Proceedings of the 7th International Conference on Applied Informatics
Eger, Hungary, January 28–31, 2007. Vol. 1. pp. 211–221.

Central limit theorems for kernel type density estimators*

István Fazekas

Faculty of Informatics, University of Debrecen,
e-mail: fazekasi@inf.unideb.hu

Abstract

Kernel type density estimators are studied for random fields. It is proved that the estimators are asymptotically normal if the set of locations of observations become more and more dense in an increasing sequence of domains. It turns out that in our setting the covariance structure of the limiting normal distribution can be a combination of those of the continuous parameter and the discrete parameter cases. The proof is based on a central limit theorem for α-mixing random fields. Simulation results support our theorems. A functional version of the limit theorem is also presented.

Keywords: Asymptotic normality of estimators, central limit theorem, functional central limit theorem, density estimator, increasing domain asymptotics, infill asymptotics, kernel, random field, α-mixing

MSC: 60F05, 60F17, 62G07, 62M30

1. Introduction

The main result of this paper is Theorem 3.2. It states asymptotic normality of the kernel type density estimator when the set of locations of observations become more and more dense in an increasing sequence of domains. It turns out, that the covariance structure of the limiting normal distribution depends on the ratio of the bandwidth of the kernel estimator and the diameter of the subdivision. This is an important issue when we approximate the integral in the estimator $f_{T_n}(x) = \frac{1}{|T_n|}\frac{1}{h_n}\int_{T_n} K\left(\frac{x-\xi_{\mathbf{t}}}{h_n}\right) d\mathbf{t}$ by a sum, i.e. in applications we use an estimator of the form $f_{\mathcal{D}_n}(x) = \frac{1}{|\mathcal{D}_n|}\frac{1}{h_n}\sum_{\mathbf{i}\in\mathcal{D}_n} K\left(\frac{x-\xi_{\mathbf{i}}}{h_n}\right)$.

This approach can be useful in geosciences, meteorology, environmental studies, image processing, etc. In these sciences several processes varying continuously in space are studied. However, in practice, we can not observe the processes continuously in space. So we have to use finite data sets and discrete approximations.

*Supported by the Hungarian Foundation of Scientific Researches under Grant No. OTKA T047067/2004 and OTKA T048544/2005.

Moreover, the theoretical analyses of statistical models often need simulation methods. In computer simulations always discrete approximations are applied. So we have to know if the limiting behaviour of the continuous model is the same as that of its discrete counterpart.

Kernel type density estimators are widely studied, see e.g. Prakasa Rao [19], Devroye and Györfi [9]. Several papers are devoted to the density estimators for weakly dependent stationary sequences (see e.g. Castellana and Leadbetter [6], Bosq et al. [5]). However, in most of the papers the goal is to find weak dependence conditions of asymptotic normality. Only few papers study the relation of the rate of dependence and the asymptotic behaviour (see e.g. Csörgő and Mielniczuk [8], and Wu and Mielniczuk [22]). In [22] a detailed description is given for the limit laws for linear random sequences.

The asymptotic normality of the kernel type density estimator is well known for weakly dependent continuous time processes (see e.g. [5]). The paper [15] gives an estimator for the asymptotic variance. However, when we calculate numerically the kernel type density estimator, its asymptotic variance can be different from that of the theoretical one. To point out this phenomenon is the goal of our paper, and therefore we turn to so called infill-increasing setup.

In statistics, most asymptotic results concern the increasing domain case, i.e. when the random process (or field) is observed in an increasing sequence of domains T_n, with $|T_n| \to \infty$. However, if we observe a random field in a fixed domain and intend to prove an asymptotic theorem when the observations become dense in that domain, we obtain the so called infill asymptotics (see Cressie [7]). It is known that several estimators being consistent for weakly dependent observations in the increasing domain setup, are not consistent if the infill approach is considered. In this paper we combine the infill and the increasing domain approaches. We call infill-increasing approach if our observations become more and more dense in an increasing sequence of domains. In the theory of kernel type estimators the infill-increasing approach is applied when optimal sampling is studied (see Masry [18], Bosq [4] and the discussion in Remark 3.3 in the present paper).

Our approach and results fit well to recent researches. We shall compare our main theorem with recent results on optimal sampling by Biau [2] and by Bosq [4] in Section 3. We mention that in [20] the kriging, while in [17] the empirical distribution functions were considered using infill-increasing approach.

