Ann. Inst. Statist. Math.
Vol. 52, No. 4, 630–645 (2000)

LAWS OF ITERATED LOGARITHM AND RELATED ASYMPTOTICS FOR ESTIMATORS OF CONDITIONAL DENSITY AND MODE*

K. L. MEHRA[1], Y. S. RAMAKRISHNAIAH[2] AND P. SASHIKALA[3]

[1] *Department of Mathematical Sciences, University of Alberta, Edmonton, Canada T6G 2G1*
[2] *Department of Statistics, Osmania University, Hyderabad-500007, India*
[3] *Department of Statistics, B.B.C.I.T. Kachiguda, Hyderabad-500027, India*

(Received September 29, 1997; revised February 4, 1999)

Abstract. Let (X_i, Y_i) be a sequence of i.i.d. random vectors in $R^{(2)}$ with an absolutely continuous distribution function H and let $g_x(y)$, $y \in \mathbb{R}^{(1)}$ denote the conditional density of Y given $X = x \in \Lambda(F)$, the support of F, assuming that it exists. Also let $M(x)$ be the (unique) conditional mode of Y given $X = x$ defined by $M(x) = \arg\max_y(g_x(y))$. In this paper new classes of smoothed rank nearest neighbor (RNN) estimators of $g_x(y)$, its derivatives and $M(x)$ are proposed and the laws of iterated logarithm (pointwise), uniform a.s. convergence over $-\infty < y < \infty$ and $x \in$ a compact $C \subseteq \Lambda(F)$ and the asymptotic normality for the proposed estimators are established. Our results and proofs also cover the Nadayara-Watson (NW) case. It is shown using the concept of the relative efficiency that the proposed RNN estimator is superior (asymtpotically) to the corresponding NW type estimator of $M(x)$, considered earlier in literature.

Key words and phrases: Conditional density, conditional mode, smooth rank nearest neighbor estimators, law of iterated logarithm, uniform strong convergence.

1. Introduction

Let $\{(X_i, Y_i) : i \geq 1\}$ be a sequence of bivariate random vectors with a common continuous distribution function (d.f.) H with an absolutely continuous density h and marginal d.f.'s F and G, respectively. Further let g_x (assuming that it exists) denote the conditional density of Y given $X = x \in \Lambda(F)$, the support of X, and for each $x \in \Lambda(F)$, let $M(x)$ defined by

$$M(x) = \arg\sup_y [g_x(y)] \tag{1.1}$$

denote the conditional mode of Y given $X = x$. Assume that $M(x)$ is the unique mode of $g_x(y)$, $-\infty < y < \infty$.

While many researchers have studied the estimation of the unconditional distribution, density and quantile functions and the mode, work on the estimation of the corresponding conditional functionals started only recently with the pioneering work of Stone (1977) and the papers of Stute (1986) on the weak convergence and a.s. convergence rates of conditional empirical processes (see also Horvath and Yandell (1988) and Hardle *et al.* (1988)). The estimators considered in the preceding papers were all of the usual unsmoothed type. However, for the estimation of conditional mode $M(x)$ as defined above,

* Research supported in part by the National Science and Engineering Council of Canada Grant A-3061.

one needs a smooth estimator of the conditional density $g_x(y)$, $-\infty < y < \infty$, which itself must be smooth, at least in some neighbourhood of $M(x)$. Additionally, when the underlying functions to be estimated are smooth, it seems natural to devise smooth estimators for their estimation. A smooth estimator of smooth g_x, of the Nadaraya-Watson (NW) smooth type, was employed by Samanta and Thavaneswaran (1990) to construct a class of smooth estimators for $M(x)$. They also proved the a.s. consistency and asymptotic normality for these estimators.

The object of the present paper is to propose a new class of smoothed RNN estimators for the conditional density $g_x(y)$, its derivatives $g_x^{(j)}(y)$, $-\infty < y < \infty$, $j = 1, 2, \ldots$, and the conditional mode $M(x)$ and establish their respective laws of iterated logarithm (LIL), asymptotic normality, uniform a.s. convergence rates over $-\infty < y < \infty$ and $x \in$ any compact $C \subseteq \Lambda(F)$, as well as derive their asymptotic relative efficiencies w.r. to their NW-type competitors. As will be demonstrated below, the proposed smooth RNN estimators of $g_x^{(j)}$, $j = 0, 1, 2, \ldots$ and $M(x)$ are generally superior to their NW counterparts studied in Samanta and Thavaneswaran (1990).

Consider the following smoothed estimators of the conditional density $g_x(y) = g_x^{(0)}(y)$, $-\infty < y < \infty$, and its j-th derivative $g_x^{(j)}(y)$ —assumed to exist and be continuous— $j = 1, 2, \ldots$ defined by

$$g_{nx}^{(j)}(y) = (na_n h_n^{1+j} t_n(x))^{-1} \sum_{i=1}^{n} k_{1n}(x, F_n(X_i)) k_{2n}^{(j)}(y, Y_i), \tag{1.2}$$

$j = 0, 1, 2, \ldots$, $-\infty < y < \infty$, where $F_n(x) = n^{-1} \sum_{i=1}^{n} I_{[X_i \leq x]}$, $k_{1n}(x, F_n(u)) = k_1((F_n(x) - F_n(u))/a_n)$, $k_{2n}^{(j)}(y, v) = k_{2n}^{(j)}((y - v)/h_n)$, $t_n(x) = (na_n)^{-1} \sum_{i=1}^{n} k_{1n}(x, F_n(X_i))$, k_1 and k_2 are suitable "smooth" kernel functions and $\{a_n\}$, $\{h_n\}$ are sequences of band-widths with a_n, $h_n \to 0$ but $na_n \to \infty$ and $nh_n \to \infty$, as $n \to \infty$. The estimators $g_{nx}^{(j)}$, $j = 0, 1, 2, \ldots$ defined by (1.2) above, denote the successive derivatives of the corresponding smooth RNN conditional empirical d.f. and shall be referred to in the sequel as the RNN estimators of g_x and its derivatives. Define now the RNN estimator $M_n(x)$ of the conditional mode $M(x)$, based on g_{nx}, by

$$M_n(x) = \arg\sup_y [g_{nx}(y)]. \tag{1.3}$$

In the case of the NW type estimator $M_n^*(x)$ of $M(x)$, based on the NW type estimator g_{nx}^* of g_x (obtained from $g_{nx}(y)$ by replacing $F_n(x)$ and $F_n(X_i)$ with x and X_i, respectively; see Samanta and Thavaneswaran (1990)), the study of its asymptotics requires the explicit existence of the marginal density f of X, while no such assumption is needed in the case of the corresponding RNN estimators. The general superiority of RNN estimators over the corresponding NW estimators results from the fact that, in addition to their relative robustness, the ratio of the aympotic variance of an RNN estimator to that of the corresponding NW estimator equals $f(x)$ in all cases, a quantity which is usually less than one for most values of x. Additionally in situations where the observed X-values are available only in terms of their ranks —relative to each other and the value x instead of their precise values, the RNN estimators g_{nx} and (therefore) $M_n(x)$ are definable while the corresponding NW types are not.

The paper is organized as follows: In Section 2, the assumptions and the main theorems along with certain standard results needed in the sequel are stated, while in Section 3 proofs of the main results are presented. The RNN estimators $g_{nx}^{(j)}$, $j = 0, 1, 2, \ldots$ and $M_n(x)$ are compared with the corresponding NW estimators $g_{nx}^{(j)*}$, $j = 0, 1, 2, \ldots$ and $M_n^*(x)$ in Section 4.

2 Assumptions and preliminaries

In this section, the required assumptions and some standard results are stated, to be referred to in the sequel as and when their need arises in proofs.

ASSUMPTIONS. For given $j = 0, 1, 2, \ldots$

A.I. (i) the conditional density $g_x(y)$, $-\infty < y < \infty$, of Y, given $X = x \in \Lambda(F)$, has $(j+3)$ bounded continuous non-zero derivatives $g_x^{(j)} = g_x^{(0,j)}$ in a neighbourhood of any specified y; and $g_x^{(0,\ell)}(y)$, $\ell = 1, 2$, is bouded continuous in a neighbourhood of $M(x)$ and $g_x^{(0,2)} \circ M(x) \neq 0$ for all $x \in$ compact $C \subseteq \Lambda(F)$, where $g_x^{(i,j)}(y) = \frac{\partial^{i+j}}{\partial x^i \partial y^j} g_x(y)$; (ii) $E|Y|^r < \infty$, for some $r > 1$; (iii) the uniqueness condition: for a compact interval $C \subseteq \Lambda(F)$, $M' : C \to \mathbb{R}$ and any $\varepsilon > 0$, there exists $\delta > 0$ such that $\sup_{x \in C} |M(x) - M'(x)| > \varepsilon$ implies $\sup_{x \in C} |g_x \circ M(x) - g_x \circ M'(x)| > \delta$ holds. (iv) there exist δ_1, $\delta_2 > 0$ such that $\inf_{x \in C} f(x) > \delta_1$, $\inf_{x \in C} |g_x^{(2)} \circ M(x)| > \delta_2$;

A.II. the kernel functions k_ℓ, $\ell = 1, 2$, (i) are bounded, symmetric around zero and satisfy $\int k_\ell(t)dt = 1$, $\int t k_\ell(t)dt = 0$ and $\int t^2 k_\ell(t)dt < \infty$; (ii) have bounded continuous non-zero derivatives upto third and $(j+3)$-th orders, respectively; (iii) have compact support, (say), [-1,1] and satisfy $Lt_{t \to \pm 1} k_\ell^{(j')}(t) = 0$, $0 \leq j' \leq j + 3$;

A.III. the bandwidth sequences $\{a_n\}$, $\{h_n\}$ satisfy (i) $0 < a_n, h_n \to 0$ with $h_n = O(a_n)$ and $n a_n^2 h_n^{1+2j} (\log n)^{-1} \to \infty$ as $n \to \infty$; but (ii) $n a_n^5 h_n^{1+2j} = O(1)$ and $n a_n h_n^{5+2j} = O(1)$ as $n \to \infty$; (iii) $n a_n h_n^{(5/2)+2j} (\log n)^{-1} \to \infty$ as $n \to \infty$; and

A.IV $n(a_n h_n)^\lambda \to \infty$ for some $\lambda > 1$.

