Annales Mathematicae et Informaticae
***39** (2012) pp. 45–56*

Proceedings of the Conference on Stochastic Models and their Applications
Faculty of Informatics, University of Debrecen, Debrecen, Hungary, August 22–24, 2011

Joint asymptotic normality of the kernel type density estimator for spatial observations

István Fazekas[a*], **Zsolt Karácsony**[b†], **Renáta Vas**[a‡]

[a]Faculty of Informatics, University of Debrecen, Hungary
fazekasi@inf.unideb.hu, vas.renata@inf.unideb.hu

[b]Department of Applied Mathematics, University of Miskolc, Hungary
matkzs@uni-miskolc.hu

Dedicated to Mátyás Arató on his eightieth birthday

Abstract

The Central Limit Theorem is considered for m-dependent random fields. The random field is observed in a sequence of irregular domains. The sequence of domains is increasing and at the same time, the locations of the observations become more and more dense in the domains. The Central Limit Theorem is applied to obtain asymptotic normality of kernel type density estimators. It turns out that the covariance structure of the limiting normal distribution can be a combination of those of the continuous parametric and the discrete parametric results. Numerical evidence is presented.

Keywords: Asymptotic normality, central limit theorem, random field, kernel, infill-increasing setup

MSC: 60F05,62M30

*The work is supported by the Hungarian Scientific Research Fund under Grant No. OTKA T079128/2009.

†This author was supported by TAMOP-4.2.1.B-10/2/KONV-2010-0001 project with support by the European Union, co-financed by the European Social Fund.

‡The work is supported by the TÁMOP-4.2.2/B-10/1-2010-0024 project. The project is co-financed by the European Union and the European Social Fund.

1. Introduction

Consider a domain D in $\mathbb{R}^d$. We observe a random field $\xi(\cdot)$ in certain points of the domain D and we assume the following setup. Suppose that the random field $\xi(\cdot)$ is observed at finitely many locations i.e. at the elements $\mathbf{s}_{n1}, \dots, \mathbf{s}_{nn}$ lying in the sampling region $D_n \subset D$. Let $\mathcal{R}_n = \{\mathbf{s}_{n1}, \dots, \mathbf{s}_{nn}\}$ denote the n-th set of the locations of the observations. We shall use the notion of the mixed (or nearly infill or infill-increasing) domain sampling which means that the sampling region D_n increases and at the same time, the data sites $\{\mathbf{s}_{n1}, \dots, \mathbf{s}_{nn}\}$ fill in any given sub-region of D_n increasingly densely as $n \to \infty$. (Increasing domains means that $D_n \subseteq D_{n+1}$ and the size of D_n goes to infinity as $n \to \infty$.) This approach was studied e.g. by Lahiri [4], Lahiri, Kaiser, Cressie and Hsu [5], Fazekas and Chuprunov [2], Park, Kim, Park and Hwang [6] and Karácsony and Filzmoser [3]. It can be useful in geostatistics, environmental sciences etc.

To obtain asymptotic normality, we assume that the n-th set of observations is $\xi_n(\mathbf{s}_{n1}), \dots, \xi_n(\mathbf{s}_{nn})$, where $\xi_n(\cdot), n = 1, 2, \dots$ is a sequence of stationary random fields and $\xi_n(\cdot)$ is weakly dependent for any fixed n. For the sake of simplicity we suppose that $\xi_n(\cdot)$ is m-dependent. It is a restriction but it has an advantage namely that we can easily obtain a central limit theorem (CLT) for irregular domains. We mention that similar results can be obtained for mixing random fields as well (see e.g. Fazekas and Chuprunov [1], but there the domain is regular and the conditions are quite difficult to check). The main objective of Park, Kim, Park and Hwang [6] is to provide central limit theorems that could be applied easily in practice. In our paper we discuss some consequences of the results of Park, Kim, Park and Hwang [6].

The article is organized as follows. In Section 2, we introduce our notations and we recall the CLT for stationary random fields of Park, Kim, Park and Hwang [6]. In Section 3, we turn to the density estimator, we quote Theorem 3 of Park, Kim, Park and Hwang [6]. It states that under mild conditions the kernel type density estimator is asymptotic normal. In Section 4, we deal with the multidimensional extension of this theorem. Simulation evidence is presented here, too. The numerical examples show the unusual covariance structure of the limiting normal distribution. This covariance structure was first presented in Fazekas and Chuprunov [2]. That is, the asymptotic covariance of the kernel type density estimator for nearly infill sampling can be a combination of the covariances of the discrete and the continuous parameter models. Similar result is valid for the regression estimator (see Karácsony and Filzmoser [3]).

