Weak convergence of local random fields of kernel density estimators [*]

Yoichi Nishiyama
The Institute of Statistical Mathematics and Utrecht University

December 2, 1997

Abbreviated Title: LOCAL KERNEL DENSITY ESTIMATOR

Abstract

It is well-known that kernel density estimators for i.i.d. data have point-wise asymptotic normality. This note extends it to the functional sense with respect to a local parameter. The localizing constants should be the same as the bandwidth.

1 Introduction and results

It is well-known that kernel density estimators for i.i.d. data have point-wise asymptotic normality; this fact is so elementary that we may safely omit a bibliographical review. However, since the density f is originally defined as a Radon-Nikodym derivative with respect to Lebesgue measure, the value $f(x)$ at each point x does not intrinsically make sense. Thus, an assertion in some functional sense is preferable in order for, e.g., the construction of confidence intervals. Although the bandwidth processes or/and deviation processes as random elements taking values in the space C were studied by Krieger and Pickands, III (1981) and Müller and Prewitt (1992, 1993), the weak convergence of local random fields of normalized residuals was not considered in the literature up to the present.

The purpose of this note is to extend the asymptotic normality of kernel density estimators to the functional sense with respect to a local parameter. The localizing constants should be chosen to be the same as the bandwidth. The results are shown by using Ossiander's central

---

[*] This research was performed in the Department of Mathematics, Utrecht University, while the author held a JSPS Fellowship for Research Abroad from the Japan Society for the Promotion of Science.

AMS 1991 subject classifications: 60G07, 62G05, 62G20.

Keywords and phrases. Weak convergence, central limit theorem, density, kernel estimator.

limit theorem for empirical processes indexed by classes of functions; see e.g. Van der Vaart and Wellner (1996) for this theory.

Let $\{X_i\}_{i\in\mathbb{N}}$ be an i.i.d. sequence of $\mathbb{R}^d$-value random variables with Lebesgue density f. Let $x_0 \in \mathbb{R}^d$ be a fixed point, and let $\{b_n\}_{n\in\mathbb{N}}$ be a sequence of positive constants such that $b_n \downarrow 0$ as $n \to \infty$. We are interested in estimating the local function $u \rightsquigarrow f(x_0 + b_n u)$, where the parameter u runs through a subset U of $\mathbb{R}^d$. We consider the kernel estimator

$$\widehat{f}_n(x_0 + b_n u) = \frac{1}{nb_n^d}\sum_{i=1}^n K\left(\frac{X_i - x_0}{b_n} - u\right) \qquad \forall u \in U,$$

where $K(x)$ is a kernel function on $\mathbb{R}^d$. We make two kinds of conditions either of which the kernel function should satisfy.

Condition 1 (smooth kernel) *The function* $K : \mathbb{R}^d \to \mathbb{R}$ *satisfies that:*

(i) $\int_{\mathbb{R}^d} K(x)dx = 1$, $K(x) = K(-x)$ *for every* $x \in \mathbb{R}^d$, *and* K *has a compact support;*

(ii) *there exist* $\alpha > 0$ *and* $L > 0$ *such that* $|K(x) - K(y)| \le L|x - y|^\alpha$ *for every* $x, y \in \mathbb{R}^d$.

Condition 2 (monotone kernel) *The function* $K : \mathbb{R}^d \to \mathbb{R}$ *is of the product form* $K(x) = \prod_{p=1}^d K_p(x^{(p)})$ *of some functions* $K_p : \mathbb{R} \to \mathbb{R}$, $p = 1, ..., d$. *The functions* K_p *need not be the same, but each of them satisfies:*

(i) $\int_{\mathbb{R}} K_p(x)dx = 1$, $K_p(x) = K_p(-x)$ *for every* $x \in \mathbb{R}$, *and* K_p *has a compact support;*

(ii) *the function* $x \rightsquigarrow K_p(x)$ *is non-increasing on* $[0, \infty)$.