In this paper we give an overview of some results. In Section 2 we present the basic central limit theorem (CLT) for α-mixing random fields (Theorem 2.1). It is analogous to Theorem 1.1 in [5]. Using Theorem 2.1, we can prove asymptotic normality of the kernel type density estimator (3.2) in the infill-increasing case (Theorem 3.2). The conditions are similar to those of Theorem 2.2 (continuous time process) and Theorem 3.1 (discrete time process) of [5]. In some sense, our result is between the discrete and the continuous time cases. Example 3.5 is a new numerical evidence for the behaviour of the estimator. Our simulation results support the covariance structure of the limiting distribution presented in Theorem 3.2. Theorem 4.1 is a functional version of the ordinary CLT, i.e. of Theorem 3.2. It is an unusual one as the convergence is proved in $L_2[0,1]$, i.e. in the space of square integrable functions defined in the interval $[0,1]$. We have to mention that most of the functional limit theorems are given in the space C of continuous functions, or in the Skorohod space D, see [3]. However, there are papers establishing criteria for functional limit theorems in L_p and containing applications of such theorems (see [14], [16]).

2. A CLT for α-mixing random fields

The following notation is used. $\mathbb{N}$ is the set of positive integers, $\mathbb{Z}$ is the set of all integers, $\mathbb{N}d$ and $\mathbb{Z}^d$ are d-dimensional lattice points, where d is a fixed positive integer. $\mathbb{R}$ is the real line, $\mathbb{R}^d$ is the d-dimensional space with the usual Euclidean norm $\|\mathbf{x}\|$. In $\mathbb{R}^d$ we shall also consider the distance corresponding to the maximum norm: $\varrho(\mathbf{x},\mathbf{y}) = \max_{1\leqslant i\leqslant d}|x^{(i)} - y^{(i)}|$, where $\mathbf{x} = (x^{(1)},\dots,x^{(d)})$, $\mathbf{y} = (y^{(1)},\dots,y^{(d)})$. The distance of two sets in $\mathbb{R}^d$ corresponding to the maximum norm is also denoted by ϱ: $\varrho(A,B) = \inf\{\varrho(\mathbf{a},\mathbf{b}) \,:\, \mathbf{a}\in A,\, \mathbf{b}\in B\}$.

$|\mathcal{D}|$ denotes the cardinality of the finite set $\mathcal{D}$ and at the same time $|T|$ denotes the volume of the domain T.

We shall suppose the existence of an underlying probability space $(\Omega,\mathcal{F},\mathbb{P})$. The σ-algebra generated by a set of events or by a set of random variables will be denoted by $\sigma\{.\}$. Sign $\mathbb{E}$ stands for the expectation. The variance and the covariance are denoted by var(.) and cov(.,.), respectively. Sign $\Rightarrow$ denotes convergence in distribution. $\mathcal{N}(m,\Sigma)$ stands for the (vector) normal distribution with mean (vector) m and covariance (matrix) Σ.

Describe the scheme of observations. For simplicity we restrict ourselves to rectangles as domains of observations. Let $\Lambda > 0$ be fixed. By $(\mathbb{Z}/\Lambda)^d$ we denote the Λ-lattice points in $\mathbb{R}^d$ i.e. lattice points with distance $1/\Lambda$:

$$\left(\mathbb{Z}/\Lambda\right)^d = \left\{\left(k_1/\Lambda,\dots,k_d/\Lambda\right) \,:\, (k_1,\dots,k_d)\in\mathbb{Z}^d\right\}.$$

T will be a bounded, closed rectangle in $\mathbb{R}^d$ with edges parallel to the axes and $\mathcal{D}$ will denote the Λ-lattice points belonging to T, i.e. $\mathcal{D} = T\cap(\mathbb{Z}/\Lambda)^d$. To describe the limit distribution we consider a sequence of the previous objects. I.e. let $T_1, T_2, \dots$ be bounded, closed rectangles in $\mathbb{R}^d$. Suppose that

$$T_1\subset T_2\subset T_3\subset\dots,\qquad \bigcup_{i=1}^{\infty} T_i = T_\infty. \tag{2.1}$$

We assume that the length of each edge of T_n is integer and converges to ∞, as $n\to\infty$ (e.g. $T_\infty = \mathbb{R}^d$ or $T_\infty = [0,\infty)^d$). Let $\{\Lambda_n\}$ be an increasing sequence of positive integers (the non-integer case is essentially the same) and $\mathcal{D}_n$ be the Λ_n-lattice points belonging to T_n.

Let $\xi_{\mathbf{t}}$, $\mathbf{t}\in T_\infty$, be a random field. The n-th set of observations involves the values of the random field $\xi_{\mathbf{t}}$ taken at each point $\mathbf{k}\in\mathcal{D}_n$. Actually, each $\mathbf{k} = \mathbf{k}^{(n)}\in\mathcal{D}_n$ depends on n but to avoid complicated notation we often omit superscript (n). By our assumptions, $\lim_{n\to\infty}|\mathcal{D}_n| = \infty$.

