Far East Journal of Theoretical Statistics
Volume 30, Number 2, 2010, Pages 89-105
Published Online: March 11, 2010
This paper is available online at http://www.pphmj.com

RECURSIVE KERNEL HAZARD ESTIMATION OF STRONG MIXING DATA

A. QUINTELA-DEL-RÍO

Departamento de Matemáticas
Facultad de Informática
Universidad de A Coruña
A Coruña, Spain
e-mail: aquintela@ucd.es

Abstract

We consider the nonparametric estimation of the hazard function, using a recursive kernel estimator of the density and the distribution function. Assuming that the data proceed from a strong mixing stationary process, the strong consistency of the proposed estimator is obtained. The rate of convergence is the same as that in the independence case. Asymptotic normality of the estimator is also proven.

1. Introduction

Let X be a real random variable, with probability density function f and distribution function F, with respect to some σ-finite measure on $\mathbb{R}$. The failure rate or hazard function is defined by

$$r(x) = \lim_{\Delta_x \to 0} \frac{P(x \leq X < x + \Delta_x / X \geq x)}{\Delta_x}.$$

By the definition of conditional probability, we have that

$$r(x) = \frac{f(x)}{1 - F(x)} = \frac{f(x)}{\overline{F}(x)},$$

2010 Mathematics Subject Classification: 62G05.

Keywords and phrases: kernel, recursive estimation, hazard, strong mixing processes, consistency, asymptotic normality.

Received January 8, 2007

considered when $\overline{F}(x) > 0$, that is, $r(\cdot)$ is defined in the set $\{x \in \mathbb{R}/\overline{F}(x) > 0\}$.

If we consider that the random variable X measures the "failure time", then $r(x)\Delta_x$ can be interpreted as the approximate probability that one subject "fails" in the time interval $[x, x + \Delta_x)$, given the subject has survived time x, i.e., the instantaneous probability of failure at x, given survival to x.

The estimation of the hazard function is a problem of considerable interest in many applied fields, such as inventory theory, reliability, medicine and seismology. In this last case, the hazard function might be thought as the instantaneous risk of the occurrence of an earthquake at time x, knowing that the last earthquake has happened at time zero (Rice and Rosenblatt [15]). Such approximation has been considered and applied to real data sets in recent papers (Estévez et al. [4, 6]).

A practical question is the estimation of $r(\cdot)$ based on a random sample $X_1, \ldots, X_n$. In the case of i.i.d. random variables we can cite the works of, for example, Watson and Leadbetter [24, 25], Ahmad [1], Hollander and Proschan [9], Prakasa Rao and Van Ryzin [13] or Hassani et al. [8]. In a context of dependent data (usually, strong mixing processes), we highlight the papers of Estévez and Quintela [5], Izenman and Tran [11], Roussas [17, 18] and Sarda and Vieu [20].

Roussas [17, 18] considers a nonparametric estimator defined by

$$r_h(x) = \frac{f_h(x)}{1 - \hat{F}(x)}, \tag{1}$$

with $f_h(x) = \frac{1}{nh}\sum_{i=1}^{n} K\left(\frac{x - X_i}{h}\right)$ the classical Parzen-Rosenblatt kernel estimator of the density, and $F(\cdot)$ the empirical distribution function. Estévez and Quintela [5] use a kernel estimator of the distribution function in the denominator, that is, $F_h(x) = \int_{-\infty}^{x} f_h(t)dt$. Roussas and Tran [19] compare (1) with a recursive version, using

$$r_n(x) = \frac{f_n(x)}{1 - \hat{F}(x)}, \tag{2}$$

where

$$f_n(x) = \frac{1}{n} \sum_{i=1}^{n} \frac{1}{b_i} K\left(\frac{x - X_i}{b_i}\right) \tag{3}$$

is a recursive estimator of the density function, with $\{b_n\} \in \mathbb{R}^+$ a sequence of smoothing parameters. Asymptotic optimality properties for this estimator under mixing conditions can be seen in Masry [12], Tran [21] or Roussas and Tran [19].