2. CLT for stationary random fields

Let us consider a zero mean strictly stationary random field $\{\xi(\mathbf{s}) : \mathbf{s} \in D\}$, $D \subseteq \mathbb{R}^d$. Here, the strict stationarity of the random field means that for any $\mathbf{s}_1, \dots, \mathbf{s}_k, \mathbf{t}$, the distribution of $(\xi(\mathbf{s}_1), \dots, \xi(\mathbf{s}_k))$ is the same as that of $(\xi(\mathbf{s}_1 + \mathbf{t}), \dots, \xi(\mathbf{s}_k + \mathbf{t}))$.

We assume that the random field $\xi(\cdot)$ is m-dependent. m-dependence means that m is the infimum of the numbers denoted by b such that if $\|\mathbf{s}_1 - \mathbf{s}_2\| > b$, then $\xi(\mathbf{s}_1)$ and $\xi(\mathbf{s}_2)$ are independent. Here, $\|\cdot\|$ denotes the Euclidean norm in $\mathbb{R}^d$.

For $\mathbf{u} \in \mathcal{R}_n$, let

$$I_{m,n}(\mathbf{u}) = \{\mathbf{s} \in \mathcal{R}_n : \|\mathbf{s} - \mathbf{u}\| \leq m\}$$

and $\kappa_n = \max_{\mathbf{u}\in\mathcal{R}_n} \sharp\{I_{m,n}(\mathbf{u})\}$. So κ_n denotes the number of elements of the set $I_{m,n}(\mathbf{u})$ with maximal cardinality. Therefore κ_n is an indicator of the strength of dependence. To avoid the independent case, we assume that $\kappa_n > 0$ for each n. We suppose that the measure κ_n of density of locations satisfies

$$\kappa_n \sim n^a \quad \text{with a constant } 0 < a < 1. \tag{2.1}$$

Here for any two sequences $\{t_n\}$ and $\{v_n\}$ of positive numbers, the notation $t_n \sim v_n$ means that the relation

$$0 < c_1 \leq \liminf_{n\to\infty}(t_n/v_n) \leq \limsup_{n\to\infty}(t_n/v_n) \leq c_2 < \infty$$

holds for positive constants c_1 and c_2.

For real valued sequences $\{a_n\}$ and $\{b_n\}$, the notation $a_n = \mathrm{o}(b_n)$ (resp. $a_n = \mathrm{O}(b_n)$) means that the sequence a_n/b_n converges to 0 (resp. is bounded). The sign $\mathbb{E}$ stands for expectation. Variance and covariance are denoted by var(.) and cov(.,.), respectively. The sign "$\Rightarrow$" denotes convergence in distribution. $\mathcal{N}(m,\Sigma)$ stands for the (vector) normal distribution with mean (vector) m and covariance (matrix) Σ.

First, recall the CLT for m-dependent random fields presented in Park, Kim, Park and Hwang [6].

Consider a series of strictly stationary m-dependent random fields $\{\xi_n(\mathbf{s}) : \mathbf{s} \in D\}$, $D \subseteq \mathbb{R}^d$, $n = 1, 2, \ldots$. For a fixed n, let us introduce the notation $S_n = \sum_{i=1}^n \xi_n(\mathbf{s}_{ni})$. Furthermore, let $\mathcal{T}_n = \{(i,j) : 0 < \|\mathbf{s}_{ni} - \mathbf{s}_{nj}\| \leq m\}$, $\nu_n = \mathrm{var}(\xi_n(\mathbf{s}))$ and

$$\tau_n = \frac{1}{n\kappa_n} \sum_{(i,j)\in\mathcal{T}_n} \mathrm{cov}(\xi_n(\mathbf{s}_{ni}), \xi_n(\mathbf{s}_{nj})). \tag{2.2}$$

At this point we notice that $\mathrm{var}(S_n) = n\nu_n + n\kappa_n\tau_n$ and τ_n can be negative as well.