We need the notion of α-mixing (see e.g. Doukhan [10]). Let $\mathcal{A}$ and $\mathcal{B}$ be two σ-algebras in $\mathcal{F}$. The α-mixing coefficient of $\mathcal{A}$ and $\mathcal{B}$ is defined as follows.

$$\alpha(\mathcal{A},\mathcal{B}) = \sup\{|\mathbb{P}(A)\mathbb{P}(B) - \mathbb{P}(AB)| \ : \ A \in \mathcal{A}, B \in \mathcal{B}\}.$$

The α-mixing coefficient of $\{\xi_{\mathbf{t}} \ : \ \mathbf{t} \in T_\infty\}$ is

$$\alpha(r) = \sup\{\alpha(\mathcal{F}_{I_1}, \mathcal{F}_{I_2}) \ : \ \varrho(I_1, I_2) \geqslant r\}$$

where I_i is a finite subset in T_∞ and $\mathcal{F}_{I_i} = \sigma\{\xi_{\mathbf{t}} \ : \ \mathbf{t} \in I_i\}$, $i = 1, 2$. We shall use the following condition. For some $1 < a < \infty$

$$\int_0^\infty s^{2d-1} \alpha^{\frac{a-1}{a}}(s)ds < \infty \,. \tag{2.2}$$

First we turn to the version of the central limit theorem appropriate to our sampling scheme. Our Theorem 2.1 is a modification of Theorem 1.1 of Bosq et al. [5]. The novelties of Theorem 2.1 are the infill-increasing setting and that it concerns random fields.

Define the discrete parameter (vector valued) random field $Y_n(\mathbf{k})$ as follows. For each $n = 1, 2, \ldots,$ and for each $\mathbf{k} \in \mathcal{D}_n$

$$\text{let } \ Y_n(\mathbf{k}) = Y_n(\mathbf{k}^{(\mathbf{n})}) \ \text{ be a Borel measurable function of } \ \xi_{\mathbf{k}^{(n)}}. \tag{2.3}$$

We concentrate on the case when $\xi_{\mathbf{t}}$ and $\xi_{\mathbf{s}}$ are dependent if $\mathbf{t}$ and $\mathbf{s}$ are close to each other. Therefore our theorem does not cover the case when $Y_n(\mathbf{k})$'s are independent and identically distributed. On the other hand, if $\xi_{\mathbf{t}}$ is a stationary field with continuous covariance function and positive variance, then the covariance is close to a fixed positive number inside a small hyperrectangle. We intend to cover this case.

Theorem 2.1. *Let $\xi_{\mathbf{t}}$ be a random field and let $Y_n(\mathbf{k}) = (Y_n^{(1)}(\mathbf{k}), \ldots, Y_n^{(m)}(\mathbf{k}))$ be an m-dimensional random field defined by (2.3). Let $S_n = \sum_{\mathbf{k}\in\mathcal{D}_n} Y_n(\mathbf{k})$, $n = 1, 2, \ldots$. Suppose that for each fixed n the field $Y_n(\mathbf{k})$, $\mathbf{k} \in \mathcal{D}_n$, is strictly stationary with $\mathbb{E}Y_n(\mathbf{k}) = 0$. Assume that $\|Y_n(\mathbf{k})\| \leqslant M_n$ where M_n depends only on n; $\sup_{n,\mathbf{k},r} \mathbb{E}\big(Y_n^{(r)}(\mathbf{k})\big)^2 < \infty$; for any increasing, unbounded sequence of rectangles G_n with $G_n \subseteq T_n$*

$$\lim_{n\to\infty} \frac{1}{\Lambda_n^d |\gg_n|} \mathbb{E}\left[\sum_{\mathbf{k}\in\gg_n} Y_n^{(r)}(\mathbf{k}) \sum_{\mathbf{l}\in\gg_n} Y_n^{(s)}(\mathbf{l})\right] = \sigma_{r,s}, \quad r, s = 1, \ldots, m, \tag{2.4}$$

where $\gg_n = G_n \cap (\mathbb{Z}/\Lambda_n)^d$; the matrix $\Sigma = (\sigma_{r,s})_{r,s=1}^m$ is positive definite; there exists $1 < a < \infty$ such that (2.2) is satisfied; and

$$M_n \leqslant c|T_n|^{\frac{a^2}{(3a-1)(2a-1)}} \quad \textit{for each} \quad n. \tag{2.5}$$

Then

$$S_n / \sqrt{\Lambda_n^d |\mathcal{D}_n|} \Rightarrow \mathcal{N}(0, \Sigma), \qquad \textit{as} \ \ n \to \infty. \tag{2.6}$$

The proof is based on a version of Bernstein's method applied in [5]. For details see [11].

