Available online at www.sciencedirect.com

ScienceDirect

Journal of Multivariate Analysis 98 (2007) 1337 – 1355

Journal of Multivariate Analysis

www.elsevier.com/locate/jmva

Hazard rate estimation on random fields

Jiexiang Li, Lanh Tat Tran*

Department of Mathematics, Indiana University, Bloomington, USA

Received 22 September 2005
Available online 14 March 2007

Abstract

Consider observations (representing lifelengths) taken on a random field indexed by lattice points. Our purpose is to estimate the hazard rate $r(x)$, which is the rate of failure at time x for the survivors up to time x. We estimate $r(x)$ by the nonparametric estimator constructed in terms of a kernel-type estimator for $f(x)$ and the natural estimator for $\bar{F}(x)$. Under some general mixing assumptions, the limiting distribution of the estimator at multiple points is shown to be multivariate normal. The result is useful in establishing confidence bands for $r(x)$ with x in an interval.

AMS 1991 subject classification: primary 62G05; 62E20; secondary 62M40; 62N02

Keywords: Random field; Spatial process; Hazard rate; Kernel density estimator; Asymptotic normality

1. Introduction

Denote the integer lattice points in the N-dimensional Euclidean space by Z^N, $N \geqslant 1$. Assume that $X_{\mathbf{i}}$ is a strictly stationary random field indexed by Z^N and defined on some probability space $(\Omega, \mathcal{F}, P)$. Suppose that $X_{\mathbf{i}}$, $\mathbf{i} \in Z^N$, takes values in R^d and has distribution function F and probability density function f with respect to Lebesgue measure. Our purpose is to estimate the hazard rate $r(x)$, defined as

$$r(x) = \frac{f(x)}{\bar{F}(x)}, \quad \bar{F}(x) = P(X > x), \quad \bar{F}(x) > 0.$$

The hazard rate is the rate of failure at time x for the survivors up to time x. The estimation of $r(x)$ in the case of independently and identically observed random variables has been studied extensively. For background information on this, see for example, Ahmad [1], Antoniadis et al.

* Corresponding author.
E-mail address: tran@indiana.edu (L.T. Tran).

doi:10.1016/j.jmva.2007.03.002

[2], Basu [3], Hollander and Proschan [8], Puri and Rubin [13], Rice and Rosenblatt [14], Watson and Leadbetter [19,20] and the references therein. In the classical theory of competing risks, it is assumed that competing risks are independent and that death does not result from simultaneous causes. But these assumptions are not always true. For example, lifelengths of animals in certain locations might be dependent since the animals compete for the same food sources. Landberg et al. [10] have studied the estimation of dependent lifelengths. Roussas [16] investigated the estimation of the hazard rate under various dependence conditions. In the present paper, we assume that the underlying r.v.'s come from a strictly stationary random field satisfying some weak dependence conditions.

A point $\mathbf{i}$ in Z^N will be referred to as a site and written as $\mathbf{i} = \langle i_1, i_2, \dots, i_N\rangle$. Let $\mathcal{S}$ and $\mathcal{S}'$ be two sets of sites. The Borel fields $\mathcal{B}(\mathcal{S}) = \mathcal{B}(X_\mathbf{i}, \mathbf{i} \in \mathcal{S})$ and $\mathcal{B}(\mathcal{S}') = \mathcal{B}(X_\mathbf{i}, \mathbf{i} \in \mathcal{S}')$ are the σ-fields generated by the random variables $X_\mathbf{i}$ with $\mathbf{i}$ in, respectively, $\mathcal{S}$ and $\mathcal{S}'$. Define the distance between $\mathcal{S}$ and $\mathcal{S}'$ as follows: $\hat{d}(\mathcal{S}, \mathcal{S}') = \inf(\hat{d}(\mathbf{i}, \mathbf{j}) : \mathbf{i} \in \mathcal{S}, \mathbf{j} \in \mathcal{S}')$, where $\hat{d}(\mathbf{i}, \mathbf{j})$ is the Euclidean distance between $\mathbf{i}$ and $\mathbf{j}$. We assume that $X_\mathbf{i}$, $\mathbf{i} \in Z^N$, satisfies the following mixing condition: there exists a function $\varphi(t) \downarrow 0$ as $t \to \infty$, such that whenever $\mathcal{S}, \mathcal{S}' \subset Z^N$,

$$\begin{aligned}\alpha(\mathcal{B}(\mathcal{S}), \mathcal{B}(\mathcal{S}')) &= \sup\{|P(AB) - P(A)P(B)|, A \in \mathcal{B}(\mathcal{S}), B \in \mathcal{B}(\mathcal{S}')\} \\ &\leqslant \hat{f}(\mathrm{Card}(\mathcal{S}), \mathrm{Card}(\mathcal{S}'))\varphi(\hat{d}(\mathcal{S}, \mathcal{S}')),\end{aligned} \tag{1.1}$$

where $\mathrm{Card}(\mathcal{S})$ denote the cardinality of $\mathcal{S}$ and $\hat{f}$ is a symmetric positive function nondecreasing in each variable. We assume that $\hat{f}$ satisfies

$$\hat{f}(n, m) \leqslant \min(m, n) \tag{1.2}$$

or

$$\hat{f}(n, m) \leqslant C(n + m + 1)^{\varpi} \tag{1.3}$$

for some ϖ with $\varpi > 1$ and some $C > 0$. Conditions (1.2) and (1.3) are widely used in literature. Note that condition (1.3) is more general than (1.2). Guyon [6] showed that the class of linear processes $X_\mathbf{i} = \sum_{\mathbf{j} \in Z^N} g_\mathbf{j} Z_{\mathbf{i}-\mathbf{j}}$, where $Z_\mathbf{j}$'s are independent random variables, satisfies mixing conditions (1.1) and (1.3) under general conditions. If $\hat{f} \equiv 1$, then $X_\mathbf{i}$ is called strongly mixing, for which the limit theorem is widely investigated. For detailed information on strongly mixing, the readers are referred to Bradley [4]. Let $I_\mathbf{n}$ be a rectangular region defined by $I_\mathbf{n} = \{\mathbf{i} : \mathbf{i} \in Z^N, 1 \leqslant i_k \leqslant n_k, k = 1, \dots, N\}$. Assume that we observe $X_\mathbf{i}$ on $I_\mathbf{n}$. We write $\mathbf{n} \to \infty$ if $\min(n_1, \dots, n_N) \to \infty$. Let $\hat{\mathbf{n}} = n_1 \dots n_N$. In the proof, the letter C denotes a generic constant whose values are unimportant and may vary from line to line.

Assume that the density $f(x)$ and hazard rate $r(x)$ have no parametric forms. We propose a nonparametric estimator $r_\mathbf{n}(x)$ given by

$$r_\mathbf{n}(x) = \frac{f_\mathbf{n}(x)}{\bar{F}_\mathbf{n}(x)},$$

where $\bar{F}_\mathbf{n}(x)$ is the proportion of the $X_\mathbf{i}'$s exceeding x and $f_\mathbf{n}(x)$ is the nonparametric kernel density estimator given by

$$f_\mathbf{n}(x) = (\hat{\mathbf{n}} b_\mathbf{n}^d)^{-1} \sum_{\substack{i_k=1 \\ k=1\dots N}}^{n_k} K((x - X_\mathbf{i})/b_\mathbf{n}), \tag{1.4}$$

where $b_{\mathbf{n}}$ is a sequence of bandwidths tending to zero as $\mathbf{n} \to \infty$. The nonparametric estimator $r_{\mathbf{n}}(x)$ is quite simple and has been discussed in literature under different situations. For instance, Roussas [16] investigated both pointwise convergence and uniform convergence of the estimator in nonspatial case.

Our paper is organized as follows: in Section 2, we present some preliminaries. An application example is given in Section 3. We then establish the basic asymptotic properties of $F_{\mathbf{n}}(x)$ in Section 4. The results of Section 4 are also of independent interest. In Section 5, we study the covariance of kernel density estimate at distinct points. Finally, in Section 6, the joint asymptotic normality of $r_{\mathbf{n}}(x)$ at multiple points is established. Note that, to construct confidence bands for $r(x)$ with x in an interval, the joint asymptotic normality at multiple points is needed.

2. Preliminary results

In this section, we gather the assumptions under which results in the paper hold true and then present some preliminary results.

Assumption 2.1. Suppose for any x, y $\in R^d$ and some constant $\rho > 0$,

$$|f(x) - f(y)| \leqslant \rho \|x - y\|.$$

Remark 2.1. Note that Assumption 2.1 implies that $f(x)$ is bounded.

Assumption 2.2. (i) $K(x)$ is a bounded probability density function on R^d and

$$\int_{R^d} \|x\| K(x)\, dx < \infty.$$

(ii) Assume K has an integrable radial majorant, that is, $\mathcal{Q}(x) \equiv \sup\{K(y) : \|y\| > \|x\|\}$ is integrable.