Here, and in the sequel, the notation $x^{(p)}$ means the p-th component of a vector $x \in \mathbb{R}^d$.

The goal of this note is to derive the asymptotic behavior of the sequence of (normalized) residual processes $R^n = (R^n(u)|u \in U)$ defined by

$$R^n(u) = \sqrt{nb_n^d}\left\{\widehat{f}_n(x_0 + b_n u) - f(x_0 + b_n u)\right\} \qquad \forall u \in U.$$

The key point is to investigate the processes $Z^n = (Z^n(u)|u \in U)$ given by

$$Z^n(u) = \sqrt{nb_n^d}\left\{\widehat{f}_n(x_0 + b_n u) - \widetilde{f}_n(x_0 + b_n u)\right\} \qquad \forall u \in U,$$

where

$$\begin{aligned}\widetilde{f}_n(x_0+b_n u) &= \frac{1}{b_n^d}\int_{\mathbb{R}^d} K\left(\frac{x-x_0}{b_n}-u\right) f(x)dx \\ &= \int_{\mathbb{R}^d} K(y)f(x_0+b_n(u+y))dy \qquad \forall u\in U.\end{aligned}$$

Notice that the processes R^n and Z^n are not necessarily continuous in the case of a monotone kernel. We thus treat them as $\ell^\infty(U)$-valued random elements, taking advantage of the modern theory of weak convergence; this approach is natural especially in the multi-dimensional case.

Proposition 1 *Choose a kernel function* $K:\mathbb{R}^d\to\mathbb{R}$ *satisfying either Condition 1 or 2, and let* $\{b_n\}$ *be a sequence of positive constants such that* $b_n\downarrow 0$ *and that* $nb_n^d\uparrow\infty$ *as* $n\to\infty$. *If* f *is continuous at* x_0, *and if* U *is bounded, then it holds that* $Z^n\Longrightarrow Z$ *in* $\ell^\infty(U)$, *where* $u\rightsquigarrow Z(u)$ *a centered, continuous Gaussian process such that*

$$E\left(Z(u_1)Z(u_2)\right)=f(x_0)\int_{\mathbb{R}^d}K(x-u_1)K(x-u_2)dx \qquad \forall u_1,u_2\in U. \tag{1}$$

Remark. The continuity of the limit process $u\rightsquigarrow Z(u)$ is considered with respect to the Euclidean metric.

Let $U_1\subset U_2\subset\cdots$ be a sequence of bounded subsets of $\mathbb{R}^d$ such that $\bigcup_{i=1}^\infty U_i=\mathbb{R}^d$. We denote by $\ell^\infty_{\mathrm{loc}}(\mathbb{R}^d)$ the set of all functions $z:\mathbb{R}^d\to\mathbb{R}$ that are bounded on every U_i, and equip it with the local uniform metric d defined by

$$d(z_1,z_2)=\sum_{i=1}^\infty\left(\sup_{u\in U_i}|z_1(u)-z_2(u)|\wedge 1\right)2^{-i}.$$

Using Theorem 1.6.1 of Van der Vaart and Wellner (1996), we obtain the following.

Theorem 2 *Choose a kernel function* $K:\mathbb{R}^d\to\mathbb{R}$ *and a sequence of constants* $\{b_n\}$ *as in Proposition 1.*

(i) *If* f *is continuous at* x_0, *then it holds that* $Z^n\Longrightarrow Z$ *in* $\ell^\infty_{\mathrm{loc}}(\mathbb{R}^d)$, *where* $u\rightsquigarrow Z(u)$ *a centered, continuous Gaussian process whose covariance* $E\left(Z(u_1)Z(u_2)\right)$ *is given by (1) for every* $u_1,u_2\in\mathbb{R}^d$.