In our paper, we consider the estimator

$$r_n(x) = \frac{f_n(x)}{1 - F_n(x)} = \frac{f_n(x)}{\overline{F}_n(x)}, \tag{4}$$

where $f_n(\cdot)$ is the same as in (2), and $F_n(\cdot)$ is the kernel estimator of $F(\cdot)$, defined by

$$F_n(x) = \frac{1}{n} \sum_{i=1}^{n} H\left(\frac{x - X_i}{b_i}\right) = \int_{-\infty}^{x} f_n(t)dt, \tag{5}$$

with $H(x) = \int_{-\infty}^{x} K(u)du.$

The hazard estimator can be computed recursively

$$r_n(x) = \frac{\frac{n-1}{n} f_{n-1}(x) + \frac{1}{nb_n} K\left(\frac{x - X_n}{b_n}\right)}{1 - \left[\frac{n-1}{n} F_{n-1}(x) + \frac{1}{n} H\left(\frac{x - X_n}{b_n}\right)\right]}.$$

In this way, the hazard estimator can be updated with each new observation X_{n+1}. This iterative scheme saves computer time in a practical case, whereas a nonrecursive estimator needs to be recalculated completely when a new data set is observed. This is the case of time series data (see Vilar and Vilar [22] for more discussion and references on this point). Moreover, in the case of the estimator (4), the use of a kernel distribution estimator instead of the classical empirical distribution gives it continuity properties. The continuity can reveal important features of the hazard in seismology studies (see, e.g., Estévez et al. [6]). Perhaps it is in this setting where the recursivity turns out to be an important tool: when a seismic series is taking place, we can obtain the instantaneous risk of a new earthquake more quickly than when using a nonrecursive estimate (Grazinoli and Lorenzo [7]).

In this paper, we establish uniform consistency properties and asymptotic normality for the estimator (5) in a dependence context. As a by-product, we obtain the rate of convergence and the asymptotic normality for the kernel hazard estimator. These rates are the same as those in the independent case. The dependence setting and the assumptions used are presented in Section 2. Consistency properties and the asymptotic normality of the proposed estimators can be seen in Section 3. Finally, all the proofs are reported in the Appendix.

2. Dependence Structure and Asymptotic Optimality

This work assumes the sample data to be dependent and to satisfy the "strong mixing" condition (α-mixing), introduced by Rosenblatt [16], which is defined as:

Definition 1. Let $\mathbb{N}^*$ denote the set of positive integers, and for any i and j in $\mathbb{N}^* \cup \{\infty\} (i \leq j)$ define $\mathcal{F}_i^j$ to be the σ-algebra spanned by the variables $X_i, \ldots, X_j$. The sequence $\{X_i\}$ is said to be α-*mixing* (*strong mixing*) if there exist mixing coefficients $\alpha(m)$ such that $| P(A \cap B) - P(A)P(B) | \leq \alpha(m)$, for any sets A, B that are, respectively, $\mathcal{F}_1^k$-measurable and $\mathcal{F}_{k+m}^{\infty}$-measurable ($k$, m positive integer), and $\alpha(m) \downarrow 0$.

This condition is one of the weakest used in studies of dependent samples. It is satisfied, for example, by the ARMA processes with continuous white noise (see Doukhan [3] for a more complete discussion of the strong mixing conditions).

In the results we will present below, the following assumptions will be considered:

About the random variables and the mixing coefficients

(H.1) $(X_n)_n$ is a strictly stationary sequence of α-mixing variables with $\sum_{r=1}^{\infty} \alpha(r) < +\infty$.

About the density function

(H.2) The density f is bounded and has $k(\geq 2)$ continuous derivatives.

(H.3) For all $j \geq 2$, there exists the joint density function $f_{1, j}(\cdot, \cdot)$ and

$$| f_{1, j}(u, v) - f(u)f(v) | \leq C, \quad \forall u, v \in \mathbb{R}.$$

About the kernel function K

(H.4) K is a bounded probability density function, verifying $\lim_{| x | \to 0} | x | K(x) = 0.$

(H.5) K is Lipschitz continuous (i.e., $\exists C_K$, $0 < C_K < \infty$, such that $| K(x) - K(y) | \leq C_K | x - y |, \forall x, y \in \mathbb{R}$).