Assumption 2.3. The joint probability density $f_{\mathbf{i},\mathbf{j}}(x, y)$ of $X_{\mathbf{i}}$ and $X_{\mathbf{j}}$ exists and satisfies $|f_{\mathbf{i},\mathbf{j}}(x, y) - f(x)f(y)| \leqslant C$ for some C and for all x, y and $\mathbf{i}, \mathbf{j}$.

Assumption 2.4. For some a such that $0 < a < \frac{1}{2}$, $\sum_{i=1}^{\infty} i^{N-1}\varphi(i)^a < \infty$.

Remark 2.2. Suppose $\varphi(i) = i^{-\theta}$, then Assumption 2.4 is satisfied if $\frac{1}{2} > a > \frac{N}{\theta}$.

Assumption 2.5. For some $0 < r < 1$,

(i) The bandwidth $b_{\mathbf{n}}$ tends to zero in a manner such that $\hat{\mathbf{n}} b_{\mathbf{n}}^{d(1+(1-r)2N)} \to \infty$.

(ii) There exist three sequences of positive integers with $q = q_{\mathbf{n}} \to \infty$, $s_{\mathbf{n}} \to \infty$, such that

$$\begin{aligned} q_{\mathbf{n}} &= o((\hat{\mathbf{n}} b_{\mathbf{n}}^{d(1+(1-r)2N)})^{1/(2N)}), \\ s_{\mathbf{n}} q_{\mathbf{n}} &= o((\hat{\mathbf{n}} b_{\mathbf{n}}^{d(1+(1-r)2N)})^{1/(2N)}), \\ p_{\mathbf{n}} &= p = (\hat{\mathbf{n}} b_{\mathbf{n}}^{d})^{1/(2N)}/s_{\mathbf{n}}, \\ \hat{\mathbf{n}}\varphi(q) &\to 0. \end{aligned}$$

(Aii) Assume the conditions in (ii) are satisfied except with the last condition replaced by $(\hat{\mathbf{n}}^{\varpi+1}/\hat{p})\varphi(q) \to 0$ where $\hat{p} = p_{\mathbf{n}}^N$.

(iii) $b_{\mathbf{n}}$ tends to zero in such a manner that $b_{\mathbf{n}}^{-d(1-r)} \sum_{i=q}^{\infty} i^{N-1} \varphi(i)^{1-r} \to 0$.
(iv) The bandwidth $b_{\mathbf{n}}$ tends to zero in a manner such that $\hat{\mathbf{n}} b_{\mathbf{n}}^{d+2} \to 0$.

The following example demonstrates that Assumption 2.5 is easily satisfied. In the example, we consider the case where $\varphi(i)$ decreases to zero at an exponential rate. It is generally used in mixing conditions.

Example 2.1. Let $b_{\mathbf{n}} = \widehat{\mathbf{n}}^{-a}$ with $dN(1-r) < \frac{1}{2a} - \frac{d}{2} < 1$; $q = (\widehat{\mathbf{n}} b_{\mathbf{n}}^{d(1+(1-r)2N)})^{1/2N} \widehat{\mathbf{n}}^{-b}$ with $b > \frac{1}{2N} - \frac{ad}{2N} - ad(1-r)$; $s_{\mathbf{n}} = \widehat{\mathbf{n}}^{c}$ with $0 < c < \min\{\frac{1-ad}{2N}, b\}$. Suppose $\varphi(x) = \mathcal{O}(\mathrm{e}^{-\xi x})$ for some $\xi > 0$ and

$$(\widehat{\mathbf{n}} b_{\mathbf{n}}^{d(1+(1-r)2N)})^{1/2N} / (\widehat{\mathbf{n}}^{b} \log \widehat{\mathbf{n}}) \to \infty. \tag{2.1}$$

Then Assumption 2.5 (i)–(iv) holds.

Proof. Note that $1 - ad(1+(1-r)2N) > 0$, hence $\widehat{\mathbf{n}} b_{\mathbf{n}}^{d(1+(1-r)2N)} = \widehat{\mathbf{n}}^{1-ad(1+(1-r)2N)} \to \infty$; the convergence $\widehat{\mathbf{n}} b_{\mathbf{n}}^{d+2} = \widehat{\mathbf{n}}^{1-a(d+2)} \to 0$ follows from $1 - a(d+2) < 0$. Since $b > 0$ from the fact $dN(1-r) < \frac{1}{2a} - \frac{d}{2}$, it follows that $\frac{q}{(\widehat{\mathbf{n}} b_{\mathbf{n}}^{d(1+(1-r)2N)})^{1/2N}} = \widehat{\mathbf{n}}^{-b} \to 0$. By the choice of c, $\frac{s_{\mathbf{n}} q_{\mathbf{n}}}{(\widehat{\mathbf{n}} b_{\mathbf{n}}^{d(1+(1-r)2N)})^{1/2N}} = \widehat{\mathbf{n}}^{c-b} \to 0$ and $p = (\widehat{\mathbf{n}} b_{\mathbf{n}}^{d})^{1/(2N)} / s_{\mathbf{n}} = \widehat{\mathbf{n}}^{(1-ad)/2N-c} \to \infty$. By Eq. (2.1) and the choice of q, for arbitrary $C > 0$ and sufficiently large $\widehat{\mathbf{n}}$,

$$q > C \log \widehat{\mathbf{n}}. \tag{2.2}$$

Hence, $\hat{\mathbf{n}} \varphi(q) \leqslant C \widehat{\mathbf{n}} \mathrm{e}^{-\xi q} \leqslant C \widehat{\mathbf{n}} \exp\{-\xi C \log \widehat{\mathbf{n}}\} = C \widehat{\mathbf{n}}^{1-\xi C} \to 0$ by choosing $C > \frac{1}{\xi}$. For $\zeta < \xi$,

$$\begin{aligned} \sum_{i=q}^{\infty} i^{N-1} \varphi(i)^{1-r} &\leqslant C \sum_{i=q}^{\infty} i^{N-1} \mathrm{e}^{-\xi i(1-r)} \\ &\leqslant C \sum_{i=q}^{\infty} \mathrm{e}^{-\zeta i(1-r)} \\ &\leqslant C \mathrm{e}^{-\zeta q(1-r)}. \end{aligned}$$

By Eq. (2.2), it follows that $\mathrm{e}^{-\zeta q(1-r)} < \widehat{\mathbf{n}}^{-\zeta C(1-r)}$ for sufficiently large $\widehat{\mathbf{n}}$. Also for arbitrary $C > 0$ and sufficiently large $\widehat{\mathbf{n}}$, $b_{\mathbf{n}}^{d} = \widehat{\mathbf{n}}^{-ad} \geqslant C \widehat{\mathbf{n}}^{-1}$. Therefore,

$$b_{\mathbf{n}}^{-d(1-r)} \sum_{i=q}^{\infty} i^{N-1} \varphi(i)^{1-r} \leqslant C \widehat{\mathbf{n}}^{1-r} \widehat{\mathbf{n}}^{-\zeta C(1-r)} \to 0$$

by choosing $C > \frac{1}{\zeta}$. □

Define

$$\sigma^2 := f(x) \int_{R^d} K^2(u)\, du.$$

Lemma 2.1. *Suppose that Assumptions* 2.2–2.5 (i)–(iii) *hold and* $X_{\mathbf{n}}$ *satisfies* (1.1) *and* (1.2). *Then*

$$(\hat{\mathbf{n}} b_{\mathbf{n}}^{d})^{1/2} (f_{\mathbf{n}}(x) - E[f_{\mathbf{n}}(x)]) \to N(0, \sigma^2) \text{ in distribution as } \mathbf{n} \to \infty.$$

Lemma 2.2. *Suppose that Assumptions* 2.2–2.5 (i), (Aii), (iii) *hold and* $X_{\mathbf{n}}$ *satisfies* (1.1) *and* (1.3). *Then*

$$(\hat{\mathbf{n}}b_{\mathbf{n}}^d)^{1/2}(f_{\mathbf{n}}(x) - E[f_{\mathbf{n}}(x)]) \to N(0, \sigma^2) \text{ in distribution as } \mathbf{n} \to \infty.$$

The proofs of the two lemmas above are presented in Tran [18].

3. Convergence result of $\bar{F}_{\mathbf{n}}(x)$

Recall that for each x in R^d,

$$\bar{F}(x) = P(X > x), \quad \bar{F}_{\mathbf{n}}(x) = \hat{\mathbf{n}}^{-1} \sum_{\substack{i_k=1\\k=1\ldots N}}^{n_k} Y_{\mathbf{i}}(x),$$

where $Y_{\mathbf{i}}(x) = 1$ if $X_{\mathbf{i}} > x$ and $Y_{\mathbf{i}}(x) = 0$ otherwise.