Theorem 2.1 (Theorem 2 of Park, Kim, Park and Hwang [6]). *Let $\{\xi_n\}$ be a sequence of strictly stationary random fields on $D \subset \mathbb{R}^d$ with $\mathbb{E}\xi_n(\mathbf{s}) = 0$. Assume that $\sup_{\mathbf{s}\in D} |\xi_n(\mathbf{s})|$ is bounded with probability one and $\mathbb{E}\left|\prod_{j=1}^l \xi_n(\mathbf{s}'_{nj})\right| = \mathrm{O}\left(\nu_n^l\right)$ holds uniformly for all the different points $\mathbf{s}'_{nj} \in \{\mathbf{s}_{n1}, \ldots, \mathbf{s}_{nn}\}$. If $\nu_n + \kappa_n\tau_n \geq \delta\kappa_n\nu_n^2$ for some $\delta > 0$, then we have*

$$\frac{S_n}{\sqrt{\mathrm{var}(S_n)}} \Rightarrow \mathcal{N}(0,1).$$

3. Application to density estimation

In Park, Kim, Park and Hwang [6], the CLT was applied to obtain asymptotic normality of the kernel type density estimator.

Let $\{Z(\mathbf{s}) : \mathbf{s} \in D\}$ be a strictly stationary m-dependent random field, $D \subseteq \mathbb{R}^d$. For each $z \in \mathbb{R}$, let $F(z) = P(Z(\mathbf{s}) \le z)$. We call the function F marginal distribution function. Assume that there exist the appropriate marginal density function f. Suppose that we observe the values of Z at the points $\mathbf{s}_{n1}, \dots, \mathbf{s}_{nn}$ in D. In this section we study the nonparametric estimation of the marginal density function. Consider the kernel type density estimator

$$\hat{f}_n(z) = \frac{1}{nh_n} \sum_{i=1}^{n} K\left(\frac{z - Z(\mathbf{s}_{ni})}{h_n}\right).$$

Here K is a kernel. We say that the function $K : \mathbb{R} \to [0, \infty)$ is a kernel if it is a bounded, continuous, symmetric density function (with respect to the Lebesgue measure) and

$$\lim_{|u| \to \infty} |u| K(u) = 0. \tag{3.1}$$

Let $f_{s_{ni}, s_{nj}}$ be the joint density function of $Z(\mathbf{s}_{ni})$ and $Z(\mathbf{s}_{nj})$. Let $z \in \mathbb{R}$ be fixed. Consider the following assumptions.

(1) (a) $f(z) > 0$,

(b) f is continuous at z,

(c) $f_{s_{ni}, s_{nj}}$ are equicontinuous at (z, z), i.e. if $(z_1, z_2) \to (z, z)$, then

$$\sup_{i,j} |f_{s_{ni}, s_{nj}}(z_1, z_2) - f_{s_{ni}, s_{nj}}(z, z)| \to 0,$$

(d) all finite dimensional densities of $Z(\mathbf{s}_{n1}), Z(\mathbf{s}_{n2}), \dots$ exist and are bounded and continuous,

(e) if $n \to \infty$, then

$$\frac{1}{n\kappa_n} \sum_{(i,j) \in \mathcal{T}_n} \{f_{s_{ni}, s_{nj}}(z, z) - f(z)^2\} \to \tau,$$

where τ is a nonnegative constant depending on z,

(f) $h^2 n^a$, $0 < a < 1$ is bounded.

(2) The kernel K is bounded, nonnegative on $\mathbb{R}$ and satisfies $\int_{\mathbb{R}} K = 1$; $|z|K(z) \to 0$ as $|z| \to \infty$.

(3) $h_n > 0$ is a sequence satisfying $h_n \to 0$ and $nh_n \to \infty$ as $n \to \infty$.

(4) There exists a constant $\delta > 0$ such that

$$f(z)\int_{\mathbb{R}} K^2 + \tau\kappa_n h_n \geq \delta\kappa_n h_n.$$

Theorem 3.1 (Theorem 3 of Park, Kim, Park and Hwang [6]). *Let us suppose that the assumptions* (1)–(4) *hold.*

1. *Then*

$$\left\{n^{-1}h_n^{-1}f(z)\int_{\mathbb{R}} K^2 + n^{-1}\kappa_n\tau\right\}^{-\frac{1}{2}}\{\hat{f}_n(z) - \mathbb{E}\hat{f}_n(z)\} \Rightarrow \mathcal{N}(0,1).$$

2. *Suppose that f is twice differentiable in a neighbourhood of z and $\int uK(u)du = 0$. Moreover, assume that f'' is continuous, bounded and $nh_n^5 \to 0$, $n\kappa_n^{-1}h_n^4 \to 0$. Then*

$$\left\{n^{-1}h_n^{-1}f(z)\int_{\mathbb{R}} K^2 + n^{-1}\kappa_n\tau\right\}^{-\frac{1}{2}}\{\hat{f}_n(z) - f(z)\} \Rightarrow \mathcal{N}(0,1).$$

4. Joint asymptotic normality for the density estimator

In Park, Kim, Park and Hwang [6], the multivariate asymptotic normality was not considered.