The conditions A.II imposed on k_1 and k_2 are satisfied by a variety of kernels and we give two examples below for either k_1 of k_2: (a) $k(s) = (2\pi)^{-1/2}(1 - s^2)^{-3/2} \cdot \exp\{-s^2/2(1-s^2)\}$ if $|s| < 1$ and $= 0$ otherwise; (b) for an $m > j$, $k(s) = c_m(1 - s^2)^m$ if $|s| < 1$ and $= 0$ otherwise, where c_m is chosen so that $\int k(s)ds = 1$. They are probability kernels and the one in (a) is just a transform of the standard normal density.

We need the following for proving the main results in Section 3. Let $\{\kappa_{\boldsymbol{r}(n)} : n \geq 1\}$ be a sequence of bivariate kernel functions, each element of which is a function of bounded variation on $[\boldsymbol{a}, \boldsymbol{b}]$, $\boldsymbol{a} = (a_1, a_2)$, $\boldsymbol{b} = (b_1, b_2)$ with the n-th element indexed by $\boldsymbol{r}(n) = (r_1(n), r_2(n)) \in R^{(2)}$ such that the j-th component of $\boldsymbol{r}(n)$ is associated with the j-th component in the domain of $\kappa_{\boldsymbol{r}(n)}$, $j = 1, 2$. Now, set $S_n(\boldsymbol{r}(n)) = \sum_{i=1}^n [\kappa_{\boldsymbol{r}(n)}(X_i, Y_i) - E\kappa_{\boldsymbol{r}(n)}(X_i, Y_i)]$ with $\sigma_{\boldsymbol{r}(n), \boldsymbol{s}(n)} = \mathrm{Cov}(\kappa_{\boldsymbol{r}(n)}(X_1, Y_1), \kappa_{\boldsymbol{s}(n)}(X_1, Y_1))$, $\sigma^2_{\boldsymbol{r}(n)} = \sigma_{\boldsymbol{r}(n), \boldsymbol{r}(n)}$ and $\varphi(n) = (2n\sigma^2_{\boldsymbol{r}(n)} \log_2 n)^{1/2}$ with $\log_2 n = \log \log n$.

For proof of the following Proposition 2.1 —an extension to the bivariate case of a theorem of Hall (1981)— the reader is referred to Theorem 3.1 of Mehra and Rama Krishnaiah (1997):

PROPOSITION 2.1. *Suppose* (i) $\log^4 n \int |d\kappa_{\boldsymbol{r}(n)}(\boldsymbol{\xi})| / n\sigma^2_{\boldsymbol{r}(n)} \log_2 n \to 0$ *as* $n \to \infty$ *and*

(ii) $\lim_{\varepsilon \to 0} \limsup_{n \to \infty} \sum_{m \in \Gamma_\varepsilon} |\sigma_{\boldsymbol{r}(m), \boldsymbol{r}(n)} / \sigma^2_{\boldsymbol{r}(n)} - 1| = 0$ *where* $\Gamma_\varepsilon = \{m : |m - n| \leq n\varepsilon\}$. *Then,* $\limsup_{n \to \infty} \pm\{\varphi(n)\}^{-1} S_n(\boldsymbol{r}(n)) \underset{\text{a.s.}}{=} 1$.

Main results. We state below the main results, the proof of which are deferred to Section 3.

THEOREM 2.1. *Under the conditions* A.I(i), A.II *and* A.III(i) *and* (iii),
(i) *for each* $x \in \Lambda(F)$ *and* $y \in \mathbb{R}$,

$$[g_{nx}^{(j)}(y) - g_x^{(j)}(y)] \underset{\text{a.s.}}{=} \sigma_{xj}(y)\tau_{nj}^* I_n + c_{xj}(y)a_n^2 + c_{xj}^*(y)h_n^2 + o(\tau_{nj}^*) \tag{2.1}$$

for $j = 0, 1, 2, \ldots$ *where* $\tau_{nj}^* = [\log_2 n / n a_n h_n^{1+2j}]^{1/2}$, $\sigma_{xj}(y) = \|k_1\|_2 \|k_2^{(j)}\|_2 \sqrt{g_x(y)}$, $\|k\|_2^2 = \int k^2(t)dt$, $c_{xj}(y)$ *and* $c_{xj}^*(y)$ *as in Lemma* 3.2, *below and* $\operatorname{Lim\,sup}_n \pm I_n =_{\text{a.s.}} 1$; (ii) *additionally, if* A.I(ii), (iv) *and* A.III(ii) *also hold, then*

$$\sup_{x \in C, y \in \mathbb{R}^{(1)}} |g_{nx}^{(j)}(y) - g_x^{(j)}(y)| \underset{\text{a.s.}}{=} O(\tau_{nj}), \tag{2.1a}$$

where $\tau_{nj} = (\log n / n a_n h_n^{1+2j})^{1/2}$, *for each* $j = 0, 1, 2, \ldots$.

From Theorem 2.1(i), we can immediately conclude the following:

COROLLARY 2.1. *Under the conditions of Theorem* 2.1(i) *and* A.III(ii),

$$\operatorname*{Lim\,sup}_{n \to \infty} \pm \left(\frac{2 \log_2 n}{n a_n h_n^{1+2j}} \right)^{-1/2} |g_{nx}^{(j)}(y) - g_x^{(j)}(y)| \underset{\text{a.s.}}{=} \sigma_{xj}(y)$$

for each $j = 0, 1, 2, \ldots$, $x \in \Lambda(F)$ *and* $y \in \mathbb{R}^{(1)}$.

THEOREM 2.2. *Under the conditions of Theorem* 2.1 *with* A.III(i) *for* $j = 2$, A.III(ii) *for* $j = 1$ *and additionally* A.I(iii), (i) *for each* $x \in C$,

$$\operatorname*{Lim\,sup}_{n \to \infty} \pm \left(\frac{2 \log_2 n}{n a_n h_n^3} \right)^{-1/2} [M_n(x) - M(x)] \underset{\text{a.s.}}{=} \sigma_x^* \circ M(x); \tag{2.2}$$

and (ii) *further if*, A.I(ii), (iv) *also hold, then* $\sup_{x \in C} |M_n(x) - M(x)| =_{\text{a.s.}} O(\tau_{n1})$, *where* $\tau_{n1} = (\log n / n a_n h_n^3)^{1/2}$ *and* $\sigma_x^*(y) = \sigma_{x1}(y) / |g_x^{(2)} \circ M(x)|$.

THEOREM 2.3. (i) *Under the conditions of Theorem* 2.1(i), *and* A.IV, *for each* $x \in C$, $y \in \mathbb{R}^{(1)}$ *and* $j = 0, 1, 2, \ldots$

$$(n a_n h_n^{1+2j})^{1/2} (g_{nx}^{(j)}(y) - g_x^{(j)}(y)) \to_{\mathcal{L}} N(b_x^{(j)}(y), \sigma_{xj}^2(y)), \tag{2.3}$$

where $b_x^{(j)}(y) = \operatorname{Lim}_{n \to \infty} (n a_n h_n^{1+2j})^{1/2} [a_n^2 c_{xj}(y) + h_n^2 c_{xj}^*(y)]$, *provided the limit exists*, $-\infty < b_x^{(j)}(y) < \infty$ (*cf.* A.III(ii)); *and* (ii) *under the conditions of Theorem* 2.2(i) *and* A.IV, *for each* $x \in C$

$$(n a_n h_n^3)^{1/2} (M_n(x) - M(x)) \to_{\mathcal{L}} N(b_x^*, \sigma_x^{*2} \circ M(x)), \tag{2.3a}$$

as $n \to \infty$, *where* c_{xj}, c_{xj}^*, $\sigma_{xj}(\cdot)$, $\sigma_x^*(\cdot)$ *and* $b_x^* = b_x^{(1)} \circ M(x) / g_x^{(2)} \circ M(x)$ *are as in Theorems* 2.1 *and* 2.2.

Remark 2.1. Theorems 2.1 and 2.3 have been stated explicitly only for the RNN estimators of conditional density, its derivatives and the conditonal mode. These results, however, hold also for the NW versions of these estimators. The proofs are implicit in the ones provided in this paper for the corresponding RNN estimators. This is so because in the NW case precisely the same arguments (with appropriate modifications) apply without the need to handle the "higher order" terms (see the proofs in Section 3) which appear while dealing with the RNN estimators.