Lemma 3.1. (1) *Suppose* (1.1) *holds. Let* $\mathcal{L}_r(\mathcal{F})$ *denote the class of* $\mathcal{F}$*-measurable r.v.'s X satisfying* $\|X\|_r = (E|X|^r)^{1/r} < \infty$. *Let* $X \in \mathcal{L}_r(\mathcal{B}(\mathcal{S}))$ *and* $Y \in \mathcal{L}_s(\mathcal{B}(\mathcal{S}'))$. *Suppose* $1 \leqslant r, s, h < \infty$ *and* $r^{-1} + s^{-1} + h^{-1} = 1$. *Then*

$$|EXY - EXEY| \leqslant C\|X\|_r\|Y\|_s\{\hat{f}(\mathrm{Card}(\mathcal{S}), \mathrm{Card}(\mathcal{S}'))\varphi(\hat{d}(\mathcal{S}, \mathcal{S}'))\}^{1/h}. \tag{3.1}$$

(2) *For r.v.'s bounded with probability* 1, *the right-hand side of* (3.1) *can be replaced by*

$$C\hat{f}(\mathrm{Card}(\mathcal{S}), \mathrm{Card}(\mathcal{S}'))\varphi(\hat{d}(\mathcal{S}, \mathcal{S}')).$$

For more information on this lemma, see Ibragimov and Linnik [9] or Deo [5].

Lemma 3.2. *Suppose that Assumption* 2.4 *holds and* $X_{\mathbf{n}}$ *satisfies* (1.1) *and* (1.3). *Then* $(\hat{\mathbf{n}}b_{\mathbf{n}}^d)^{1/2}(\bar{F}_{\mathbf{n}}(x) - \bar{F}(x)) \to 0$ *in probability as* $\mathbf{n} \to \infty$.

Proof. Clearly, $EY_{\mathbf{i}}(x) = P(X_{\mathbf{i}} > x) = \bar{F}(x)$ by definition of $Y_{\mathbf{i}}(x)$ and stationarity of $X_{\mathbf{i}}$. Therefore,

$$\bar{F}_{\mathbf{n}}(x) - \bar{F}(x) = \hat{\mathbf{n}}^{-1} \sum_{\substack{i_k=1\\k=1\ldots N}}^{n_k} (Y_{\mathbf{i}}(x) - EY_{\mathbf{i}}(x)).$$

For any given $\varepsilon > 0$,

$$\begin{aligned} &P((\hat{\mathbf{n}}b_{\mathbf{n}}^d)^{1/2}|\bar{F}_{\mathbf{n}}(x) - \bar{F}(x)| > \varepsilon)\\ &\quad = P((\hat{\mathbf{n}}^{-1}b_{\mathbf{n}}^d)^{1/2}|\sum_{\substack{i_k=1\\k=1\ldots N}}^{n_k} (Y_{\mathbf{i}}(x) - EY_{\mathbf{i}}(x))| > \varepsilon) \end{aligned}$$

$$\leqslant \frac{b_{\mathbf{n}}^d}{\hat{\mathbf{n}}\varepsilon^2} \mathrm{Var}\left(\sum_{\substack{i_k=1 \\ k=1\ldots N}}^{n_k} Y_{\mathbf{i}}(x) \right)$$

$$\leqslant \frac{b_{\mathbf{n}}^d}{\hat{\mathbf{n}}\varepsilon^2} \left(\sum_{\substack{i_k=1 \\ k=1\ldots N}}^{n_k} \mathrm{Var}(Y_{\mathbf{i}}(x)) + \sum_{\substack{i_k=1 \\ k=1\ldots N \\ i_k \neq j_k}}^{n_k} \sum_{\substack{j_k=1 \\ k=1\ldots N \\ \text{for some } k}}^{n_k} |\mathrm{Cov}(Y_{\mathbf{i}}(x), Y_{\mathbf{j}}(x))| \right).$$

Set

$$K_{\mathbf{n}} = \sum_{\substack{i_k=1 \\ k=1\ldots N}}^{n_k} \mathrm{Var}(Y_{\mathbf{i}}(x))$$

and

$$R_{\mathbf{n}} = \sum_{\substack{i_k=1 \\ k=1\ldots N \\ i_k \neq j_k}}^{n_k} \sum_{\substack{j_k=1 \\ k=1\ldots N \\ \text{for some } k}}^{n_k} |\mathrm{Cov}(Y_{\mathbf{i}}(x), Y_{\mathbf{j}}(x))|.$$

Note that $\mathrm{Var}(Y_{\mathbf{i}}(x)) \leqslant EY_{\mathbf{i}}^2(x) = P(X_{\mathbf{i}} > x) = \bar{F}(x)$. Thus,

$$K_{\mathbf{n}} \leqslant \hat{\mathbf{n}} \bar{F}(x)$$

and subsequently,

$$\frac{b_{\mathbf{n}}^d K_{\mathbf{n}}}{\hat{\mathbf{n}}\varepsilon^2} \leqslant \frac{b_{\mathbf{n}}^d \bar{F}(x)}{\varepsilon^2} \to 0 \quad \text{as } \mathbf{n} \to \infty. \tag{3.2}$$

Since $Y_{\mathbf{i}}$ equals either 0 or 1, by Lemma 3.1 for bounded random variables,

$$|\mathrm{Cov}(Y_{\mathbf{i}}(x), Y_{\mathbf{j}}(x))| \leqslant C \hat{f}(1, 1) \varphi(\hat{d}(\mathbf{i}, \mathbf{j})) \leqslant C \varphi(\|\mathbf{i} - \mathbf{j}\|).$$

A simple computation shows that

$$R_{\mathbf{n}} \leqslant C \sum_{\substack{i_k=1 \\ k=1\ldots N \\ i_k \neq j_k}}^{n_k} \sum_{\substack{j_k=1 \\ k=1\ldots N \\ \text{for some } k}}^{n_k} \varphi(\|\mathbf{i} - \mathbf{j}\|)$$

$$\leqslant C \, \hat{\mathbf{n}} \sum_{\substack{i_k=1 \\ k=1\ldots N}}^{n_k} \varphi(\|\mathbf{i}\|)$$

$$\leqslant C\ \hat{\mathbf{n}} \sum_{j=1}^{\infty} \sum_{j \leqslant \|\mathbf{i}\| < j+1} \varphi(\|\mathbf{i}\|)$$

$$\leqslant C\ \hat{\mathbf{n}} \sum_{j=1}^{\infty} \varphi(j) \sum_{j \leqslant \|\mathbf{i}\| < j+1} 1$$

$$\leqslant C\ \hat{\mathbf{n}} \sum_{j=1}^{\infty} j^{N-1} \varphi(j). \tag{3.3}$$

Also since $\varphi(t) \downarrow 0$ as $t \to \infty$, there exists some i_0 such that for $i > i_0$, $0 \leqslant \varphi(i) \leqslant 1$, which implies that $\varphi(i) \leqslant \varphi(i)^a$. Recall that $0 < a < \frac{1}{2}$. Hence

$$\begin{aligned}\sum_{i=1}^{\infty} i^{N-1} \varphi(i) &= \sum_{i=1}^{i_0} i^{N-1} \varphi(i) + \sum_{i=i_0+1}^{\infty} i^{N-1} \varphi(i) \\ &\leqslant \sum_{i=1}^{i_0} i^{N-1} \varphi(i) + \sum_{i=i_0+1}^{\infty} i^{N-1} \varphi(i)^a \\ &< \infty. \end{aligned} \tag{3.4}$$

The last inequality follows from Assumption 2.4. By (3.3) and (3.4),

$$R_{\mathbf{n}} \leqslant C \hat{\mathbf{n}}.$$

Hence,

$$\frac{b_{\mathbf{n}}^d R_{\mathbf{n}}}{\hat{\mathbf{n}} \varepsilon^2} \leqslant \frac{C\ b_{\mathbf{n}}^d}{\varepsilon^2} \to 0 \quad \text{as } \mathbf{n} \to \infty. \tag{3.5}$$

By (3.2) and (3.5), $(\hat{\mathbf{n}} b_{\mathbf{n}}^d)^{1/2} (\bar{F}_{\mathbf{n}}(x) - \bar{F}(x)) \to 0$ in probability. □

4. Covariance of density estimates at multiple points

In this section, we establish that for any two distinct points x_r and x_s,

$$\hat{\mathbf{n}} b_{\mathbf{n}}^d \operatorname{Cov}(f_{\mathbf{n}}(x_r), f_{\mathbf{n}}(x_s)) \to 0.$$

We make the following assumption:

Assumption 4.1. Suppose that

$$\lim_{\|x\| \to \infty} K(x) = 0.$$

Define $K_{r\mathbf{i}} = \frac{1}{b_{\mathbf{n}}^d} K\left(\frac{x_r - X_{\mathbf{i}}}{b_{\mathbf{n}}}\right)$.