Our aim is to study the multidimensional version of Theorem 3.1, i.e. the joint asymptotic normality of the kernel type density estimator.

Proposition 4.1. *Let $z_1, z_2, \ldots, z_q$ be given distinct real numbers. We assume that*

$$\frac{1}{n\kappa_n}\sum_{i,j\in\mathcal{T}_n}\left(f_{s_{ni},s_{nj}}(z_r,z_t) - f(z_r)f(z_t)\right) \to \tau_{rt} \text{ if } n \to \infty.$$

Let $W = \left(\frac{\tau_{ij}\kappa_n}{n}\right)_{1\leq i,j\leq q}$ and let V be a diagonal matrix with diagonal elements $\frac{1}{nh_n}f(z_i)\int_{-\infty}^{\infty}K^2(t)dt$, $i = 1,\ldots,q$. Let $\Sigma = V + W$.

Then under certain conditions, $(\hat{f}_n(z_i) - f(z_i), i = 1,\ldots,q)$ is asymptotically $\mathcal{N}(0,\Sigma)$. The structure of Σ is the following:

$$\Sigma = \frac{1}{nh_n}\begin{bmatrix} f(z_1)\int K^2(t)dt + \tau_{11}\kappa_n h_n & \tau_{12}\kappa_n h_n & \ldots & \tau_{1q}\kappa_n h_n \\ \tau_{21}\kappa_n h_n & f(z_2)\int K^2(t)dt + \tau_{22}\kappa_n h_n & \ldots & \tau_{2q}\kappa_n h_n \\ \vdots & & \ddots & \vdots \\ \tau_{q1}\kappa_n h_n & \ldots & \ldots & f(z_q)\int K^2(t)dt + \tau_{qq}\kappa_n h_n \end{bmatrix}.$$

To obtain this result one has to apply Theorem 2.1 and the Cramér-Wold device.

We can see that the asymptotic covariance matrix Σ has a special structure. In the diagonal, the expressions $f(z_i)\int K^2(t)dt$ come from the asymptotic covariance matrix of the discrete parameter model. On the other hand, the elements $\tau_{ij}\kappa_n h_n$ correspond to the asymptotic covariance matrix of the continuous parameter model. We mention that the asymptotic covariance matrices are well-known both for the discrete time and the continuous time models. The combination of the two covariance structures was first pointed out in Fazekas and Chuprunov [2] for the kernel type density estimator and then in Karácsony and Filzmoser [3] for the regression estimator. To underline the importance of the covariance structure, we mention the following. When calculating numerically the density estimator for a continuous time model, we approximate the estimator with a one corresponding to an infill-increasing model. However, the limiting covariance structures of those models can be distinct.

We present examples that give numerical evidence for the phenomena described in the above proposition. First we consider a one-dimensional regular domain D.

Example 1. Moving average on the real line.

We consider the process on the l-lattice points of the domain $D = [0, t]$ with $l = 0.1$ and $t = 200$. It means that the distance between two neighbours is $l = 0.1$.

That is, the sample is $z_1 = \xi(1/10), \dots, z_n = \xi(2000/10)$ with $n = 2000$. The data generation for the simulation is easy. Let $y_1, \dots, y_{n+4}$ be i.i.d. standard normal random variables and choose

$$z_i = 0.05 \cdot y_i + 0.2 \cdot y_{i+1} + 0.5 \cdot y_{i+2} + 0.2 \cdot y_{i+3} + 0.05 \cdot y_{i+4}, \ i = 1, \dots, n.$$

So $\xi(s)$ is a moving average process. We can see that the data is m-dependent with $m = 5$. The marginal density is $f(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-x^2}{2\sigma^2}\right)$ where $\sigma = 0.5788$.