It is worth noting that our methods do bring out some improvements in the rates of a.s. convergence over those achieved in Samanta and Thavaneswaran (1990). They showed using $a_n = h_n$ and higher order kernels k_ℓ, $\ell = 1, 2$ which satisfy $Lt_{|s|\to\infty} |s^2 k_\ell^{(j')}(s)| = 0$ for $0 \le j' \le j$, that $\sup_y |g_{nx}^{(2)}(y) - g_x^{(2)}(y)| =_{\text{a.s.}} O(\tau'_{n2})$, as $n \to \infty$, where $\tau'_{nj} = (\log n/nh_n^{2+3j})^{1/2}$. On the other hand, the methods of the present paper use only probability kernels, with compact support, and only k_2 satisfying a similar condition $Lt_{|s|\to 1} k_2^{(j')}(s) = 0$, $0 \le j' \le j$, but establish a better uniform a.s. rate of convergence by showing $\sup_y |g_{nx}^{(j)}(y) - g_x^{(j)}(y)| =_{\text{a.s.}} O(\tau_{nj})$, as $n \to \infty$, for every $j = 0, 1, 2, \ldots$, where $\tau_{nj} = (\log n/nh_n^{2+2j})^{1/2}$, whose rate of convergence to zero is faster than that of τ'_{nj}, $j = 0, 1, 2$.

Remark 2.2. The choice of $\{a_n\}$ and $\{h_n\}$. From the representation obtained in Theorem 2.1, one may determine the "optimum" bandwidth sequences $\{a_n\}$ and $\{h_n\}$ so as to maximize the a.s. rate of convergence of $[g_{nx}^{(j)}(y) - g_x^{(j)}(y)]$ or equivalently minimize its (so termed) asymptotic "span", defined by

$$S(a_n, h_n) = \sigma_{xj}(y)\tau^*_{nj} + |c_{xj}(y)|a_n^2 + |c^*_{xj}(y)|h_n^2, \tag{2.4}$$

where $\tau^*_{nj} = |\log_2 n/na_n h_n^{1+2j}]^{1/2}$ and $c_{xj}(y)$ and $c^*_{xj}(y)$ are given by (3.10a) and (3.10b). Using the standard differentiation technique, one obtains

$$a_{n,\text{opt}} = \left(\frac{4|c^*_{xj}(y)|}{(1+2j)\sigma_{xj}(y)}\right)^{(1+2j)/4(3+j)} \left(\frac{\sigma_{xj}(y)}{4|c_{xj}(y)|}\right)^{(5+2j)/4(3+j)} \lambda_n^{1/(3+j)} \tag{2.4a}$$

$$h_{n,\text{opt}} = \left(\frac{(1+2j)\sigma_{xj}(y)}{4|c^*_{xj}(y)|}\right)^{5/4(3+j)} \left(\frac{4|c_{xj}(y)|}{\sigma_{xj}(y)}\right)^{1/4(3+j)} \lambda_n^{1/3+j}, \tag{2.4b}$$

where $\lambda_n = [\log_2 n/n]^{1/2}$. Thus the optimum bandwidths $a_{n,\text{opt}}$ and $h_{n,\text{opt}}$ are both of the same order $\lambda_n^{1/(3+j)}$, as $n \to \infty$. The above "optimum" values have been obtained under the conditions of Theorem 2.1(i), without the Assumption A.III(ii). This last condition is necessary for Theorem 2.3, otherwise the limiting mean $b_x^{(j)}(y)$ in these would tend to ∞ invalidating the assertion of the theorem. Thus the slowest assumable order for $\{a_n\}$ and $\{h_n\}$ for each Theorem 2.1–Theorem 2.3, and also for Corollary 2.1, to be valid is $\approx [\lambda_n(\log_2 n)^{-1/2}]^{1/(3+j)} = n^{-1/(6+2j)}(n^{-1/8})$.

3. Proofs

We need the following lemmas to establish the main results. Consider the following Taylor's expansion which is valid in view of the smoothness Assumption A.II(ii). For each $x \in \Lambda(F)$ and $y \in \mathbb{R}$, let $\nu_{nx}^{(j)}(y) = t_n(x)(g_{nx}^{(j)}(y) - g_x^{(j)}(y))$ and $k_{1n}^{(j)}(x, F(u)) = k_1^{(j)}((F(x) - F(u))/a_n)$. Then by setting $p_{nx}(y, v) = k_2^{(j)}((y - v/h_n) - h_n^{1+j} g_x^{(j)}(y)$,

$$\nu_{nx}^{(j)}(y) = a_n^{-1} h_n^{-(1+j)} \iint_{A_n} k_{1n}(x, F_n(u)) p_{nx}(y, v)\, dH_n(u, v) \tag{3.1}$$

$$
\begin{aligned}
&= a_n^{-1}h_n^{-(1+j)} \iint_{A_n} k_{1n}(x, F(u))p_{nx}(y,v)dH_n(u,v) \\
&\quad + n^{-1/2}a_n^{-2}h_n^{-(1+j)} \iint_{A_n} \alpha_n(x,u)k_{1n}^{(1)}(x,F(u))p_{nx}(y,v)dH_n(u,v) \\
&\quad + (2na_n^3h_n^{1+j})^{-1} \iint_{A_n} \alpha_n(x,u)^2 k_1^{(2)}(\Delta_n)p_{nx}(y,v)dH_n(u,v) \\
&= J_{nx1}^{(j)}(y) + J_{nx2}^{(j)}(y) + J_{nx3}^{(j)}(y), \qquad \text{(say)}
\end{aligned}
$$

where $H_n = n^{-1}\sum_{\ell=1}^n I_{[X_\ell \le x, Y_\ell \le y]}$, $U_n(x) = n^{1/2}[F_n(x) - F(x)]$, $\alpha_n(x,u) = U_n(x) - U_n(u)$, $\Delta_n a_n$ lies between $F(x)-F(u)$ and $F_n(x)-F_n(u)$ and $A_n = \{u : |F_n(x)-F_n(u)| \le a_n\}$. We shall first prove a lemma below concerning the rates of a.s. convergence of $J_{nx\ell}^{(j)}(y)$, $\ell = 2,3$, $j = 0,1,2,\ldots$ uniformly in $x \in C$ and $y \in \mathbb{R}$:

LEMMA 3.1. *Suppose the conditions of Theorem* 2.1(i) *hold. Then, as* $n \to \infty$, *uniformly in* $y \in \mathbb{R}^{(1)}$ *and* $x \in C \subseteq \Lambda(F)$, (a) $|J_{nx\ell}^{(j)}(y)| =_{\text{a.s.}} o(\tau_{nj}^*)$, (b) $J_{nx\ell}^{(j)}(y) = o_p(\tau_{nj}'')$, $j = 0,1,2,\ldots,\ell = 2,3$ *where* $\tau_{nj}^* = (na_nh_n^{1+2j})^{-1/2}(\log_2 n)^{1/2}$ *and* $\tau_{nj}'' = \tau_{nj}^*(\log_2 n)^{-1/2}$.

PROOF. We first consider $J_{nx3}^{(j)}(y)$. Since $k_1(\cdot)$ vanishes for values outside $[-1,1]$, the expansion (3.1) holds with integration restricted to $A_n = \{u : |F_n(x) - F_n(u)| \le a_n\}$ and further by Stute (1982) we have a.s. on A_n

$$
|F(x) - F(u)| \le |F_n(x) - F_n(u)| + n^{-1/2}|\alpha_n(x,u)| \le ca_n \tag{3.2}
$$

for some positive constant (generic) c, not depending on $x \in C$, and sufficiently large n, the last inequality following in view of the assumption that $a_n \ge (n^{-1}\log a_n^{-1})$, so that $A_n \subset B_n = \{u : |F(x) - F(u)| < ca_n\}$. Writing $V_n(u,v) = n^{1/2}(H_n - H) \circ (u,v)$, $\beta_n(u,v;x,y) = V_n(u,v) - V_n(u,y) - V_n(x,v) + V_n(x,y)$ and $w_n(a_n,h_n;\beta_n) = \sup_{\substack{|x-u|\le c_1a_n \\ |y-v|\le c_2h_n}} |\beta_n(u,v;x,y)|$, we have from (3.1),

$$
\begin{aligned}
|J_{nx3}^{(j)}(y)| &= (2n^{3/2}a_n^3h_n^{1+j})^{-1} \iint_{A_n} \alpha_n^2(x,u)k_1^{(2)}(\Delta_n) \\
&\quad \cdot \left[\int k_{2n}(y,v)dV_n(u,v) - h_n^{1+j}g_x^{(j)}U_n(u)\right] \\
&\quad + (2na_n^3h_n^{1+j})^{-1} \iint_{A_n} \alpha_n^2(x,u)k_1^{(2)}(\Delta_n)p_{nx}(y,v)dH(u,v) \\
&=: J_{nx31}^{(j)}(y) + J_{nx32}^{(j)}(y) = o(\tau_{nj}^*) \qquad \text{for} \quad j = 0,1,2,
\end{aligned} \tag{3.3}
$$

since from (3.1), by Stute (1982, 1984) and the use of (3.11b) as for (3.8) below,

$$
\begin{aligned}
\sup_{x,y}|J_{nx31}^{(j)}(y)| &\underset{\text{a.s.}}{\le} c_1(n^{3/2}a_n^3h_n^{1+j})^{-1}\sup_{B_n}\alpha_n^2(x,u) \\
&\quad \cdot [w_n(a_n,h_n;\beta_n) + h_n^{1+j}|g_x^{(j)}(y)|\sup_{B_n}|\alpha_n(x,u)|] \underset{\text{a.s.}}{=} o(\tau_{nj}^*),
\end{aligned} \tag{3.4}
$$

