R}_n(\mathbf{x}) = O(n^{-1}\ln_2 n)$ a.e. We obtain also:

$$R_n(\mathbf{x}) = O(n^{-1}h_n^{-r/2}\ln n) \qquad \text{a.e.} \tag{2.8}$$

(c) If J is empty, the conditions i and ii can be removed.

PROPOSITION 2.2. *If λ_I and H are continuous on τ, and if K satisfies $(K1)$ then, for all $i = 1, ..., n$ and $(\mathbf{x}, \mathbf{y})\in\tau^2$, $\xi_{iI}(\mathbf{x})$ is centered and, if $\mathbf{x}_I = \mathbf{y}_I$,*

$$\operatorname{Cov}(\xi_{iI}(\mathbf{x}), \xi_{iI}(\mathbf{y})) = h_n^{-r}\Phi(\mathbf{x}, \mathbf{y}) + o(h_n^{-r}) \tag{2.9}$$

where

$$\Phi(\mathbf{x}, \mathbf{y}) = \int K^2(\mathbf{u}_I)\, d\mathbf{u}_I \cdot \frac{H(\cdot -, \mathbf{x}_J \vee \mathbf{y}_J)\, \lambda_I(\cdot, \mathbf{x}_J \vee \mathbf{y}_J)}{H(\cdot -, \mathbf{x}_J)\, H(\cdot -, \mathbf{y}_J)}(\mathbf{x}_I). \tag{2.10}$$

Denote $Y_{iI}(\mathbf{x}_J) = \mathbf{1}\{\vec{\delta}_{iI} = \mathbf{1}_I, \mathbf{X}_{iJ} > \mathbf{x}_J\}\, H^{-1}(\mathbf{X}_{iI}-, \mathbf{x}_J)$.

Note that $\bar{\xi}(\mathbf{x}) = n^{-1}\sum_{i=1}^n \xi_{iI}(\mathbf{x})$ is the sum of two terms: $\bar{\xi}_1(\mathbf{x})$, the numerator of a kernel estimator in the regression of $Y_I(\mathbf{x}_J)$ on $\mathbf{X}_I$ (less its expectation), and $\bar{\xi}_2(\mathbf{x})$ which is of order $(n^{-1}\ln_2 n)^{1/2}$ a.e. uniformly on τ (if λ_I is bounded on τ). That is why well-known properties of regression kernel estimators can be obtained. Particularly

PROPOSITION 2.3. *Under the conditions* iii, iv, *if the c.d.f. of C_k is continuous on τ_k for each $k\in J$ and if $nh_n^r/\ln n\to\infty$, then*

$$\left(\frac{nh^r}{\ln n}\right)^{1/2} \sup_{\mathbf{x}\in\tau} |\bar{\xi}(\mathbf{x})| \qquad \textit{is bounded a.e.} \tag{2.11}$$

Notice that the previous result is not standard because the conditional distribution of $Y_I(\mathbf{x}_J)$ with respect to $\mathbf{X}_I$ is not continuous. See the proof in the appendix. Thanks to a first order expansion, we deduce

Proposition 2.4. *Under the conditions* iii, iv, *if G is continuous on τ and if $nh^r/\ln n \to +\infty$, then*

$$\left(\frac{nh^r}{\ln n}\right)^{1/2} \sup_{\mathbf{x}\in\tau} |\hat{\lambda}_I(\mathbf{x}) - E(\hat{\lambda}_I)(\mathbf{x})| \qquad \text{is bounded a.e.} \tag{2.12}$$

Denoting $\tau' = \prod_{i\in I} [\varepsilon_i, \tau_i] \times \tau_J$ with $\varepsilon_i > 0$ for all i, if moreover λ_I is uniformly continuous on τ, then

$$\left(\frac{nh^r}{\ln n}\right)^{1/2} \sup_{\mathbf{x}\in\tau'} |\hat{\lambda}_I(\mathbf{x}) - \lambda_I(\mathbf{x})| \qquad \text{is bounded a.e.} \tag{2.13}$$

We only need to consider a compact $\tau'_I \subset \tau_I$ because of the nonnullity of the bias when a component of $\mathbf{x}_I$ is zero.

Let $\Phi(\mathbf{x}) = \Phi(\mathbf{x}, \mathbf{x})$. The mean squared error (or MSE) is approximated in the following way.

Proposition 2.5. *If K satisfies ($K1$) and ($K3$), f_I is continuous on τ_I, $\lambda_I(\cdot, \mathbf{x}_J)$ is C^2 on τ_I for all $\mathbf{x}_J \in \tau_J$, $nh_n^r/\ln n \to \infty$, then*

$$E((\hat{\lambda}_I - \lambda_I)^2 (\mathbf{x})) = (nh_n^r)^{-1} \Phi(\mathbf{x}) + h_n^4 \chi^2(\mathbf{x})/4 + \bar{\varepsilon}_n(\mathbf{x}) \tag{2.14}$$

with $\sup_{\mathbf{x}_I \in \tau_I} |\bar{\varepsilon}_n(\mathbf{x})| = o((nh_n^r)^{-1} + h_n^4)$ *and*

$$\chi(\mathbf{x}) = \sum_{i\in I} \partial_i^2 \lambda_I(\mathbf{x}) \cdot \int v_i^2 K(\mathbf{v})\, d\mathbf{v}$$

If λ_I is C^2 on $\tau_I \times \bar{\tau}_J$ with $\bar{\tau}_J \subset \tau_J$, then the rest is uniform on $\tau_I \times \bar{\tau}_J$.

The asymptotic mean squared error at $\mathbf{x}$ is then $AMSE(\mathbf{x}) = (nh_n^r)^{-1} \Phi(\mathbf{x}) + h_n^4 \chi^2(\mathbf{x})/4$. The problem of the practical choice for h_n is classical: see e.g. Härdle (1990) or Scott (1992). By a plug-in method, the *AMSE* is optimized for the following choice of h_n:

$$h_n^* = \left(\frac{r\Phi(\mathbf{x})}{n\chi^2(\mathbf{x})}\right)^{1/(4+r)} \tag{2.15}$$

With this choice, the *AMSE* is of order $n^{-4/(4+r)}$. Note that $\Phi(\mathbf{x})$ can be estimated by

$$\hat{\Phi}(\mathbf{x}) = \int K^2 \cdot \hat{\lambda}_I(\mathbf{x}) \cdot \frac{1}{\hat{H}(x)} \tag{2.16}$$

If $\partial_i^2 \lambda_I(\mathbf{x})$ is estimated, it is possible to estimate $\chi(\mathbf{x})$. An obvious choice is

$$\widehat{\partial_i^2 \lambda_I}(\mathbf{x}) = h_n^{-2} \int \partial_i^2 K_h(\mathbf{x}_I - \mathbf{u}_I)\, \hat{\Lambda}_I(d\mathbf{u}_I, \mathbf{x}_J) \tag{2.17}$$

Thus a two-step procedure could be used to approach the asymptotically optimal bandwidth: a first estimation of $\lambda_I(\mathbf{x})$ allows the estimation of $\Phi(\mathbf{x})$; then use in a second step

$$h_n^{**} = \left(\frac{r\hat{\Phi}(\mathbf{x})}{n\hat{\chi}^2(\mathbf{x})} \right)^{1/(4+r)} \tag{2.18}$$

Finally, the central-limit theorem for the kernel-regression yields:

PROPOSITION 2.6. *For all $\mathbf{x} \in \tau$ such that f_I is continuous at $\mathbf{x}_I$, if $h_n^r \ln_2 n \to 0$, and there exists $\varepsilon > 0$ such that $nh_n^{r(1+\varepsilon)} \to +\infty$, under the conditions of Proposition 2.2,*

$$(nh_n^r)^{1/2} (\hat{\lambda}_I(\mathbf{x}) - E(\hat{\lambda}_I(\mathbf{x}))) \xrightarrow{\text{law}} \mathcal{N}(0, \Phi(\mathbf{x})) \tag{2.19}$$

If moreover $\lambda_I(\cdot, \mathbf{x}_J)$ is C^2 in a neighborhood of $\mathbf{x}_I$, and if $nh_n^{r+4} = o(1)$, then

$$(nh_n^r)^{1/2} (\hat{\lambda}_I(\mathbf{x}) - \hat{\lambda}_I(\mathbf{x})) \xrightarrow{\text{law}} \mathcal{N}(0, \Phi(\mathbf{x})) \tag{2.20}$$

3. SIMULATIONS

We have selected a bivariate exponential distribution for (T_1, T_2):

$$F(x_1, x_2) = (\exp(x_1) + \exp(x_2) - 1)^{-1}\, \mathbf{1}\{x_1 \geqslant 0, x_2 \geqslant 0\}.$$

The censoring variables are independently exponentially distributed:

$$G(x_1, x_2) = \exp(-x_1/5)\exp(-x_2/5)\, \mathbf{1}\{x_1 \geqslant 0, x_2 \geqslant 0\}.$$

For each point we have made 100 simulations of size 1000. The bandwidth is equal to $0.3 \simeq n^{-1/6}$. The optimal asymptotic bandwidth is often larger than 1; this one gives unsatisfactory results since the hypothesis $h_n \to 0$ is obviously unrealistic. We need larger sample sizes to use the plug-in method and to obtain a more precise and optimal bandwidth. Taking computation time into consideration, we will restrict ourselves to this crude choice for h_n.

We have chosen the multidimensional product Epanechnikov's kernel because of its well-known good properties of optimality:

$$K(\mathbf{u}_I) = \left(\frac{3}{4\sqrt{5}}\right)^r \prod_{i \in I} \left(1 - \frac{u_i^2}{5}\right) \mathbf{1}\{u_i \in [-\sqrt{5}, \sqrt{5}]\}$$

In each box of the following arrays, the first number is the true value of the hazard function at the considered point; the second number is the mean of the simulations and the third one (into brackets) is the standard error of the simulations. The first array display the results for $\lambda_{12}(s, t)$ and the second one the results for $\lambda_1(s, t)$.

Results for $\lambda_{12}(s,t)$

s \ t	0.5	1	1.5	2	2.5	3
0.5	1.030 0.897 (0.056)	0.791 0.747 (0.065)	0.561 0.541 (0.065)	0.377 0.380 (0.069)	0.244 0.253 (0.065)	0.154 0.158 (0.075)
1	0.791 0.742 (0.057)	0.751 0.725 (0.072)	0.634 0.620 (0.086)	0.484 0.472 (0.085)	0.343 0.337 (0.096)	0.230 0.242 (0.103)
1.5	0.561 0.537 (0.067)	0.634 0.614 (0.085)	0.633 0.592 (0.101)	0.560 0.551 (0.100)	0.445 0.428 (0.127)	0.324 0.324 (0.133)
2	0.377 0.382 (0.077)	0.484 0.488 (0.086)	0.560 0.547 (0.111)	0.536 0.566 (0.125)	0.522 0.488 (0.142)	0.423 0.403 (0.165)
2.5	0.244 0.243 (0.072)	0.343 0.345 (0.094)	0.445 0.464 (0.120)	0.522 0.548 (0.189)	0.544 0.512 (0.216)	0.501 0.482 (0.241)
3	0.154 0.173 (0.080)	0.230 0.233 (0.097)	0.324 0.332 (0.146)	0.423 0.421 (0.248)	0.501 0.464 (0.229)	0.526 0.494 (0.276)

Results for $\lambda_1(s,t)$

s \ t	0.5	1	1.5	2	2.5	3
0.5	0.718 0.682 (0.042)	0.490 0.470 (0.039)	0.321 0.312 (0.046)	0.205 0.203 (0.049)	0.128 0.138 (0.050)	0.079 0.079 (0.055)
1	0.791 0.742 (0.057)	0.613 0.611 (0.065)	0.438 0.449 (0.072)	0.298 0.300 (0.076)	0.196 0.189 (0.071)	0.125 0.137 (0.081)
1.5	0.873 0.875 (0.075)	0.723 0.701 (0.085)	0.563 0.555 (0.083)	0.412 0.412 (0.091)	0.286 0.286 (0.093)	0.190 0.177 (0.085)
2	0.919 0.919 (0.105)	0.811 0.816 (0.116)	0.680 0.690 (0.120)	0.536 0.567 (0.124)	0.398 0.376 (0.124)	0.279 0.285 (0.156)
2.5	0.949 0.943 (0.146)	0.876 0.881 (0.172)	0.778 0.780 (0.160)	0.656 0.672 (0.172)	0.521 0.506 (0.169)	0.389 0.386 (0.184)
3	0.969 0.955 (0.198)	0.921 0.922 (0.231)	0.852 0.873 (0.244)	0.759 0.740 (0.254)	0.642 0.627 (0.288)	0.513 0.537 (0.296)

Notice the relatively good performances of these estimations as long as $P(T_1 > s, T_2 > t)$ is not too small. Changing the frequency of the censoring variables (the standard deviation particularly) degrades the result.

The larger are the values of s and t, the smaller λ (and the bias) and the larger is the variance. Because of this trade-off, the mean of our simulations is not necessarily worse for large values of the coordinates.

Other usual sample sizes have been used, particularly $n = 50$ and $n = 100$. With these choices, 100 and 1000 simulations have been made; we have used a crude bandwidth selector: multiply $n^{-1/(4+r)}$ by the empirical standard error of the simulated distribution. The conclusions are the same in every cases: the estimators of $\lambda_{12}(s, t)$ and $\lambda_1(s, t)$ suffer from a bias somewhere between -10 and -60 percent of the true value. This proportion is most of the time smaller for λ_1 than for λ_{12}. The result is much worse for large values of s and t, i.e. values larger than 2. This last point is certainly due to the lack of simulated points in the neighborhood of (s, t). Moreover, it follows from these small sample sizes that the standard errors are very often larger than in the case $n = 1000$; more precisely, they range from one to three times the previous standard errors unless s and t are larger than 2; in this case, smaller standard errors can occur because of the strong previous bias.