(H.6) K has order k $\left(0 = \int xK(x)dx = \int x^2 K(x)dx = \cdots = \int x^{k-1} K(x)dx\right.$ and $\left. 0 < \int x^k K(x)dx < \infty\right).$

About the bandwidth sequence

(H.7) $\{b_n\}$ is a monotone nonincreasing sequence verifying $b_n \to 0$ and $nb_n^k \to \infty$ as $n \to \infty$.

(H.8) $\frac{1}{n} \sum_{j=1}^{n} \left(\frac{b_j}{b_n}\right)^l \to \beta_l < \infty,\ l = 1, \ldots, k + 1.$

3. Results

3.1. Uniform strong consistency

First, we present two lemmas that show the asymptotic behavior of the kernel estimators of the density and distribution functions. In the sequel, $C > 0$ will denote a positive constant, and all limits are taken as $\{n\}$, or subsequences thereof, tend to infinity. Also, we will consider S as any compact subset contained in the interior of $\{x \in \mathbb{R}/\overline{F}(x) > 0\}$.

Lemma 1 (Tran [21]). *Under* (H.1)-(H.8) *and assuming that the mixing coefficients have the form* $\alpha(n) = O(n^{-a})$ *with* $a > \frac{11}{2}$, *we have*

$$\sup_{x \in S} |(f_n(x) - f(x))| = O\left(\left(\frac{\log n}{nb_n}\right)^{1/2}\right), \quad a.s.$$

Lemma 2. *Under* (H.1)-(H.8) *and assuming that*

(H.9) *the mixing coefficients have the form* $\alpha(n) = O(n^{-a})$ *with* $a + 1 > \frac{11}{2}$,

(H.10) *F is k-times continuously differentiable*, *and*

(H.11) *the bandwidth sequence verifies* $\sum_{n=1}^{\infty} nb_n^{k(a+1)-1} < \infty$, *we have*

$$\sup_{x \in S} |(F_n(x) - F(x))| = O(b_n^k), \quad a.s.$$

Now, to study the asymptotic properties of $r_n(\cdot)$, we write, for each $x \in S$,

$$r_n(x) - r(x) = \frac{1}{D_n(x)}[(1 - F(x)(f_n(x) - f(x)) - f(x)(F(x) - F_n(x)))], \quad (6)$$

with $D_n(x) = F(x)\overline{F}_n(x) = (1 - F(x))(1 - F_n(x))$. It is easy to obtain that $|D_n(x)| > C > 0, \forall x \in S$. Considering the results of Lemmas 1 and 2, and using (6), we can write the following theorem for the estimator of the hazard function:

Theorem 3. *Under the assumptions of Lemma* 2, *we have*

$$\sup_{x \in S} |r_n(x) - r(x)| = O(b_n^k), \quad a.s.$$

Remark 1. If we suppress, in Theorem 3 (resp. Lemma 2), the assumptions related to the kth order of the kernel function K and the existence of k derivatives of the function f, we simply obtain the convergence toward zero of the bias of the estimator $r_n(\cdot)$ and the strong consistency (pointwise and uniform) of the same.

Remark 2. Note that the rate of convergence obtained in Theorem 3 (resp. Lemma 2) is the same as that in the case of independent samples. This rate is achieved by means of the reconstruction techniques used by Ango-Nzé and Doukhan [2] (see Quintela and Vieu [14]).

3.2. Asymptotic normality

To establish the asymptotic normality of the estimator $r_n(\cdot)$ we need to include the following assumptions:

(H.12) There exists a subsequence of positive integers $\{m_n\}$, $1 \le m_n \le n$, such that

$$m_n b_n \to 0 \quad \text{and} \quad \frac{1}{b_n^{1-\gamma}} \sum_{j=m_n}^{\infty} \alpha^{1-\gamma}(j) \to 0 \quad \text{for some} \quad \gamma \in (0, 1).$$

(H.13) The sequence $n^{-1}\left(\sum_{k=1}^{n}\left(\frac{b_k}{b_n}\right)^2\right)$ has a finite limit.