Lemma 4.1. *Suppose that Assumptions* 2.1, 2.2 *and* 4.1 *hold. Then*,

$$\hat{\mathbf{n}}^{-1} b_{\mathbf{n}}^d \sum_{\substack{i_k=1 \\ k=1\ldots N}}^{n_k} \operatorname{Cov}(K_{r\mathbf{i}}, K_{s\mathbf{i}}) \to 0.$$

Proof. Now for x_r and x_s in R^d with $x_r \neq x_s$.

$$\begin{aligned} b_{\mathbf{n}}^d \; E[K_{r\mathbf{i}}K_{s\mathbf{i}}] &= \frac{1}{b_{\mathbf{n}}^d}\int K\left(\frac{x_r-u}{b_{\mathbf{n}}}\right)K\left(\frac{x_s-u}{b_{\mathbf{n}}}\right)f(u)\,du \\ &= \frac{1}{b_{\mathbf{n}}^d}\int K\left(\frac{u}{b_{\mathbf{n}}}\right)K\left(\frac{x_s-x_r+u}{b_{\mathbf{n}}}\right)f(x_r-u)\,du \\ &:= \frac{1}{b_{\mathbf{n}}^d}\int S\,du. \end{aligned}$$

Note,

$$\frac{1}{b_{\mathbf{n}}^d}\int_{\|u\|\leqslant\delta} S\,du \leqslant M\int K(u)K\left(\frac{x_s-x_r}{b_{\mathbf{n}}}+u\right)du \to 0. \tag{4.1}$$

The inequality follows from uniform boundedness of $f(x)$ by Remark 2.1 after Assumption 2.1. The integrand in (4.1) converges to 0 by Assumption 4.1 and is bounded by a constant multiple of $K(u)$ by the uniform boundedness of K. Therefore Eq. (4.1) follows from Lebesgue Dominated Convergence Theorem. We know that

$$\begin{aligned} &\frac{1}{b_{\mathbf{n}}^d}\int_{\|u\|\geqslant\delta} S\,du \\ &\leqslant C\frac{1}{b_{\mathbf{n}}^d}\int_{\|u\|\geqslant\delta} K\left(\frac{u}{b_{\mathbf{n}}}\right)du \\ &\leqslant C\int_{\|u\|\geqslant\delta}\frac{\|u\|}{b_{\mathbf{n}}^d}K\left(\frac{u}{b_{\mathbf{n}}}\right)du \\ &\leqslant C\int\frac{\|u\|}{b_{\mathbf{n}}^d}K\left(\frac{u}{b_{\mathbf{n}}}\right)du \\ &\leqslant C\int b_{\mathbf{n}}\|u\|\,K(u)\,du \\ &\leqslant Cb_{\mathbf{n}} \to 0. \end{aligned} \tag{4.2}$$

The first inequality follows from the uniform boundedness of $K(x)$ and $f(x)$. The last inequality follows from Assumption 2.2 (i). By (4.1) and (4.2),

$$b_{\mathbf{n}}^d \; E[K_{r\mathbf{i}}K_{s\mathbf{i}}] \to 0.$$

Similarly,

$$E[K_{r\mathbf{i}}] \to f(x_r)\int_{R^d} K(u)\,du.$$

Hence,

$$b_{\mathbf{n}}^d\,\mathrm{Cov}(K_{r\mathbf{i}}, K_{s\mathbf{i}}) \to 0.$$

Therefore,

$$\hat{\mathbf{n}}^{-1}b_{\mathbf{n}}^d\sum_{\substack{i_k=1\\k=1\ldots N}}^{n_k}\mathrm{Cov}(K_{r\mathbf{i}}, K_{s\mathbf{i}}) = b_{\mathbf{n}}^d\,\mathrm{Cov}(K_{r\mathbf{1}}, K_{s\mathbf{1}}) \to 0. \qquad \square$$

Lemma 4.2. *Suppose that Assumptions* 2.2–2.4 *hold and* $X_{\mathbf{n}}$ *satisfies* (1.1) *and* (1.3). *Then*

$$S := \hat{\mathbf{n}}^{-1} b_{\mathbf{n}}^d \sum_{\substack{i_k=1\\ k=1\ldots N\\ i_k \neq j_k \text{ for some } k}}^{n_k} \sum_{\substack{j_k=1\\ k=1\ldots N}}^{n_k} |\mathrm{Cov}(K_{r\mathbf{i}}, K_{s\mathbf{j}})| \to 0.$$

Let $c_{\mathbf{n}} = b_{\mathbf{n}}^{-d(1-\gamma)/\upsilon}$ where $\upsilon = -N - \varepsilon + (1-\gamma)Na^{-1}$ with γ and ε being small positive numbers such that $a^{-1} - (N+\varepsilon)(N(1-\gamma))^{-1} > 1$. This can be done since $0 < a < \frac{1}{2}$. Thus, $\upsilon > N(1-\gamma)$. Note $c_{\mathbf{n}} \to \infty$ as $\mathbf{n} \to \infty$. Define

$$S_1 = \{\mathbf{i}, \mathbf{j} \in I_{\mathbf{n}} | 0 < \hat{d}(\mathbf{i}, \mathbf{j}) \leqslant c_{\mathbf{n}}\},$$
$$S_2 = \{\mathbf{i}, \mathbf{j} \in I_{\mathbf{n}} | \hat{d}(\mathbf{i}, \mathbf{j}) > c_{\mathbf{n}}\}.$$

We can decompose the sum S in the following way:

$$S = \hat{\mathbf{n}}^{-1} b_{\mathbf{n}}^d \sum_{\mathbf{i},\mathbf{j} \in S_1} |\mathrm{Cov}(K_{r\mathbf{i}}, K_{s\mathbf{j}})| + \hat{\mathbf{n}}^{-1} b_{\mathbf{n}}^d \sum_{\mathbf{i},\mathbf{j} \in S_2} |\mathrm{Cov}(K_{r\mathbf{i}}, K_{s\mathbf{j}})| := J_1 + J_2. \tag{4.3}$$

Since

$$b_{\mathbf{n}}^d |\mathrm{Cov}(K_{r\mathbf{i}}, K_{s\mathbf{j}})| \leqslant \int\int K\left(\frac{x_r - u}{b_{\mathbf{n}}}\right) K\left(\frac{x_s - \upsilon}{b_{\mathbf{n}}}\right) \frac{1}{b_{\mathbf{n}}^d} |f(u, \upsilon) - f(u) f(\upsilon)| \, du \, d\upsilon,$$

it follows that

$$J_1 \leqslant C \hat{\mathbf{n}}^{-1} b_{\mathbf{n}}^d \left\{ \int K(u) \, du \right\}^2 \sum_{\mathbf{i},\mathbf{j} \in S_1} 1$$
$$\leqslant C b_{\mathbf{n}}^d c_{\mathbf{n}}^N \leqslant C b_{\mathbf{n}}^{d(1 - N(1-\gamma)/\upsilon)} \to 0. \tag{4.4}$$

The convergence follows from the fact that $\upsilon > N(1-\gamma)$.

Now let $\delta = 2(1-\gamma)/\gamma$. Note $\gamma = 2/(2+\delta)$ and $\delta/(2+\delta) = 1-\gamma$. To prove $J_2 \to 0$, we apply Lemma 3.1 with $r = s = 2+\delta$ and $h = (2+\delta)/\delta$.

$$|\mathrm{Cov}(K_{r\mathbf{i}}, K_{s\mathbf{j}})| \leqslant C\{E K_{r\mathbf{i}}^{2+\delta}\}^{1/(2+\delta)} \{E K_{s\mathbf{j}}^{2+\delta}\}^{1/(2+\delta)} \{\hat{f}(1,1) \varphi(\hat{d}(\mathbf{i}, \mathbf{j}))\}^{1-\gamma}.$$

Note that, for any $x \in c(f)$,

$$\left\{ \int \left[\frac{1}{b_{\mathbf{n}}^d} K\left(\frac{x-u}{b_{\mathbf{n}}}\right) \right]^{2+\delta} f(u) \, du \right\}^{1/(2+\delta)}$$
$$\leqslant b_{\mathbf{n}}^{-d(1+\delta)/(2+\delta)} \left\{ \int \frac{1}{b_{\mathbf{n}}^d} \left[K\left(\frac{x-u}{b_{\mathbf{n}}}\right) \right]^{2+\delta} f(u) \, du \right\}^{1/(2+\delta)}$$
$$\leqslant C b_{\mathbf{n}}^{-d(1+\delta)/(2+\delta)}.$$