4. CONCLUDING REMARKS

Our estimators suffer from the well-known drawbacks of kernel estimators: we need large samples (at least several hundred observations in the bivariate case) to apply our asymptotic results. The larger the dimension of the model, the worse the results. But the parameter of interest is here r, i.e. the cardinality of I, that sets the number of simultaneous failure times. The other failure times are conditional variables; they do not affect the rate of convergence.

In practice, $r = 1$ or $r = 2$ are the only cases, because they alone can be displayed. Since the results of our estimations are hypersurfaces, only sections can be displayed for larger r. Nevertheless, the problems of convergence in dimensions larger than 2 are essentially theoretical.

An important improvement on our results would be the removal of the hypothesis on C in Theorem 2.2. The existence of a density for a censoring variable is often not satisfied in surveys (consider the end of the observations in a panel data for example). It seems to us that, if λ_I has a finite number of discontinuities of the first kind, an "uniform Bochner's lemma" is still valid, and our main results remain true in this framework (but suffer from an heavier presentation: see technical details in Major and Rejtö (1988)).

The study would be completed if we had $\hat{\lambda}_I$'s rate of convergence. In our approach, a law of the iterated logarithm for multivariate kernel regression seems to be the relevant tool. But, although convergence rates are known for multivariate kernel density estimation (see Stute (1982) or a nonstandard approach in Deheuvels and Mason (1992)), it is not yet the case for the regression.

The previous estimators can be used when there exists a dependence between recurrent or different events for the same individual (onset of disease and subsequent death, successive spells in the labour market), or between events for different individuals (twins, litters, husband and wife).

The direct display of λ_1 and λ_2 provides some ideas about the dependence (or independence) between both failure times. See e.g. Pons (1986) for a "true" test of independence that is based on the bivariate integrated hazard function. Moreover, the estimation of λ_{12} helps us to build the association of the two (or more) failure times and to test the parametric or semiparametric model if the sample is big enough. See e.g. Clayton (1978) or Oakes (1982) who introduce a bivariate parametric distribution of (T_1, T_2) that can be explained by their common dependence on an unobserved gamma random variable (frailty) through a proportional hazard structure. Other references can be found in Andersen *et al.* (1993, p. 674). Note also a natural extension of our results for the bivariate (semiparametric) Cox model of Pons (1989): the introduction of the estimated parameter of interest in our estimators allows for the estimation of the baseline (bivariate) hazard function. Nevertheless, the asymptotic theory of this new estimator has yet to be undertaken.

APPENDIX

1. *Two Preliminary Lemmas*

Using the same arguments as Gijbels and Wang (1993) (Lemma 1), it is easy to prove

LEMMA A.1. *For each $\mathbf{x} \in \tau$, each integer p and each bounded function $\phi(\cdot, \cdot, \mathbf{x})$: $\mathbb{R}^{2r} \to \mathbb{R}$ there exists a constant $C(p, \mathbf{x})$ such that, for all n,*

$$E\left[\left\{\int \phi(\mathbf{u}_I, \mathbf{v}_I, \mathbf{x}) \cdot (\hat{S}_I - S_I)(d\mathbf{u}_I, \mathbf{x}_J) \cdot (\hat{H} - H)(d\mathbf{v}_I, \mathbf{x}_J)\right\}^p\right] \leqslant C(p, \mathbf{x})\, n^{-p}$$

In addition, if $\phi(\cdot, \cdot, \mathbf{x})$ is uniformly bounded in $\mathbf{x} \in \tau$ the majorations are uniforms on τ and there exists a constant C_0 such that for all n,

$$\sup_{\mathbf{x} \in \tau} E\left[\left\{\int \phi(\mathbf{u}_I, \mathbf{v}_I, \mathbf{x}) \cdot (\hat{S}_I - S_I)(d\mathbf{u}_I, \mathbf{x}_J) \cdot (\hat{H} - H)(d\mathbf{v}_I, \mathbf{x}_J)\right\}^p\right] \leqslant (C_0\, p/n)^p$$

For each triplet of integers (m_1, m_2, m_3), consider a function ω: $\mathbb{R}^{(m_1+m_2)r+d} \to \mathbb{R}$, and a function ϕ: $(1, ..., m_3\} \to \{1, ..., m_1+m_2\}$. Denote $m = m_1 + m_3$, $\bar{m} = m + m_2$ and the (m_1+m_2) r-dimensional vector $\vec{\mathbf{u}} = (\mathbf{u}_1, ..., \mathbf{u}_{m_1+m_2})$. Let

$$R(\mathbf{x}) = \int \omega(\vec{\mathbf{u}}, \mathbf{x}) \cdot \prod_{\nu=1}^{m_1} K_h(\mathbf{x}_I - \mathbf{u}_\nu) \cdot (\hat{S}_I - S_I)(d\mathbf{u}_\nu, \mathbf{x}_J)$$
$$\cdot \prod_{\nu=m_1+1}^{m_1+m_2} K_h(\mathbf{x}_I - \mathbf{u}_\nu)\, d\mathbf{u}_\nu \cdot \prod_{\mu=1}^{m_3} (\hat{H} - H)(\mathbf{u}_{\phi(\mu)}, \mathbf{x}_J)$$
$$= \int \omega(\vec{\mathbf{u}}, \mathbf{x}) \cdot \prod_{\nu=1}^{m_1} K_h(\mathbf{x}_I - \mathbf{u}_\nu) \cdot (\hat{S}_I - S_I)(d\mathbf{u}_\nu, \mathbf{x}_J)$$
$$\cdot \prod_{\nu=m_1+1}^{m_1+m_2} K_h(\mathbf{x}_I - \mathbf{u}_\nu)\, d\mathbf{u}_\nu \cdot \prod_{\mu=1}^{m_3} \mathbf{1}\{\mathbf{v}_\mu \geqslant \mathbf{u}_{\phi(\mu)}\} (\hat{H} - H)(d\mathbf{v}_\mu, \mathbf{x}_J)$$

LEMMA A.2. *If $\omega(\cdot, \mathbf{x})$ and f_I are bounded, if $h^r < \|K\|$ and if K is bounded compactly supported, then there exist constants (C_1, C_2) independent from (n, h, p, m) such that*

- *If* $\liminf_n nh^r/(2pm) > \|K\|$,

$$E[R^{2p}(\mathbf{x})] \leqslant C_1^{p\bar{m}} (pm)^{pm}\, n^{-pm} h^{-rpm_1} \tag{A.1}$$

- *If* $\limsup_n nh^r/(2pm) < \|K\|$,

$$E[R^{2p}(\mathbf{x})] \leqslant C_2^{p\bar{m}} (pm)^{pm}\, n^{-2pm} h^{-rp(2m_1+m_3)} \tag{A.2}$$

Remark A.1. (a) The obtained inequality is true uniformly on $\tau \subset \mathbb{R}^d$ when $\omega(\cdot, \mathbf{x})$ is uniformly bounded on $\mathbf{x} \in \tau$.

(b) Suppose that K is of order o. If, for each vector $\mathbf{u} \in \mathbb{R}^{m_1 r}$, the function $\omega(\mathbf{u}_1, ..., \mathbf{u}_{m_1 r}, \cdot, \mathbf{x})$ is $C^o(\mathbb{R}^{m_2 r})$ then the previous inequalities can be multiplied by h^{2pom_2}.

Proof.

$$R(\mathbf{x}) = \frac{1}{n^m} \sum_{\{i_\nu, j_\mu\}} \int \omega(\vec{\mathbf{u}}, \mathbf{x})$$
$$\cdot \prod_{\nu=1}^{m_1} K_h(\mathbf{x}_I - \mathbf{u}_\nu) \cdot d(\mathbf{1}\{\mathbf{X}_{i_\nu I} \geqslant \mathbf{u}_\nu, \mathbf{X}_{i_\nu J} > \mathbf{x}_J, \vec{\delta}_{i_\nu I} = \mathbf{1}_I\} - S_I(\mathbf{u}_\nu, \mathbf{x}_J))$$
$$\cdot \prod_{\nu=m_1+1}^{m_1+m_2} K_h(\mathbf{x}_I - \mathbf{u}_\nu)\, d\mathbf{u}_\nu$$
$$\cdot \prod_{\mu=1}^{m_3} \mathbf{1}\{\mathbf{v}_\mu \geqslant \mathbf{u}_{\phi(\mu)}\}\, d(\mathbf{1}\{\mathbf{X}_{j_\mu I} \geqslant \mathbf{v}_\mu, \mathbf{X}_{j_\mu J} > \mathbf{x}_J\} - H(\mathbf{v}_\mu, \mathbf{x}_J))$$

where we sum over all the m-tuples $\{(i_\nu, j_\mu)_{\nu=1,...,m_1;\,\mu=1,...,m_3} \mid 1 \leqslant i_\nu, j_\mu \leqslant n\}$. Thus

$$E[R^{2p}(\mathbf{x})] = n^{-2pm} \sum_{\mathscr{I}_{2p}} \int \prod_{k=1}^{2p} \Bigg\{ \omega(\vec{\mathbf{u}}_k, \mathbf{x}) \cdot \prod_{\nu_k=1}^{m_1} K_h(\mathbf{x}_I - \mathbf{u}_{\nu_k}) \cdot \prod_{\nu_k=m_1+1}^{m_1+m_2} K_h(\mathbf{x}_I - \mathbf{u}_{\nu_k})\, d\mathbf{u}_{\nu_k} \cdot \prod_{\mu_k=1}^{m_3} \mathbf{1}\{\mathbf{v}_{\mu_k} \geqslant \mathbf{u}_{\phi(\mu_k)}\} \Bigg\} \cdot E[d\Psi]$$

where $\mathscr{I}_{2p}$ is the following set of $2pm$-tuples of elements in $\{1, ..., n\}$:

$$\mathscr{I}_{2p} = \{((i_{\nu_k}, j_{\mu_k})_{\nu_k=1,...,m_1;\,\mu_k=1,...,m_3})_{k=1,...,2p}\}$$

and we have set

$$\Psi = \prod_{k=1}^{2p} \left\{ \prod_{\nu_k=1}^{m_1} (\mathbf{1}\{\mathbf{X}_{i_{\nu_k} I} \geqslant \mathbf{u}_{\nu_k}, \mathbf{X}_{i_{\nu_k} J} > \mathbf{x}_J, \vec{\delta}_{i_{\nu_k} I} = \mathbf{1}_I\} - S_I(\mathbf{u}_{\nu_k}, \mathbf{x}_J)) \right\} \cdot \left\{ \prod_{\mu_k=1}^{m_3} (\mathbf{1}\{\mathbf{X}_{j_{\mu_k} I} \geqslant \mathbf{v}_{\mu_k}, \mathbf{X}_{j_{\mu_k} J} > \mathbf{x}_J\} - H(\mathbf{v}_{\mu_k}, \mathbf{x}_J)) \right\}.$$

The differentiation of Ψ is taken over all the variables $\mathbf{u}_{\nu_k}$ and $\mathbf{v}_{\mu_k}$ (Ψ is a product of $4pm_1m_3$ functions, so $d\Psi$ is the product of $4pm_1m_3$ differentiations).

By Fubini, $E[d\Psi] = dE[\Psi]$. If one index in Ψ is different from the others, $E[\Psi]$ is zero, so we restrict ourselves to matched indices inside $\mathscr{I}_{2p}$; more precisely, there exist a (resp. b) equalities between indices $i_.$ (resp. $j_.$), and c between indices $i_.$ and $j_.$ that are non redundant with the previous ones.

This implies that $a \leqslant 2pm_1 - 1$, $b \leqslant 2pm_3 - 1$, $pm \leqslant a + b + c \leqslant 2pm - 1$ and $c \leqslant 2p(m_1 \wedge m_3)$.

Ψ is the sum of 4^{2pm} terms; each term is indexed by $\vec{\varepsilon} = (\varepsilon_1, ..., \varepsilon_{2pm_1})$ and $\vec{\kappa} = (\kappa_1, ..., \kappa_{2pm_3})$, each ε (resp. κ) being 0 or 1.

Each of these terms is the product of a deterministic function *Det* that is less than 1, by the indicator function

$$Ind = \prod\nolimits_{\theta \in \Theta} \mathbf{1}\{\mathbf{X}_{\theta I} > \mathbf{u}_{\alpha_1} \vee \cdots \vee \mathbf{u}_{\alpha_q} \vee \mathbf{v}_{\beta_1} \vee \cdots \vee \mathbf{v}_{\beta_s}, \mathbf{X}_{\theta J} > \mathbf{x}_J, \vec{\delta}^{(q)}_{\theta I} = \mathbf{1}_I\}.$$

If $q = s = 0$, we set $Ind = 1$. We denote $\vec{\delta}^{(q)}_{\theta I}$ for $\vec{\delta}_{\theta I}$ if $q > 0$, and $\mathbf{1}_I$ if $q = 0$. The product that defines *Ind* is over the (different) indices that supply independent variables $\mathbf{X}_.$ i.e. $i_{\alpha_1} = \cdots = i_{\alpha_q} = j_{\beta_1} = \cdots = j_{\beta_s} = \theta$.

Notice that each index α_l (resp. $\beta_{l'}$) comes from a different index k. Denote $\alpha(\theta)$ (resp. $\beta(\theta)$) the set of indices α_l (resp. $\beta_{l'}$) corresponding to θ. Each set Θ depends on $(\vec{\varepsilon}, \vec{\kappa})$ and its cardinal $\imath$ is less than $2pm - (a + b + c)$. Each element θ of Θ is associated with an integer $q(\theta)$ (resp. $s(\theta)$), or simpler q (resp. s), that is the cardinality of $\alpha(\theta)$ (resp. $\beta(\theta)$).