(H.14) Let $\rho = \rho(n) \in \mathbb{N}$ and let $\beta = \beta(n) \in \mathbb{N}$ subsequences of $n \in \mathbb{N}$ such that $\rho, \beta \to \infty$ and $\rho + \beta \le n$. Let $\mu = \mu(n)$, the largest positive integer for which $\mu(\rho + \beta) \le n$. Then, it verifies that $\beta/\rho \to 0$, $\mu\alpha(\beta) \to 0$ and $\frac{\rho^2}{nb_n} \to 0$.

Masry [12] proves the asymptotic normality for the estimator (3) in the case of strong mixing and asymptotically uncorrelated processes (for the definition of processes of this type see (3.2) of Masry [12]). Roussas and Tran [19] show a similar result under (H.1.)-(H.8) and (H.12)-(H.14). Because expression (6) and Lemma 4 below, we can establish the asymptotic normality for the recursive hazard estimator.

Lemma 4. *Under* (H.1)-(H.4) *and* (H.7), *we have*

$$nVar(F_n(x)) \to F(x)\overline{F}(x).$$

Theorem 5. *Under* (H.1)-(H.10) *and* (H.12)-(H.14), *we have*

$$(nb_n)^{1/2}(r_n(x) - r(x)) \xrightarrow{d} N(0, \tau(x)),$$

with $\tau^2(x) = r(x)\dfrac{\int K^2(x)dx}{1 - F(x)}$, *for each point of continuity* $x \in S$.

Remark 3. Except (H.5), the assumptions (H.1)-(H.8) and (H.12)-(H.14) are used by Roussas and Tran [19] to obtain the asymptotic normality of the recursive kernel estimator of the density function *f*. Because the proof of Theorem 5 makes use of equation (6), we will use the result of Roussas and Tran [19] and therefore, we need to use their assumptions. Also, because (6), we need to get the asymptotic normality of the kernel estimator of the distribution function.

Remark 4. Because our assumptions are the same as those used in Roussas and Tran [19], their comments on these remain valid (see pages 351-354 of the mentioned paper). In this work, we find specific examples of the selection of possible values in (H.14).

Remark 5. The popular choice for the bandwidth parameter $b_n = Cn^{-s}$, $0 < s < 1$ is contemplated in our case, because all the conditions about the bandwidth sequence are fulfilled.

4. Appendix: Proofs

Proof of Lemma 2. Let us calculate the expectation

$$E(F_n(x)) = \frac{1}{n}\sum_{i=1}^{n} E\left[H\left(\frac{x - X_i}{b_i}\right)\right] = \frac{1}{n}\sum_{i=1}^{n}\int_{-\infty}^{+\infty} F(x - b_i u)K(u)du, \qquad (7)$$

using integration by parts. Now, a Taylor's expansion of order k of $F(x - b_i u)$ provides that ((H.6) and (H.10)):

$$E(F_n(x)) = \frac{1}{n}\sum_{i=1}^{n}\left(F(x) + \frac{(-1)^k}{k!} b_i^k F^k(x)\int u^k K(x)dx\right)$$

$$= F(x) + \frac{(-1)^k}{k!} F^k(x)\left(\int u^k K(x)dx\right) b_n^k \left(\frac{1}{n}\sum_{i=1}^{n}\left(\frac{b_i}{b_n}\right)^k\right) + o(b_n^k).$$

By (H.8), we have that $\sup_{x\in S}| E(F_n(x)) - F(x)| = O(b_n^k)$.

Now set $A_n(x) = F_n(x) - E(F_n(x))$. We are going to use the reconstruction techniques (see Doukhan [3]): consider a sequence $\{(q(n), l(n))\}_{n\in\mathbb{N}^*}$ such that $n - q(n) \le 2l(n)q(n) \le n$, and set

$$A_n(x) = \frac{1}{n}\sum_{j=1}^{2l(n)} Z_j(x) + \Delta_n(x),$$

where $Z_j(x) = \sum_{i=(j-1)q(n)+1}^{jq(n)}\left[H\left(\frac{x - X_i}{b_i}\right) - EH\left(\frac{x - X_i}{b_i}\right)\right]$ and note that we have

$$\sup_{x\in S}| \Delta_n(x)| \le \frac{\varepsilon}{3} \quad \text{a.s.,} \tag{8}$$