$$
\sup_{x,y}|J_{nx32}^{(j)}(y)| \underset{\text{a.s.}}{\le} c(na_n^3h_n^{1+j})^{-1}\sup_{B_n}\alpha_n^2(x,u)h_n^{2+j}a_n \underset{\text{a.s.}}{=} o(\tau_{nj}^*), \tag{3.5}
$$

as $n \to \infty$. Now again from (3.1), $J_{nx2}^{(j)}(y)$ can be expressed as

$$(3.6)\quad n^{-1}a_n^{-2}h_n^{-(1+j)}\iint_{A_n}\alpha_n(x,u)k_{1n}^{(1)}(x,F(u))p_{nx}(y,v)dV_n(u,v) + n^{-1/2}a_n^{-2}h_n^{-(1+j)}\iint_{A_n}\alpha_n(x,u)k_{1n}^{(1)}(x,F(u))p_{nx}(y,v)dH(u,v) =: J_{nx21}^{(j)}(y) + J_{nx22}^{(j)}(y) \underset{\text{a.s.}}{=} o(\tau_{nj}^*),$$

where the last equality follows since, by using the same reasoning as for $J_{nx31}^{(1)}(y)$,

$$(3.7)\quad \sup_{x,y}|J_{nx21}^{(j)}| \underset{\text{a.s.}}{\le} n^{-1}a_n^{-2}h_n^{-(1+j)}\sup_{B_n}|\alpha_n(x,u) \cdot [w_n(a_n,h_n;\beta_n) + h_n^{1+j}|g_x^{(j)}(y)|\sup_{B_n}|\alpha_n(x,u)|] \underset{\text{a.s.}}{=} o(\tau_{nj}^*),$$

and, upon setting $x_n(t) = F^{-1}(F(x) - ta_n)$,

$$(3.8)\quad \sup_{x,y}|J_{nx22}^{(j)}(y)| \le n^{-1/2}a_n^{-1}h_n^{-1-j}\sup_{B_n}|\alpha_n(x,x_n(t)| \times \int|k_1^{(1)}(t)|\left|\int p_{nx}(y,y-h_ns)dG_{x_n(t)}(y-h_ns)\right|dt \underset{\text{a.s.}}{=} o(\tau_{nj}^*),$$

as $n \to \infty$, in view of the fact that $\int p_{nx}(y,y-h_ns)dG_{x_n(t)}(y-h_ns) = O(h_n^{2+j})$ for $j = 0,1,2,\ldots$ (see (3.11b) below). This completes the proof of part (a). The part (b) of the lemma can be proved by using the following in-probability bounds

$$(3.9)\quad n^{1/2}\sup_{u,v}[H_n(u,v) - H(u,v)] = O_p(1), \qquad \sup_{A_n}|U_n(x) - U_n(u)| = O_p(a_n^{1/2})$$

and steps similar to those in (3.4), (3.5), (3.7) and (3.8). This completes the proof of the lemma. □

We now establish the asymptotic behaviour of $J_{nx1}^{(j)}(y)$ in (3.1) for each case $j = 0,1,2,\ldots$ in the following lemma:

LEMMA 3.2. *Under the conditions of Lemma* 3.1, *we have, for each* $x \in C \subseteq \Lambda(F)$, $y \in \Lambda(G_x) \subseteq \mathbb{R}^{(1)}$ *and each* $j = 0,1,2$, *as* $n \to \infty$,

$$(3.10)\quad \text{(i)}\quad J_{nx1}^{(j)}(y) \underset{\text{a.s.}}{=} J_n^{(j)}(x,y) + c_{xj}(y)a_n^2 + c_{xj}^*(y)h_n^2 + o(\tau_{nj}^*),$$

$$\text{(i')}\quad = J_n^{(j)}(x,y) + c_{xj}(y)a_n^2 + c_{xj}^*(y)h_n^2 + o_p(\tau_{nj}''), \qquad j = 0,1,2$$

$$\text{(ii)}\quad t_n(x) \underset{\text{a.s.}}{=} 1 + t_{(n)}(x) + O((\log_2 n/na_n^2)^{1/2}),$$

where the last order terms in (3.10)(i), (i') *and* (ii) *are uniform in* $x \in C$ *and* $y \in \mathbb{R}^{(1)}$, $J_{nx1}^{(j)}(y)$ *is given by* (3.1),

$$(3.10\text{a})\quad J_n^{(j)}(x,y) = (a_nh_n^{1+j})^{-1}\iint k_{1n}(x,F(u))p_{nx}(y,v)d(H_n - H)\circ(u,v),$$

$$c_{xj}(y) = 2^{-1}f^{-3}(x)G_x^{(2,j+1)}(y)\mu_2(k_1) \quad \text{and} \quad c_{xj}^*(y) = 2^{-1}g_x^{(j+2)}(y)\mu_2(k_2),$$

$$t_{(n)}(x) = a_n^{-1}\int k_{1n}(x,F(u))d(F_n - F)\circ(u), \qquad \mu_2(k) = \int s^2k(s)ds.$$

PROOF. To establish part (i), we first prove the equivalence of $J^{(j)}_{nx1}$ to $\bar{J}^{(j)}_{nx1}$, i.e.

$$|J^{(j)}_{nx1}(y) - \bar{J}^{(j)}_{nx1}(y)| \underset{\text{a.s.}}{=} o(\tau^*_{nj}), \tag{3.10b}$$

as $n \to \infty$, where $\bar{J}^{(j)}_{nx1}(y)$ is $J^{(j)}_{nx1}(y)$ with integration over the whole space $R^{(2)}$ instead of A_n. To achieve this, note that on the set $A^c_n \cap B_n$ for large n

$$a_n > |F(x) - F(u)| \geq |F_n(x) - F_n(u)| - (|\alpha_n(x,u)|/\sqrt{n}) \underset{\text{a.s.}}{>} a_n(1 - ch_n^{(1/2)+j}\tau_{nj})$$

for some $c > 0$, not depending on $x \in C$, by Stute (1982), so that using the transformations $u = F^{-1}(F(x) - a_n t) = x_n(t)$ and $\nu = y - h_n s = y_n(s)$, we have $1 \geq t > (1 - ch_n^{(1/2)+j}\tau_{nj})$ on $A^c_n \cap B_n$, $j = 0, 1, 2, \ldots$. We thus obtain

$$\begin{aligned} &|J^{(j)}_{nx1}(y) - \bar{J}^{(j)}_{nx1}(y)| \\ &\quad\leq n^{-1/2}a_n^{-1}h_n^{-1-j}\left|\int_{A^c_n\cap B_n} k_1(t)\int p_{nx}(y, y_n(s))d[V_n(x_n(t), y_n(s))]\right| \\ &\quad + a_n^{-1}h_n^{-1-j}\left|\int_{A^c_n\cap B_n} k_1(t)\left\{\int p_{nx}(y, y_n(s))dG_{x_n(t)}(y_n(s))\right\}a_n dt\right| \\ &\quad = \xi^{(j)}_{nx1}(y) + \xi^{(j)}_{nx2}(y) \quad \text{(say).} \end{aligned} \tag{3.11}$$

Now for the first term in (3.11), note that by the boundedness of the integrand, the Assumptions A.II and Stute (1984), we have for some constants c_1, $c_2 > 0$,

$$\sup_{x,y} \xi^{(j)}_{nx1}(y) \leq cn^{-1/2}a_n^{-1}h_n^{-1/2}\tau_{nj}w_n(a_n, h_n; \beta_n) \underset{\text{a.s.}}{=} o(\tau^*_{nj}). \tag{3.11a}$$

Further, for the socond term $\xi^{(j)}_{nx2}$, upon using integration by parts and then applying Taylor's Theorem to $G_{x_n(t)}(y - h_n s)$, the middle integral equals

$$\begin{aligned} &\int k_2^{(j)}(s)d\{G_{x_{n(t)}}(y - h_n s) - G_{x_{n(t)}}(y)\} \\ &\quad = \sum_{\ell=1}^{j+1}(-1)^\ell(h_n^\ell/\ell!)G^{(0,\ell)}_{x_n(t)}(y)\int s^\ell dk_2^{(j)}(s) + O(h_n^{j+2}) \\ &\quad = h_n^{j+1}g^{(j)}_{x_n(t)}(y) + O(h_n^{j+2}), \end{aligned} \tag{3.11b}$$

as $n \to \infty$, the result following by noting that $\int s^{j+1}dk_2^{(j)}(s) = (-1)^{j+1}(j+1)!$ and $\int s^{j'}dk_2^{(j)}(s) = 0$, $0 \leq j' \leq j$ by the Assumptions A.II(i)–(iii). Substituting (3.11b) results in the expression for $\xi^{(j)}_{nx2}$ in (3.11) and using the fact that $A^c_n \cap B_n \subset \lambda^*_n = \{t : 1 \geq t > (1 - ch_n^{(1/2)+j}\tau_{nj}\}$, we obtaing for some constant $c_2 > 0$

$$\sup_{x,y}|\xi^{(j)}_{nx2}(y)| \underset{\text{a.s.}}{\leq} \frac{c_2}{h_n^{1+j}}\left(\int_{\lambda^*_n} dt\right)h_n^{2+j} = O(h_n^{1+j}\tau^*_{nj}) = o(\tau^*_{nj}), \tag{3.11c}$$