Obviously, $q \in \{0, ..., a+1\}$ and $s \in \{0, ..., b+1\}$. If $q=0$ (resp. $s=0$), there is no variable $\mathbf{u}$ (resp. $\mathbf{v}$) as argument. In any case, $q+s \geqslant 2$.

Endly, *Det* and *Ind* are functions of different variables. Let

$$A_\Theta = \bigcup_{\theta \in \Theta} \alpha(\theta)$$

$$B_\Theta = \bigcup_{\theta \in \Theta} \beta(\theta)$$

Let $F_I^{(q)} = S_I$ if $q>0$ and H if $q=0$. We obtain

$$E(Det \cdot Ind) = E(Det) \cdot E(Ind)$$
$$= Det \cdot \prod_{\theta \in \Theta} F_I^{(q)} \left(\bigvee_{k=1}^{q} \mathbf{u}_{\alpha_k} \vee \bigvee_{l=1}^{s} \mathbf{v}_{\beta_l}, \mathbf{x}_J \right)$$

Deduce that $dE(\Psi)$ is the sum of 4^{2pm} terms like

$$dDet \cdot \prod_\theta \left\{ (-1)^{(q+s-1)\,r} F_I^{(q)}(d\mathbf{u}_\theta, \mathbf{x}_J) \prod_{l=1}^{q} \delta\{\mathbf{u}_{\alpha_l} = \mathbf{u}_\theta\} \prod_{l'=1}^{s} \delta\{\mathbf{v}_{\beta_{l'}} = \mathbf{u}_\theta\} \right\}.$$

We notice that $F_I^{(0)}(d\mathbf{u}_\theta, \mathbf{x}_J) = H(d\mathbf{u}_\theta, \mathbf{x}_J)$ and that, if $q \neq 0$,

$$F_I^{(q)}(d\mathbf{u}_\theta, \mathbf{x}_J) = G(\mathbf{u}_\theta -, \mathbf{x}_J) \cdot f_I(\mathbf{u}_\theta \mid \mathbf{x}_J)\, d\mathbf{u}_\theta.$$

In this last case, we will execute a change of variables in the following formulas. Denoting $u^+ = \max(u, 0)$, for each $\mathbf{x} \in \tau$, we have

$$E[R^{2p}(\mathbf{x})] = n^{-2pm} \sum_{\mathscr{I}_{2p}} \int \prod_{k=1}^{2p} \left\{ \omega(\vec{\mathbf{u}}_k, \mathbf{x}) \cdot \prod_{v_k=1}^{m_1} K_h(\mathbf{x}_I - \mathbf{u}_{v_k}) \right.$$
$$\left. \cdot \prod_{v_k=m_1+1}^{m_1+m_2} K_h(\mathbf{x}_I - \mathbf{u}_{v_k})\, d\mathbf{u}_{v_k} \cdot \prod_{\mu_k=1}^{m_3} \mathbf{1}\{\mathbf{v}_{\mu_k} \geqslant \mathbf{u}_{\phi(\mu_k)}\} \right\}$$
$$\cdot \sum_{\vec{\varepsilon}, \vec{\kappa}} dDet \prod_\theta \left\{ (-1)^{(q+s-1)r} \prod_{l=1}^{q} \delta\{\mathbf{u}_{\alpha_l} = \mathbf{u}_\theta\} \right.$$
$$\left. \cdot \prod_{l'=1}^{s} \delta\{\mathbf{v}_{\beta_{l'}} = \mathbf{u}_\theta\} F_I^{(q)}(d\mathbf{u}_\theta, \mathbf{u}_J) \right\}$$
$$= n^{-2pm} \sum_{\mathscr{I}_{2p}} \sum_{\vec{\varepsilon}, \vec{\kappa}} \int \prod_{k=1}^{2p} \left\{ \omega'(\vec{\mathbf{u}}_k, \mathbf{x}) \cdot \prod_{v_k \notin A_\Theta,\, v_k \leqslant m_1} K_h(\mathbf{x}_I - \mathbf{u}_{v_k}) \right.$$
$$\cdot \prod_{v_k=m_1+1}^{m_1+m_2} K_h(\mathbf{x}_I - \mathbf{u}_{v_k})\, d\mathbf{u}_{v_k} \cdot \prod_\theta \prod_{\mu_k \in B_\Theta} \mathbf{1}\{\mathbf{u}_\theta \geqslant \mathbf{u}_{\phi_\Theta(\mu_k)}\}\, K_h^q(\mathbf{x}_I - \mathbf{u}_{\boldsymbol{\theta}})$$
$$\left. \cdot \prod_{\mu_k \notin B_\Theta} \mathbf{1}\{\mathbf{v}_{\mu_k} \geqslant \mathbf{u}_{\phi_\theta(\mu_k)}\}\, dDet \right\} \cdot (-1)^{(q+s-1)\,rt} \prod_{\theta \in \Theta} F_I^{(q)}(d\mathbf{u}_\theta, \mathbf{x}_J)$$

We have denoted $\phi_\Theta(\mu_k)=\phi(\mu_k)$ if $\phi(\mu_k)\notin A_\Theta$ and θ if $v_k\in\alpha(\theta)$. The function $\omega(\vec{\mathbf{u}}_k,\mathbf{x})$ has begun

$$\omega'(\vec{\mathbf{u}}_k,\mathbf{x})=\omega(\mathbf{u}_{v_k},v_k\notin A_\Theta;\mathbf{u}_\theta,\theta\in\Theta;\mathbf{x}).$$

Note that each previous argument $\mathbf{u}_\theta$ appears $q(\theta)$ times inside the brackets of K_h and ω'. The integral of $\mathbf{1}\{\mathbf{v}_{\mu_k}\geqslant\mathbf{u}_{\phi_\Theta(\mu_k)}\}\,dDet$ with respect to a variable $\mathbf{v}_{\mu_k}$ such that μ_k is not in B_Θ gives us $H(\mathbf{u}_{\phi_\Theta(\mu_k)},\mathbf{x}_J)$. We make the change of variables $\mathbf{x}_I-\mathbf{u}_{v_k}=h\mathbf{t}_{v_k}$ for all $v_k\in\{m_1+1,...,m_1+m_2\}$. Note that, if K is a kernel of order o, we can make a limited expansion in $\mathbf{t}_{v_k}$ of order o after all other integrations; this would add an extra factor h^{2pom_2} in our majoration.

In any case, make the change of variables $\mathbf{x}_I-\mathbf{u}_\theta=h\mathbf{z}_\theta$ for each θ. This last change is necessary only if $q\neq 0$. When $q=0$, the corresponding term is $H(d\mathbf{u}_\theta,\cdot)$ thus after integration, it is less than 1. Therefore, we obtain

$$\begin{aligned}
&E[R^{2p}(\mathbf{x})]\\
&\quad\leqslant Cst^{p\bar{m}}n^{-2pm}\sum_{\mathscr{I}_{2p}}\sum_{\vec{\varepsilon},\vec{\kappa}}\int\prod_\theta\,[h^{-r(q-1)^+}\,|K|^q\,(\mathbf{z}_\theta)\,f_I(\mathbf{x}_I-h\mathbf{z}_\theta)\,d\mathbf{z}_\theta\,\mathbf{1}\{q>0\}\\
&\qquad+\mathbf{1}\{q=0\}\,]\\
&\quad\leqslant\frac{Cst^{p\bar{m}}}{n^{2pm}}\sum_{\mathscr{I}_{2p}}\sum_{\vec{\varepsilon},\vec{\kappa}}\int\prod_\theta\,[h^{-r(q-1)^+}\,\|K\|^{q-1}\\
&\qquad\cdot\|f_I\|\cdot|K|\,(\mathbf{z}_\theta)\,d\mathbf{z}_\theta\,\mathbf{1}\{q>0\}+\mathbf{1}\{q=0\}\,]
\end{aligned}$$

Note that $\sum_\theta(q-\mathbf{1})^+=a$ and denote $S=a+b+c$. Consider n-tuples $(p_1,...,p_n)$ of nonnegative integers such that $\sum_{i=1}^n p_i=2pml$. Also the sum over $\mathscr{I}_{2p}$ can be written

$$\sum_{\mathscr{I}_{2p}}=\sum_{\substack{(p_1,...,p_n)\\ \sum p_i=2pm}}\;\sum_{\substack{(i_{v_k},j_{\mu_k})\in\mathscr{I}_{2p}\\ \sharp\{k\mid i_{v_k}=i\}\cup\{k\mid j_{\mu_k}=i\}=p_i}}$$

Following Dehling, Denker and Philipp (1987), we use some combinatorial arguments: For a fixed $S\in\{pm,...,2pm-1\}$, we consider at most $2pm-S$ r.v. $\mathbf{X}_\theta$ which give at most $2pm-S$ nonzero indices p_i. The sum over these p_i has C_n^{2pm-S} summands. Moreover, there are C_{2pm-1}^S ways to find positive integers (q_j) such that $\sum_{j=1}^{2pm-S}q_j=2pm$. It is easy to see that $\max_S C_{2pm-1}^S$ is obtained with $S=pm$ and prove by recurrence over n that $C_{2n-1}^n\leqslant 4^n$. This is why the first sum has at most $4^{pm}C^{2pm-S}$ summands. At last, the inner sum has less than $(2pm)!$ summands.

Also denoting $\bar{h}^{-r} = \|K\| \, h^{-r} > 1$, we deduce

$$E[R^{2p}(\mathbf{x})]$$
$$\leqslant Cst^{p\bar{m}} n^{-2pm} \sum_{\mathscr{I}_{2p}} \sum_{\vec{\varepsilon}, \vec{\kappa}} \left(1 \vee \int |K| \cdot \|f_I\|\right)^{2pm-S} \bar{h}^{-ra}$$
$$\leqslant Cst^{p\bar{m}} n^{-2pm} \sum_{S=pm}^{2pm-1} 4^{pm} C_n^{2pm-S} (2pm)! \cdot 2^{2pm} \sup_{\{(a,b,c) \mid a+b+c=S\}} \bar{h}^{-ra}$$
$$\leqslant Cst^{p\bar{m}} n^{-2pm} \sum_{S=pm}^{2pm-1} \frac{n^{2pm-S}}{(2pm-S)!} (2pm)! \sup_{\{(a,b,c) \mid a+b+c=S\}} \bar{h}^{-ra}$$

Since $\bar{h} < 1$, $\sup_{\{(a,b,c) \mid a+b+c=S\}} \bar{h}^{-ra}$ is obtained with the largest a among the considered triplets; the minimum of $b+c$ is reached if the $j_.$ are matched together (two by two). This gives pm_3 necessary equalities between the $j_.$, and the considered a is at most $S - pm_3$. Also,

$$E[R^{2p}(\mathbf{x})] \leqslant Cst^{p\bar{m}} \bar{h}^{rpm_3} \sum_{S=pm}^{2pm-1} \left(\frac{2pm}{n\bar{h}^r}\right)^S$$

Hence the result. ∎

The case $\bar{h} \geqslant 1$ is not detailed because the previous inequalities will be available for n sufficiently large.

The latter proof can be adapted to include several kernels and several bandwidths. See Fermanian (1996).

2. *Proofs*

Proof of Theorem 2.1. Simple algebraic calculations give us the decomposition above with:

$$r_n(\mathbf{x}) = (-1)^r \int_{(0, \mathbf{x}_I]} \frac{(\hat{H} - H)^2 (\mathbf{u}_I -, \mathbf{x}_J)}{H^2 \cdot \hat{H}(\mathbf{u}_I -, \mathbf{x}_J)} \hat{S}_I(d\mathbf{u}_I, \mathbf{x}_J)$$
$$+ (-1)^{r+1} \int_{(0, \mathbf{x}_I]} \frac{(\hat{H} - H)(\mathbf{u}_I -, \mathbf{x}_J)}{H^2(\mathbf{u}_I -, \mathbf{x}_J)} (\hat{S}_I - S_I)(d\mathbf{u}_I, \mathbf{x}_J)$$
$$\overset{\text{not}}{=} \bar{r}_n(\mathbf{x}) + \tilde{r}_n(\mathbf{x})$$

As $\|(\hat{H} - H)^2\|_\infty = O(\ln_2 n/n)$ a.e., we have a.e.

$$\sup_{\mathbf{x} \in \tau} |\bar{r}_n(\mathbf{x})| = O(\ln_2 n/n)$$

By Lemma A.1, and Markov's inequality, we obtain the majoration of $\tilde{r}_n$ for all integer p:

$$P(|\tilde{r}_n(\mathbf{x})| > z_0) \leqslant \frac{E[(\tilde{r}_n(\mathbf{x}))^{2p}]}{z_0^{2p}} \leqslant \left(\frac{C_0 p}{n z_0}\right)^{2p}$$

Set $p = [\ln n]$ and $z_0 = C \ln n/n$; also $\sum_n P(|n\tilde{r}_n(\mathbf{x})| > C \ln n) < +\infty$ for all C larger than C_0. By Borel–Cantelli, we deduce for all $\mathbf{x} \in \tau$, that $\tilde{r}_n(\mathbf{x}) = O(n^{-1} \ln n)$ a.e.