(ii) *If* f *is twice continuously differentiable in a neighborhood of* x_0, *and if*

$$\lim_{n\to\infty}nb_n^{5d}=h<\infty,$$

then it holds that $R^n \Longrightarrow z_0 + Z$ *in* $\ell^\infty_{\mathrm{loc}}(\mathbb{R}^d)$, *where*

$$z_0 = \sqrt{h}\sum_{p=1}^{d}\sum_{q=1}^{d}\int_{\mathbb{R}^d} y^{(p)}y^{(q)}K(y)dy \left.\frac{\partial^2 f(x)}{\partial x^{(p)}\partial x^{(q)}}\right|_{x=x_0}.$$

This result can be applied to construct a confidence band, substituting estimators for $f(x_0)$ in the covariance of the limit process Z and for the second derivatives of f at x_0 in the constant z_0. Other applications will be presented elsewhere. Notice that the assumptions appearing above are exactly the same as those in the context of point-wise asymptotic normality, and thus are quite reasonable. Our conclusion is that the local smoothness of the density f implies not only the point-wise asymptotic normality but also the weak convergence of local residual processes R^n.

2 Proofs

Let us begin with stating a version of Ossiander's central limit theorem; see e.g. Theorem 2.11.9 of Van der Vaart and Wellner (1996) or Corollary 4.4 of Nishiyama (1997b) for the proof [the latter is a martingale version of the theorem below]. Notice that no metric is required a-priori there.

$\Pi = \{\Pi(\varepsilon)\}_{\varepsilon\in(0,1]}$ is called a Decreasing series of Finite Partitions (DFP) of an arbitrary set U if:

(i) each $\Pi(\varepsilon) = \{U(\varepsilon;k) : 1 \le k \le N_\Pi(\varepsilon)\}$ is a finite partition of U;

(ii) $\varepsilon \rightsquigarrow N_\Pi(\varepsilon)$ is decreasing, and $\lim_{\varepsilon\downarrow 0} N_\Pi(\varepsilon) = \infty$.

For every $n \in \mathbb{N}$, let $(\Omega^n, \mathcal{F}^n, P^n)$ be a probability space. Given a mapping $Z : \Omega^n \to \mathbb{R}\cup\{\infty\}$ we denote by $[Z]_{\mathcal{F}^n}$ any $\mathcal{F}^n$-measurable majorant of Z; see Lemma 1.2.1 of Van der Vaart and Wellner (1996). Given an $\ell^\infty(U)$-valued row-independent array $\{\xi_i^n\}_{i\in\mathbb{N}} = \{(\xi_i^n(u)|u \in U)\}_{i\in\mathbb{N}}$ on $(\Omega^n, \mathcal{F}^n, P^n)$, we define

$$\overline{\xi_i^n} = \left[\sup_{u\in U}|\xi_i^n(u)|\right]_{\mathcal{F}^n} \qquad \forall i \in \mathbb{N},$$

and given a DFP Π,

$$\|\xi^n\|_\Pi = \sup_{\varepsilon\in(0,1]\cap\mathbb{Q}}\ \max_{1\le k\le N_\Pi(\varepsilon)} \frac{\sqrt{\sum_{i=1}^n E^n|\xi_i^n(U(\varepsilon;k))|^2}}{\varepsilon},$$

where

$$\xi_i^n(U') = \left[\sup_{u_1, u_2 \in U'} |\xi_i^n(u_1) - \xi_i^n(u_2)| \right]_{\mathcal{F}^n} \qquad \forall i \in \mathbb{N},\ \forall U' \subset U.$$

Theorem 3 *Let U be an arbitrary set. For every $n \in \mathbb{N}$, let $\{\xi_i^n\}_{i\in\mathbb{N}}$ be an $\ell^\infty(U)$-valued, centered row-independent array on a probability space $(\Omega^n, \mathcal{F}^n, P^n)$. Assume the following conditions* (A), (B) *and* (C)*:*