provided that $q(n) = o(n\varepsilon)$. By Rio's lemma (see Ango-Nzé and Doukhan [2]), we can change the variables $Z_j(x)$ with j odd (resp. with j even) by independent variables $Z'_j(x)$ with the same distribution of $Z_j(x)$ and such that

$$P(| Z_j(x) - Z'_j(x)| \ne 0) = \alpha(q(n)). \tag{9}$$

Because of (8) we get:

$$I_n = P[\sup_{x\in S}| A_n(x)| > \varepsilon]$$

$$\le P\left[\sup_{x\in S}\frac{1}{n}\sum_{j \text{ even}} Z_j(x) > \frac{\varepsilon}{3}\right] + P\left[\sup_{x\in S}\frac{1}{n}\sum_{j \text{ odd}} Z'_j(x) > \frac{\varepsilon}{3}\right]. \tag{10}$$

Using (9) and classical Lipschitz arguments ((H.5)), we write the compact S as $S = \bigcup_{k=1}^{l_n}[\gamma_k, \gamma_{k+1}]$ with $l_n = O(r_n^{-1})$ intervals with longitude $r_n = b_n$. We get

$$P\left[\sup_{x\in S}\frac{1}{n}\sum_{j\text{ even}} Z_{j,y} > \frac{\varepsilon}{3}\right]$$

$$\leq C\frac{1}{b_n}\sup_k P\left[\frac{1}{n}\sum_{j\text{ even}} Z'_{j,\gamma_k}1_{\{Z'_{j,\gamma_k}\neq Z_{j,\gamma_k}\}} > \frac{\varepsilon}{3}\right] + \frac{n\alpha(q(n))}{2b_n q(n)}. \tag{11}$$

The other term in r.h.s. of (10) can be treated similarly, and then, by Bernstein's inequality for i.i.d. variables, we get

$$I_n \leq 2C\frac{1}{b_n}\exp\left\{\frac{-n^2\varepsilon^2/q^2(n)}{18l(n)\left(\frac{\sigma^2}{q(n)} + \frac{n\varepsilon}{9l(n)q(n)}\right)}\right\} + \frac{n\alpha(q(n))}{b_n q(n)}, \tag{12}$$

where $\sigma^2 = Var\left[H\left(\frac{x - X_i}{b_i}\right)\right]$. By Lemma 4, we get $\sigma^2 \leq C$.

Now, we take $\varepsilon = \varepsilon_o b_n^k$ and $q(n) = o(b_n^{-k})$. We have, for some $l < +\infty$,

$$\sum_{n=1}^{\infty} I_n \leq l + \sum_{n=1}^{\infty}\frac{n\alpha(q(n))}{b_n q(n)}.$$

Because (H.9) and (H.11), we have that $\sum_{n=1}^{\infty} I_n < +\infty$ and the proof is complete.

Proof of Lemma 4. We have

$$Var(F_n(x)) = \frac{1}{n^2}\sum_{i=1}^{n} Var\left[H\left(\frac{x - X_i}{b_i}\right)\right]$$

$$+ \frac{2}{n^2}\sum\sum_{1\leq i<j\leq n} Cov\left[H\left(\frac{x - X_i}{b_i}\right), H\left(\frac{x - X_j}{b_j}\right)\right]. \tag{13}$$

For each i, the variance

$$Var\left[H\left(\frac{x-X_i}{b_i}\right)\right] = E\left[H\left(\frac{x-X_i}{b_i}\right)^2\right] - \left(E\left[H\left(\frac{x-X_i}{b_i}\right)\right]\right)^2, \tag{14}$$

with $E\left[H\left(\frac{x-X_i}{b_i}\right)\right] = F(x) + o(1)$ (see (7)). An integration by parts argument can be applied to the moment of order two, to obtain that $E\left[H\left(\frac{x-X_i}{b_i}\right)^2\right] = F(x) + o(1)$. For the sum of covariances in (13), we can use (22) below to get that this term tends to zero. □

Proof of Theorem 5. We will use a typical argument (Masry [12]) which consists in decomposing the sum of dependent random variables into a sum of large and small blocks. It will be proven that the contribution of the small blocks is negligible and that the large blocks are approximately independent. After this, we will use Lindeberg-Feller's central limit theorem to achieve asymptotic normality.