We need sharper inequalities to obtain a uniform bound over τ. By continuity of the X_k's distribution functions, we can find the following decomposition of τ into squares: $\tau = \bigcup_{i=1}^{m^d} C_i$ such that for all i, i' in $\{1, ..., m^d\}$,

- $C_i = \prod_{j=1}^{d} (s_{i,j}; t_{i,j}]$ denoting "(" for "[" if $s_{i,j} = 0$, and "]" else
- $C_i \neq \varnothing$ and $C_i \cap C_{i'} = \varnothing$ if $i \neq i'$
- $P(T_k \in (s_{i,j}; t_{i,j}]) \leqslant m^{-1}$ for all $k \in I$
- $P(X_k \in (s_{i,j}; t_{i,j}]) \leqslant m^{-1}$ for all $k \in J$.

Denote $\mathbf{t}_i = (t_{i1}, ..., t_{id})$ and $\mathbf{s}_i = (s_{i1}, ..., s_{id})$.

For each $\mathbf{z} \in \tau$ there exists a unique $i \in \{1, ..., m^d\}$ such that $\mathbf{z} \in C_i$; we denote $\Delta(\mathbf{z}) = [0, \mathbf{t}_{iI}] \backslash [0, \mathbf{z}_I]$ and $\Delta_i = [0, \mathbf{t}_{iI}] \backslash [0, \mathbf{s}_{iI}]$. So,

$$\begin{aligned}
&P(\sup_{\mathbf{x} \in \tau} |\tilde{r}_n(\mathbf{x})| > z) \\
&\quad \leqslant P(\max_{1 \leqslant i \leqslant m^d} |\tilde{r}_n(\mathbf{t}_i)| > z_1) \\
&\qquad + P\left(\max_i \sup_{\mathbf{z} \in C_i} \left| \int_{\Delta(\mathbf{z})} \frac{(\hat{H} - H)}{H^2} (\mathbf{u}_I -, \mathbf{z}_J) \cdot (\hat{S}_I - S_I)(d\mathbf{u}_I, \mathbf{z}_J) \right| > z_2 \right) \\
&\qquad + P\left(\max_i \sup_{\mathbf{z} \in C_i} \left| \int_{[0, \mathbf{t}_{iI}]} \frac{(\hat{H} - H)}{H^2} (\mathbf{u}_I -, \mathbf{z}_J) \cdot (\hat{S}_I - S_I)(d\mathbf{u}_I, \mathbf{z}_J) \right.\right. \\
&\qquad \left.\left. - \frac{(\hat{H} - H)}{H^2} (\mathbf{u}_I -, \mathbf{t}_{iJ}) \cdot (\hat{S}_I - S_I)(d\mathbf{u}_I, \mathbf{t}_{iJ}) \right| > z_3 \right) \\
&\quad \leqslant p_1 + p_2 + p_3
\end{aligned}$$

with $z_1 + z_2 + z_3 = z$. If $J = \varnothing$, there is no term p_3. First, for all integer p,

$$p_1 \leqslant \sum_{i=1}^{m^d} P(|\tilde{r}_n(\mathbf{t}_i)| > z_1) \leqslant m^d (C_0 p/n z_1)^{2p} \tag{A.1}$$

For estimating p_2, we remark that

$$\sup_{\mathbf{z}\in C_i}\left|\int_{\Delta(\mathbf{z})}\frac{(\hat{H}-H)}{H^2}(\mathbf{u}_I-,\mathbf{z}_J)\cdot(\hat{S}_I-S_I)(d\mathbf{u}_I,\mathbf{z}_J)\right|$$

$$\leqslant \sup_{\mathbf{x}\in\tau}\left|\frac{(\hat{H}-H)}{H^2}(\mathbf{x})\right|\cdot\sup_{\mathbf{z}\in C_i}\int_{\Delta(\mathbf{z})}|\hat{S}_I(d\mathbf{u}_I,\mathbf{z}_J)|+|S_I(d\mathbf{u}_I,\mathbf{z}_J)|$$

$$\leqslant \sup_{\mathbf{x}\in\tau}\left|\frac{(\hat{H}-H)}{H^2}(\mathbf{x})\right|\cdot\sup_{\mathbf{z}\in C_i}(\hat{S}_I+S_I)(\Delta(\mathbf{z}),\mathbf{z}_J)$$

$$\leqslant H^{-2}(\tau)\cdot\sup_{\mathbf{x}\in\tau}|(\hat{H}-H)(\mathbf{x})|\cdot\sup_{\mathbf{z}\in C_i}(\hat{S}_I+S_I)(\Delta(\mathbf{z}),0)$$

So, if $z_4\cdot z_5=z_2$,

$$p_2\leqslant P(\sup_{\mathbf{x}\in\tau}|\hat{H}-H|\,(\mathbf{x})>z_4H^2(\tau))+P(\max_i\ \sup_{\mathbf{z}\in C_i}(\hat{S}_I+S_I)(\Delta(\mathbf{z}),0)>z_5)$$

Now, we have (cf. Bosq and Lecoutre (1987), p. 48)

$$P(\sup_{\mathbf{x}\in\tau}|(\hat{H}-H)(\mathbf{x})|>z_4H^2(\tau))$$

$$\leqslant 4\exp(4H^2(\tau)\,z_4+4H^4(\tau)\,z_4^2)\cdot(1+n^2)^d\cdot\exp(-2nH^4(\tau)\,z_4^2)$$

Moreover

$$\max_i\ \sup_{\mathbf{z}\in C_i}(\hat{S}_I+S_I)(\Delta(\mathbf{z}),0)$$

$$\leqslant\max_i(\hat{S}_I+S_I)(\Delta_i,0)$$

$$\leqslant\max_i(|\hat{S}_I-S_I|+2S_I)\left(\bigcup_{j\in I}(s_{i,j},t_{ij}]\times\tau_{I\setminus\{j\}},0\right)$$

$$\leqslant\max_i|\hat{S}_I-S_I|\left(\bigcup_{j\in I}(s_{i,j},t_{ij}]\times\tau_{I\setminus\{j\}},0\right)+2d/m$$

Thanks to Hoeffding's inequality (Bosq and Lecoutre (1986), p. 41), we have for each $s>0$

$$P\left(|\hat{S}_I-S_I|\left(\bigcup_{j\in I}(s_{i,j},t_{ij}]\times\tau_{I\setminus\{j\}},0\right)>s\right)\leqslant 2\exp(-2ns^2)$$

So, if $z_5 - 2d/m > 0$,

$$p_2 \leqslant 4 \exp(4H^2(\tau)\, z_4 + 4H^4(\tau)\, z_4^2) \cdot (1+n^2)^d \cdot \exp(-2nH^4(\tau)\, z_4^2) + 2m^d \exp(-2n(z_5 - 2d/m)^2) \tag{A.2}$$

We need now to estimate p_3. Note that

$$p_3 \leqslant P\left(\max_i \sup_{\mathbf{z} \in C_i} \sum_{k=1}^{d-r} |f_k(\mathbf{z}, \mathbf{t}_i)| > z_3 \right) \leqslant \sum_{k=1}^{d-r} P(\max_i \sup_{\mathbf{z} \in C_i} |f_k(\mathbf{z}, \mathbf{t}_i)| > z_3/d)$$

where, supposing that $J = \{1, ..., d-r\}$, we denote

$$\begin{aligned} f_k(\mathbf{z}, \mathbf{t}_i) = \int_{[0, \mathbf{t}_{iI}]} & \frac{(\hat{H} - H)}{H^2} (\mathbf{u}_I -;\, \mathbf{z}_{1, ..., k-1}, \mathbf{t}_{i; k, ..., d-r}) \\ & \times (\hat{S}_I - S_I)\, (d\mathbf{u}_I;\, \mathbf{z}_{1, ..., k-1}, \mathbf{t}_{i; k, ..., d-r}) \\ & - \frac{(\hat{H} - H)}{H^2} (\mathbf{u}_I -;\, \mathbf{z}_{1, ..., k}, \mathbf{t}_{i; k+1, ..., d-r}) \\ & \times (\hat{S}_I - S_I)\, (d\mathbf{u}_I;\, \mathbf{z}_{1, ..., k}, \mathbf{t}_{i; k+1, ..., d-r}) \end{aligned} \tag{A.3}$$

For simplicity, we will write only the kth coordinate of vectors in τ_J; so,

$$\begin{aligned} f_k(\mathbf{z}, \mathbf{t}_i) = \int_{[0, \mathbf{t}_{iI}]} & (-1)^r \left[(\hat{S}_I - S_I)(\mathbf{u}_I, t_{ik}) - (\hat{S}_I - S_I)(\mathbf{u}_I, z_k) \right] \\ & \times \frac{(\hat{H} - H)}{H^2} (d\mathbf{u}_I, t_{ik}) \\ & + (\hat{H} - H)(\mathbf{u}_I -, t_{ik}) \left[\frac{1}{H^2} (\mathbf{u}_I -, t_{ik}) - \frac{1}{H^2} (\mathbf{u}_I -, z_k) \right] \\ & \times (\hat{S}_I - S_I)(d\mathbf{u}_I, z_k) \\ & + \frac{1}{H^2} (\mathbf{u}_I -, z_k) [(\hat{H} - H)(\mathbf{u}_I -, t_{ik}) - (\hat{H} - H)(\mathbf{u}_I -, z_k)] \\ & \times (\hat{S}_I - S_I)(d\mathbf{u}_I, z_k) \end{aligned}$$

Denoting Cst constants independent from n, h, m, and from all points and indices, this implies

$$\sup_{\mathbf{z}\in C_i} |f_k(\mathbf{z}, \mathbf{t}_i)| \leqslant Cst \cdot \sup_{\mathbf{z}\in C_i,\, \mathbf{u}_I\in\tau_I} |(\hat{S}_I - S_I)(\mathbf{u}_I, t_{ik}) - (\hat{S}_I - S_I)(\mathbf{u}_I, z_k)|$$
$$+ Cst\cdot \|\hat{H} - H\|_\infty \cdot P(X_k \in (s_{ik}, t_{ik}])$$
$$+ Cst\cdot \sup_{\mathbf{z}\in C_i,\, \mathbf{u}_I\in\tau_I} |(\hat{H} - H)(\mathbf{u}_I, t_{ik}) - (\hat{H} - H)(\mathbf{u}_I, z_k)|$$

Let us examine the first term on the right; let $\mathbf{Z}$ the d-dimensional r.v. $\mathbf{X}$ knowing that $\vec{\delta}_I = \mathbf{1}_I$. Let μ (resp. μ_n) the probability measure (resp. empirical measure) induced by $\mathbf{Z}$; the survival function of $\mathbf{Z}$ is $\tilde{S}_I = S_I/P(\vec{\delta}_I = \mathbf{1}_I)$; we set $X'_k = \tilde{S}_I(Z_k, 0_{-k})$; all the similar notions for X'_k are denoted with primes. Notice that

$$(\hat{S}_I - S_I)(\mathbf{u}_I, z_k) - (\hat{S}_I - S_I)(\mathbf{u}_I, t_{ik})$$
$$= P_n(\vec{\delta}_I = \mathbf{1}_I)\cdot(\mu_n - \mu)$$
$$\times([\mathbf{u}_I, +\infty[,]z_{1,\ldots,k-1}, +\infty[,]z_k, t_{ik}],]t_{i;\, k+1,\ldots,d-r}, +\infty[)$$
$$+ \left(\frac{P_n - P}{P}\right)(\vec{\delta}_I = \mathbf{1}_I)\cdot O(1/m)$$

The second term is $O(m^{-1} n^{-1/2} \ln_2 n)$ a.e.; for the first one, we obtain

$$\sup_{\mathbf{z}\in C_i,\, \mathbf{u}_I\in\tau_I} |\mu_n - \mu|$$
$$\times([\mathbf{u}_I, +\infty[,]z_{1,\ldots,k-1}, +\infty[,]z_k, t_{ik}],]t_{i;\, k+1,\ldots,d-r}, +\infty[)$$
$$= \sup_{\mathbf{z}'\in C'_i,\, \mathbf{u}'_I\in\tau'_I} |\mu'_n - \mu'|$$
$$\times([0, \mathbf{u}'_I], [0, z'_{1,\ldots,k-1}[, [t'_{ik}, z'_k], [0, t'_{i;\, k+1,\ldots,d-r}[)$$
$$\leqslant \sup_{\mathbf{z}'\in C'_i,\, \mathbf{u}'_I\in[0,1]_I} |\mu'_n - \mu'|$$
$$\times([0, \mathbf{u}'_I], [0, z'_{1,\ldots,k-1}[, [t'_{ik}, z'_k], [0, t'_{i;\, k+1,\ldots,d-r}[)$$

The r.v. $\mathbf{Z}'$ is distributed on $[0,1]^d$, with uniform marginals and a continuous c.d.f.; we can also use the results of Stute (1984, p. 366 and discussion p. 364).

$$P\Big(\sup_{\mathbf{z}'\in C'_i,\, \mathbf{u}'_I\in[0,1]_I} |\mu'_n - \mu'|$$
$$\times([0, \mathbf{u}'_I], [0, z'_{1,\ldots,k-1}[, [t'_{ik}, z'_k[, [0, t'_{i;\, k+1,\ldots,d-r}[) > z_6\Big)$$
$$\leqslant Cst\cdot \exp(-3nz_6^2(1-2\delta)^{2d}/(6\bar{p} + 2z_6(1-2\delta)^d))$$

if (setting $\bar{p}=P(Z'_k \in [t'_{ik}, s'_{ik}])=P(Z_k \in [s_{ik}; t_{ik}])$)

$$|t'_{ik}-s'_{ik}|<\delta/4 \qquad 2\leqslant nz_6 \qquad 32\bar{p}\leqslant n(z_6\delta(1-2\delta))^2$$