- (A) $\lim_{n\to\infty} \sum_{i=1}^n E^n \xi_i^n(u_1)\xi_i^n(u_2) = C(u_1, u_2)$ *[some constant] for every $u_1, u_2 \in U$;*
- (B) $\lim_{n\to\infty} \sum_{i=1}^n E^n |\overline{\xi_i^n}|^2 1_{\{\overline{\xi_i^n} > \varepsilon\}} = 0$ *for every $\varepsilon > 0$;*
- (C) *there exists a DFP Π of U such that*

$$\limsup_{n\to\infty} \|\xi^n\|_\Pi < \infty \quad \textit{and} \quad \int_0^1 \sqrt{\log N_\Pi(\varepsilon)}\, d\varepsilon < \infty.$$

Then, it holds that $\sum_{i=1}^n \xi_i^n \Longrightarrow X$ in $\ell^\infty(U)$, where each marginal $(X(u_1), ..., X(u_d))$ has the normal distribution $N(0, \Sigma)$ with $\Sigma = \{C(u_i, u_j)\}_{ij}$. Furthermore, the formula

$$\rho(u_1, u_2) = \sqrt{C(u_1, u_1) + C(u_2, u_2) - 2C(u_1, u_2)} \qquad \forall u_1, u_2 \in U$$

defines a pseudo-metric on U such that U is totally bounded with respect to ρ, and that almost all paths of X are uniformly ρ-continuous.

Proof of Proposition 1. We can write $Z^n(u) = \sum_{i=1}^n \xi_i^n(u)$ where

$$\xi_i^n(u) = \frac{1}{\sqrt{nb_n^d}} \left\{ K\left(\frac{X_i - x_0}{b_n} - u \right) - \int_{\mathbb{R}^d} K\left(\frac{x - x_0}{b_n} - u \right) f(x) dx \right\}.$$

We will check the conditions of Theorem 3. For every $u_1, u_2 \in U$, since

$$\begin{aligned} E\xi_i^n(u_1)\xi_i^n(u_2) &= \frac{1}{nb_n^d} \left\{ \int_{\mathbb{R}^d} K\left(\frac{x - x_0}{b_n} - u_1 \right) K\left(\frac{x - x_0}{b_n} - u_2 \right) f(x) dx \right. \\ &\qquad \left. - \int_{\mathbb{R}^d} K\left(\frac{x - x_0}{b_n} - u_1 \right) f(x) dx \int_{\mathbb{R}^d} K\left(\frac{x - x_0}{b_n} - u_2 \right) f(x) dx \right\} \\ &= \frac{1}{n} \left\{ \int_{\mathbb{R}^d} K(y - u_1) K(y - u_2) f(x_0 + b_n y) dy \right. \\ &\qquad \left. - b_n^d \int_{\mathbb{R}^d} K(y - u_1) f(x_0 + b_n y) dy \int_{\mathbb{R}^d} K(y - u_2) f(x_0 + b_n y) dy \right\}, \end{aligned}$$

we easily obtain

$$\lim_{n\to\infty} \sum_{i=1}^n E\xi_i^n(u_1)\xi_i^n(u_2) = f(x_0) \int_{\mathbb{R}^d} K(y - u_1) K(y - u_2) dy.$$

The Lindeberg condition (B) follows from the assumption $nb_n^d \uparrow \infty$. In the following, we will show that (C) of Theorem 3 is satisfied under either Condition 1 or 2, and that the limit process $u \rightsquigarrow Z(u)$ is continuous with respect to the Euclidean metric.