Using these data, we calculated the estimation of the marginal density function of the random field at the points $x_1 = -1.0$, $x_2 = -0.5$, $x_3 = 0.0$, $x_4 = 0.5$ and $x_5 = 1.0$. We used two values of the bandwidth, $h_1 = 0.10$ and $h_2 = 0.01$, and applied the standard normal density function as kernel K.

The simulations were performed with MATLAB, 5000 repetitions of the procedure were made. The data sets for both bandwidths h_1 and h_2 were the same. The theoretical values of the density function and the average of their estimators are shown in Table 1. For both values of the bandwidths we can see a close similarity of the theoretical and the empirical values.

We calculated the empirical covariance matrices Σ_1 (corresponding to bandwidth h_1) and Σ_2 (corresponding to bandwidth h_2) for our estimators

$$(\widehat{f}_n(x_1), \dots, \widehat{f}_n(x_5)).$$

$$\Sigma_1 = \begin{bmatrix} 0.3078 & 0.0516 & -0.1107 & -0.1475 & -0.0624 \\ 0.0516 & 0.8053 & -0.1524 & -0.3343 & -0.1540 \\ -0.1107 & -0.1524 & 0.9289 & -0.1485 & -0.1221 \\ -0.1475 & -0.3343 & -0.1485 & 0.7853 & 0.0632 \\ -0.0624 & -0.1540 & -0.1221 & 0.0632 & 0.3195 \end{bmatrix} \cdot 10^{-3};$$

$$\Sigma_2 = \begin{bmatrix} 2.2605 & 0.0244 & -0.1598 & -0.0875 & -0.0631 \\ 0.0244 & 6.7115 & -0.1994 & -0.3860 & -0.1832 \\ -0.1598 & -0.1994 & 9.8334 & -0.1701 & -0.2003 \\ -0.0875 & -0.3860 & -0.1701 & 6.8598 & 0.0881 \\ -0.0631 & -0.1832 & -0.2003 & 0.0881 & 2.2602 \end{bmatrix} \cdot 10^{-3}.$$

The difference in the diagonals of Σ_1 and Σ_2 is clearly visible. The off-diagonal elements are almost the same.

x	-1.0	-0.5	0.0	0.5	1.0
$f(x)$	0.1549	0.4746	0.6892	0.4746	0.1549
$\overline{\hat{f}_n(x)}$ with $h_1 = 0.10$	0.1590	0.4726	0.6794	0.4728	0.1599
$\overline{\hat{f}_n(x)}$ with $h_2 = 0.01$	0.1543	0.4747	0.6876	0.4763	0.1564

Table 1: Theoretical values of the density function and the average of their estimators for the data of Example 1.

Now calculate the additional terms in the diagonals of the covariance matrices described by Σ defined in Proposition 4.1. In our case the elements of the diagonal matrix V_k for the bandwidth h_k $(k = 1, 2)$ are

$$\frac{1}{n}\frac{1}{h_k} f(x_i) \int_{-\infty}^{\infty} K^2(u)du = \frac{1}{2000}\frac{1}{h_k} f(x_i) \frac{1}{2\sqrt{\pi}}.$$

Since 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 ratio between the diagonals of the difference of the empirical covariance matrices, $diag(\Sigma_2 - \Sigma_1)$, and of the theoretical covariance matrices, $diag(V_2 - V_1)$.

x	-1.0	-0.5	0.0	0.5	1.0
$\frac{diag(\Sigma_2-\Sigma_1)}{diag(V_2-V_1)}$	0.9927	0.9803	1.0176	1.0082	0.9867

Table 2: Ratio between the diagonal of the difference of the empirical covariance matrices and that of the theoretical covariance matrices for the data of Example 1.

These are close to 1 as it is expected from the above proposition.

Finally, Figure 1 shows histograms of $\frac{1}{2}(\hat{f}_n(0.5) + \hat{f}_n(1.0))$ for the bandwidths $h_1 = 0.10$ (left picture) and $h_2 = 0.01$ (right picture). Figure 2 shows histograms of $\frac{1}{3}(\hat{f}_n(-1.0) + \hat{f}_n(-0.5) + \hat{f}_n(0.0))$ for the above bandwidths.

The histograms are presented together with the theoretical normal densities with means and variances estimated from the data used for the histograms. The approximate normality of the density estimator stated in the above proposition is reflected in these figures. Different bandwidths lead to different spreads of the normal distribution.