Choosing $\delta=1/4$, since $\bar{p}=(t'_{ik}-s'_{ik})\leqslant 1/m$, the above conditions are verified if $m>16$, if $2\leqslant nz_6$ and $m_n z_6\rightarrow +\infty$ as n goes to ∞. In this case, for n sufficiently large, when $mz_6>1$,

$$P(\sup_{\mathbf{z}\in C_i,\, \mathbf{u}_I\in\tau_I} |(\hat{S}_I-S_I)(\mathbf{u}_I, t_{ik})-(\hat{S}_I-S_I)(\mathbf{u}_I, z_k)|>z_6)$$
$$\leqslant Cst\cdot\exp(-n(z_6-1/m)\,2^{-d-1}) \tag{A.4}$$

With the same technique, it can be proved that, under similar conditions,

$$P(\sup_{\mathbf{z}\in C_i,\, \mathbf{u}_I\in\tau_I} |(\hat{H}-H)(\mathbf{u}_I-, t_{ik})-(\hat{H}-H)(\mathbf{u}_I-, z_k)|>z_8)$$
$$\leqslant Cst\cdot\exp(-Cst\cdot nz_8) \tag{A.5}$$

We obtain, with $z_3=z_6+z_7+z_8$:

$$p_3\leqslant\sum_{k=1}^{d-r}P(\max_{i=1,\dots,m^d}\ \sup_{\mathbf{z}\in C_i}|f_k(\mathbf{z},\mathbf{t}_i)|>z_3/d)$$

$$\leqslant\sum_{k,i}P(\sup_{\mathbf{z}\in C_i}|f_k(\mathbf{z},\mathbf{t}_i)|>z_3/d)$$

$$\leqslant\sum_{k,i}[P(\sup_{\mathbf{z}\in C_i,\, \mathbf{u}_I\in\tau_I}|(\hat{S}_I-S_I)(\mathbf{u}_I-, t_{ik})$$
$$-(\hat{S}_I-S_I)(\mathbf{u}_I-, z_k)|>C_6(z_6-m^{-1}))]$$

$$+P(\|\hat{H}-H\|_\infty>mC_7z_7)$$
$$+P(\sup_{\mathbf{z}\in C_i,\, \mathbf{u}_I\in\tau_I}|(\hat{H}-H)(\mathbf{u}_I-, t_{ik})-(\hat{H}-H)(\mathbf{u}_I-, z_k)|>C_8z_8)$$
$$\leqslant Cst\cdot m^d[\exp(-Cst\cdot n(z_6-m^{-1}))+\exp(-Cst\cdot nz_8)$$
$$+\exp(Cst\cdot mz_7+Cst\cdot m^2z_7^2)(1+n^2)^d\exp(-Cst\cdot nm^2z_7^2)] \tag{A.6}$$

The constants Cst and $C_.$ are independent from m, n, $z_.$, k, i.

Now, we choose the following parameters:

- $z=C^*\ln n/n$ and $m=n$
- $z_1=z_2=z_3=z/3$
- $z_4=z_5=z_2^{1/2}$ and $z_6=z_7=z_8=z_3/3$

Recalling A.1, A.2 and A.6, there exists a C^* such that, for n sufficiently large, $\sum (p_1+p_2+p_3)$ is convergent. Hence the result by Borel–Cantelli. ∎

Proof of Proposition 2.2. Consider the following lemma (see Fermanian (1996)).

LEMMA A.1. *If each r.v. X_k has a continuous distribution, the sequence of $\mathscr{D}_{2d}([0,\tau]^2)$ valued processes $n^{1/2}(\hat{S}_I - S_I, \hat{H}-H)$ converges weakly in the $\mathscr{S}$-topology to a Gaussian process $G=(G_1, G_2)$, where each process G_i has a.e. continuous paths, and for all $\mathbf{x}$ and $\mathbf{y}$ in τ, $E(G(\mathbf{x}))=0$ and*

(i) $\operatorname{Cov}(G_1(\mathbf{x}), G_1(\mathbf{y})) = S_I(\mathbf{x}\vee\mathbf{y}) - S_I(\mathbf{x})\, S_I(\mathbf{y})$

(ii) $\operatorname{Cov}(G_2(\mathbf{x}), G_2(\mathbf{y})) = H(\mathbf{x}\vee\mathbf{y}) - H(\mathbf{x})\, H(\mathbf{y})$

(iii) $\operatorname{Cov}(G_1(\mathbf{x}), G_2(\mathbf{y})) = S_I(\mathbf{x}\vee\mathbf{y}) - S_I(\mathbf{x})\, H(\mathbf{y})$.

Invoking the Skorohod–Dudley–Wichura's theorem (Shorack and Wellner (1986), p. 47), there exist a probability space $(\Omega^*, \mathscr{A}^*, P^*)$ and processes $n^{1/2}(\hat{S}_I^* - S_I^*, \hat{H}^* - H^*)$ (resp. G^*) from Ω^* to $(\mathscr{D}_{2d}([0,\tau]^2), \mathscr{S})$ that induce the same distributions than the latter ones and such that (ρ is Skorohod's metric)

$$\rho(n^{1/2}(\hat{S}_I^* - S_I^*, \hat{H}^* - H^*), G^*) \xrightarrow[n\to\infty]{} 0 \qquad \text{a.e.}$$

Since the limit G^* has a.e. continuous sample paths, this can be rewritten as

$$\| n^{1/2}(\hat{S}_I^* - S_I^*, \hat{H}^* - H^*) - G^* \|_\infty \xrightarrow[n\to\infty]{} 0 \qquad \text{a.e.}$$

We can suppose this last fact without changing our notations because we are looking for properties in law only. For all $\mathbf{x}\in\tau$, set the Gaussian process

$$W(\mathbf{x}) = (-1)^r \left[\int \mathbf{1}\{\mathbf{u}_I \leqslant \mathbf{x}_I\} \frac{G_1(d\mathbf{u}_I, \mathbf{x}_J)}{H(\mathbf{u}_I-, \mathbf{x}_J)} \right.$$
$$\left. - \int \mathbf{1}\{\mathbf{u}_I \leqslant \mathbf{x}_I\} \frac{G_2(\mathbf{u}_I-, \mathbf{x}_J)}{H^2(\mathbf{u}_I-, \mathbf{x}_J)} S_I(d\mathbf{u}_I, \mathbf{x}_J) \right]$$

The definition of W above is valid thanks to integration by parts (see Gill *et al.* (1993)). We obtain the result by integration by parts,

$$\begin{aligned}
&\rho(n^{1/2}(\hat{\Lambda}_I - \Lambda_I), W) \\
&\quad \leqslant \|n^{1/2}(\hat{\Lambda}_I - \Lambda_I) - W\|_\infty \\
&\quad \leqslant Cst \cdot \{ \|n^{1/2}(\hat{S}_I - S_I) - G_1\|_\infty + \|n^{1/2}(\hat{H}_I - H_I) - G_2\|_\infty \\
&\qquad + n^{-1/2} \ln n\} \xrightarrow[n \to \infty]{} 0 \qquad \text{a.e.} \quad \blacksquare
\end{aligned}$$

Proof of Theorem 2.2. The formula 2.5 is obvious from 2.1; we need only to prove the order of $\sup_{\mathbf{x}} |R_n(\mathbf{x})|$. With the notations of the proof of Theorem 2.1, consider

$$\begin{aligned}
R_n(\mathbf{x}) = &\int K_h(\mathbf{x}_I - \mathbf{u}_I)\, \bar{r}_n(d\mathbf{u}_I, \mathbf{x}_J) \\
&+ \int K_h(\mathbf{x}_I - \mathbf{u}_I)\, \tilde{r}_n(d\mathbf{u}_I, \mathbf{x}_J) \overset{\text{not}}{=} \bar{R}_n(\mathbf{x}) + \tilde{R}_n(\mathbf{x})
\end{aligned}$$

Suppose $h < h_0$, with h_0 such that $H(\tau + h_0 \mathbf{A}) > 0$. In the sequel, we will not distinguish between τ and $\tau + h_0 \mathbf{A}$. Denote "*Cst*" all constants independent from $\mathbf{x}, h, n, \ldots$. For all $\mathbf{x} \in \tau$, we have a.e.

$$\begin{aligned}
|\bar{R}_n(\mathbf{x})| &\leqslant Cst \cdot n^{-1}\, \|\hat{H} - H\|_\infty^2 \cdot \sum_i |K_h|\, (\mathbf{x}_I - \mathbf{X}_{iI})\, \mathbf{1}\{\vec{\delta}_{iI} = \mathbf{1}_I, \mathbf{X}_{iJ} > \mathbf{x}_J\} \\
&\leqslant Cst \cdot (\ln_2 n/n) \cdot n^{-1} \sum_i |K_h|\, (\mathbf{x}_I - \mathbf{T}_{iI}) \\
&\leqslant Cst \cdot (\ln_2 n/n) \cdot (-1)^r \left[\int |K_h|\, (\mathbf{x}_I - \mathbf{u}_I)\, F(d\mathbf{u}_I, 0) \right. \\
&\qquad \left. + \int |K_h|\, (\mathbf{x}_I - \mathbf{u}_I) \cdot (\hat{F} - F)(d\mathbf{u}_I, 0)) \right] \\
&\leqslant Cst \cdot (\ln_2 n/n) \cdot \{ \sup_{\mathbf{x} \in \tau} f_I(\mathbf{x}_I) + |\hat{f}_I(\mathbf{x}_I) - E(\hat{f}_I(\mathbf{x}_I))| \}
\end{aligned}$$

where, since $|K| / \int |K|$ is a kernel on $\mathbb{R}^r$, $\hat{f}_I$ is a kernel estimator of f_I. Since $nh_n^r / \ln n \to \infty$ and f_I is continuous on τ_I, it is well known (Bosq and Lecoutre (1986), p. 65) that $\sup_{\mathbf{x} \in \tau} |\hat{f}_I(\mathbf{x}_I) - E(\hat{f}_I(\mathbf{x}_I))| \to 0$.

So, with the conditions above, we have obtained, as $nh_n^r / \ln n \to \infty$,

$$\sup_{\mathbf{x} \in \tau} |\bar{R}_n(\mathbf{x})| = O(n^{-1} \ln_2 n) \qquad \text{a.e.} \tag{A.7}$$

To study $\sup_{\mathbf{x}\in\tau}|\widetilde{R}_n(\mathbf{x})|$, use similar arguments as in the proof of Theorem 2.1: τ is the disjoint union of m^d boxes $C_i=(\mathbf{s}_i,\mathbf{t}_i]$. When $z_1+z_2+z_3=z$,

$$\begin{aligned}
&P(\sup_{\mathbf{x}\in\tau}|\widetilde{R}_n(\mathbf{x})|>z)\\
&\quad\leqslant P(\max_{i=1,\ldots,m^d}|\widetilde{R}_n(\mathbf{t}_i)|>z_1)\\
&\qquad+P\left(\max_i\sup_{\mathbf{z}\in C_i}\left|\int(K_h(\mathbf{z}_I-\mathbf{u}_I)-K_h(\mathbf{t}_{iI}-\mathbf{u}_I))\,\tilde{r}_n(d\mathbf{u}_I,\mathbf{z}_J)\right|>z_2\right)\\
&\qquad+P\left(\max_i\sup_{\mathbf{z}\in C_i}\left|\int K_h(\mathbf{t}_{iI}-\mathbf{u}_I)[\tilde{r}_n(d\mathbf{u}_I,\mathbf{z}_J)-\tilde{r}_n(d\mathbf{u}_I,\mathbf{t}_{iJ})]\right|>z_3\right)\\
&\quad\leqslant p_1+p_2+p_3
\end{aligned}$$

Set $z_1=C^*n^{-1}h_n^{-r/2}\ln n/3$, and $p=[\ln n]$; then $\lim_n nh_n^r/p_n=+\infty$. If there exists $\rho>0$ such that $m_n<Cst\cdot n^\rho$ for all n, deduce from Lemma 5.2

$$p_1\leqslant m^dC_1^p(2p)^{2p}\,(nz_1h^{r/2})^{-2p}\leqslant Cst\cdot n^{d\rho}\cdot(C^*/2)^{-2p}\cdot C_1^p \tag{A.8}$$

which is the term of a convergent serie for all values of ρ when C^* is large enough.