as $n \to \infty$. The use of (3.11a)–(3.11c) in (3.11), yields the required equation (3.10b). Now to establish (3.10), we now express $\bar{J}^{(j)}_{nx1}(y)$ as

$$(3.12)\qquad (a_n h_n^{1+j})^{-1}\Big[\iint k_{1n}(x,F(u))p_{nx}(y,v)d(H_n-H)\circ(u,v) + \iint k_{1n}(x,F(u))p_{nx}(y,v)dG_u(v)dF(u)\Big] =: J^{(j)}_n(x,y)+T_{nxj}(y),$$

where $T_{nxj}(y)=E[\kappa_{\boldsymbol{r}(n)}(X_1,Y_1]$ is evaluated in (3.18)–(3.18c) below. Hence part (i) of the lemma follows from (3.10b) and (3.12). Further using the probability bounds given by (3.9a) in place of a.s. bounds, (3.10)(i') can be established following steps similar to those in (3.11) to (3.12). Now to prove part (ii), note from (3.1), upon obtaining a result of the type (3.10b) and followed by one term Taylor's expansion and suitable transformations and splitting of integrals, that

$$(3.13)\qquad \begin{aligned} t_n(x) &= a_n^{-1}\int k_{1n}(x,F(u))[dF_n(u)-d(F_n-F)\circ u] \\ &\quad + n^{-1/2}a_n^{-2}\int_{A_n}\alpha_n(x,u)k_1^{(1)}(\Delta_n(u))[dF(u)+d(F_n-F)\circ u] \\ &\underset{\text{a.s.}}{=}: 1+t_{(n)}(x)+\varepsilon_{n1}(x)+\varepsilon_{n2}(x) \qquad \text{(say)}, \end{aligned}$$

where, using the same reasoning as for (3.4)–(3.8), it follows that $\sup_{x\in C}[\varepsilon_{n1}(x)+\varepsilon_{n2}(x)] \underset{\text{a.s.}}{=} O((\log_2 n/na_n^2)^{1/2})$, as $n\to\infty$. The proof of part (ii) thus follows from (3.13). The proof of the lemma is complete. □

PROOF OF THEOREM 2.1. To establish part (i) we shall first prove the LIL for $J^{(j)}_n(x,y)$ given by (3.10a). In Proposition 2.1, set $\kappa^{(j)}_{\boldsymbol{r}(n)}(X_1,Y_1) = (a_n h_n^{1+j})^{-1} k_{1n}(x,F(X_1))p_{nx}(y,Y_1)$ and $\sigma_{\boldsymbol{r}(n),\boldsymbol{r}(n')} = \operatorname{cov}(\kappa^{(j)}_{\boldsymbol{r}(n)}(X_1,Y_1),\ \kappa^{(j)}_{\boldsymbol{r}(n')}(X_1,Y_1))$. Now

$$(3.14)\qquad \begin{aligned} &E\kappa^{(j)}_{\boldsymbol{r}(n)}(X_1,Y_1)\kappa^{(j)}_{\boldsymbol{r}(n')}(X_1,Y_1) \\ &\quad = (a_n h_n^{1+j} a_{n'} h_{n'}^{1+j})^{-1}\times\iint k_{1n}(x,F(u)) \\ &\qquad \cdot k_{1n'}(x,F(u))p_{nx}(y,v)p_{n'x}(y,v)dH(u,v), \end{aligned}$$

where the middle integral —say, $I_n(u)$— can be expressed as

$$(3.15)\qquad \begin{aligned} I_n(u) &= \int[k^{(j)}_{2n}(y,v)-h_n^{1+j}g^{(j)}_x(y)][k^{(j)}_{2n'}(y,v)-h_{n'}^{1+j}g^{(j)}_x(y)]dG_u(v) \\ &= \int k^{(j)}_{2n}(y,v)k^{(j)}_{2n'}(y,v)dG_u(v) - h_{n'}^{1+j}g^{(j)}_x(y)\int k^{(j)}_{2n}(y,v)dG_u(v) \\ &\quad - h_n^{1+j}g^{(j)}_x(y)\int k^{(j)}_{2n'}(y,v)dG_u(v) + h_n^{1+j}h_{n'}^{1+j}(g^{(j)}_x(y))^2 \\ &= \varepsilon_{n1}(u)+\varepsilon_{n2}(u)+\varepsilon_{n3}(u)h_n^{1+j}h_{n'}^{1+j}(g^{(j)}_x(y))^2, \qquad \text{(say)} \end{aligned}$$

with $\varepsilon_{n1}(u)$ —after the transformation $v = y - h_n s$, integration by parts, the use of Taylors expansion and the Assumption A.II(iii)— given by

$$(3.15a)\quad \begin{aligned}\varepsilon_{n1}(u) &= -\int k_{2n}^{(j)}(s)k_{2n}^{(j)}(sh_n/h_{n'})d[G_u(y-h_ns)-G_u(y)]\\ &= \sum_{\ell=1}^{j+1}(-1)^\ell\frac{h_n^\ell}{\ell!}G_u^{(0,\ell)}(y)\int_{-1}^{1}s^\ell\{k_2^{(j)}(s)+(h_n/h_{n'}-1)sk_2^{(j+1)}(s)\\ &\qquad\qquad + O(h_n/h_{n'}-1)^2\}dk_2^{(j)}(s)\\ &\quad+\sum_{\ell=1}^{j+1}(-1)^\ell\frac{h_n^\ell}{\ell!}G_u^{(0,\ell)}(y)\int_{-1}^{1}s^\ell k_2^{(j)}(s)\\ &\quad\cdot d\{k_2^{(j)}(s)+(h_n/h_{n'}-1)sk_2^{(j+1)}(s)+O(h_n/h_{n'}-1)^2\}+O(h_n^{j+2})\\ &= \sum_{\ell=1}^{j+1}\frac{(-1)^\ell h_n^\ell}{\ell!}G_u^{(0,\ell)}(y)\\ &\quad\cdot\left\{2\int s^\ell k_2^{(j)}(s)dk_2^{(j)}(s)+\left(\frac{h_n}{h_{n'}}-1\right)\right.\\ &\quad\times\left[\int s^{\ell+1}k_2^{(j)}(s)dk_2^{(j)}(s)+\int s^{\ell+1}k_2^j(s)dk_2^{(j+1)}(s)\right.\\ &\quad\left.\left.+\int s^\ell k_2^{(j)}(s)k_2^{(j+1)}(s)ds\right]+O((h_n/h_{n'}-1)^2)\right\}\\ &\quad+O(h_n^{j+2}).\end{aligned}$$

Similar treatment, coupled with the use of $\int s^{j+1}dk_2^{(j)}(s) = (-1)^{j+1}(j+1)!$ and $\int s^{j'}$ $dk_2^{(j)}(s) = 0$ for $0 \le j' \le j$, yields

$$(3.15b)\quad \begin{aligned}\varepsilon_{n2}(u) &= h_{n'}^{1+j}g_x^{(j)}(y)\int k_2^{(j)}(s)d\{G_u(y-h_ns)-G_u(y)\}\\ &= -h_{n'}^{1+j}g_x^{(j)}(y)\sum_{\ell=1}^{j+1}\frac{h_n^\ell}{\ell!}G_u^{(0,\ell)}(y)\left\{(-1)^\ell\int s^\ell dk_2^{(j)}(s)\right\}+O(h_{n'}^{1+j}h_n^{j+2})\\ &= -h_{n'}^{1+j}h_n^{1+j}g_x^{(j)}(y)g_u^{(j)}(y)+O(h_{n'}^{1+j}h_n^{2+j}),\end{aligned}$$

$$(3.15c)\quad \begin{aligned}\varepsilon_{n3}(u) &= -h_{n'}^{1+j}g_x^{(j)}(y)\sum_{\ell=1}^{n+1}\frac{h_{n'}^\ell}{\ell!}G_u^{(0,\ell)}(y)\left\{(-1)^\ell\int_{-1}^{1}s^\ell dk_2^{(j)}(s)\right\}+O(h_n^{1+j}h_{n'}^{j+2})\\ &= -h_n^{1+j}h_{n'}^{1+j}g_x^{(j)}(y)g_u^{(j)}(y)+O(h_n^{1+j}h_{n'}^{2+j}),\end{aligned}$$

From (3.14), we obtain by setting $u = x_n(t)$ and using Taylor's expansion,

$$(3.16)\quad \begin{aligned}&E[\kappa_{r(n)}^{(j)}(X_1,Y_1)\kappa_{r(n')}^{(j)}(X_1,Y_1)]\\ &\quad= (a_{n'}h_{n'}^{1+j}h_n^{1+j})^{-1}\int k_1(t)\{k_1(t)+(a_n/a_{n'}-1)tk_1^{(1)}(t)\}\\ &\qquad+O((a_n/a_{n'}-1)^2)I_n(x_n(t))dt,\end{aligned}$$

where $I_n(x_n(t))$, obtained by using (3.15a)–(3.15c) in (3.15), is given by

$$(3.16a)\quad I(x_n(t)) = \sum_{\ell=1}^{j+1} \frac{(-1)^\ell h_n^\ell}{\ell!} G_{x_n(t)}^{(0,\ell)}(y) \cdot \Big\{ 2\int s^\ell k_2^{(j)}(s)dk_2^{(j)}(s) + (h_n/h_{n'}-1) \times \Big[\int s^{\ell+1} k_2^{(j)}(s)dk_2^{(j)}(s) + \int s^{\ell+1}k_2^{(j)}(s)dk_2^{(j+1)}(s) + \int s^\ell k_2^{(j)}(s)dk_2^{(j)}(s)\Big] + O((h_n/h_{n'}-1)^2)\Big\} + O(h_n^{j+2}) - (h_{n'}h_n)^{1+j}(2g_{x_n(t)}^{(j)}g_x^{(j)}(y) - (g_x^{(j)}(y))^2) + O(h_{n'}^{1+j}h_n^{2+j}).$$