3. A CLT for kernel-type density estimators

Now assume that the random field $\xi_{\mathbf{t}}$, $\mathbf{t} \in T_\infty$, is strictly stationary with unknown continuous marginal density function f. We shall estimate f from the data $\xi_{\mathbf{k}}$, $\mathbf{k} \in \mathcal{D}_n$.

A function $K : \mathbb{R} \to \mathbb{R}$ will be called a kernel if K is a bounded, continuous, symmetric density function (with respect to the Lebesgue measure),

$$\lim_{|u|\to\infty} |u|K(u) = 0, \qquad \int_{-\infty}^{+\infty} u^2 K(u)\, du < \infty\,. \tag{3.1}$$

Let K be a kernel and let $h_n > 0$, then the kernel-type density estimator is

$$f_n(x) = \frac{1}{|\mathcal{D}_n|}\frac{1}{h_n}\sum\nolimits_{\mathbf{i}\in\mathcal{D}_n} K\left(\frac{x-\xi_{\mathbf{i}}}{h_n}\right), \qquad x \in \mathbb{R}\,. \tag{3.2}$$

Let $f_{\mathbf{u}}(x,y)$ be the joint density function of $\xi_{\mathbf{0}}$ and $\xi_{\mathbf{u}}$, $\mathbf{u} \neq \mathbf{0}$. Denote $\mathbb{R}^d_{\mathbf{0}}$ the set $\mathbb{R}^d \setminus \{\mathbf{0}\}$. Let

$$g_{\mathbf{u}}(x,y) = f_{\mathbf{u}}(x,y) - f(x)f(y), \quad \mathbf{u} \in \mathbb{R}^d_{\mathbf{0}},\ x,y \in \mathbb{R}. \tag{3.3}$$

We assume that $f_{\mathbf{u}}(x,y)$ (and therefore $g_{\mathbf{u}}(x,y)$) is continuous in x and y for each fixed $\mathbf{u}$.

Denote by $C(\mathbb{R}^2)$ the space of continuous real-valued functions over $\mathbb{R}^2$. Consider the function $\mathbf{u} \to g_{\mathbf{u}}(.,.)$, $\mathbf{u} \in \mathbb{R}^d_{\mathbf{0}}$. This is an $\mathbb{R}^d_{\mathbf{0}} \to C(\mathbb{R}^2)$ mapping. For the sake of brevity we denote this function by $g_{\mathbf{u}}$ (and we consider it as the above $\mathbb{R}^d_{\mathbf{0}} \to C(\mathbb{R}^2)$ mapping). Let $\|g_{\mathbf{u}}\| = \sup_{(x,y)\in\mathbb{R}^2} |g_{\mathbf{u}}(x,y)|$. It is the norm of $g_{\mathbf{u}}(.,.)$ in $C(\mathbb{R}^2)$. For the sake of brevity the function $\mathbf{u} \to \|g_{\mathbf{u}}\|$ (which is an $\mathbb{R}^d_{\mathbf{0}} \to \mathbb{R}$ mapping) will also be denoted by $\|g_{\mathbf{u}}\|$.

Introduce the notation

$$\sigma(x,y) = \int_{\mathbb{R}^d_{\mathbf{0}}} g_{\mathbf{u}}(x,y)\, d\mathbf{u}\,, \quad x,y \in \mathbb{R}. \tag{3.4}$$

For a fixed positive integer m and fixed distinct real numbers $x_1, \dots, x_m$ let

$$\Sigma^{(m)} = \big(\sigma(x_i,x_j)\big)_{1\leqslant i,j\leqslant m}\,. \tag{3.5}$$

Remark 3.1. In Theorem 3.2 we need approximations of the integral $\int_{\mathbb{R}^d_{\mathbf{0}}} \|g_{\mathbf{u}}\|\, d\mathbf{u}$ with Riemannian sums. This procedure is usually applied only for bounded functions defined in bounded closed domains. Therefore we turn to the idea of direct Riemann integrability. The notion of direct Riemann integrability is well-known for univariate functions (see e.g. [1], p. 118). That is somewhat stronger than Lebesgue integrability. As we did not find appropriate references for multivariate functions, below we describe the notion of direct Riemann integrability for nonnegative functions defined on $\mathbb{R}^d_{\mathbf{0}}$ and being unbounded at the origin.