We can choose the boxes C_i such that, for all $k\leqslant d$, $|t_{ik}-s_{ik}|<\tau_k\cdot m^{-1}$. Since the d.f. of C_k is γ-lipschitzian, the c.d.f. of X_k is γ'-lipschitzian for all k in J, with $\gamma'=\gamma\wedge 1$. So there exist constants c_k such that $P(X_k\in(s_{ik},t_{ik}])\leqslant c_km^{-\gamma'}$ for all k in J. Moreover, since K is lipschitzian on τ_I

$$\begin{aligned}
p_2&\leqslant P\left(\max_i\sup_{\mathbf{z}\in C_i}\left|\int(K_h(\mathbf{z}_I-\mathbf{u}_I)-K_h(\mathbf{t}_{iI}-\mathbf{u}_I))\right.\right.\\
&\qquad\left.\left.\times\frac{(\hat{H}-H)}{H^2}(\mathbf{u}_I-,\mathbf{x}_J)\cdot(\hat{S}_I-S_I)(d\mathbf{u}_I,\mathbf{x}_J)\right|>z_2\right)\\
&\leqslant P\left(\max_i\sup_{\mathbf{z}\in C_i}\|\mathbf{t}_{iI}-\mathbf{z}_I\|\cdot\|\hat{H}-H\|_\infty\right.\\
&\qquad\left.\times\left|\int(\hat{S}_I+S_I)(d\mathbf{u}_I,\mathbf{x}_J)\right|>Cst\cdot h^{r+1}z_2\right)\\
&\leqslant P(\|\hat{H}-H\|_\infty>Cst\cdot mh^{r+1}z_2)
\end{aligned}$$

So far, we have obtained:

$$p_2 \leqslant m^d \exp(Cst \cdot mh^{r+1} z_2 + Cst \cdot (mh^{r+1} z_2)^2) \cdot n^{2d} \cdot \exp(-Cst \cdot n(mh^{r+1} z_2)^2) \tag{A.9}$$

The last term p_3 is similar with the one in the proof of Theorem 2.1; with the same notations (see A.3), we remark that:

$$\begin{aligned}
p_3 &= P\left(\max_i \sup_{\mathbf{z} \in C_i} \left| \int K_h(\mathbf{t}_{iI} - \mathbf{u}_I)[\tilde{r}_n(d\mathbf{u}_I, \mathbf{z}_J) - \tilde{r}_n(d\mathbf{u}_I, \mathbf{t}_{iJ})] \right| > z_3 \right) \\
&\leqslant P\left(\max_i \sup_{\mathbf{z} \in C_i} \left| \int K_h(\mathbf{t}_{iI} - \mathbf{u}_I) \left[\frac{(\hat{H} - H)}{H^2} (\mathbf{u}_I -, \mathbf{z}_J) \cdot (\hat{S}_I - S_I)(d\mathbf{u}_I, \mathbf{z}_J) \right.\right.\right. \\
&\qquad \left.\left.\left. - \frac{(\hat{H} - H)}{H^2} (\mathbf{u}_I -, \mathbf{t}_{iJ}) \cdot (\hat{S}_I - S_I)(d\mathbf{u}_I, \mathbf{t}_{iJ}) \right] \right| > z_3 \right) \\
&\leqslant \sum_{k=1}^{d-r} P\left(\max_i \sup_{\mathbf{z} \in C_i} \left| \int K_h(\mathbf{t}_{iI} - \mathbf{u}_I)\, f_k(\mathbf{z}, d\mathbf{u}_I, \mathbf{t}_{iJ}) \right| > z_3/d \right)
\end{aligned}$$

Writing the kth coordinate only,

$$\begin{aligned}
&\left| \int K_h(\mathbf{t}_{iI} - \mathbf{u}_I)\, f_k(\mathbf{z}, d\mathbf{u}_I, \mathbf{t}_{iJ}) \right| \\
&\quad = \left| \int K_h(\mathbf{t}_{iI} - \mathbf{u}_I) \left\{ \frac{(\hat{H} - H)(\mathbf{u}_I -, t_{ik}) - (\hat{H} - H)(\mathbf{u}_I -, z_k)}{H^2(\mathbf{u}_I -, t_{ik})} \right.\right. \\
&\qquad \times (\hat{S}_I - S_I)(d\mathbf{u}_I, t_{ik}) + (\hat{H} - H)(\mathbf{u}_I -, z_k) \\
&\qquad \times \left[\frac{1}{H^2} (\mathbf{u}_I -, t_{ik}) - \frac{1}{H^2} (\mathbf{u}_I -, z_k) \right] (\hat{S}_I - S_I)(d\mathbf{u}_I, t_{ik}) \\
&\qquad \left.\left. + \frac{(\hat{H} - H)}{H^2} (\mathbf{u}_I -, z_k)[(\hat{S}_I - S_I)(d\mathbf{u}_I, t_{ik}) - (\hat{S}_I - S_I)(d\mathbf{u}_I, z_k)] \right\} \right| \\
&\quad \leqslant Cte \cdot (-1)^r \int |K_h|\, (\mathbf{t}_{iI} - \mathbf{u}_I) \cdot (\hat{S}_I + S_I)(d\mathbf{u}_I, t_{ik}) \\
&\qquad \times \{ \sup_{\mathbf{u}_I \in \tau_I,\, \mathbf{z} \in C_i} |(\hat{H} - H)(\mathbf{u}_I -, t_{ik}) - (\hat{H} - H)(\mathbf{u}_I -, z_k)| \\
&\qquad + m^{-\gamma'} \cdot \|\hat{H} - H\|_\infty \} \\
&\qquad + Cst \cdot (-1)^r \int |K_h|\, (\mathbf{t}_{iI} - \mathbf{u}_I) \cdot (\hat{S}_I + S_I)(d\mathbf{u}_I,]z_k, t_{ik}]) \cdot \|\hat{H} - H\|_\infty
\end{aligned} \tag{A.10}$$

Since $\max_i \sup_{\mathbf{z} \in C_i} |\int |K_h| (\mathbf{t}_{iI} - \mathbf{u}_I) \cdot F_I(d\mathbf{u}_I, t_{ik})| \leqslant C_3$ a.e., we deduce

$$\left| \int |K_h| (\mathbf{t}_{iI} - \mathbf{u}_I) \cdot \hat{S}_I(d\mathbf{u}_I, t_{ik}) \right|$$

$$\leqslant (-1)^r \int |K_h| (\mathbf{t}_{iI} - \mathbf{u}_I) \hat{F}_I(d\mathbf{u}_I)$$

$$\leqslant \left| \int |K_h| (\mathbf{t}_{iI} - \mathbf{u}_I)(\hat{F}_I - F_I)(d\mathbf{u}_I) \right| + (-1)^r \int |K_h| (\mathbf{t}_{iI} - \mathbf{u}_I) F_I(d\mathbf{u}_I)$$

$$\leqslant \left| n^{-1} \sum_{j=1}^{n} |K_h| (\mathbf{t}_{iI} - \mathbf{T}_{jI}) - E(|K_h| (\mathbf{t}_{iI} - \mathbf{T}_I)) \right| + C_3$$

Due to Bennett's inequality (Shorack and Wellner (1986), p. 851), for every constant $s > 2C_3$,

$$P\left(\max_i \sup_{\mathbf{z} \in C_i} \left| \int |K_h| (\mathbf{t}_{iI} - \mathbf{u}_I) \cdot (\hat{S}_I - S_I)(d\mathbf{u}_I, t_{ik}) \right| > s \right)$$

$$\leqslant P\left(\max_i n^{-1} \left| \sum_{j=1}^{n} |K_h| (\mathbf{t}_{iI} - \mathbf{T}_{jI}) - E(|K_h| (\mathbf{t}_{iI} - \mathbf{T}_{jI})) \right| > s - 2C_3 \right)$$

$$\leqslant 2m^d \exp\left(-n(s - 2C_3)^2 \psi\left(\frac{(s - 2C_3) \|K\|}{h^r \sigma^2} \right) \Big/ (2\sigma^2) \right)$$

where $\sigma^2 = \mathrm{Var}(|K_h| (\mathbf{t}_{iI} - \mathbf{T}_I))$ and $\psi(x) = (2/x^2)[(1+x) \ln(1+x) - x]$.

As $\sigma^2 = O(h^{-r})$, the term $\psi(((s - 2C_3) \|K\|)/h^r \sigma^2)$ is bounded from below. If $m_n \leqslant Cst \cdot n^\rho$ for all integers n, and since $nh_n^r/\ln n \to \infty$, we obtain for all $s > 2C_3$

$$P\left(\max_i \sup_{\mathbf{z} \in C_i} \left| \int |K_h| (\mathbf{t}_{iI} - \mathbf{u}_I) \cdot (\hat{S}_I + S_I)(d\mathbf{u}_I, t_{ik}) \right| > s \right)$$

$$\leqslant Cst \cdot n^{d\rho} \exp(-Cst \cdot nh^r (s - 2C_3)^2)$$

$$\leqslant Cst \cdot n^{d\rho} \exp(-A(s - 2C_3)^2 \ln n)$$

that is the term of a convergent sum if A is sufficiently large. By Borel–Cantelli,

$$\max_i \sup_{\mathbf{z} \in C_i} \left| \int |K_h| (\mathbf{t}_{iI} - \mathbf{u}_I) \cdot (\hat{S}_I + S_I)(d\mathbf{u}_I, t_{ik}) \right| < 3C_3 \qquad \text{a.e.} \quad \text{(A.11)}$$

The last unknown term is

$$\max_i \sup_{\mathbf{z}\in C_i} \left| \int |K_h| (\mathbf{t}_{iI}-\mathbf{u}_I)\cdot(\hat{S}_I+S_I)(d\mathbf{u}_I,]z_k, t_{ik}]) \right|$$

$$\leqslant \max_i n^{-1} \sum_{j=1}^{n} |K_h| (\mathbf{t}_{iI}-\mathbf{T}_{jI})\cdot \mathbf{1}\{X_{jk}\in(s_{ik}; t_{ik}]\}$$

$$+E(|K_h| (\mathbf{t}_{iI}-\mathbf{T}_I)\cdot \mathbf{1}\{X_k\in(s_{ik}; t_{ik}]\})$$

$$\leqslant \max_i \left(n^{-1} \sum_{j=1}^{n} |K_h| (\mathbf{t}_{iI}-\mathbf{T}_{jI})\cdot \mathbf{1}\{X_{jk}\in(s_{ik}; t_{ik}]\} \right)$$

$$-E(|K_h| (\mathbf{t}_{iI}-\mathbf{T}_{jI})\cdot \mathbf{1}\{X_{jk}\in(s_{ik}; t_{ik}]\})$$

$$+2\max_i E(|K_h| (\mathbf{t}_{iI}-\mathbf{T}_I)\cdot \mathbf{1}\{X_k\in(s_{ik}; t_{ik}]\})$$

Denoting $I'=I\cup\{k\}$, we have for all i and k,

$$E(|K_h| (\mathbf{t}_{iI}-\mathbf{T}_I)\cdot \mathbf{1}\{X_k\in(s_{ik}; t_{ik}]\})$$

$$\leqslant E(|K_h| (\mathbf{t}_{iI}-\mathbf{T}_I)\, \mathbf{1}\{T_k\in(s_{ik}; t_{ik}]\})$$

$$+E(|K_h| (\mathbf{t}_{iI}-\mathbf{T}_I)\, \mathbf{1}\{C_k\in(s_{ik}; t_{ik}]\})$$

$$\leqslant \int_{s_{ik}}^{t_{ik}} \left\{ \int |K| (\mathbf{v}_I)\, f_{I'}(\mathbf{t}_{iI}-h\mathbf{v}_I, x_k)\, d\mathbf{v}_I \right\} dx_k + Cst\cdot m^{-\gamma}$$

$$\leqslant Cst\cdot|t_{ik}-s_{ik}| + Cst\cdot m^{-\gamma} \leqslant Cst\cdot m^{-\gamma'}$$

By Bennett's inequality, for all constants $s>0$ and $0<\alpha<1$,

$$P\left(n^{-1} \left| \sum_{j=1}^{n} |K_h| (\mathbf{t}_{iI}-\mathbf{T}_{jI})\cdot \mathbf{1}\{X_{jk}\in(s_{ik}; t_{ik}]\} \right.\right.$$

$$\left.\left. -E(|K_h| (\mathbf{t}_{iI}-\mathbf{T}_{jI})\cdot \mathbf{1}\{X_{jk}\in(s_{ik}; t_{ik}]\}) \right| > sm^{-(1-\alpha)} \right)$$

$$\leqslant 2\exp\left(-\frac{ns^2}{2\sigma^2 m^{2(1-\alpha)}}\, \psi\left(\frac{s\,\|K\|}{m^{1-\alpha}\sigma^2 h^r} \right) \right) \tag{A.12}$$

where $\sigma^2=\operatorname{Var}(|K_h| (\mathbf{t}_{iI}-\mathbf{T}_{jI})\cdot \mathbf{1}\{X_{jk}\in(s_{ik}; t_{ik}]\}) \leqslant Cst\cdot(m^{\gamma'}h^r)^{-1}$.

Also $m\sigma^2h^r$ is bounded for all n, and $\psi((sm^\alpha\,\|K\|)/(m\sigma^2h^r))$ is greater than $Cst\cdot s^{-1}m^{-\alpha}\ln(sm^\alpha)$ if $m_n \xrightarrow[n]{} \infty$. So, if for all $\gamma>0$,

$$\sum_n m_n^d \exp(-\gamma n h_n^r m_n^{\alpha-1}\ln m_n) < \infty \tag{A.13}$$

then, by Borel–Cantelli, a.e.

$$\max_i n^{-1} \left| \sum_{j=1}^{n} |K_h| (\mathbf{t}_{iI} - \mathbf{T}_{jI}) \cdot \mathbf{1}\{X_{jk} \in (s_{ik}; t_{ik}]\} - E(|K_h| (\mathbf{t}_{iI} - \mathbf{T}_{jI}) \cdot \mathbf{1}\{X_{jk} \in (s_{ik}; t_{ik}]\}) \right| < Cst/m_n^{1-\alpha}$$

Recall inequality A.10. We have obtained that, if $m_n \to \infty$, then a.e.

$$\sup_{\mathbf{z} \in C_i} \left| \int K_h(\mathbf{t}_{iI} - \mathbf{u}_I) \, f_k(\mathbf{z}, d\mathbf{u}_I, \mathbf{t}_{iJ}) \right| \leqslant Cst \cdot \sup_{\mathbf{u}_I \in \tau_I,\, \mathbf{z} \in C_i} |(\hat{H} - H)(\mathbf{u}_I -, t_{ik}) - (\hat{H} - H)(\mathbf{u}_I, z_{ik})| + Cst \cdot \|\hat{H} - H\|_\infty \, [m^{-\gamma'} + m^{-(1-\alpha)}] \tag{A.14}$$

So, if $z_3 = z_4 + z_5$ and if $0 < 1 - \alpha < \gamma \wedge 1$, we obtain with the same manipulations as in Theorem 2.1 (see the development on the majorants in A.4 and A.5):

$$p_3 \leqslant \sum_{k=1}^{d-r} P(\max_i \sup_{\mathbf{z} \in C_i,\, \mathbf{u}_I \in \tau_I} |(\hat{H} - H)(\mathbf{u}_I -, t_{ik}) - (\hat{H} - H)(\mathbf{u}_I -, z_k)| > C_4 z_4) + (d-r) \, P(\|\hat{H} - H\|_\infty > m^{1-\alpha} C_5 z_5) \leqslant p_4 + p_5$$