From (3.16) and (3.16a), by setting $n' = n$ we obtain

$$(3.16b)\quad E[\kappa_{\mathbf{r}(n)}^{(j)}]^2 = (a_n h_n^{2+2j})^{-1} \int k_1^2(t) \cdot \Big[\sum_{\ell=1}^{j+1} \frac{(-1)^\ell h_n^\ell}{\ell!} G_{x_n(t)}^{(0,\ell)}(y) 2\int s^\ell k_2^{(j)}(s)dk_2^{(j)}(s) + O(h_n^{j+2}) - h_n^{2+2j}(2g_{x_n(t)}^{(j)}(y)g_x^{(j)}(y) - g_x^{(j)}(y)) + O(h_n^{3+2j})\Big] dt.$$

Now (3.16b), upon further Taylor's expansion of the integrand on the right around t and some simplification, yields

$$(3.17)\quad \begin{aligned} E[\kappa_{\mathbf{r}(n)}^{(j)}]^2 &= (a_n h_n^{1+2j})^{-1}\|k_1\|_2^2\ [-2\psi(k_2^{(j)})g_x(y) + O(h_n \vee a_n)] \\ &= (a_n h_n^{1+2j})^{-1}\|k_1\|_2^2\ \|k_2^{(j)}\|_2^2\ g_x(y) + O(h_n \vee a_n)] \end{aligned}$$

the last equality following since $\psi(k_2^{(j)}) = -\int sk_2^{(j)}(s)dk_2^{(j)}(s) = (1/2)\|k_2^{(j)}\|_2^2$. Also note that, again using transformations $u = x_n(t)$ and $v = y_n(s)$ and then Taylor's expansion of the integrand around t and s, we obtain

$$(3.18)\quad \begin{aligned} E[\kappa_{\mathbf{r}(n)}(X_1,Y_1)] &= (a_n h_n^{1+j})^{-1} \iint k_{1n}(x,F(u))[k_{2n}^{(j)}(y,v) - h_n^{1+j}g_x^{(j)}(y)]dH(u,v) \\ &= -(h_n^{1+j})^{-1}\Big[\int k_2^{(j)}(s)d\{G_x(y-h_ns) - G_x(y)\} + \frac{a_n^2\mu_2(k_1)}{2f^3(x)} \cdot \int k_2^{(j)}(s)d\{G_x^{(2,0)}(y-h_ns) - G_x^{(2,0)}(y)\} + O(a_n^3)\Big] - g_x^{(j)}(y) \\ &=: \varepsilon_{n1}^{(j)} + \varepsilon_{n2}^{(j)} - g_x^{(j)}(y), \end{aligned}$$

where proceeding as for (3.11b) and using $\int sk_2(s)ds = 0$

$$(3.18a)\quad \begin{aligned} \varepsilon_{n1}^{(j)} &= h_n^{-(j+1)}\Big[g_x^{(j)}(y)\frac{(-1)^{j+1}}{(j+1)!}h_n^{j+1}\int s^{(j+1)}k_2^{(j+1)}(s)ds + g_x^{(j+2)}(y)\frac{(-1)^{j+3}}{(j+3)!}h_n^{j+3}\int s^{(j+3)}k_2^{(j+1)}(s)ds + O(h_n^3)\Big] \\ &= g_x^{(j)}(y) + h_n^2 c_{xj}^*(y) + O(h_n^3), \end{aligned}$$

$$(3.18b)\quad \varepsilon_{n2}^{(j)} = a_n^2 c_{xj}(y) + O(a_n^2 h_n^2),$$

so that from (3.18) to (3.18b), we obtain

$$E[\kappa_{\mathbf{r}(n)}(X_1, Y_1)] = h_n^2 c_{xj}^*(y) + a_n^2 c_{xj}(y) + O(h_n^3) + O(a_n^2 h_n^2). \tag{3.18c}$$

From (3.16), (3.16a), (3.17) and (3.18c), it follows forthwith upon simplification that

$$|(\sigma_{\mathbf{r}(n),\mathbf{r}(n')}/\sigma^2_{\mathbf{r}(n)}) - 1| = O(|h_n/h_{n'} - 1| \vee |a_n/a_{n'} - 1|),$$

as $n \to \infty$. Consequently, the conditions of Proposition 2.1 are satisfied, yielding the following LIL for $J_n^{(j)}(x,y)$ given by (3.10a):

$$\operatorname*{Lim\,sup}_{n\to\infty} \pm (2\log_2 n/na_n h_n^{1+2j})^{-1/2} J_n^{(j)}(x,y) \underset{\text{a.s.}}{=} \|k_1\|_2 \, \|k_2^{(j)}\|_2 \sqrt{g_x(y)}. \tag{3.19}$$

Further, by Dvoretzky *et al.* (1956), we obtain from (3.10) and (3.10a) that $\sup_{x\in C}$ $|t_{(n)}(x)| \le \sup_x |n^{-1/2} a_n^{-1} \int k_1(t) d[U_n \circ x_n(t)]| =_{\text{a.s.}} O(\log_2 n/n\, a_n^2)^{1/2}$, which yields

$$\sup_{x\in C} |t_n(x) - 1| = O(\log_2 n/na_n^2)^{1/2} \quad \text{a.s.} \quad n \to \infty. \tag{3.19a}$$

Now using notations of (3.1) and (3.10a) consider the identity in (3.19b) below: for each $x \in C$ and $y \in \mathbb{R}$

$$\begin{aligned} g_{nx}^{(j)}(y) - g_x^{(j)}(y) &\equiv \nu_{nx}^{(j)}(y) - J_n^{(j)}(x,y)(t_n(x) - 1) - (\nu_{nx}^{(j)}(y) - J_n^{(j)}(x,y)) \\ &\quad \cdot (t_n(x) - 1)/t_n(x) + J_n^{(j)}(x,y) \cdot (t_n(x) - 1)^2/t_n(x) \\ &=: \nu_{nx}^{(j)}(y) + \varepsilon_{n1}(x,y) + \varepsilon_{n2}(x,y) + \varepsilon_{n3}(x,y), \end{aligned} \tag{3.19b}$$

where by Lemma 3.1(a), Lemma 3.2(i), (3.19), (3.19a) and $\inf_{x\in C} |t_n(x)| > 0$ a.s., $\varepsilon_{np}(x,y)$ $p = 1,2,3$ are all almost surely $o(\tau_{nj}^*)$, as $n \to \infty$; in fact, they are uniformly so in $x \in C$ and $y \in \mathbb{R}^{(1)}$, in view of (3.21a) below. Thus from (3.19b) and Lemmas 3.1 and 3.2(i), it follows that

$$g_{nx}^{(j)}(y) - g_x^{(j)}(y) \underset{\text{a.s.}}{=} J_n^{(j)}(x,y) + c_{xj}(y)a_n^2 + c_{xj}^*(y)h_n^2 + o(\tau_{nj}^*), \tag{3.19c}$$

the order term being uniform in $x \in C$ and $y \in \mathbb{R}^{(1)}$, as $n \to \infty$, form which in view of (3.19) it follows that

$$g_{nx}^{(j)}(y) - g_x^{(j)}(y) \underset{\text{a.s.}}{=} \sigma_{xj}(y)\tau_{nj}^* I_n + c_{xj}(y)a_n^2 + c_{xj}^*(y)h_n^2 + o(\tau_{nj}^*), \tag{3.20}$$

where $\operatorname{Lim\,sup}_{n\to\infty} \pm I_n = 1$ with either of the signs. Thus, part (i) is proved.

Part (ii): In view of Lema 3.2(i), the boundedness of c_{xj} and c_{xj}^* and the identity (3.19b), it suffices to establish

$$\sup_{x\in C, y\in\Lambda(G_x)} |J_n^{(j)}(x,y)| \underset{\text{a.s.}}{=} O(\tau_{nj}) \quad \text{and} \tag{3.21}$$

$$\sup_{x\in C, y\in\Lambda(G_x)} |\varepsilon_{np}(x,y)| \underset{\text{a.s.}}{=} o(\tau_{nj}) \quad \text{as} \quad n \to \infty, \; p = 1,2,3; \tag{3.21a}$$

but (3.21a) is implied by (3.21) in view of (3.19a). The proof is thus reduced to the establishment of (3.21). To this end, let $\{b_n\}$ and $\{c_n^*\}$ be two sequences of reals such that $b_n \to \infty$ and $c_n^* \to 0$ as $n \to \infty$, say, $b_n = (a_n h_n^{1+j}/n)^{-1/r}$ (see A.I(ii)) and