[The case of smooth kernel.] Assume Condition 1. First notice that for any $u_1, u_2 \in U$

$$|K(y-u_1) - K(y-u_2)| \leq L|u_1 - u_2|^\alpha \qquad \forall y \in \mathbb{R}^d. \tag{2}$$

We can take a compact set S which is a common support of the functions $y \rightsquigarrow K(y-u)$ for all $u \in U$. Now, for every $\varepsilon > 0$, choose a finite partition $\Pi(\varepsilon) = \{U(\varepsilon; k) : 1 \leq k \leq N_\Pi(\varepsilon)\}$ of U such that the diameter of each partitioning set is not bigger than $\varepsilon^{1/\alpha}$. This can be done with $N_\Pi(\varepsilon) \leq$ const. $\varepsilon^{-d/\alpha}$; thus it holds that $\int_0^1 \sqrt{\log N_\Pi(\varepsilon)} d\varepsilon < \infty$. On the other hand, it follows from (2) that if $|u_1 - u_2| \leq \varepsilon^{1/\alpha}$ then

$$|\xi_i^n(u_1) - \xi_i^n(u_2)| \leq \frac{L\varepsilon}{\sqrt{nb_n^d}} \left\{ 1_S\left(\frac{X_i - x_0}{b_n}\right) + \int_{\mathbb{R}^d} 1_S\left(\frac{x - x_0}{b_n}\right) f(x)dx \right\}.$$

We thus have

$$\begin{aligned}
\|\xi^n\|_\Pi^2 &\leq \frac{L^2}{nb_n^d} \sum_{i=1}^n E\left| 1_S\left(\frac{X_i - x_0}{b_n}\right) + \int_{\mathbb{R}^d} 1_S\left(\frac{x - x_0}{b_n}\right) f(x)dx \right|^2 \\
&\leq \frac{4L^2}{b_n^d} \int_{\mathbb{R}^d} 1_S\left(\frac{x - x_0}{b_n}\right) f(x)dx \\
&= 4L^2 \int_{\mathbb{R}^d} 1_S(y) f(x_0 + b_n y)dy \\
&\leq 4L^2 \cdot \mathrm{Leb}(S) \cdot \sup_{x \in N} f(x) \qquad \text{for all sufficiently large } n \in \mathbb{N},
\end{aligned}$$

where N is a neighborhood of x_0. The condition (C) of Theorem 3 has been established.

[The case of monotone kernel.] Assume Condition 2. For every $p = 1, ..., d$, choose a constant $c_p > 0$ such that $[-c_p, c_p]$ is a support of K_p and that $U \subset \prod_{p=1}^d (-c_p, c_p]$.

For every $\varepsilon > 0$ and every $p = 1, ..., d$, we introduce a finite partition $(-c_p, c_p] = \bigcup_{k_p=1}^{N_p(\varepsilon)} I_p(\varepsilon; k_p)$ where $I_p(\varepsilon; k_p) = (\gamma_p(\varepsilon; k_p - 1), \gamma_p(\varepsilon; k_p)]$, such that

$$0 < \gamma_p(\varepsilon; k_p) - \gamma_p(\varepsilon; k_p - 1) \leq \varepsilon^2, \qquad k_p = 1, ..., N_p(\varepsilon).$$

This can be done with $N_p(\varepsilon) \leq [2c_p\varepsilon^{-2}] + 1$. Now, to check the condition (C) of Theorem 3, we consider the DFP $\Pi = \{\Pi(\varepsilon)\}_{\varepsilon\in(0,1]}$ of U given by

$$\Pi(\varepsilon) = \left\{ U \cap \prod_{p=1}^{d} I_p(\varepsilon;k_p) : \quad 1 \le k_p \le N_p(\varepsilon),\ 1 \le p \le d \right\}.$$

Then, since

$$N_\Pi(\varepsilon) = \#\Pi(\varepsilon) \le \prod_{p=1}^{d} N_p(\varepsilon) \le \prod_{p=1}^{d}\left(\left[\frac{2c_p}{\varepsilon^2}\right] + 1 \right)$$

we have $\int_0^1 \sqrt{\log N_\Pi(\varepsilon)}\, d\varepsilon < \infty$.