Let $l : \mathbb{R}^d_{\mathbf{0}} \to [0,\infty)$ be given. For an $h > 0$ consider a subdivision of $\mathbb{R}^d$ into (right closed and left open) d-dimensional cubes $\Delta_{\mathbf{i}}$ with edge length h such that the center of $\Delta_{\mathbf{0}}$ is the origin $\mathbf{0} \in \mathbb{R}^d$. If $\mathbf{i} \neq \mathbf{0}$, for $\mathbf{x} \in \Delta_{\mathbf{i}}$ let $\overline{l}_h(\mathbf{x}) = \sup\{l(\mathbf{y}) : \mathbf{y} \in \Delta_{\mathbf{i}}\}$, $\underline{l}_h(\mathbf{x}) = \inf\{l(\mathbf{y}) : \mathbf{y} \in \Delta_{\mathbf{i}}\}$, while $\overline{l}_h(\mathbf{x}) = \underline{l}_h(\mathbf{x}) = 0$ if $\mathbf{x} \in \Delta_{\mathbf{0}}$. If

$$\lim_{h\to 0} \int_{\mathbb{R}^d} \overline{l}_h(\mathbf{x})\, d\mathbf{x} = \lim_{h\to 0} \int_{\mathbb{R}^d} \underline{l}_h(\mathbf{x})\, d\mathbf{x} = I$$

and this common value is finite, then l is called directly Riemann integrable and I is its direct Riemann integral.

Theorem 3.2. *Assume that $g_{\mathbf{u}}$ is Riemann integrable on each bounded closed d-dimensional rectangle $R \subset \mathbb{R}^d_{\mathbf{0}}$, moreover $\|g_{\mathbf{u}}\|$ is directly Riemann integrable. Let $x_1, \dots, x_m$ be given distinct real numbers and assume that $\Sigma^{(m)}$ in (3.5) is positive definite. Suppose that there exists $1 < a < \infty$ such that (2.2) is satisfied and*

$$(h_n)^{-1} \leqslant c|T_n|^{\frac{a^2}{(3a-1)(2a-1)}} \quad \text{for each} \quad n\,. \tag{3.6}$$

Assume that $\lim_{n\to\infty} \Lambda_n = \infty$ *and* $\lim_{n\to\infty} h_n = 0$. *If*

$$\lim_{n\to\infty} 1/(\Lambda_n^d h_n) = L < \infty, \tag{3.7}$$

then

$$\sqrt{|\mathcal{D}_n|/\Lambda_n^d}\left\{\left(f_n(x_i) - \mathbb{E}f_n(x_i)\right), i = 1, \dots, m\right\} \Rightarrow \mathcal{N}(0, \Sigma^{'(m)}), \quad \text{as} \quad n \to \infty, \tag{3.8}$$

where

$$\Sigma^{'(m)} = \Sigma^{(m)} + D\,, \tag{3.9}$$

and D is a diagonal matrix with diagonal elements $Lf(x_i)\int_{-\infty}^{+\infty} K^2(u)\,du$, $i = 1, \dots, m$.

If, moreover, $f(x)$ has bounded second derivative and $\lim_{n\to\infty} |T_n| h_n^4 = 0$, *then in (3.8) $\mathbb{E}f_n(x_i)$ can be changed for $f(x_i)$, $i = 1, \dots, m$, and the above statement remains valid, i.e.*

$$\sqrt{|\mathcal{D}_n|/\Lambda_n^d}\left\{\left(f_n(x_i) - f(x_i)\right), i = 1, \dots, m\right\} \Rightarrow \mathcal{N}(0, \Sigma^{'(m)}), \quad \text{as} \quad n \to \infty. \tag{3.10}$$

Proof. We have to check the conditions of Theorem 2.1. Let $x_1, \dots, x_m$ be fixed distinct real numbers and define the m-dimensional random vector $X_n(\mathbf{i})$ with the following coordinates:

$$X_n^{(r)}(\mathbf{i}) = \frac{1}{h_n} K\left(\frac{x_r - \xi_{\mathbf{i}}}{h_n}\right) - \frac{1}{h_n}\mathbb{E}K\left(\frac{x_r - \xi_{\mathbf{i}}}{h_n}\right), \tag{3.11}$$

for $r = 1, \dots, m$, and $\mathbf{i} \in \mathcal{D}_n$. Divide T_n into d-dimensional unit cubes (having Λ_n^d points from $\mathcal{D}_n$ in each of them). Denote by $\mathcal{D}'_n$ the set of these cubes. Let $Y_n(\mathbf{k})$ be the arithmetical mean of variables $X_n(\mathbf{i})$ having indices $\mathbf{i}$ in the $\mathbf{k}$-th cube. Then for each fixed n the field $Y_n(\mathbf{k})$, $\mathbf{k} \in \mathcal{D}'_n$, is strictly stationary with $\mathbb{E}Y_n(\mathbf{k}) = 0$. We shall apply Theorem 2.1 to $Y_n(\mathbf{k})$, $\mathbf{k} \in \mathcal{D}'_n$, i.e. we shall use a non infill form of that theorem. For details see [12]. □