$c_n^* = (a_n h_n^{1+j})^{3/2}$. Now set $A_{nj} = \{B((x_\ell, y_\ell); c_n^*); \ell = 1, 2, \ldots, N_j\}$ for a finite set of suitably selected balls with centers at (x_ℓ, y_ℓ) and radii c_n^*, covering the compact set $C \times [-b_n, b_n]$, where $N_j \simeq [4b^* b_n / \pi c_n^{*2}]$ with $C =: (-b^*, b^*)$ (say), $[\cdot]$ denoting the integer part. Now expressing

$$\text{(3.22a)} \quad \sup_{x\in C, y\in\mathbb{R}^{(1)}} |J_n^{(j)}(x,y)| = \sup_{x\in C, |y|\le b_n} |J_n^{(j)}(x,y)| + \sup_{x\in C, |y|>b_n} |J_n^{(j)}(x,y)| =: J_{n1} + J_{n2},$$

$$\text{(3.22b)} \quad J_{n1} \le \max_{1\le \ell \le N_j} \sup_{(x,y)\in B(x_\ell, y_\ell)} \{|J_n^{(j)}(x,y) - J_n^{(j)}(x_\ell, y_\ell)| + |J_n^{(j)}(x_\ell, y_\ell)|\} =: J_{n11} + J_{n12},$$

where, from (3.10a), $|J_n^{(j)}(x,y) - J_n^{(j)}(x_\ell, y_\ell)|$ can be split up as

$$\begin{aligned} \text{(3.22c)} \quad & |(a_n h_n^{1+j})^{-1} \iint k_{1n}(x, F(u))[\{p_{nx}(y,v) - p_{nx}(y_\ell, v)\} \\ & \qquad + \{p_{nx}(y_\ell, v) - p_{n,x_\ell}(y_\ell, v)\}] \\ & \cdot d(H_n - H)\circ(u,v) + (a_n h_n^{1+j})^{-1} \\ & \cdot \iint [k_{1n}(x, F(u)) - k_{1n}(x_\ell, F(u))] p_{x_\ell}(y_\ell, v) d(H_n - H)\circ(u,v) \\ & \le \{c_1 (a_n h_n^{2+j})^{-1} |y - y_\ell| + c_2 (a_n^2 h_n^{1+j})^{-1} |x - x_\ell|\} \sup_{u,v} |H_n(u,v) - H(u,v)| \\ & \le c c_n^* (a_n^2 h_n^{2+j})^{-1} \sup_{u,v} |H_n(u,v) - H(u,v)|, \end{aligned}$$

and hence by the above construction, (3.22b)-(3.22c), Kiefer and Wolfowitz (1958) and the fact that the last quantity in (3.22c) does not depend on ℓ,

$$\begin{aligned} \text{(3.23)} \quad P(J_{n11} > c_1 \tau_{nj}) & \le P[\sup_{u,v} |H_n(u,v) - H(u,v)| > (a_n h_n^{1+2j})^{1/2} c_1 \tau_{nj}] \\ & \le c \exp[-c_1^2 n a_n h_n^{1+2j} \tau_{nj}^2] \\ & \le c n^{-1-\delta_1} \end{aligned}$$

for some suitable $\delta_1 > 0$, upon adjustment of constant c_1. Further, upon setting $Z_i(x,y) = \kappa_{\mathbf{r}(n)}^{(j)}(X_i, Y_i)$ noting that $|Z_i(x,y)| < c(a_n h_n^{1+j})^{-1}$ a.s., $\mathrm{Var}(Z_i(x,y)) = (a_n h_n^{1+2j})^{-1} \cdot [\|k_1\|_2^2 \cdot \|k_2^{(j)}\|_2^2 \, g_x(y) + O(a_n \vee h_n)]$ (see (3.17)-(3.18c)), and

$$\text{(3.23a)} \quad J_{n12} \le \max_{1\le\ell\le N_j} \left| n^{-1} \sum_{i=1}^n \{Z_i(x_\ell, y_\ell) - EZ_i(x_\ell, y_\ell)\} \right| + O(h_n^2 \vee a_n^2),$$

it follows by applying Bernstein inequality to $\{Z_i(x_\ell, y_\ell); 1 \le i \le n\}$ that

$$\text{(3.23b)} \quad P(J_{n12} > c_2 \tau_{nj}) \le 2N_j \exp\left[-\frac{c_6 n a_n h_n^{1+2j} (c_2 \tau_{nj} - c_7 a_n^2)^2}{2c_8 + 2/3 \cdot c_9 (c_2 \tau_{nj} - c_7 a_n^2) h_n} \right].$$

Now, setting $c_2^2 = 2\delta(c_8/c_6)$ for a $\delta > 0$ and adjusting constant c_2 so that $\delta > [1 + \{(2 + 3r)/r\}] =: (1 + \eta)$, denote $\delta_2 = (\delta - \eta - 1) > 0$. Then from (3.23b),

$$\text{(3.23c)} \quad P[J_{n12} > \tau_2] \le c N_j n^{-\delta} = c n^{-1-\delta_2},$$

the last inequality following since $N_j = O(n^\eta)$ as $n \to \infty$, in view of $(a_n h_n^{1+j})^{-1} = O(n)$ by the Assumption A.III(i).

Now we deal with the term J_{n2} in (3.19a): Since $k_2^{(j)}$ has compact support on $[-1,1]$, $|y - Y_1| \leq h_n$ implies $y - h_n < Y_1 < y + h_n$, so that, whenever $|y| > b_n$, $Y_1 < -b_n + h_n < -b_n/2$ or $Y_1 > b_n - h_n > b_n/2$ for sufficiently large $n \geq n_0$ (say). Thus, $|k_2^{(j)}((y - Y_1)/h_n)| \leq \|k_2^{(j)}\| \cdot I(|Y_1| \geq b_n/2)$ a.s. and $J_{n2} = \sup_{x,|y|>b_n} |J_n^{(j)}(x,y)| \leq \|k_1\| \cdot \|k_2^{(j)}\| \cdot (na_n h_n^{1+j})^{-1} \sum_{i=1}^n I(|Y_i| \geq b_n/2)$. Now setting $Y_i^* = I(|Y_i| \geq b_{n/2})$ and $p_n = EY_i^* = P[|Y_i| \geq b_n/2] \leq 2^r b_n^{-r} E|Y_i|^r < cb_n^{-r}$ by Markov inequality, $\theta = \|k_1\|\|k_2^{(2)}\|$ and applying again Bernstein inequality to $\{Y_i^*\}$,

$$
\begin{aligned}
(3.24)\qquad P[J_{n2} > c_3\tau_{nj}] &\leq P\left[n^{-1}\sum_{i=1}^n (Y_i^* - p_n) \geq (a_n h_n^{1+j}\theta^{-1}c_3\tau_{nj} - cb_n^{-r})\right] \\
&\leq 2\exp\left[-\frac{c_{11}na_n h_n^{1+j}\{c_3\tau_{nj} - (c\theta/n)\}^2}{2\theta^2 c_{12} + c_{13}\theta\{c_3\tau_{nj} + c\theta/n\}}\right] \\
&\leq c\exp[-c_{13}na_n h_n^{1+2j}\{c_3\tau_{nj} - O(n^{-1})\}^2] \leq cn^{-1-\delta_3},
\end{aligned}
$$

for some $\delta_3 > 0$ (by adjusting constant c_3), in view of the above construction with $b_n \simeq (a_n h_n^{1+j}/n)^{-1/r}$. From (3.22a), (3.22b), (3.23), (3.23c) and (3.24), it follows that $\sum_{n\geq 1} P[\sup_{x\in C, y\in\mathbb{R}^{(1)}} |J_n^{(j)}(x,y)| > c_0\tau_{nj}] < \infty$, where $c_0 = c_1 + c_2 + c_3$. Thus by B.C. Lemma, (3.21) follows. The proof of Theorem 2,1 is complete. □

PROOF OF THEOREM 2.2. In view of uniqueness of mode $M_n(x)$, by Taylor's expansion of $g_{nx}^{(1)} \circ M_n(x)$ around $M(x)$ we obtain

$$(3.25)\qquad 0 = g_{nx}^{(1)} \circ M_n(x) = g_{nx}^{(1)} \circ M(x) + (M_n(x) - M(x))g_{nx}^{(2)} \circ M_n'(x)$$

for some $M_n'(x)$ lying between $M(x)$ and $M_n(x)$. This leads to

$$
\begin{aligned}
(3.25a)\qquad M_n(x) - M(x) &= -\frac{g_{nx}^{(1)} \circ M(x)}{g_n^{(2)} \circ M(x)}\left[1 + \frac{(g_x^{(2)} \circ M(x) - g_{nx}^{(2)} \circ M_n'(x))}{g_{nx}^{(2)} \circ M_n'(x)}\right] \\
&= -\frac{g_{nx}^{(1)} \circ M(x)}{g_x^{(2)} \circ M(x)}(1 + o(1)) \qquad \text{a.s.},
\end{aligned}
$$

as $n \to \infty$, the last equality following since $|[g_x^{(2)} \circ M(x) - g_{nx}^{(2)} \circ M_n'(x)]| \leq \|g_x^{(3)}\|_\infty |M_n(x) - M(x)| + \sup_{x\in C, y\in\mathbb{R}^{(1)}} |g_{nx}^{(2)}(y) - g_x^{(2)}(y) \underset{\text{a.s.}}{\to} 0$ by Theorem 2.1(ii) for $j = 2$ provided we establish (3.25d) below, the order term in (3.25a) being clearly uniform in $x \in C$, since in view of the Assumption A.I(iv), Theorem 2.1(ii) with $j = 2$ implies $\operatorname{Lim\,inf}_{n\to\infty} \inf_{x\in C} |g_{nx}^{(2)} \circ M_n'(x)| > 0$ a.s. The assertion of part (i) of the theorem is then an immediate consequence of A.I(i) and Theorem 2.1(i) with $j = 1$, since $g_x^{(1)} \circ M(x) = 0$ by definition.