Write

$$\begin{aligned} n^{1/2}[F_n(x) - E(F_n(x))] &= n^{1/2}\left\{\frac{1}{n}\sum_{i=1}^{n}\left(H\left(\frac{x-X_i}{b_i}\right) - E\left[H\left(\frac{x-X_i}{b_i}\right)\right]\right)\right\} \\ &= n^{-1/2}\sum_{i=1}^{n} L_i(x). \end{aligned}$$

Now, we split

$$n^{-1/2}\sum_{i=1}^{n} L_i(x) = n^{-1/2}(S_n + T_n + T_n'),$$

where

$$S_n = \sum_{m=1}^{\mu} y_m, \quad T_n = \sum_{m=1}^{\mu} y_m', \quad T_n' = y_{\mu+1}', \tag{15}$$

and

$$y_m = \sum_{i=k_m}^{k_m+\rho-1} L_i(x), \quad \text{with } k_m = (m-1)(\rho+\beta)+1; \tag{16}$$

$$y'_m = \sum_{i=l_m}^{l_m+\beta-1} L_i(x), \quad \text{with } l_m = (m-1)(\rho+\beta)+\rho+1; \tag{17}$$

$$y'_{\mu+1} = \sum_{i=\mu(\rho+\beta)+1}^{n} L_i(x). \tag{18}$$

The asymptotic normality is proven showing that:

$$n^{-1}E[T_n^2] \to 0, \tag{19}$$

$$n^{-1}E[(T'_n)^2] \to 0, \tag{20}$$

and

$$n^{-1/2}S_n \to N(0, \sigma(x)), \tag{21}$$

where $\sigma^2(x) = F(x)(1-F(x))$.

Proof of (19). Because the variables $L_i(x)$ have mean zero, we obtain

$$\frac{1}{n}E[T_n^2] = \frac{1}{n}Var\left(\sum_{m=1}^{\mu} y'_m\right)$$

$$= \frac{1}{n}\sum_{m=1}^{\mu} E(y'^2_m) + \frac{2}{n}\sum\sum_{1\le i<j\le\mu} E(y'_i y'_j)$$

$$= A_1 + A_2. \tag{22}$$

By (16),

$$E(y'^2_m) = \sum_{t=0}^{\beta-1} Var(L_{l_n+t}(x)) + \sum_{t=0,}^{\beta-1}\sum_{l=0;\, t\ne l}^{\beta-1} Cov(L_{l_n+t}(x),\, L_{l_n+l}(x))$$

$$\leq \sum_{t=0}^{\beta-1} Var(L_{l_n+t}(x)) + 4\sum_{t=0,\, l=0;\, t\neq l}^{\beta-1}\sum^{\beta-1} \alpha(|t-l|)\,|L_{l_n+t}(x)|\,|L_{l_n+l}(x)|, \quad (23)$$

because $|L_{l_n+t}(x)| \leq C$ and using Ibragimov's inequality [10]. Moreover, $Var(L_{l_n+t}(x))$ is also bounded (see (12)). Therefore,

$$A_1 \leq \frac{1}{n}\mu\beta C(x) + \frac{1}{n}16\mu\sum_{t=0,\, l=0;\, t\neq l}^{\beta-1}\sum^{\beta-1} \alpha(|t-l|)$$

$$\leq \frac{1}{n}\mu\beta C(x) + \frac{1}{n}16\mu\beta\sum_{r=0}^{\infty}\alpha(r). \quad (24)$$

Because (H.14), $\mu\beta/n \to 0$, and by (H.1), $A_1 \to 0$.

Let us consider now A_2 in (22): because $i \neq j$ and the definition of the y'_s we can write

$$\sum\sum_{1\leq i<j\leq\mu} E(y'_i y'_j) \leq 2\sum_{l_1=1}^{n-\rho}\sum_{l_2=l_1+\rho}^{n}\alpha(l_2-l_1) \leq 2n\sum_{r=\rho}^{\infty}\alpha(r).$$

Then, by (H.14), we have that $A_2 \to 0$.