Remark 3.3. (1) In [5] (for $d = 1$, i.e. for onedimensional parameter space) it is shown that in the continuous time case the asymptotic covariance matrix is $\Sigma^{(m)}$, while in the discrete time case it is a diagonal matrix with diagonal elements $f(x_i) \int_{-\infty}^{+\infty} K^2(u)\, du$, $i = 1, \dots, m$. Therefore in the infill-increasing case the asymptotic covariance matrix can be the same as in the continuous time case (if $L = 0$) or a special linear combination of the ones of the continuous time case and of the discrete time case.

(2) Formula (3.10) can be used to construct (asymptotic) confidence regions.

(3) For kernel type density estimators the magnitude of the mean square error (MSE) and the optimal sampling are widely studied (see e.g. Bosq [4], Biau [2]). Our aim is different, i.e. we prove asymptotic normality. However, our theorem fits well to the results of the above mentioned papers. Their typical result (for onedimensional state space and in our notation) is the following. Under suitable mixing and analytical conditions

$$\mathbb{E}\left(f_n(x) - f(x)\right)^2 = \mathrm{O}\left(|T_n|^{-1}\right), \tag{3.12}$$

if $\Lambda_n = |T_n|^{1/2dr}$, $h_n = |T_n|^{-1/2dr}$ (see Proposition 7.1 in [4] for processes and $r = 2$, and see Theorem 5.1 in [2] for fields and $r > 0$). So the rates in (3.8) and in (3.12) are the same (besause $|T_n| = |\mathcal{D}_n|/\Lambda_n^d$). We mention that for the above Λ_n and h_n we have $\Lambda_n h_n = 1$, so (3.7) is satisfied.

Remark 3.4. Our assumptions on $g_{\mathbf{u}}$ seem to be quite strong. However, we mention that in Bosq et al. [5], Theorem 2.2 the (Lebesgue) integrability of $\|g_{\mathbf{u}}\|$ is assumed, which (assuming strong measurability of $g_{\mathbf{u}}$) is equivalent to the Bochner integrability of $g_{\mathbf{u}}$.

Now we present a simple example that gives numerical evidence for the phenomenon described in Theorem 3.2.

Let $\xi_{\mathbf{u}}$, $\mathbf{u} \in \mathbb{R}^d$, be a stationary Gaussian random field with mean value function zero and with covariance function $r_{\mathbf{u}}$. Assume that $r_{\mathbf{u}}$ is continuous and $r_{\mathbf{0}} = 1$. In [12] it is proved that the improper Riemann integral $\int_{\mathbb{R}^d_{\mathbf{0}}} \|g_{\mathbf{u}}\|\, d\mathbf{u}$ exists and is finite if the following three conditions are satisfied.

$$\int_O \frac{1}{\sqrt{1-r_{\mathbf{u}}^2}}\, d\mathbf{u} < \infty, \int_N \left[\frac{1}{\sqrt{1-r_{\mathbf{u}}^2}} - 1\right] d\mathbf{u} < \infty, \int_N \left|\frac{-r_{\mathbf{u}}^2 \pm r_{\mathbf{u}}}{1-r_{\mathbf{u}}^2}\right| d\mathbf{u} < \infty \tag{3.13}$$

for a (closed, bounded) domain O containing a neigbourhood of the origin and for N being the complement of a bounded neigbourhood of the origin.

Example 3.5. Consider the Gaussian process ξ_u, $u \in \mathbb{R}$, with mean zero and covariance function $r_u = e^{-|u|}$, $u \in \mathbb{R}$. This function satisfies the above mentioned conditions (3.13). The direct Riemann integrability of $\|g_{\mathbf{u}}\|$ is also satisfied. The mixing condition (2.2) is also fulfilled, see Rozanov [21], Ch. IV, Sect. 11. (As the parameter space in this example is onedimensional we can apply the notion of the α-mixing coefficient and the CLT given in [5].) Using this model, we obtained simulation evidence for Theorem 3.2.

We observe this process in the $1/\Lambda$-lattice points of the domain $T = [0, t]$ with $\Lambda = 200$ and $t = 100$. That is the sample is $z_1 = \xi(1/200), \dots, z_s = \xi(20000/200)$ with $s = 20000$. Now the covariance matrix of this data vector is $(r^{|i-j|})_{i,j=1}^{s}$, where $r = e^{-1/\Lambda}$. Therefore the data generation for the simulation is easy. Let $y_1, \dots, y_s$ be i.i.d. standard normal and choose $z_i = r^{i-1}y_1 + \sqrt{1-r^2}\sum_{j=2}^{i} r^{i-j}y_j$, $i = 1, \dots, s$.