Next, for every $u^{(p)} \in (-c_p, c_p]$ we define

$$K_p^{n,u^{(p)}}(x^{(p)}) = \frac{1}{\sqrt{b_n}} K_p\left(\frac{x^{(p)} - x_0^{(p)}}{b_n} - u^{(p)} \right), \qquad \forall x^{(p)} \in \mathbb{R}.$$

Then it holds that

$$|K_p^{n,u_1^{(p)}} - K_p^{n,u_2^{(p)}}| \le \overline{K}_p^{n,\varepsilon,k_p} \qquad \text{whenever } \ u_1^{(p)}, u_2^{(p)} \in I_p(\varepsilon;k_p), \tag{3}$$

where

$$\overline{K}_p^{n,\varepsilon,k_p}(x^{(p)}) = \begin{cases} \dfrac{K_p(0)}{\sqrt{b_n}}, & \text{if } \ \dfrac{x^{(p)} - x_0^{(p)}}{b_n} \in I_p(\varepsilon;k_p), \\ |K_p^{n,\gamma_p(\varepsilon;k_p)} - K_p^{n,\gamma_p(\varepsilon;k_p-1)}|(x^{(p)}), & \text{otherwise.} \end{cases}$$

The key points are the following:

$$\int_{\mathbb{R}} |\overline{K}_p^{n,\varepsilon,k_p}(x_0^{(p)} + b_n y^{(p)})|^2 dy^{(p)} \le \frac{3|K_p(0)|^2\varepsilon^2}{b_n}; \tag{4}$$

$$\mathrm{Support}(\overline{K}_p^{n,\varepsilon,k_p}) \subset [x_0^{(p)} - 2b_n c_p, x_0^{(p)} + 2b_n c_p]; \tag{5}$$

$$f_0 = \limsup_{n\to\infty} \sup_{y\in S} f(x_0 + b_n y) < \infty, \quad \text{where } S = \prod_{p=1}^{d} [-2c_p, 2c_p]. \tag{6}$$

The fact (4) will be proved later, while (5) and (6) are trivial.

Let us proceed with the main part of the proof. It follows from (3) that

$$\begin{aligned}
&\frac{1}{\sqrt{b_n^d}} \left| K\left(\frac{x - x_0}{b_n} - u_1 \right) - K\left(\frac{x - x_0}{b_n} - u_2 \right) \right| \\
&= \left| \prod_{p=1}^{d} K_p^{n,u_1^{(p)}}(x^{(p)}) - \prod_{p=1}^{d} K_p^{n,u_2^{(p)}}(x^{(p)}) \right| \\
&= \left| \sum_{p=1}^{d} \left(\prod_{q=1}^{p-1} K_q^{n,u_2^{(q)}}(x^{(q)}) \right) \left(\prod_{q=p+1}^{d} K_q^{n,u_1^{(q)}}(x^{(q)}) \right) \left\{ K_p^{n,u_1^{(p)}}(x^{(p)}) - K_p^{n,u_2^{(p)}}(x^{(p)}) \right\} \right| \\
&\le \sum_{p=1}^{d} \left(\prod_{q\ne p} \frac{K_q(0)}{\sqrt{b_n}} \right) \left| K_p^{n,u_1^{(p)}}(x^{(p)}) - K_p^{n,u_2^{(p)}}(x^{(p)}) \right| \\
&\le \sum_{p=1}^{d} \left(\prod_{q\ne p} \frac{K_q(0)}{\sqrt{b_n}} \right) \overline{K}_p^{n,\varepsilon,k_p}(x^{(p)}), \qquad \text{if } \ u_1, u_2 \in \prod_{p=1}^{d} I_p(\varepsilon;k_p).
\end{aligned}$$