It remains to prove (3.25d). For this note that by definition $g_{nx} \circ M_n(x) = \sup_{y\in\Lambda(G_x)} g_{nx}(y)$ and $g_x \circ M(x) = \sup_{y\in\Lambda(G_x)} g_x(y)$, so that

$$
\begin{aligned}
(3.25b)\qquad &|g_x \circ M_n(x) - g_x \circ M(x)| \\
&= |-g_{nx} \circ M_n(x) + g_x \circ M_n(x) - g_x \circ M(x) + g_{nx} \circ M_n(x)| \\
&\leq \sup_{y\in\Lambda(G_x)} |g_{nx}(y) - g_x(y)| + \left|\sup_{y\in\Lambda(G_x)} g_{nx}(y) - \sup_{y\in\Lambda(G_x)} g_x(y)\right| \\
&\leq 2\sup_{y\in\Lambda(G_x)} |g_{nx}(y) - g_x(y)|.
\end{aligned}
$$

Also by using two term Taylor's expansion around $M(x)$ and $g_x^{(1)} \circ M(x) = 0$, $g_x \circ M_n(x) - g_x \circ M(x) = \frac{1}{2} g_x^{(2)} \circ M_n''(x)[M_n(x) - M(x)]^2$, so that in conjunction with (3.25b), this implies

$$\text{(3.25c)} \qquad \sup_{x \in C} |M_n(x) - M(x)|^2 \leq 4 \sup_{x \in C, y \in \mathbb{R}^{(1)}} |g_{nx}(y) - g_x(y)| / \inf_{x \in C} |g_x^{(2)}(M_n''(x))|.$$

Upon taking limit, as $n \to \infty$, on both sides of (3.25c), by virtue of the Assumption A.I(iv) and Theorem 2.1(ii) for $j = 0$, it follows that

$$\text{(3.25d)} \qquad \sup_{x \in C} |M_n(x) - M(x)| = O(\tau_{n0}^{1/2}),$$

as $n \to \infty$. This completes the proof of part (i). In fact, (3.25d) proves part (ii) also partially (in establishing a rate of convergence to zero, but a rate that is slower than the rate achieved in part (ii)).

To prove part (ii) beyond (3.25d), namely, concerning the stated faster a.s. rate of convergence to zero of $\sup_{x \in C} |M_n(x) - M(x)|$, note from (3.25a) that, in view of $g_x^{(1)} \circ M(x) = 0$ and the Assumption A.I(iv), this a.s. rate equals that of $\sup_{x \in C} |g_{nx}^{(1)} \circ M(x)| = \sup_{x \in C} |g_{nx}^{(1)} \circ M(x) - g_x^{(1)} \circ M(x)|$, which by Theorem 2.1(ii) (for $j = 1$) is $O(\tau_{n1})$ a.s., as $n \to \infty$. This proves part (ii). The proof is complete. □

PROOF OF THEOREM 2.3. Note from (3.1), Lemmas 3.1(b) and 3.2(i') and (3.19) that, for sufficiently large n and $j = 0, 1, 2$,

$$t_n(x)[g_{nx}^{(j)}(y) - g_x^{(j)}(y)] = J_n^{(j)}(x, y) + c_{xj}(y) a_n^2 + c_{xj}^*(y) h_n^2 + o_p((n a_n h_n^{1+2j})^{-1/2}),$$

or equivalently,

$$\text{(3.26)} \quad t_n(x) \sqrt{n a_n h_n^{1+2j}} [g_{nx}^{(j)}(y) - g_x^{(j)}(y)] = \frac{1}{\sqrt{n}} \sum_{i=1}^{n} (Z_{ni}^{(j)} - E Z_{ni}^{(j)}) + b_x^{(j)}(y) + o_p(1)$$

by A.III(ii), where from (3.10a), $Z_{ni}^{(j)} = (a_n h_n)^{-1/2} k_{1n}(x, F(X_i)) p_{nx}(y, Y_i)$ with

$$\begin{aligned}
\sigma_{nxj}^2(y) &= \operatorname{Var} Z_{n1}^{(j)} \\
&= \int k_1^2(t) \int [k_2^{(j)}(s) - h_n^{1+j} g_x^{(j)}(y)]^2 g_{x_n(t)}(y - h_n s) ds dt - E Z_{n1}^{(j)^2} \\
&= g_x(y) \int k_1^2(t) dt \cdot \int k_2^{(j)}(s)^2 ds + O(a_n^2) + O(h_n^{2+j}) \to \sigma_{xj}^2(y),
\end{aligned}$$

as $n \to \infty$. Therefore, for $\delta = 2/(\lambda - 1) > 0$, we have (see A.IV)

$$\frac{\sum_{i=1}^n E|Z_{ni}^{(j)} - E Z_{ni}^{(j)}|^{2+\delta}}{(\sum_1^n \operatorname{Var} Z_{ni}^{(j)})^{1+\delta/2}} \leq \frac{n^{-(1+\delta/2)} \cdot n (a_n h_n)^{-(1+\delta/2)}}{(n^{-1} \cdot \sum_1^n \operatorname{Var} Z_{ni}^{(j)})^{1+\delta/2}} = O([n (a_n h_n)^{\lambda}]^{-\delta/2})$$

which tends to zero, as $n \to \infty$, by the Assumption A.IV, so that the Liapunov condition is satisfied. Hence the assertion of part (i) is proved in view of (3.10) and (3.26),

Part(ii) is an immediate consequence of part (i) with $j = 1$ and (3.25a). The proof of the theorem is complete. □

4. Concluding remarks

In this section we find the asymptotic relative efficiency of the RNN estimator $M_n(x)$ of conditional mode $M(x)$ with respect to the competing estimator $M_n^*(x)$ of Samanta and Thavaneswaran (1990). For this we compare the asymptotic variances of $M_n^*(x)$ and $M_n(x)$ and obtain the asymptotic relative efficiency $\mathrm{ARE}(M_n^*, M_n)$. Now from Theorem 2.3 (ii) we have upon setting $M(x) = y_0$, that for sufficiently large n,

$$\mathrm{Var}(M_n(x)) \simeq (na_nh_n^3)^{-1} g_x(y_0) \int k_1^2(t)dt \int k_2^{(1)}(s)^2 ds/(g_x^{(2)}(y_0))^2,$$

and from the results of Samanta and Thavaneswaran (1990), we have for sufficiently large n,

$$\mathrm{Var}(M_n^*(x)) \simeq (na_nh_n^3)^{-1} h(x,y_0) \int k_1^2(t)dt \int k_2^{(1)}(s)^2 ds/(h^{(0,2)}(x,y_0))^2,$$

Since $g_x(y_0) = h(x,y_0)/f(x)$ and $g_x^{(2)}(y_0)^2 = (h^{(0,2)}(x,y_0))^2/(f(x))^2$, we obtain

$$\mathrm{ARE}(M_n^*, M_n) = \mathrm{Var}(M_n)/\,\mathrm{Var}(M_n^*) = f(x).$$

Since $f(x)$ is usually less than 1 for unimodal densities $f(x)$ when x is not too close to the centre, for such values of x the RNN smooth estimator of $M(x)$ studied in this paper, in addition to being "robust", is also superior in the sense of ARE to the corresponding NW smooth estimator studied by Samanta and Thavaneswaran (1990). The preceding conclusion clearly applies also to the RNN vs. NW estimators of g_x and its derivatives $g_x^{(j)}$, $j = 1, 2, \ldots$.

References

Dvoretzky, A., Kiefer, J. and Wolfowitz, J. (1956). Asymptotic minimax character of the sample distribution function and of the classical multinomial estimator, *Ann. Math. Statist.*, **27**, 642–669.

Hall, P. (1981). Laws of iterated logarithm for nonparametric density estimators, *Z. Wahrsch. Verw. Gebiete.*, **56**, 47–61.

Hardle, W., Janssen, P. and Serfling, R. (1988). Strong uniform convergence rates for estimators of conditional functions, *Ann. Statist.*, **16**, 1428–1499.

Horvath, L. and Yandell, B. S. (1988). Asymptotics of conditional empirical processes, *J. Multivariate Anal.*, **26**, 184–206.

Kiefer, J. and Walfowitz, J. (1958). On the derivations of the empiric distribution function of vector chance variables, *Trans. Amer. Math. Soc.*, **87**, 173–186.

Mehra, K. L. and Rama Krishnaiah, Y. S. (1997). Bahadur representation and the law of the iterated logarithm for smooth conditional quantile estimators, Tech. Report, No. 2, Department of Mathematical Sciences, University of Alberta Edmonton, Canada T6G 2G1.

Samanta, M. and Thavaneswaran, A. (1990). Nonparametric estimation of the conditional mode, *Comm. Statist. Theory Methods*, **19** (12), 4515–4524.

Stone, C. J. (1977). Consistent nonparametric regression, *Ann. Statist.*, **5**, 595–645.

Stute, W. (1982). The oscillation behavior of empirical process, *Ann. Probab.*, **10**, 86–107.

Stute, W. (1984). The oscillation behaviour of empirical processes: The multivariate case, *Ann. Probab.*, **12**, 361–379.

Stute, W. (1986). On almost sure convergence of conditional empirical distribution functions, *Ann. Probab.*, **4**, 891–901.