Hence,

$$\begin{aligned} J_2 &\leqslant C\hat{\mathbf{n}}^{-1} b_{\mathbf{n}}^{d} b_{\mathbf{n}}^{-2d(1+\delta)/(2+\delta)} \sum_{\mathbf{i},\mathbf{j}\in S_2} \varphi(\|\mathbf{i}-\mathbf{j}\|)^{1-\gamma} \\ &\leqslant C\hat{\mathbf{n}}^{-1} b_{\mathbf{n}}^{-d(1-\gamma)} \sum_{\mathbf{i},\mathbf{j}\in S_2} \varphi(\|\mathbf{i}-\mathbf{j}\|)^{1-\gamma} \\ &\leqslant C b_{\mathbf{n}}^{-d(1-\gamma)} \sum_{\|\mathbf{i}\|>c_{\mathbf{n}}} \varphi(\|\mathbf{i}\|)^{1-\gamma}. \end{aligned} \tag{4.5}$$

By Assumption 2.4, $i^{N-1}\varphi(i)^a = o(1/i)$ or $\varphi(i) = o(i^{-N/a})$ as $i \to \infty$. Since φ is a nonincreasing function, we have $\varphi(x) = o(x^{-N/a})$ as $x \to \infty$. Therefore,

$$\begin{aligned} &\|\mathbf{i}\|^{\upsilon}\varphi(\|\mathbf{i}\|)^{1-\gamma} \\ &\quad = \|\mathbf{i}\|^{-N-\varepsilon+(1-\gamma)Na^{-1}}\varphi(\|\mathbf{i}\|)^{1-\gamma} \\ &\quad = \|\mathbf{i}\|^{-N-\varepsilon}\|\mathbf{i}\|^{(1-\gamma)Na^{-1}}\varphi(\|\mathbf{i}\|)^{1-\gamma}. \end{aligned}$$

The last equality implies that

$$\|\mathbf{i}\|^{\upsilon}\varphi(\|\mathbf{i}\|)^{1-\gamma} = o(\|\mathbf{i}\|^{-N-\varepsilon}).$$

Turning to (4.5), we have

$$\begin{aligned} J_2 &\leqslant C b_{\mathbf{n}}^{-d(1-\gamma)} \sum_{\|\mathbf{i}\|>c_{\mathbf{n}}} \varphi(\|\mathbf{i}\|)^{1-\gamma} \\ &\leqslant C b_{\mathbf{n}}^{-d(1-\gamma)} c_{\mathbf{n}}^{-\upsilon} \sum_{\|\mathbf{i}\|>c_{\mathbf{n}}} \|\mathbf{i}\|^{\upsilon}\varphi(\|\mathbf{i}\|)^{1-\gamma} \\ &\leqslant C \sum_{\|\mathbf{i}\|>c_{\mathbf{n}}} \|\mathbf{i}\|^{-N-\varepsilon} \\ &\leqslant C \sum_{j=c_{\mathbf{n}}}^{\infty} \sum_{j\leqslant\|\mathbf{i}\|<j+1} \|\mathbf{i}\|^{-N-\varepsilon} \\ &\leqslant C \sum_{j=c_{\mathbf{n}}}^{\infty} j^{-N-\varepsilon} j^{N-1} \to 0 \quad \text{as } c_{\mathbf{n}} \to \infty. \end{aligned} \tag{4.6}$$

The convergence is from the following:

$$C\sum_{j=1}^{\infty} j^{-N-\varepsilon} j^{N-1} = C\sum_{j=1}^{\infty} j^{-1-\varepsilon} < \infty.$$

By (4.4) and (4.6), lemma is established. □

Lemma 4.3. *Suppose that Assumptions* 2.1–2.4 and 4.1 *hold and* $X_{\mathbf{n}}$ *satisfies* (1.1) *and* (1.3). *Then*

$$\hat{\mathbf{n}} b_{\mathbf{n}}^{d} \operatorname{Cov}(f_{\mathbf{n}}(x_r), f_{\mathbf{n}}(x_s)) \to 0.$$

Proof. By Lemmas 4.1 and 4.2,

$$\hat{\mathbf{n}} b_{\mathbf{n}}^d \operatorname{Cov}(f_{\mathbf{n}}(x_r), f_{\mathbf{n}}(x_s)) \\ = \hat{\mathbf{n}}^{-1} b_{\mathbf{n}}^d \sum_{\substack{i_k=1 \\ k=1\ldots N}}^{n_k} \operatorname{Cov}(K_{r\mathbf{i}}, K_{s\mathbf{i}}) + \hat{\mathbf{n}}^{-1} b_{\mathbf{n}}^d \sum_{\substack{i_k=1 \\ k=1\ldots N \\ i_k \neq j_k \text{ for some } k}}^{n_k} \sum_{\substack{j_k=1 \\ k=1\ldots N}}^{n_k} \operatorname{Cov}(K_{r\mathbf{i}}, K_{s\mathbf{j}}) \to 0. \qquad \square$$

5. Asymptotic normality at multiple points

In this section, we first want to establish that for any points $x_1, \ldots, x_t$ of $f(x)$, the vector $(f_{\mathbf{n}}(x_1), \ldots, f_{\mathbf{n}}(x_t))$ is asymptotically normal after proper normalization. By the Cramer-Wold device, it suffices to prove the asymptotic normality of $\psi_{\mathbf{n}} = \sum_{r=1}^t c_r \psi_{\mathbf{n}r}$ for arbitrary but fixed real constants $c_1, \ldots, c_t$, where

$$\psi_{\mathbf{n}r} = (\hat{\mathbf{n}} b_{\mathbf{n}}^d)^{1/2} (f_{\mathbf{n}}(x_r) - E f_{\mathbf{n}}(x_r)).$$

Note that,

$$\operatorname{Var}(\psi_{\mathbf{n}}) = \sum_{r=1}^t c_r^2 \operatorname{Var}(\psi_{\mathbf{n}r}) + 2 \sum_{1 \leqslant r < s \leqslant t} c_r c_s \operatorname{Cov}(\psi_{\mathbf{n}r}, \psi_{\mathbf{n}s}).$$

Following the proof of Lemmas 4.1 and 4.2, we have

$$\operatorname{Var}(\psi_{\mathbf{n}r}) = \hat{\mathbf{n}} b_{\mathbf{n}}^d \operatorname{Var}(f_{\mathbf{n}}(x_r)) \to f(x_r) \int K^2(u)\, du := \tau_r^2,$$

Also

$$\operatorname{Cov}(\psi_{\mathbf{n}r}, \psi_{\mathbf{n}s}) = \hat{\mathbf{n}} b_{\mathbf{n}}^d \operatorname{Cov}(f_{\mathbf{n}}(x_r), f_{\mathbf{n}}(x_s)) \to 0,$$

where the convergence follows from Lemma 4.3. Thus,

$$\operatorname{Var}(\psi_{\mathbf{n}}) \to \sum_{r=1}^t c_r^2 \tau_r^2.$$

Set

$$Z_{r\mathbf{i}} = b_{\mathbf{n}}^{d/2} (K_{r\mathbf{i}} - E K_{r\mathbf{i}}), \quad S_{\mathbf{n}r} = \sum_{\substack{i_k=1 \\ k=1\ldots N}}^{n_k} Z_{r\mathbf{i}},$$

$$S_{\mathbf{n}} = \sum_{r=1}^t c_r S_{\mathbf{n}r}, \quad S_{\mathbf{n}r} = \hat{\mathbf{n}}^{1/2} \psi_{\mathbf{n}r}, \quad S_{\mathbf{n}} = \hat{\mathbf{n}}^{1/2} \psi_{\mathbf{n}}.$$

Next we set $S_{\mathbf{n}r}$ into large and small blocks. Define

$$U(1, \mathbf{n}, \mathbf{x}_r, \mathbf{j}) := \sum_{\substack{i_k=j_k(p+q)+1\\ k=1,\dots,N}}^{j_k(p+q)+p} Z_{r\mathbf{i}},$$

$$U(2, \mathbf{n}, \mathbf{x}_r, \mathbf{j}) := \sum_{\substack{i_k=j_k(p+q)+1\\ k=1,\dots,N-1}}^{j_k(p+q)+p} \sum_{i_N=j_N(p+q)+p+1}^{(j_N+1)(p+q)} Z_{r\mathbf{i}},$$

$$U(3, \mathbf{n}, \mathbf{x}_r, \mathbf{j}) := \sum_{\substack{i_k=j_k(p+q)+1\\ k=1,\dots,N-2}}^{j_k(p+q)+p} \sum_{i_{N-1}=j_{N-1}(p+q)+p+1}^{(j_{N-1}+1)(p+q)} \sum_{i_N=j_N(p+q)+1}^{j_N(p+q)+p} Z_{r\mathbf{i}},$$

$$U(4, \mathbf{n}, \mathbf{x}_r, \mathbf{j}) := \sum_{\substack{i_k=j_k(p+q)+1\\ k=1,\dots,N-2}}^{j_k(p+q)+p} \sum_{i_{N-1}=j_{N-1}(p+q)+p+1}^{(j_{N-1}+1)(p+q)} \sum_{i_N=j_N(p+q)+p+1}^{(j_N+1)(p+q)} Z_{r\mathbf{i}},$$

and so on. Note that

$$U(2^N-1, \mathbf{n}, \mathbf{x}_r, \mathbf{j}) := \sum_{\substack{i_k=j_k(p+q)+p+1\\ k=1,\dots,N-1}}^{(j_k+1)(p+q)} \sum_{i_N=j_N(p+q)+1}^{j_N(p+q)+p} Z_{r\mathbf{i}}$$

and

$$U(2^N, \mathbf{n}, \mathbf{x}_r, \mathbf{j}) := \sum_{\substack{i_k=j_k(p+q)+p+1\\ k=1,\dots,N}}^{(j_k+1)(p+q)} Z_{r\mathbf{i}}.$$