Proof of (20). Decomposing $E[(T'_n)^2]$ into the sum of variances and covariances, we obtain that

$$\frac{1}{n}E[(T'_n)^2] \leq C\frac{n-\mu(\rho+\beta)}{n} + \frac{2}{n}\sum\sum_{1\leq i<j\leq n}|E[L_i(x)L_j(x)]|. \quad (25)$$

By definition of $L_k(x)$, we get

$$|E[L_k(x)L_l(x)]|$$

$$\leq b_k b_l\left|\int_{\mathbb{R}^2} H(u)H(v)[f_{k,l}(x-b_k u,\, x-b_l v) - f(x-b_k u)f(x-b_l v)]dudv\right|$$

$$\leq b_k b_l \alpha(|k-l|), \quad (26)$$

integrating by substitution and using (H.3).

Now, we consider the subsets:

$$S_1 = \{(i,\ j) \in \{1,\ ...,\ n\}^2 / 1 \le j - i \le m_n\},$$

$$S_2 = \{(i,\ j) \in \{1,\ ...,\ n\}^2 / m_n + 1 \le j - i \le n - 1\},$$

where m_n is given in (H.12). The sum in the second term at r.h.s. of (25) can be divided into sets S_1 and S_2. By definition of S_2 and by (26), we obtain

$$\sum_{(i,\ j) \in S_2} |E[L_i(x) L_j(x)]| \le \sum_{j = m_n + 1}^{\beta - 1} \alpha(j) \sum_{i=1}^{\beta - j} b_{j+i} b_i. \tag{27}$$

Now, using Schwarz's inequality,

$$\sum_{i=1}^{\beta - j} b_{j+i} b_i \le b_n^2 \sum_{k=1}^{n} \frac{b_k^2}{b_n^2}. \tag{28}$$

Because (H.1), (H.7) and (H.13), we obtain that

$$n^{-1} \sum_{(i,\ j) \in S_2} |E[L_i(x) L_j(x)]| \to 0.$$

A similar argument can be applied for the set S_1. By (H.14), the proof of (20) is complete.

Proof of (21). We write $n^{-1/2} S_n = \sum_{m=1}^{\mu} n^{-1/2} y_m = \sum_{m=1}^{\mu} Z_m$. To prove (21) we only have to prove that

$$\left| E[e^{iun^{-1/2} S_n}] - \prod_{m=1}^{\mu} E[e^{iun^{-1/2} y_m}] \right| \to 0 \tag{29}$$

and

$$\sum_{m=1}^{\mu} Z_m \to N\left(0,\ \sqrt{Var\left(\sum_{m=1}^{\mu} Z_m\right)}\right). \tag{30}$$

(29) results from Volkonskii and Rozanov [23]:

$$\left| E[e^{iun^{-1/2}S_n}] - \prod_{m=1}^{\mu} E[e^{iun^{-1/2}y_m}] \right| \leq c(\mu - 1)\alpha(\beta + 1),$$

that tends to zero, by (H.14).

For (30), consider $Z'_m = Z_m / s_n$, where $s_n^2 = \sum_{m=1}^{\mu} Var(Z_m)$. Using analogous arguments to those used in (22), we obtain that

$$s_n^2 \to F(x)\overline{F}(x). \tag{31}$$

Now, $\{Z'_m\}$ is a sequence of independent random variables with mean zero and variance one. Therefore, to show (30), we only have to check that

$$\sum_{m=1}^{\mu} Z'_m \to N(0, 1). \tag{32}$$

Let us consider

$$g_n(\varepsilon) = \sum_{m=1}^{\mu} E[(Z'_m)^2 I_{\{Z'_m \geq \varepsilon\}}]$$

$$= s_n^{-2} n^{-1} \sum_{m=1}^{\mu} E[y_m^2 I_{\{|\, y_m \,| \geq \varepsilon s_n n^{1/2}\}}] \leq C \frac{\mu}{n s_n^2} \rho \max_{1 \leq m \leq \mu} P[|\, y_m \,| \geq \varepsilon s_n n^{1/2}]$$

because $|\, y_m \,| \leq c\rho$, for all m. Using (H.14) and (31) we have that $g_n(\varepsilon)$ tends to zero. This proves the asymptotic normality (32). This fact and (30) prove (21).