Using this data, we gave kernel estimation for the density function of the process (i.e. the standard normal density function). We calculated the estimator at points $x_1 = -2$, $x_2 = -1$, $x_3 = 0$, $x_4 = 1$, $x_5 = 2$. We used values of the bandwidth: $h_1 = 0.01$ and $h_2 = 0.001$. We applied the standard normal density function as kernel K.

The simulations were performed with MATLAB. 500 repetitions were made. The data sets for $h_1 = 0.01$ and $h_2 = 0.001$ were the same. The theoretical values of the density function and the averages of their estimators are shown in Table 1.

	x	-2	-1	0	1	2
	$f(x)$	0.0540	0.2420	0.3989	0.2420	0.0540
$h = 0.01$	estimators' mean	0.0543	0.2418	0.4003	0.2414	0.0518
$h = 0.001$	estimators' mean	0.0528	0.2422	0.4026	0.2445	0.0492

Table 1

Then we calculated the empirical covariance matrices of our standardized estimators (according to equation (3.8), the standardization factor is $\sqrt{|\mathcal{D}|/\Lambda} = 10$).

$$\Sigma_1 = \begin{bmatrix} +0.0683 & +0.0320 & -0.0308 & -0.0444 & -0.0135 \\ +0.0320 & +0.2307 & -0.0258 & -0.1356 & -0.0453 \\ -0.0308 & -0.0258 & +0.2442 & -0.0355 & -0.0284 \\ -0.0444 & -0.1356 & -0.0355 & +0.2421 & +0.0388 \\ -0.0135 & -0.0453 & -0.0284 & +0.0388 & +0.0598 \end{bmatrix};$$

$$\Sigma_2 = \begin{bmatrix} +0.1339 & +0.0379 & -0.0296 & -0.0451 & -0.0162 \\ +0.0379 & +0.5481 & -0.0085 & -0.1544 & -0.0482 \\ -0.0296 & -0.0085 & +0.7092 & -0.0416 & -0.0205 \\ -0.0451 & -0.1544 & -0.0416 & +0.6002 & +0.0459 \\ -0.0162 & -0.0482 & -0.0205 & +0.0459 & +0.1116 \end{bmatrix}.$$

Covariance Σ_1 corresponds to bandwidth h_1 while Σ_2 corresponds to bandwidth h_2. The difference of the diagonals of Σ_2 and Σ_1 seems to be significant.

Now calculate the additional terms of the covariance matrices described in Theorem 3.2. In our case

$$\frac{1}{\Lambda}\frac{1}{h}f(x_i)\int_{-\infty}^{+\infty} K^2(u)\,du = \frac{1}{200}\frac{1}{h}f(x_i)\,\frac{1}{2\sqrt{\pi}}.$$

Therefore the diagonal elements of the matrix D in Theorem 3.2 for $h_1 = 0.01$ and $h_2 = 0.001$ are the following:

$$\begin{aligned} \mathrm{diag}D_1 &= \left[\begin{array}{ccccc} 0.0076 & 0.0341 & 0.0563 & 0.0341 & 0.0076 \end{array}\right]; \\ \mathrm{diag}D_2 &= \left[\begin{array}{ccccc} 0.0762 & 0.3413 & 0.5627 & 0.3413 & 0.0762 \end{array}\right]. \end{aligned}$$

As in the infill-increasing case only the diagonals of the limit covariance matrices can be different for different values of the bandwidth, we show in Table 2 the diagonal of the differences of the empirical covariance matrices and that of the theoretical covariance matrices.

$\mathrm{diag}(D_2 - D_1)$	0.0686	0.3072	0.5064	0.3072	0.0686
$\mathrm{diag}(\Sigma_2 - \Sigma_1)$	0.0656	0.3174	0.4649	0.3582	0.0517

Table 2

The results show that the diagonal matrix D of Theorem 3.2 explains well the dependence of the limit covariance matrix on the bandwidth.

4. Functional CLT for kernel type density estimator

We shall present a functional central limit theorem in the space $L_2[0,1]$. In this section we suppose that both f and f_n are equal to 0 outside of the interval $[0,1]$. If we restrict our study to densities and kernel functions with compact supports, by appropriate transformation, this condition can be realized.