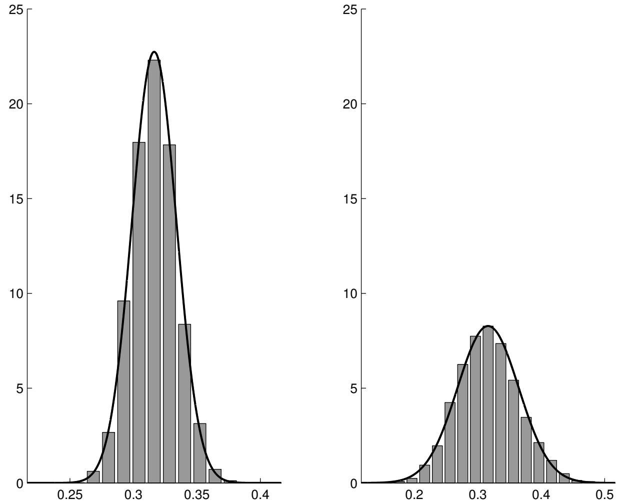

Figure 1: Histograms of $\frac{1}{2}(\hat{f}_n(0.5) + \hat{f}_n(1.0))$ for the bandwidths $h_1 = 0.10$ (left) and $h_2 = 0.01$ (right), together with the theoretical densities of the normal distribution for the data of Example 1.

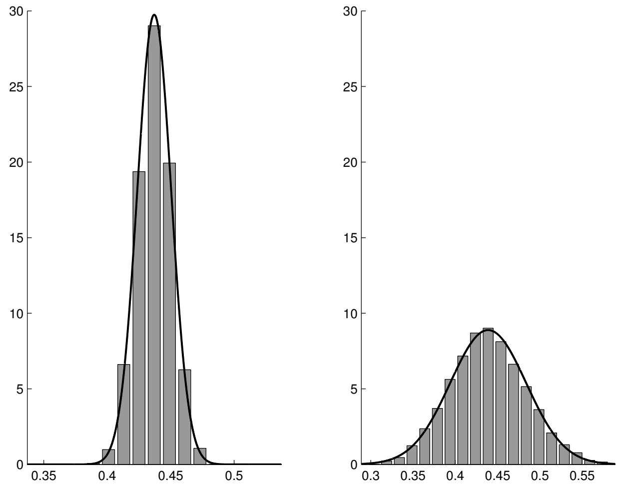

Figure 2: Histograms of $\frac{1}{3}(\hat{f}_n(-1.0) + \hat{f}_n(-0.5) + \hat{f}_n(0.0))$ for the bandwidths $h_1 = 0.10$ (left) and $h_2 = 0.01$ (right), together with the theoretical densities of the normal distribution for the data of Example 1.

Now we consider a two-dimensional domain with fractal-like shape.

Example 2. Two-dimensional moving average.

Now the locations will be the l-lattice points of the domain $D = [0,t]^2$ with $l = 0.1$ and $t = 10$. Thus the random field is $z_{(i,j)} = \xi_{(i/10,j/10)}, i,j = 1,\ldots,100$. Let $y_{k,l}$, $k,l = 1,\ldots,102$, be i.i.d. standard normal random variables, and let

$$z_{(i,j)} = \frac{1}{9}\sum_{k=i}^{i+2}\sum_{l=j}^{j+2} y_{k,l}, \quad i,j = 1,\ldots,100.$$

Therefore the random field is m-dependent with $m = 3$. The marginal density is $f(x) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(\frac{-x^2}{2\sigma^2}\right)$ where $\sigma = 0.3333$.

Some points from the locations were omitted. In Figure 3, the small squares where the locations were deleted are marked with dark. We can see that in each white small square we have 16 sites of observations. Denote the set of the remaining locations by D. So the observations are $z_{(i,j)}$, $i,j \in D$. Therefore the actual sample size is 7056.