References

[1] I. A. Ahmad, Uniform strong convergence of the generalized failure rate estimate, Bull. Math. Statist. 17 (1976), 77-84.

[2] P. Ango-Nzé and P. Doukhan, Functional estimation for time series: uniform convergence properties, J. Statist. Plann. Inference 68 (1998), 5-29.

[3] P. Doukhan, Mixing: properties and examples, Lecture Notes in Statist. 85, Springer-Verlag, 1994.

[4] G. Estévez, H. Lorenzo and A. Quintela, Nonparametric analysis of the time structure of seismicity in a geographic region, Ann. Geophys.-Italy 45 (2002a), 497-512.

[5] G. Estévez and A. Quintela, Nonparametric estimation of the hazard function under dependence conditions, Commun. Statist. Theor. Method 28(10) (1999), 2297-2331.

[6] G. Estévez, A. Quintela and P. Vieu, Convergence rate for crossvalidatory bandwidth in kernel hazard estimation from dependent samples, J. Statist. Plann. Inference 104 (2002b), 1-30.

[7] P. Grazinoli and H. Lorenzo, Avaliação do Radar de Penetração no Solo (GPR) em Uua investigação do subsolo, Proceedings of the 7th International Congress of the Brazilian Geophysical Society, 2001.

[8] S. Hassani, P. Sarda and P. Vieu, Approche non paramétrique en théorie de la fiabilité, Rev. Statist. Appl. 34(4) (1986), 27-41.

[9] M. Hollander and F. Proschan, Nonparametric concepts and methods in reliability, Handbook of Statist., P. R. Krishnaiah and P. K. Sen, eds., Vol. 4., North-Holland, Amsterdam, 1984, pp. 613-655.

[10] I. A. Ibragimov, Some limit theorems for stationary processes, Theo. Probab. Appl. 7 (1962), 349-382.

[11] A. J. Izenman and L. T. Tran, Kernel estimation of the survival function and hazard rate under weak dependence, J. Statist. Plann. Inference 24 (1990), 234-247.

[12] E. Masry, Recursive probability density estimation for weakly dependent stationary processes, IEEE Trans. Inform. Theory 32 (1986), 254-267.

[13] B. L. S. Prakasa Rao and J. Van Ryzin, Asymptotic theory for two estimators of the generalized failure rate, Statistical Theory and Data Analysis, K. Matusita, ed., North-Holland, Amsterdam, 1985.

[14] A. Quintela-Del-Rio and Ph. Vieu, A nonparametric conditional mode estimate, J. Nonparametr. Stat. 8 (1997), 253-266.

[15] J. Rice and M. Rosenblatt, Estimation of the log survivor function and hazard function, Sankhyä, Ser. A 38 (1976), 60-78.

[16] M. Rosenblatt, A central limit theorem and a strong mixing condition, Proc. Natl. Acad. Sci. 42 (1956), 43-47.

[17] G. Roussas, Hazard rate estimation under dependence conditions, J. Statist. Plann. Inference 22 (1989), 81-93.

[18] G. Roussas, Asymptotic normality of the kernel estimate under dependence conditions: application to hazard rate, J. Statist. Plann. Inference 25 (1990), 81-104.

[19] G. Roussas and L. H. Tran, Joint asymptotic normality of kernel estimates under dependence conditions, with application to hazard rate, J. Nonparametr. Stat. 1 (1992), 335-355.

[20] P. Sarda and P. Vieu, Empirical distribution function for mixing random variables, Appl. Nonparametr. Hazard Estimation Statist. 20 (1989), 559-571.

[21] L. H. Tran, Recursive density estimation under dependence, IEEE Trans. Inform. Theory 35 (1989), 1103-1108.

[22] J. M. Vilar and J. A. Vilar, Recursive local polynomial regression under dependence conditions, Test 9 (2000), 209-232.

[23] V. A. Volkonskii and Y. A. Rozanov, Some limit theorems for random functions I, Theory Probab. Appl. 4 (1959), 178-197.

[24] G. S. Watson and M. R. Leadbetter, Hazard analysis I, Biometrika 51 (1964a), 175-184.

[25] G. S. Watson and M. R. Leadbetter, Hazard analysis II, Sankhyä, A 26 (1964b), 110-116.