For each integer $1 \leqslant i \leqslant 2^N$, denote

$$T(\mathbf{n}, \mathbf{x}_r, i) := \sum_{\substack{j_k=0\\ k=1,\dots,N}}^{r_k-1} U(i, \mathbf{n}, \mathbf{x}_r, \mathbf{j}).$$

Clearly $S_{\mathbf{n}r} = \sum_{i=1}^{2^N} T(\mathbf{n}, \mathbf{x}_r, i)$. Define

$$T(\mathbf{n}, \mathbf{x}, i) := \sum_{r=1}^{t} c_r T(\mathbf{n}, \mathbf{x}_r, i).$$

Thus, $S_{\mathbf{n}} = \sum_{i=1}^{2^N} T(\mathbf{n}, \mathbf{x}, i)$. By Eq. (3.9) in Tran [18], we have for $r = 1, \dots, t$ and $2 \leqslant i \leqslant 2^N$,

$$\widehat{\mathbf{n}}^{-1} ET^2(\mathbf{n}, \mathbf{x}_r, i) \longrightarrow 0\,. \tag{5.1}$$

By Minkowski inequality, it follows that

$$\begin{aligned} \|T(\mathbf{n}, \mathbf{x}, i)\|_2 &:= \| \sum_{r=1}^{t} c_r T(\mathbf{n}, \mathbf{x}_r, i)\|_2, \\ &\leqslant \sum_{r=1}^{t} |c_r| \|T(\mathbf{n}, \mathbf{x}_r, i)\|_2. \end{aligned} \tag{5.2}$$

Multiply $\widehat{\mathbf{n}}^{-1/2}$ on both sides of (5.2), we have

$$\sqrt{\widehat{\mathbf{n}}^{-1} ET^2(\mathbf{n}, \mathbf{x}, i)} \leqslant \sum_{r=1}^{t} |c_r| \sqrt{\widehat{\mathbf{n}}^{-1} ET^2(\mathbf{n}, \mathbf{x}_r, i)}. \tag{5.3}$$

By (5.1) and (5.3), for each $2 \leqslant i \leqslant 2^N$,

$$\widehat{\mathbf{n}}^{-1} ET^2(\mathbf{n}, \mathbf{x}, i) \longrightarrow 0. \tag{5.4}$$

Define

$$Q_2 \equiv \sum_{i=2}^{2^N} T(\mathbf{n}, \mathbf{x}, i).$$

Then by the same argument, we have

$$\widehat{\mathbf{n}}^{-1} E Q_2^2 \to 0.$$

Now,

$$T(\mathbf{n}, \mathbf{x}, 1) = S_{\mathbf{n}} - Q_2 = \sum_{r=1}^{t} \sum_{\substack{j_k=0 \\ k=1,\dots,N}}^{r_k-1} c_r U(1, \mathbf{n}, \mathbf{x}_r, \mathbf{j}).$$

Let $U(1, \mathbf{n}, r, \mathbf{j})$, with $r = 1, \dots, t$ and $j_k = 0, \dots, r_k - 1$, where $k = 1, \dots, N$, be independent random variables having the same distribution as $c_r \widehat{\mathbf{n}}^{-1/2} U(1, \mathbf{n}, \mathbf{x}_r, \mathbf{j})$. Set $V(1, \mathbf{n}, r, \mathbf{j}) = U(1, \mathbf{n}, r, \mathbf{j})/t_{\mathbf{n}}$ where

$$\begin{aligned} t_{\mathbf{n}}^2 &:= \sum_{r=1}^{t} \sum_{\substack{j_k=0 \\ k=1,\dots,N}}^{r_k-1} \mathrm{Var}(U(1, \mathbf{n}, r, \mathbf{j})) \\ &= \sum_{r=1}^{t} c_r^2 \sum_{\substack{j_k=0 \\ k=1,\dots,N}}^{r_k-1} \mathrm{Var}(\widehat{\mathbf{n}}^{-1/2} U(1, \mathbf{n}, \mathbf{x}_r, \mathbf{j})) \\ &\to \sum_{r=1}^{t} c_r^2 \tau_r^2. \end{aligned}$$

Thus

$$EV(1, \mathbf{n}, r, \mathbf{j}) = 0 \quad \text{and} \quad \sum_{r=1}^{t} \sum_{\substack{j_k=0 \\ k=1,\dots,N}}^{r_k-1} \operatorname{Var}(V(1, \mathbf{n}, r, \mathbf{j})) = 1.$$

By Lindeberg condition,

$$\sum_{r=1}^{t} \sum_{\substack{j_k=0 \\ k=1,\dots,N}}^{r_k-1} V(1, \mathbf{n}, r, \mathbf{j}) \;\to\; N(0, 1) \tag{5.5}$$

if and only if $g(\varepsilon) \to 0$ where

$$g(\varepsilon) := \sum_{r=1}^{t} \sum_{\substack{j_k=0 \\ k=1,\dots,N}}^{r_k-1} EV^2(1, \mathbf{n}, r, \mathbf{j}) I(|V(1, \mathbf{n}, r, \mathbf{j})| > \varepsilon).$$

Observe that

$$|U(1, \mathbf{n}, \mathbf{x}_r, \mathbf{j})| \leqslant \sum_{\substack{i_k=j_k(p+q)+1 \\ k=1,\dots,N}}^{j_k(p+q)+p} |Z_{r\mathbf{i}}|, \quad |Z_{r\mathbf{i}}| = b_{\mathbf{n}}^{d/2} |K_{r\mathbf{i}} - EK_{r\mathbf{i}}| \leqslant \frac{C}{b_{\mathbf{n}}^{d/2}}. \tag{5.6}$$

By (5.6), we have

$$|U(1, \mathbf{n}, \mathbf{x}_r, \mathbf{j})| \leqslant \frac{Cp^N}{b_{\mathbf{n}}^{d/2}}.$$

Hence,

$$\begin{aligned} g(\varepsilon) &\leqslant \frac{Cp^{2N}}{\varepsilon^2 \widehat{\mathbf{n}} b_{\mathbf{n}}^d t_{\mathbf{n}}^2} \sum_{r=1}^{t} \sum_{\substack{j_k=0 \\ k=1,\dots,N}}^{r_k-1} \operatorname{Var}(V(1, \mathbf{n}, r, \mathbf{j})) \\ &= \frac{Cp^{2N}}{\varepsilon^2 \widehat{\mathbf{n}} b_{\mathbf{n}}^d t_{\mathbf{n}}^2} \to 0, \end{aligned}$$

where the convergence follows from the choice of p and the fact $s_{\mathbf{n}} \to \infty$ in Assumption 2.5.

The convergence of (5.5) implies that

$$\widehat{\mathbf{n}}^{-1/2} T(\mathbf{n}, \mathbf{x}, 1) = \sum_{r=1}^{t} \sum_{\substack{j_k=0 \\ k=1,\dots,N}}^{r_k-1} c_r \widehat{\mathbf{n}}^{-1/2} U(1, \mathbf{n}, \mathbf{x}_r, \mathbf{j}) \to N\left(0, \sum_{r=1}^{t} c_r^2 \tau_r^2\right)$$

provided the following is true:

$$\begin{aligned}
&\left| E\exp\left\{ it\sum_{r=1}^{t}\sum_{\substack{j_k=0\\k=1,\dots,N}}^{r_k-1} c_r\widehat{\mathbf{n}}^{-1/2}U(1,\mathbf{n},\mathbf{x}_r,\mathbf{j})\right\} - E\exp\left\{ it\sum_{r=1}^{t}\sum_{\substack{j_k=0\\k=1,\dots,N}}^{r_k-1} U(1,\mathbf{n},r,\mathbf{j})\right\}\right| \\
&= \left| E\prod_{\substack{j_k=0\\k=1,\dots,N}}^{r_k-1}\exp\left\{ it\sum_{r=1}^{t} c_r\widehat{\mathbf{n}}^{-1/2}U(1,\mathbf{n},\mathbf{x}_r,\mathbf{j})\right\}\right. \\
&\quad\left. - \prod_{\substack{j_k=0\\k=1,\dots,N}}^{r_k-1} E\exp\left\{ it\sum_{r=1}^{t} c_r\widehat{\mathbf{n}}^{-1/2}U(1,\mathbf{n},\mathbf{x}_r,\mathbf{j})\right\}\right| \\
&\to 0.
\end{aligned} \tag{5.7}$$

Let

$$I(1,\mathbf{n},\mathbf{j}) := \{\mathbf{i} : j_k(p+q)+1 \leqslant i_k \leqslant j_k(p+q)+p,\ k=1,\dots,N\}.$$

Note that $\sum_{r=1}^{t} c_r\widehat{\mathbf{n}}^{-1/2}U(1,\mathbf{n},\mathbf{x}_r,\mathbf{j})$ is measurable with respect to $\{X_{\mathbf{i}}, \mathbf{i} \in I(1,\mathbf{n},\mathbf{j})\}$. The distance between two distinct sets $I(1,\mathbf{n},\mathbf{j})$ and $I(1,\mathbf{n},\mathbf{j}')$ is at least q and $I(1,\mathbf{n},\mathbf{j})$ contains p^N sites. Following the proof of Lemma 5.3 and (5.18) in Hallin et al. [7], we obtain that under (1.2), (5.7) is bounded by $\hat{\mathbf{n}}\varphi(q)$, which is convergent to 0 by (ii) in Assumption (2.5).