But, using the results of Stute (1984),

$$p_4 \leqslant Cst \cdot m_n^d \exp(-Cst \cdot n z_4) \tag{A.15}$$

if $nz_4 \to \infty$ and $nm_n z_4^2 \to \infty$. Moreover,

$$p_5 \leqslant Cst \cdot \exp(Cst \cdot m^{1-\alpha} z_5 + Cst \cdot m^{2(1-\alpha)} z_5^2) \, n^{2d} \exp(-Cst \cdot n m^{2(1-\alpha)} z_5^2) \tag{A.16}$$

We can set $m = [(nh^r/\ln n)^{1/(1-\alpha)}]$; we remark that the conditions A.13 is satisfied since, for all $\gamma > 0$:

$$\sum_n m_n^d \exp(-\gamma n h_n^r m_n^{-(1-\alpha)} \ln m_n) \leqslant \sum_n n^{d/(1-\alpha)} \exp(-\gamma \ln n \cdot \ln m_n) \leqslant \infty$$

Set $z = C^* n^{-1} h^{-r/2} \ln n$, $z_1 = z_2 = z_3 = z/3$ and $z_4 = z_5 = z/6$. Choose α sufficiently close to 1 so that $n m_n z_4^2 \xrightarrow[n]{} \infty$ (so A.15 is true) and so that

$$\frac{3-\alpha}{1+\alpha} r + 2 \frac{1-\alpha}{1+\alpha} \leqslant r(1+\varepsilon) \tag{A.17}$$

Recalling A.9, the last condition implies

$$\begin{aligned}\sum_n p_2(n) &\leqslant Cst \cdot \sum n^{2d + d/(1-\alpha)} \\ &\qquad \times \exp(-Cst \{ n h^{[r(3-\alpha) + 2(1-\alpha)]/(1+\alpha)} / \ln n \}^{(1+\alpha)/(1-\alpha)} \cdot \ln n) \\ &\leqslant Cst \cdot \sum n^{2d + d/(1-\alpha)} \exp(-Cst \{ n h^{r(1+\varepsilon)} / \ln n \}^{(1+\alpha)/(1-\alpha)} \cdot \ln n) < \infty\end{aligned}$$

Recalling A.15 and A.16, we obtain

$$\begin{aligned}\sum_n p_3(n) &\leqslant Cst \cdot \sum_n n^{d/(1-\alpha)} \exp(-Cst \cdot h_n^{-r/2} \ln n) \\ &\qquad + Cst \cdot \sum_n (2n)^{2d} \exp(-Cst \cdot n (n h^r / \ln n)^2 \cdot (\ln n / n h^{r/2})^2) \leqslant +\infty\end{aligned}$$

Endly $\sum_n p_1(n) < \infty$ due to our choice of z_1, if C^* sufficiently large. By Borel–Cantelli, the proof is completed. ∎

Proof of Proposition 2.2.

$$\mathrm{Cov}(\xi_{iI}(\mathbf{x}), \xi_{iI}(\mathbf{y})) = \int K_h(\mathbf{x}_I - \mathbf{u}_I)\, K_h(\mathbf{y}_I - \mathbf{v}_I)\, d \cdot \mathrm{Cov}(\eta_{iI}(\mathbf{u}_I, \mathbf{x}_J), \eta_{iI}(\mathbf{v}_I, \mathbf{y}_J))$$

where the differentiation is with respect to $\mathbf{u}_I$ and $\mathbf{v}_I$. Using Proposition 2.1,

$$\begin{aligned}&\mathrm{Cov}(\xi_{iI}(\mathbf{x}), \xi_{iI}(\mathbf{y})) \\ &\quad = (-1)^r \int K_h(\mathbf{x}_I - \mathbf{u}_I)\, K_h(\mathbf{y}_I - \mathbf{u}_I) \frac{S_I(d\mathbf{u}_I, \mathbf{x}_J \vee \mathbf{y}_J)}{H(\mathbf{u}_I -, \mathbf{x}_J)\, H(\mathbf{u}_I -, \mathbf{y}_J)} \\ &\qquad + \int K_h(\mathbf{x}_I - \mathbf{u}_I)\, K_h(\mathbf{y}_I - \mathbf{v}_I) \\ &\qquad \times \left[\frac{S_I(d\mathbf{u}_I, \mathbf{x}_J)\, S_I(d\mathbf{v}_I, \mathbf{y}_J)}{H^2(\mathbf{u}_I -, \mathbf{x}_J)\, H^2(\mathbf{v}_I -, \mathbf{y}_J)} H(\mathbf{u}_I \vee \mathbf{v}_I -, \mathbf{x}_J \vee \mathbf{y}_J) \right. \\ &\qquad - \frac{S_I(d\mathbf{u}_I, \mathbf{x}_J)\, S_I(d\mathbf{v}_I, \mathbf{x}_J \vee \mathbf{y}_J)}{H^2(\mathbf{u}_I -, \mathbf{x}_J)\, H(\mathbf{v}_I -, \mathbf{y}_J)} \mathbf{1}\{\mathbf{u}_I \leqslant \mathbf{v}_I\} \\ &\qquad \left. - \frac{S_I(d\mathbf{v}_I, \mathbf{y}_J)\, S_I(d\mathbf{u}_I, \mathbf{x}_J \vee \mathbf{y}_J)}{H^2(\mathbf{v}_I -, \mathbf{y}_J)\, H(\mathbf{u}_I -, \mathbf{x}_J)} \mathbf{1}\{\mathbf{v}_I \leqslant \mathbf{u}_I\} \right] = C_1 + C_2\end{aligned}$$

C_2 is convergent because it is the convolution of a kernel in $\mathbb{R}^{2r}$ by a function with discontinuities of the first kind (see Fermanian (1996)). Moreover,

$$C_1 = (-1)^r \int K_h(\mathbf{x}_I - \mathbf{u}_I)\, K_h(\mathbf{y}_I - \mathbf{u}_I)\, \frac{S_I(d\mathbf{u}_I, \mathbf{x}_J \vee \mathbf{y}_J)}{H(\mathbf{u}_I -, \mathbf{x}_J)\, H(\mathbf{u}_I -, \mathbf{y}_J)}$$

$$= h^{-2r} \int K^2((\mathbf{x}_I - \mathbf{u}_I)/h)\, \frac{H(\mathbf{u}_I -, \mathbf{x}_J \vee \mathbf{y}_J)\, \lambda_I(\mathbf{u}_I, \mathbf{x}_J \vee \mathbf{y}_J)}{H(\mathbf{u}_I -, \mathbf{x}_J)\, H(\mathbf{u}_I -, \mathbf{y}_J)}\, d\mathbf{u}_I$$

$$= h^{-r} N_h * \phi_{(\mathbf{x}_J, \mathbf{y}_J)}(\mathbf{x}_I)$$

where N is a kernel of $\mathbb{R}^r$ and $\phi_{(\mathbf{x}_J, \mathbf{y}_J)}$ is defined by

$$\phi_{(\mathbf{x}_J, \mathbf{y}_J)}(\mathbf{u}_I) = \int K^2 \cdot \frac{H(\mathbf{u}_I -, \mathbf{x}_J \vee \mathbf{y}_J)\, \lambda_I(\mathbf{u}_I, \mathbf{x}_J \vee \mathbf{y}_J)}{H(\mathbf{u}_I -, \mathbf{x}_J)\, H(\mathbf{u}_I -, \mathbf{y}_J)}$$

Due to Bochner's lemma, it is clear that $C_1 \sim h^{-r} \Phi(\mathbf{x}, \mathbf{y})$. ∎

Remark A.1. Note the rest is uniform on $\mathbf{u}_I \in \tau_I$ if $\phi(\mathbf{x}, \mathbf{y})$ is uniformly continuous on τ_I.

Proof of Proposition 2.3. It is enough to prove the convergence of $\bar{\xi}_1(\mathbf{x})$ uniformly on τ. We use the same decomposition of τ as in Theorem 2.2's proof: $\tau = \bigcup_{i=1}^{m^d} C_i$ with $C_i = (\mathbf{s}_i, \mathbf{t}_i]$ such that, setting $\gamma' = \gamma \wedge \mathbf{1}$, for all $i = \mathbf{1}, ..., m^d$,

$$\|\mathbf{t}_{iI} - \mathbf{s}_{iI}\| \leqslant Cst \cdot m^{-1} \qquad P(X_k \in (s_{ik}, t_{ik}]) \leqslant m^{-\gamma'} \qquad \text{for all} \quad k \in J$$

Each $\mathbf{x} \in \tau$ belongs to a unique $C_{i(\mathbf{x})} = (\mathbf{s}_{i(\mathbf{x})}, \mathbf{t}_{i(\mathbf{x})}]$. The index $i(\mathbf{x})$ will be denoted by i. We prove first that $\sup_{\mathbf{x} \in \tau} |\bar{\xi}_1(\mathbf{x}) - \bar{\xi}_1(\mathbf{t}_i)| \xrightarrow[n \to \infty]{} 0$ a.e.

Notice that $\bar{\xi}_1(\mathbf{x}) - \bar{\xi}_1(\mathbf{t}_i) = (-1)^r\, (A_1(\mathbf{x}) + A_2(\mathbf{x}) + A_3(\mathbf{x}))$ where

$$A_1(\mathbf{x}) = \int \frac{(\hat{S}_I - S_I)(d\mathbf{u}_I, \mathbf{x}_J)}{H(\mathbf{u}_I -, \mathbf{x}_J)}\, (K_h(\mathbf{x}_I - \mathbf{u}_I) - K_h(\mathbf{t}_{iI} - \mathbf{u}_I))$$

$$A_2(\mathbf{x}) = \int \left(\frac{1}{H(\mathbf{u}_I -, \mathbf{x}_J)} - \frac{1}{H(\mathbf{u}_I -, \mathbf{t}_{iJ})} \right) K_h(\mathbf{t}_{iI} - \mathbf{u}_I) \cdot (\hat{S}_I - S_I)(d\mathbf{u}_I, \mathbf{x}_J)$$

$$A_3(\mathbf{x}) = \int \frac{K_h(\mathbf{t}_{iI} - \mathbf{x}_I)}{H(\mathbf{u}_I -, \mathbf{t}_{iJ})}\, [(\hat{S}_I - S_I)(d\mathbf{u}_I, \mathbf{x}_J) - (\hat{S}_I - S_I)(d\mathbf{u}_I, \mathbf{t}_{iJ})]$$

Since K is lipschitzian, $|A_1(\mathbf{x})| \leqslant Cst \cdot \|\mathbf{x}_I - \mathbf{t}_{iI}\| \cdot h^{-r-1}$ and we have

$$\sup_{\mathbf{x} \in \tau} |A_1(\mathbf{x})| \leqslant Cst \cdot (mh^{r+1})^{-1} \qquad \text{(A.18)}$$

Since $H(\mathbf{u}_I-, \mathbf{x}_J)-H(\mathbf{u}_I-, \mathbf{t}_{iJ}) \leqslant P(\exists p\in J \quad X_p \in (\mathbf{x}_p, t_{ip}]) \leqslant dm^{-\gamma'}$, we have, using A.11,

$$\sup_{\mathbf{x}\in\tau} |A_2(\mathbf{x})| \leqslant Cst\cdot m^{-\gamma'} \tag{A.19}$$

We suppose also that $\lim_{n\to\infty} m_n h_n^{r+1} = +\infty$. Moreover,

$$|A_3(x)| \leqslant Cste\cdot \frac{\|K\|}{h^r}\cdot (\hat{S}_I+S_I)\left(0_I, \bigcup_{k\in J} (z_k, t_{ik}]\right)$$

Then, the same developments as in Theorem 2.2's proof imply

$$\sup_{\mathbf{x}\in\tau} |A_3(\mathbf{x})| = O(m^{-\gamma'}h^{-r}) \tag{A.20}$$

The inequalities A.18, A.19 and A.20 give

$$\sup_{\mathbf{x}\in\tau} |A_1(\mathbf{x})+A_2(\mathbf{x})+A_3(\mathbf{x})| = O(m^{-\gamma'}h^{-r-1}) \qquad \text{a.e.} \tag{A.21}$$

Choose $m=n^\rho$ with ρ is sufficiently large such that a.e.

$$\left(\frac{nh^r}{\ln n}\right)^{1/2} \sup_{\mathbf{x}\in\tau} |A_1(\mathbf{x})+A_2(\mathbf{x})+A_3(\mathbf{x})| = o(1)$$

Since $\sup_{\mathbf{x}\in\tau} |\bar{\xi}_1(\mathbf{x})-\bar{\xi}_1(\mathbf{t}_i)| \xrightarrow[n\to\infty]{} 0$ a.e., it remains to show that

$$\sup_{1\leqslant i\leqslant m^d} |\bar{\xi}_1(\mathbf{t}_i)| \xrightarrow[n\to\infty]{} 0 \qquad \text{a.e.} \tag{A.22}$$

Write $\bar{\xi}_1(\mathbf{x}) = (nh^r)^{-1}\sum_{k=1}^n Z_k(\mathbf{x})$ with

$$Z_k(\mathbf{x}) = Y_{kI}(\mathbf{x}_J)\, K\left(\frac{\mathbf{x}_I-\mathbf{X}_{kI}}{h}\right) - E\left(Y_{kI}(\mathbf{x}_J)\, K\left(\frac{\mathbf{x}_I-\mathbf{X}_{kI}}{h}\right)\right).$$