Here, for every $p = 1, ..., d$, we obtain from (4), (5) and (6) that for all sufficiently large $n \in \mathbb{N}$

$$\begin{aligned}
&\left(\prod_{q\neq p} \frac{K_q(0)}{\sqrt{b_n}}\right)^2 \int_{\mathbb{R}^d} |\overline{K}_p^{n,\varepsilon,k_p}(x^{(p)})|^2 f(x)dx \\
= \;& \left(\prod_{q\neq p} |K_q(0)|^2\right) b_n \int_{\mathbb{R}^d} |\overline{K}_p^{n,\varepsilon,k_p}(x_0^{(p)} + b_n y^{(p)})|^2 f(x_0 + b_n y)dy \\
\leq \;& \left(\prod_{q\neq p} |K_q(0)|^2\right) b_n (f_0 + 1) \cdot \int_S |\overline{K}_p^{n,\varepsilon,k_p}(x_0^{(p)} + b_n y^{(p)})|^2 dy \\
\leq \;& \left(\prod_{q\neq p} |K_q(0)|^2\right) b_n (f_0 + 1) \cdot \left(\prod_{q\neq p}^{d} 4c_q\right) \frac{3|K_p(0)|^2\varepsilon^2}{b_n} \\
= \;& \varepsilon^2 D_p, \qquad \text{where} \qquad D_p = \left(\prod_{q=1}^{d} 4c_q |K_q(0)|^2\right) \frac{3(f_0+1)}{4c_p},
\end{aligned}$$

which implies that

$$\int_{\mathbb{R}^d} \left| \sum_{p=1}^{d} \left(\prod_{q\neq p} \frac{K_q(0)}{\sqrt{b_n}}\right) \overline{K}_p^{n,\varepsilon,k_p}(x^{(p)}) \right|^2 f(x)dx \leq \varepsilon^2 d \sum_{p=1}^{d} D_p.$$

We therefore have

$$\limsup_{n\to\infty} \|\xi^n\|_{\Pi} \leq \sqrt{4d \sum_{p=1}^{d} D_p}.$$

It remains to prove (4). Observe that

$$\int_{\mathbb{R}} |\overline{K}_p^{n,\varepsilon,k_p}(x_0^{(p)} + b_n y^{(p)})|^2 dy^{(p)} = (I) + \frac{|K_p(0)|^2}{b_n}\varepsilon^2 + (II),$$

where:

$$\begin{aligned}
(I) &= \frac{1}{b_n} \int_{-\infty}^{\gamma_p(\varepsilon;k_p-1)} \left| K_p(y^{(p)} - \gamma_p(\varepsilon;k_p)) - K_p(y^{(p)} - \gamma_p(\varepsilon;k_p-1)) \right|^2 dy^{(p)}; \\
(II) &= \frac{1}{b_n} \int_{\gamma_p(\varepsilon;k_p)}^{\infty} \left| K_p(y^{(p)} - \gamma_p(\varepsilon;k_p)) - K_p(y^{(p)} - \gamma_p(\varepsilon;k_p-1)) \right|^2 dy^{(p)}.
\end{aligned}$$

Further, it holds that

$$\begin{aligned}
(II) &= \frac{1}{b_n} \int_0^{\infty} \left| K_p(y^{(p)}) - K_p(y^{(p)} + \gamma_p(\varepsilon,k_p) - \gamma_p(\varepsilon,k_p-1)) \right|^2 dy^{(p)} \\
&\leq \frac{1}{b_n} \int_0^{\infty} \left| K_p(y^{(p)}) - K_p(y^{(p)} + \varepsilon^2) \right|^2 dy^{(p)} \\
&\leq \frac{K_p(0)}{b_n} \int_0^{\infty} \left| K_p(y^{(p)}) - K_p(y^{(p)} + \varepsilon^2) \right| dy^{(p)} \\
&= \frac{K_p(0)}{b_n} \int_{-\varepsilon^2}^{0} K_p(y^{(p)} + \varepsilon^2) dy^{(p)} \\
&\leq \frac{K_p(0)}{b_n} \cdot K_p(0)\varepsilon^2.
\end{aligned}$$

Since the same bound holds also for (I), we get (4).