Theorem 4.1. *Assume that $g_{\mathbf{u}}$ is Riemann integrable on each bounded closed d-dimensional rectangle $R \subset \mathbb{R}^d_{\mathbf{0}}$, moreover $\|g_{\mathbf{u}}\|$ is directly Riemann integrable. Let the function $\sigma(x,y)$ defined in (3.4) be positive definite. Suppose that there exists $1 < a < \infty$ such that (2.2) and (3.6) are satisfied. Assume that* $\lim_{n\to\infty} \Lambda_n = \infty$, $\lim_{n\to\infty} h_n = 0$, *and* $\lim_{n\to\infty} 1/(\Lambda_n^d h_n) = 0$. *Assume that $f(x)$ has bounded second derivative and* $\lim_{n\to\infty} |T_n| h_n^4 = 0$. *Then, as $n \to \infty$,*

$$L_n(x) = \sqrt{|\mathcal{D}_n|/\Lambda_n^d}\,[f_n(x) - f(x)] \Rightarrow G(x) \tag{4.1}$$

in $L_2[0,1]$ *where G is a Gaussian process with mean* 0 *and with covariance function* $\sigma(\,.\,,\,.\,)$.

To prove our theorem we have to check the conditions given in [14] and [16]. The convergence of the finite dimensional distributions is a consequence of Theorem 2.1. For the details see [13].

References

[1] Asmussen, S., Applied Probability and Queues, *Wiley,* New York, (1987).

[2] Biau, G., Spatial kernel density estimation, *Math. Methods Statist.,* 12 (2003), no. 4, (2004), 371–390.

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

[4] Bosq, D., Parametric rates of nonparametric estimators and predictors for continuous time processes, *Ann. Statist.,* 25, no. 3, (1997), 982–1000.

[5] Bosq, D., Merlevède, F., Peligrad, M., Asymptotic normality for density kernel estimators in discrete and continuous time, *J. Multivariate Analysis,* 68, (1999), 78–95.

[6] Castellana, J. V., Leadbetter, M. R., On smoothed probability density estimation for stationary processes, *Stochastic Process. Appl.,* 21, no. 2, (1986), 179–193.

[7] Cressie, N. A. C., Statistics for Spatial Data, *Wiley,* New York, (1991).

[8] Csörgő, S., Mielniczuk, J., Density estimation under long-range dependence. *Ann. Statist.,* 23, (1995), 990–999.

[9] Devroye, L., Györfi, L., Nonparametric Density Estimation, The L_1 View, *Wiley,* New York, (1985).

[10] Doukhan, P., Mixing. Properties and Examples, *Lecture Notes in Statistics,* 85, Springer, New York, (1994).

[11] Fazekas, I., Chuprunov, A., A central limit theorem for random fields, *Acta Mathematica Academiae Paedagogicae Nyiregyhaziensis*, 20, no. 1, (2004), 93–104, http://www.emis.de/journals/AMAPN.

[12] Fazekas, I., Chuprunov, A., Asymptotic normality of kernel type density estimators for random fields, *Stat. Inf. Stoch. Proc.,* 9, (2006), 161–178.

[13] Fazekas, I., Filzmoser, P., A functional central limit theorem for kernel type density estimators, *Austrian J. Statist.,* 35, no. 4, (2006), 409–418.

[14] Grinblat, L. S., A limit theorem for measurable random processes and its applications, *Proc. Amer. Math. Soc.,* 61, (1976), 371–376.

[15] Guillou, A., Merlevède, F., Estimation of the asymptotic variance of kernel density estimators for continuous time processes, *J. Multivariate Analysis*, 79, (2001), 114–137.

[16] Ivanov, A. V., On convergence of distributions of functionals of measurable random fields, *Ukrainean J. Mathematics*, 32, no. 1, (1980), 27–34.

[17] Lahiri, S. N., Kaiser, M. S., Cressie, N., Hsu, N-J., Prediction of spatial cumulative distribution functions using subsampling, *J. Amer. Statist. Assoc.*, 94, no. 445, (1999), 86–110.

[18] Masry, E., Probability density estimation from sampled data, *IEEE Trans. Inform. Theory*, 29, no. 5, (1983), 696–709.

[19] Prakasa Rao, B. L. S., Nonparametric Functional Estimation, *Academic Press,* New York, (1983).

[20] Putter, H., Young, G. A., On the effect of covariance function estimation on the accuracy of kriging predictors, *Bernoulli*, 7, no. 3, (2001), 421–438.

[21] Rozanov, Yu. A., Stationary Random Processes, *Holden-Day,* San Francisco, (1967).

[22] Wu, W. B., Mielniczuk, J., Kernel density estimators for linear processes, *Ann. Statist.,* 30, (2002), 1441–1459.

István Fazekas
Faculty of Informatics,
University of Debrecen,
P.O. Box 12, 4010 Debrecen,
Phone (36 52) 512900-22825,
Hungary