Define

$$\Sigma = \mathrm{diag}(\tau_1^2,\dots,\tau_t^2).$$

By Lemmas 2.1 and 4.3, the above result in the section is summarized in the following.

Theorem 5.1. *Suppose Assumptions* 2.1–2.5 (i)–(iii) *and Assumption* 4.1 *hold and $X_{\mathbf{n}}$ satisfies* (1.1) *and* (1.2). *Then, for distinct points $x_1, x_2, \dots, x_t$,*

$$(\hat{\mathbf{n}}b_{\mathbf{n}}^d)^{1/2}(f_{\mathbf{n}}(x_1)-Ef_{\mathbf{n}}(x_1),\dots,f_{\mathbf{n}}(x_t)-Ef_{\mathbf{n}}(x_t))' \to N(0,\Sigma) \quad \text{as } \mathbf{n}\to\infty.$$

Theorem 5.2. *Suppose Assumptions* 2.1–2.5 (i)–(iv) *and Assumption* 4.1 *hold and $X_{\mathbf{n}}$ satisfies* (1.1) *and* (1.2). *Then for distinct points $x_1, x_2, \dots, x_t$,*

$$(\hat{\mathbf{n}}b_{\mathbf{n}}^d)^{1/2}(f_{\mathbf{n}}(x_1)-f(x_1),\dots,f_{\mathbf{n}}(x_t)-f(x_t))' \to N(0,\Sigma) \quad \text{as } \mathbf{n}\to\infty.$$

Proof. Let

$$\xi_{\mathbf{n}r} = (\hat{\mathbf{n}}b_{\mathbf{n}}^d)^{1/2}(f_{\mathbf{n}}(x_r)-f(x_r)),$$

$$\xi_{\mathbf{n}} = \sum_{r=1}^{t} c_r\xi_{\mathbf{n}r}.$$

Employing Assumptions 2.1, (i) of Assumption 2.2 and (iv) of Assumption 2.5, it follows that

$$\psi_{\mathbf{n}} - \xi_{\mathbf{n}} \to 0. \tag{5.8}$$

By Theorem 5.1 and (5.8), the Cramer–Wold device yields the desired result. □

Define

$$\Gamma = \operatorname{diag}\left(\frac{\tau_1^2}{\bar{F}^2(x_1)}, \ldots, \frac{\tau_t^2}{\bar{F}^2(x_t)}\right).$$

Theorem 5.3. *Suppose Assumptions* 2.1–2.5 (i)–(iv) *and Assumption* 4.1 *hold. In addition, suppose* (1.1) *and* (1.2) *hold. Then for distinct points* $x_1, x_2, \ldots, x_t$,

$$(\hat{\mathbf{n}} b_{\mathbf{n}}^d)^{1/2}(r_{\mathbf{n}}(x_1) - r(x_1), \ldots, r_{\mathbf{n}}(x_t) - r(x_t))' \to N(0, \Gamma) \quad \textit{as } \mathbf{n} \to \infty.$$

Proof. Define

$$Q_{\mathbf{n}} = \sum_{r=1}^{t} c_r (\hat{\mathbf{n}} b_{\mathbf{n}}^d)^{1/2}(r_{\mathbf{n}}(x_r) - r(x_r)) = U_{\mathbf{n}} - V_{\mathbf{n}},$$

where

$$U_{\mathbf{n}} = \sum_{r=1}^{t} c_r \frac{1}{\bar{F}_{\mathbf{n}}(x_r)} (\hat{\mathbf{n}} b_{\mathbf{n}}^d)^{1/2}[f_{\mathbf{n}}(x_r) - f(x_r)],$$
$$V_{\mathbf{n}} = \sum_{r=1}^{t} c_r \frac{f(x_r)}{\bar{F}_{\mathbf{n}}(x_r)\bar{F}(x_r)} (\hat{\mathbf{n}} b_{\mathbf{n}}^d)^{1/2}[\bar{F}_{\mathbf{n}}(x_r) - \bar{F}(x_r)]. \tag{5.9}$$

By Lemma 3.2 and continuous mapping theorem,

$$V_{\mathbf{n}} \to 0$$

in probability as $\mathbf{n} \to \infty$. Define,

$$U_{\mathbf{n}1} = \left(\frac{c_1}{\bar{F}_{\mathbf{n}}(x_1)}, \ldots, \frac{c_t}{\bar{F}_{\mathbf{n}}(x_t)}\right),$$
$$U_{\mathbf{n}2} = (\hat{\mathbf{n}} b_{\mathbf{n}}^d)^{1/2}([f_{\mathbf{n}}(x_1) - f(x_1)], \ldots, [f_{\mathbf{n}}(x_t) - f(x_t)]),$$
$$U_{\mathbf{n}} = U_{\mathbf{n}1} \cdot U_{\mathbf{n}2}'.$$

By Theorem 5.2 and Lemma 3.2, it follows that $U_{\mathbf{n}}$ is asymptotically normal with expectation 0 and variance $\sum_{r=1}^{t} c_r^2 \tau_r^2 / \bar{F}^2(x_r)$. Hence, it follows that $Q_{\mathbf{n}}$ is asymptotically normal with expectation 0 and variance $\sum_{r=1}^{t} c_r^2 \tau_r^2 / \bar{F}^2(x_r)$. Therefore, Theorem 5.3 is established by an application of Cramer–Wold device. □

Similarly, we have the following.

Theorem 5.4. *Suppose Assumptions* 2.1–2.5 (i), (Aii), (iii), (iv) *and Assumption* 4.1 *hold. In addition, suppose* (1.1) *and* (1.3) *hold. Then for distinct points* $x_1, x_2, \ldots, x_t$,

$$(\hat{\mathbf{n}} b_{\mathbf{n}}^d)^{1/2}(r_{\mathbf{n}}(x_1) - r(x_1), \ldots, r_{\mathbf{n}}(x_t) - r(x_t))' \to N(0, \Gamma) \quad \textit{as } \mathbf{n} \to \infty.$$

6. Applications

Consider the case $N = 2$ and $d = 1$. Assume that $\{Y_{\mathbf{i}}, i_1 \geqslant 1, i_2 \geqslant 1\}$ is a spatial process fitting the model:

$$Y_{i_1,i_2} = \rho Y_{i_1-1,i_2} + \tau Y_{i_1,i_2-1} + Z_{i_1,i_2}, \tag{6.1}$$

with $Z_{\mathbf{i}}$'s being independent standard normal random variables and $|\rho| + |\tau| < 1$. Then model (6.1) can be written as

$$Y_{i_1,i_2} = \sum_{j_1=0}^{\infty} \sum_{j_2=0}^{\infty} \binom{j_1 + j_2}{j_1} \rho^{j_1} \tau^{j_2} Z_{i_1-j_1,i_2-j_2}.$$

It is not hard to show that $\{Y_{\mathbf{i}}, i_1 \geqslant 1, i_2 \geqslant 1\}$ is a strictly stationary spatial process satisfying the mixing conditions (1.1) and (1.3). We choose $\rho = \tau = 0.25$. Suppose $X_{i_1,i_2} = F_1^{-1}(F_2(Y_{i_1,i_2}))$ (representing lifelengths), where F_1 is an exponential distribution function with mean 1 and F_2 is a standard normal distribution function. Then $X_{\mathbf{i}}$ is also a strictly stationary spatial process satisfying the mixing conditions (1.1) and (1.3). In particular, all $X_{\mathbf{i}}$'s are exponential random variables with mean 1. We generate observations of the random field X_{i_1,i_2} at 180 sites on a lattice of size 9 by 20 (Table 1). We choose $b_{\mathbf{n}} = \frac{1}{10}$ and $x = 0.9$. Take K as the standard normal density function. Simple computation shows that $f_{\mathbf{n}}(0.9)$ is 0.445 and $\bar{F}_{\mathbf{n}}(0.9)$ is 0.433. Our estimate for the hazard rate function at $x = 0.9$ is 1.028, which is fairly close to the actual value. Note that if the distribution of lifelengths is exponential with mean 1, then the hazard rate is 1. Repeat the above procedure 20 times, we can obtain the normal probability plot of the estimated hazard rate at 0.9 with $b_{\mathbf{n}} = 0.1$ (Fig. 1). We also provide a reasonable scatter plot of the estimated hazard rate and