[Continuity of the limit process.] Theorem 3 says that the process $u \rightsquigarrow Z(u)$ is continuous with respect to the pseudo-metric ρ on U defined by

$$\rho(u_1, u_2) = \sqrt{\int_{\mathbb{R}^d} |K(x - u_1) - K(x - u_2)|^2 dx} \qquad \forall u_1, u_2 \in U.$$

Hence it suffices to show that $u_1 \rightsquigarrow \rho(u_1, u_2)$ is continuous at u_2 with respect to the Euclidean metric for every $u_2 \in U$. This is immediate from (2) in the case of a smooth kernel. On the other hand, in the case of a monotone kernel, the claim follows from the inequality

$$|K(x - u_1) - K(x - u_2)| \le \sum_{p=1}^{d} \left(\prod_{q \ne p} K_q(0) \right) \left| K_p(x^{(p)} - u_1^{(p)}) - K_p(x^{(p)} - u_2^{(p)}) \right|$$

which can be easily shown by the same argument as above. □

Proof of Theorem 2. The assertion (i) is immediate from Proposition 1 and Theorem 1.6.1 of Van der Vaart and Wellner (1996). Next, observe that

$$\widetilde{f}_n(x_0 + b_n u) - f(x_0 + b_n u) = \int_{\mathbb{R}^d} K(y) \left\{ f(x_0 + b_n(u + y)) - f(x_0 + b_n u) \right\} dy$$

and that

$$\begin{aligned} f(x_0 + b_n(u + y)) - f(x_0 + b_n u) &= b_n^d \sum_{p=1}^{d} y^{(p)} \left. \frac{\partial f(x)}{\partial x^{(p)}} \right|_{x = x_0 + b_n u} \\ &\quad + b_n^{2d} \sum_{p=1}^{d} \sum_{q=1}^{d} y^{(p)} y^{(q)} \left. \frac{\partial^2 f(x)}{\partial x^{(p)} \partial x^{(q)}} \right|_{x = \tilde{x}_n}, \end{aligned}$$

where $\tilde{x}_n$ is a point on the segment connecting $x_0 + b_n u$ and $x_0 + b_n(u + y)$. We can obtain the assertion (ii) using the assumption that the kernel function K is symmetric. □

3 Remark

As mentioned in Section 2, Ossiander's central limit theorem has been generalized to some martingale cases including continuous-time ones; see Nishiyama (1997a,b). Thus it is possible to obtain similar results for some dependent cases.

Acknowledgments. I am grateful to Nakahiro Yoshida for a stimulating discussion about this subject at Le Mans in January of 1997, and to Boris Y. Levit for a valuable comment. I would like to thank Richard D. Gill for his helpful remarks on a preliminary manuscript and constant encouragement.

References

[1] Krieger, A.M. and Pickands, J., III (1981). Weak convergence and efficient density estimation at a point. *Ann. Statist.* **9** 1066-1078.

[2] Müller, H.G. and Prewitt, K.A. (1992). Weak convergence and adaptive peak estimation for spectral densities. *Ann. Statist.* **20** 1329-1349.

[3] Müller, H.G. and Prewitt, K.A. (1993). Multiparameter bandwidth processes and adaptive surface smoothing. *J. Multivariate Anal.* **47** 1-21.

[4] Nishiyama, Y. (1997a). Some central limit theorems for ℓ^∞-valued semimartingales and their applications. *Probab. Theory Relat. Fields* **108** 459-494.

[5] Nishiyama, Y. (1997b). Weak convergence of multivariate point processes and its applications. *Preprint nr. 1036: Dept. Math., Utrecht Univ.*

[6] Ossiander, M. (1987). A central limit theorem under metric entropy with L_2 bracketing. *Ann. Probab.* **15** 897-919.

[7] van der Vaart, A.W. and Wellner, J.A. (1996). *Weak Convergence and Empirical Processes: With Applications to Statistics.* Springer, New York.

The Institute of Statistical Mathematics
4-6-7 Minami-Azabu
Minato-ku
Tokyo 106
Japan

Mathematical Institute
Utrecht University
Budapestlaan 6
3584 CD Utrecht
The Netherlands