For all integer $p\geqslant 2$,

$$\begin{aligned} &|E(\{Y_{kI}(\mathbf{x}_J)\, K((\mathbf{x}_I-\mathbf{X}_{kI})/h)\}^p)| \\ &\quad\leqslant Cst\cdot(\|K\|/H(\tau))^{p-2}\, h^r\cdot \int K^2(\mathbf{v}_I)\, f_I(\mathbf{x}_I-h\mathbf{v}_I)\, d\mathbf{v}_I \\ &\quad\leqslant Cst^{p-2}h^r\cdot \sup_\tau f_I\cdot \int K^2 \end{aligned}$$

Using Newton's formula, we deduce easily that, for all integers p,

$$E\,|Z_k(\mathbf{x})|^p \leqslant Cst^p h^r.$$

Due to the Bernstein's inequality (Shorack and Wellner, p. 855), we obtain for all $\varepsilon > 0$,

$$P(\sup_{1 \leqslant i \leqslant m^d} |\bar{\xi}_1(\mathbf{t}_i)| > \varepsilon) \leqslant m^d P\left((nh_n^r)^{-1}\left|\sum_{k=1}^{n} Z_k(\mathbf{t}_k)\right| > \varepsilon\right)$$
$$\leqslant 2m^d \exp(-Cst \cdot nh^r \varepsilon^2)$$

Choose $\varepsilon = C^*(\ln n/(nh^r))^{1/2}$. Then $P(\sup_{1 \leqslant i \leqslant m^d} |\bar{\xi}_1(\mathbf{t}_i)| > \varepsilon)$ is the term of a convergent sum for C^* sufficiently large. By Borel–Cantelli, A.22 is proved, hence the result. ∎

Proof of Proposition 2.5. To study *MSE* and *AMSE of* $\hat{\lambda}_I$, it is clear that:

$$E((\hat{\lambda}_I - \lambda_I)^2\,(\mathbf{x})) = \mathrm{Var}(\hat{\lambda}_I(\mathbf{x})) + (K_h * \lambda_I - \lambda_I)^2\,(\mathbf{x})$$
$$= \mathrm{Var}(\bar{\xi}(\mathbf{x})) + (K_h * \lambda_I - \lambda_I)^2\,(\mathbf{x}) + \varepsilon_n(\mathbf{x})$$

where $\varepsilon_n(\mathbf{x}) = \mathrm{Var}(R_n(\mathbf{x})) + 2 \cdot \mathrm{Cov}(R_n(\mathbf{x}), \bar{\xi}(\mathbf{x}))$.

But, due to Lemma A.2, if f_I is bounded on τ_I then $\mathrm{Var}(\tilde{R}_n(\mathbf{x})) = O(n^{-2}h_n^{-r})$ and this order is true uniformly on τ (we use here the notations introduced in the proof of Theorem 2.2). Moreover, if $nh_n^r/\ln n \to 0$, if f_I is continuous on τ_I, then $\sup_{\mathbf{x} \in \tau} |\bar{R}_n(\mathbf{x})| = O(n^{-1} \ln_2 n)$ a.e. Also $\mathrm{Var}(\bar{R}_n(\mathbf{x})) = O((n^{-1} \ln_2 n)^2)$ uniformly on τ. Using Schwartz's inequality,

$$\sup_{\tau} |\varepsilon_n(\mathbf{x})| = O\left(\left(\frac{1}{nh^r}\right)^{1/2} \cdot \frac{\ln_2 n}{n}\right)$$

By a Taylor–Lagrange's expansion, we find clearly for all $\mathbf{x} \in \tau$

$$(K_h * \lambda_I - \lambda_I)(\mathbf{x}) = h_n^2/2 \cdot \sum_{i \in I} \partial_i^2 \lambda_I(\mathbf{x}) \cdot \int v_i^2 K(\mathbf{v}_I)\, d\mathbf{v}_I + O(h_n^3)$$
$$= h_n^2 \chi(\mathbf{x})/2 + O(h_n^3)$$

and the rest is uniform on τ_I. Hence the result. ∎

Proof of Proposition 2.6. Recalling Remark 2.2(b), for all $\mathbf{x} \in \tau$

$$(nh_n^r/\Phi(\mathbf{x}))^{1/2}\,(\hat{\lambda}_I(\mathbf{x}) - E(\hat{\lambda}_I(\mathbf{x})))$$
$$= \tilde{\xi}(\mathbf{x}) + O((h_n^r \ln_2 n)^{1/2} + n^{-1/2} \ln n \qquad \text{a.e.}$$

Write $\tilde{\xi}(\mathbf{x}) = \sum_{k=1}^{n} \tilde{\xi}_{nk}$, where the r.v. $\tilde{\xi}_{nk}$ are i.i.d. and

$$\tilde{\xi}_{nk} = h_n^{r/2}(n\Phi(\mathbf{x}))^{-1/2} \{ Y_{kI}(\mathbf{x}_J)\, K_{h_n}(\mathbf{x}_I - \mathbf{X}_{kI}) - E(Y_I(\mathbf{x}_J)\, K_{h_n}(\mathbf{x}_I - \mathbf{X}_I))\}$$

We notice that $(n^{1/2}h_n^{r/2}\tilde{\xi}_{nk})_n$ is a bounded sequence and that $\mathrm{Var}(\tilde{\xi}_{nk}(\mathbf{x})) \sim n^{-1}(\mathbf{1} + o(\mathbf{1}))$. So there exists a constant $\delta = 2/\varepsilon$ such that

$$(\mathrm{Var}(\tilde{\xi}(\mathbf{x}))^{-2-\delta} \sum_{k=1}^{n} E\, |\tilde{\xi}_{nk}|^{2+\delta} = O(n^{-\delta/2} h_n^{-r(1+\delta/2)}) \to 0$$

By the Lyapunov's sufficient condition (cf. Billingsley (1968)), $\tilde{\xi}(\mathbf{x})$ converges in law towards a r.v. $\mathcal{N}(0, 1)$. Hence the result. ∎

ACKNOWLEDGMENT

I thank Michel Delecroix for his constant support. D. Fougere, F. Kramarz, G. Laroque, C. P. Robert, M. Visser, and especially I. Gijbels offered me helpful comments.

REFERENCES

[1] Andersen, P. K., Borgan, Ø., Gill, R. D., and Keiding, N. (1993). *Statistical Models Based on Counting Processes*. Springer-Verlag, New York/Berlin.

[2] Bickel, P. J., and Wichura, M. J. (1971). Convergence criteria for multiparameter stochastic processes and some applications. *Ann. Math. Statist.* **42**, 1656–1670.

[3] Billingsley, P. (1968). *Convergence of Probability Measures.* Wiley, New York.

[4] Bosq, D., and Lecoutre, J.-P. (1987). *Théorie de l'estimation fonctionnelle*. Economica, Paris.

[5] Breslow, N., and Crowley, J. (1974). A large sample study of the life table and product limit estimates under random censorship. *Ann. Statist.* **2**, 437–453.

[6] Campbell, G. (1981). Nonparametric bivariate estimation with randomly censored data. *Biometrika* **68**, 417–422.

[7] Campbell, G., and Földes, A. (1982). Large sample properties of nonparametric bivariate estimators with censored data. In *Nonparametric Statistical Inference*, Colloquia Methematica-Societatis János Bolyai (B. V. Gnedenko, M. L. Puri, and I. Vincze, Eds.), pp. 103–122. North-Holland, Amsterdam.

[8] Clayton, D. G. (1978). A model for association in bivariate life tables and its applications in epidemiological studies of familial tendency in chronic disease incidence. *Biometrika* **65**, 141–151.

[9] Dabrowska, D. (1988). Kaplan–Meier estimate on the plane. *Ann. Statist.* **16**, 1475–1489.

[10] Dabrowska, D. (1989). Kaplan–Meier on the plane: Weak convergence, LIL, and the bootstrap. *J. Multivariate Anal.* **29**, 308–325.

[11] Deheuvels, P., and Mason, D. (1992). Fonctionals laws of the iterated logarithm for the increments of empirical and quantile processes. *Ann. Probab.* **20**, 1248–1287.

[12] Dehling, H., Denker, M., and Philipp, W. (1987). The almost sure invariance principle for the empirical process of *U*-statistic structure. *Ann. Inst. Henri Poincaré* **23**, 121–134.

[13] Fermanian, J.-D. (1996). *Multivariate Hazard Rate under Random Censorship*, Working paper **9603**. CREST.
[14] Fermanian, J.-D. (1996). *A Root n Bandwidth Selector in Hazard Estimation*, Working paper **9638**. CREST.
[15] Gijbels, I., and Wang, J. L. (1993). Strong representation of the survival function estimator for truncated and censored data with applications. *J. Multivariate Anal.* **47**, 210–229.
[16] Gill, R. D. (1989). Non- and semi-parametric maximum likelihood estimators and the Von Mises method, 1. *Scand. J. Statist.* **16**, 97–128.
[17] Gill, R. D., Van Der Laan, M. J., and Wellner, J. A. (1993). Inefficient estimators of the bivariate survival function for three models, Preprint 767. Department of Mathematics, University of Utrecht.
[18] Hanley, J. A., and Parnes, M. N. (1983). Nonparametric estimation of a multivariate distribution in the presence of censoring. *Biometrics* **39**, 129–139.
[19] Härdle, W. (1990). *Applied Nonparametric Regression.* Econometric Society Monographs, Vol. 19. Cambridge University Press, Cambridge.
[20] Kaplan, E. L., and Meier, P. (1958). Nonparametric estimation from incomplete observation. *J. Amer. Statist. Assoc.* **53**, 457–481.
[21] Landberg, N. A., and Shaked, M. (1982). On the identifiability of multivariate life distribution functions. *Ann. Probab.* **10**, 773–779.
[22] Lin, D. Y., and Ying, Z. (1993). A simple nonparametric estimator of the bivariate survival function under univariate censoring. *Biometrika* **80**, 573–581.
[23] Lo, S. H., Mack, Y. P., and Wang, J. L. (1989). Density and hazard rate estimation for censored data via strong representation of the Kaplan–Meier estimator. *Probab. Theory Related Fields* **80**, 461–473.
[24] Lo, S. H., and Singh, K. (1986). The product-limit estimator and the bootstrap: some asymptotic representation. *Probab. Theory Related Fields* **71**, 455–465.
[25] Major, P., and Rejtö, L. (1988). Strong embedding of the estimator of the distribution function under random censorship. *Ann. Statist.* **16**, 1113–1132.
[26] Neuhaus, G. (1971). On weak convergence of stochastic processes with multidimensional time parameter. *Ann. Math. Statist.* **42**, 1285–1295.
[27] Oakes, D. (1982). A model for association in bivariate survival data. *J. Roy. Statist. Soc. B* **44**, 414–422.
[28] Pons, O. (1986). A test of independence for two censored survival times. *Scand. J. Statist.* **13**, 173–185.
[29] Pons, O. (1989). *Nonparametric Model and Cox Model for Bivariate Survival Data*, Rapport technique de biométrie 89-02. Laboratoire de biométrie, INRA, Jouy-en-Josas.
[30] Prentice, R. L., and Cai, J. (1992). Covariance and survival function estimation using censored multivariate failure time data. *Biometrika* **79**, 495–512.
[31] Pruitt, R. C. (1991). *Strong Consistency of Self-Consistent Estimators: General Theory and an Application to Bivariate Survival Analysis*, Technical Report 543. University of Minnesota.
[32] Ramlau-Hansen, H. (1983). Smoothing counting process intensities by means of kernel functions. *Ann. Statist.* **11**, 453–466.
[33] Rice, J., and Rosenblatt, M. (1976). Estimation of the log survivor function and hazard functions. *Sankhya Ser. A* **38**, 60–78.
[34] Ruymgaart, F. H. (1989). Some properties of bivariate empirical hazard processes under random censoring. *J. Multivariate Anal.* **28**, 271–281.
[35] Scott, D. (1992). *Multivariate Density Estimation*. Wiley, New York.
[36] Shorack, G. R., and Wellner, J. A. (1986). *Empirical Processes with Applications to Statistics.* Wiley, New York.

[37] Stute, W. (1982). A law of the iterated logarithm for kernel density estimators. *Ann. Probab*. **10**, 414–422.
[38] Stute, W. (1984). The oscillation behavior of empirical processes: the multivariate case. *Ann. Probab*. **12**, 361–369.
[39] Stute, W. (1994). Strong and weak representations of cumulative hazard function and Kaplan–Meier estimators on increasing sets. *J. Statist. Planning Inferences* **42**, 315–329.
[40] Tanner, M. A., and Wong, W. H. (1983). The estimation of the hazard function from randomly censored data by the kernel method. *Ann. Statist*. **11**, 989–993.
[41] Tsai, W. Y., Leurgans, S., and Crowley, J. (1986). Nonparametric estimation of a bivariate survival function in the presence of censoring. *Ann. Statist*. **14**, 1351–1365.
[42] Van Zuijlen, M. C. A. (1978). Properties of the empirical distributions function for independent nonidentically distributed random variables. *Ann. Probab*. **6**, 250–266.
[43] Watson, G. S., and Leadbetter, M. R. (1964a). Hazard analysis, I. *Biometrika* **51**, 175–184.
[44] Watson, G. S., and Leadbetter, M. R. (1964b). Hazard analysis, II. *Sankhya A* **26**, 101–116.
[45] Yandell, S. (1983). Nonparametric inference for rates with censored survival data. *Ann. Statist*. **11**, 1119–1135.
[46] Xiang, X. (1994). Law of the logarithm for density and hazard rate estimation for censored data. *J. Multivariate Anal*. **49**, 278–286