Table 1
Observations of lifelengths

$X_{\mathbf{i}}$	1	2	3	4	5	6	7	8	9
1	3.85	1.41	.05	1.55	0.66	0.36	0.29	0.15	2
2	2.33	1.34	3.44	1.22	1.11	1.78	1	0.16	1.18
3	0.43	0.01	0.99	0.08	6.01	0.91	0.81	1.03	0.09
4	1.95	0.16	0.48	1.23	0.05	0.84	1.08	0.17	0.85
5	0.52	1.07	0.13	0.29	0.08	0.21	0.1	0.08	1.04
6	0.54	0.79	0.95	1.93	0.46	0.01	2.82	0.14	0.61
7	0.04	0.22	0.99	0.1	0.2	1.75	3.39	1.05	3.11
8	0.69	0.22	0.3	2.33	0.61	0.29	0.04	0.72	0.17
9	3.6	0.43	0.16	0.88	0.96	1.03	1.39	1.85	0.001
10	0.59	1.12	0.5	0.96	1.08	1.96	0.4	0.54	0.13
11	2.57	0.01	1.36	0.07	0.06	1.98	1.07	0.08	0.14
12	1.49	1.04	3.84	0.16	0.37	0.07	4.22	0.48	1.16
13	0.77	1.65	0.98	0.24	0.41	1.19	0.2	0.32	1.86
14	3.21	0.01	1.66	0.82	0.04	0.01	0.12	0.21	0.64
15	0.14	1.05	3.33	0.92	0.64	1.08	1.21	0.12	0.33
16	1.56	1.57	0.69	0.04	0.77	2.94	0.23	0.69	1.13
17	3	1.04	0.45	5.4	0.04	0.35	0.02	0.52	0.3
18	0.61	1.61	0.74	1.41	0.19	0.28	0.73	0.46	1.42
19	0.92	2.41	0.32	0.1	1.13	0.14	0.1	1.35	0.28
20	0.08	4.35	0.99	0.26	0.24	2.53	0.51	0.69	1.03

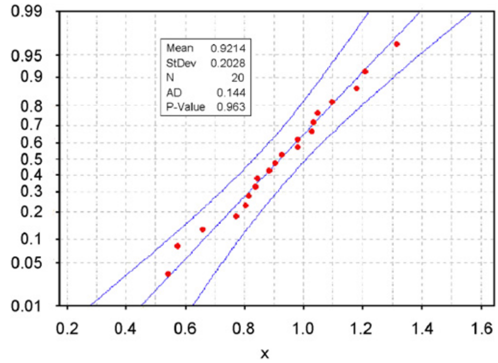

Fig. 1. Normal probability plot of hazard rate estimate at 0.9 with $b_{\mathbf{n}} = 0.1$ with 95% confidence interval.

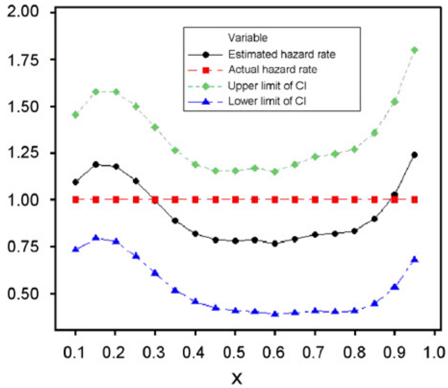

Fig. 2. Scatter plot with $b_{\mathbf{n}} = 0.1$.

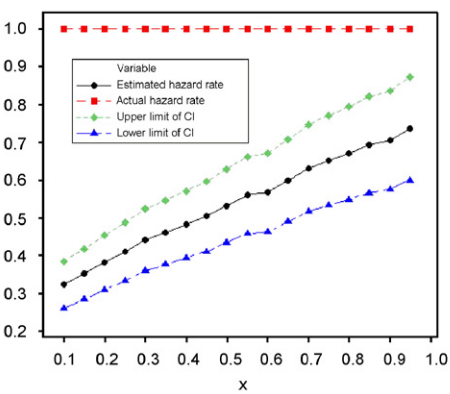

Fig. 3. Scatter plot with $b_{\mathbf{n}} = 1$.

the actual hazard rate at 18 points in interval (0, 1) with appropriate $b_{\mathbf{n}} = 0.1$ (Fig. 2). We ask that $b_{\mathbf{n}}$ goes to zero not too slow or too fast in the theorem. The scatter plot of the estimated hazard rate and the actual hazard rate at 18 points in interval (0, 1) with inappropriate $b_{\mathbf{n}}$ (Figs.

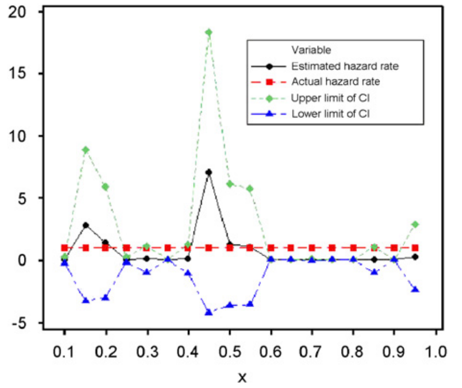

Fig. 4. Scatter plot with $b_{\mathbf{n}} = 0.001$.

3, 4) shows the estimation is unreasonable when $b_{\mathbf{n}}$ goes to zero too slow ($b_{\mathbf{n}} = 1$) (Fig. 3) or too fast ($b_{\mathbf{n}} = 0.001$) (Fig. 4). The pointwise 99% confidence intervals for the hazard rate at those 18 points are also included in the scatter plot (Figs. 2–4).

Acknowledgments

The authors are grateful to the referees for useful comments that have significantly improved the presentation of the paper.

References

[1] I. Ahmad, Uniform strong convergence of a generalized failure rate estimate, Bull. Math. Statist. 17 (1976) 77–84.
[2] A. Antoniadis, G. Grégoire, G. Nason, Density and hazard rate estimation for right-censored data by using wavelet methods, J. Roy. Statist. Soc. B 61 (1999) 63–84.
[3] A.P. Basu, Bivariate failure rate, J. Amer. Statist. Assoc. 66 (1971) 103–104.
[4] R.C. Bradley, Introduction to Strong Mixing Conditions, 2005.
[5] C.M. Deo, A note on empirical processes of strong mixing sequences, Ann. Probab. 1 (1973) 870–875.
[6] X. Guyon, Estimation d'un champ par pseudo-vraisemblale conditionelle: étude asymptotique et application au cas Markovien, in: J.-J. Droesbeke, et al. (Eds.), Spatial Processes and Spatial Time Series Analysis, Proceedings of the Sixth Franco-Begian Meeting of Statisticians, FUSL, Brussels, 1987.
[7] M. Hallin, Z. Lu, L.T. Tran, Local linear spatial regression, Ann. Statist. 32 (2004) 2469–2500.
[8] M. Hollander, F. Proschan, Nonparametric concepts and methods in reliability, in: P.R. Krishnaiah, P.K. Sen (Eds.), Handbook of Statistics, vol. 4, North-Holland, Amsterdam, 1984, pp. 613–655.
[9] I.A. Ibragimov, Y.V. Linnik, Independent and Stationary Sequence of Random Variables, Wolters-Noordhoff Publishing, Groningen, 1971, 443pp.
[10] N. Landberg, F. Proschan, A.J. Quinzi, Estimating dependent life lengths, with applications to the theory of competing risks, Ann. Statist. 9 (1981) 157–167.
[13] P.S. Puri, H. Rubin, On a characterization of the family of distributions with constant multivariate failure rates, Ann. Probab. 2 (1974) 738–740.
[14] J. Rice, M. Rosenblatt, Estimation of the log survivor function and hazard function, Sankhya A 38 (1976) 60–78.
[16] G.G. Roussas, Hazard rate estimation under dependence conditions, J. Statist. Plann. Inference 22 (1989) 81–93.
[18] L.T. Tran, Kernel density estimation on random fields, J. Multivariate Anal. 34 (1990) 37–53.
[19] G.S. Watson, M.R. Leadbetter, Hazard analysis, I, Biometrika 51 (1964) 175–184.
[20] G.S. Watson, M.R. Leadbetter, Hazard analysis, II, Sankhya A 26 (1964) 110–116.