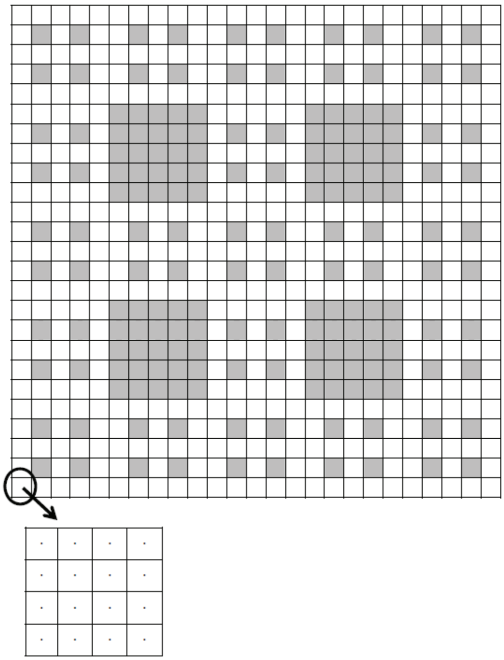

Figure 3: Sampling sites

It can be seen that the resulted domain is not convex. In the above proposition the asymptotic properties of the estimator remain true. It is clearly shown by the following numerical results.

As in the previous example, we calculated the density estimator $\hat{f}_n$ at the points $x_1 = -1.0$, $x_2 = -0.5$, $x_3 = 0.0$, $x_4 = 0.5$, $x_5 = 1.0$. We used the bandwidths

$h_1 = 0.10$ and $h_2 = 0.01$ and applied the standard normal density function as kernel K. The data sets for both bandwidths were the same, and 5000 repetitions were performed. Table 3 shows that the theoretical values of the density function and the average of their estimators are very similar.

x	-1.0	-0.5	0.0	0.5	1.0
$f(x)$	0.3886	0.9034	1.1968	0.9034	0.3886
$\overline{\hat{f}_n(x)}$ with $h = 0.10$	0.4087	0.8852	1.1460	0.8858	0.4085
$\overline{\hat{f}_n(x)}$ with $h = 0.01$	0.3907	0.9032	1.1965	0.9029	0.3895

Table 3: Theoretical values of the density function and the average of their estimators for the data of Example 2.

The empirical covariance matrices are

$$\Sigma_1 = \begin{bmatrix} 0.5124 & 0.3246 & -0.1801 & -0.4534 & -0.2921 \\ 0.3246 & 0.7406 & 0.0403 & -0.5479 & -0.4382 \\ -0.1801 & 0.0403 & 0.5769 & 0.0194 & -0.1941 \\ -0.4534 & -0.5479 & 0.0194 & 0.7785 & 0.3362 \\ -0.2921 & -0.4382 & -0.1941 & 0.3362 & 0.5089 \end{bmatrix} \cdot 10^{-3};$$

$$\Sigma_2 = \begin{bmatrix} 1.9357 & 0.2898 & -0.1783 & -0.5075 & -0.2852 \\ 0.2898 & 4.0989 & -0.0694 & -0.6534 & -0.5137 \\ -0.1783 & -0.0694 & 4.9750 & -0.1292 & -0.2899 \\ -0.5075 & -0.6534 & -0.1292 & 4.2037 & 0.3005 \\ -0.2852 & -0.5137 & -0.2899 & 0.3005 & 1.9322 \end{bmatrix} \cdot 10^{-3}$$

for the bandwidths h_1 and h_2, respectively. Again, the agreement of the off-diagonal elements and the difference in the diagonal becomes visible.

Similarly to the previous example, we show the ratios $\frac{diag(\Sigma_2-\Sigma_1)}{diag(V_2-V_1)}$ in Table 4. These are close to 1 as it was expected from our proposition.

x	-1.0	-0.5	0.0	0.5	1.0
$\frac{diag(\Sigma_2-\Sigma_1)}{diag(V_2-V_1)}$	1.0181	1.0331	1.0213	1.0537	1.0180

Table 4: Ratio between the diagonal of the difference of the empirical covariance matrices and that of the theoretical covariance matrices for the data of Example 2.

Finally, Figure 4 shows histograms of $\frac{1}{2}(\hat{f}_n(0.0) + \hat{f}_n(0.5))$ for the bandwidths $h_1 = 0.10$ (left picture) and $h_2 = 0.01$ (right picture). Figure 5 shows histograms of $\frac{1}{3}(\hat{f}_n(-1.0) + \hat{f}_n(-0.5) + \hat{f}_n(0.0))$ for the above bandwidths.

The histograms are presented together with the theoretical normal densities with means and variances estimated from the data used for the histograms. The approximate normality of the density estimator stated in the above proposition is reflected in these figures. Different bandwidths lead to different spreads of the normal distribution.

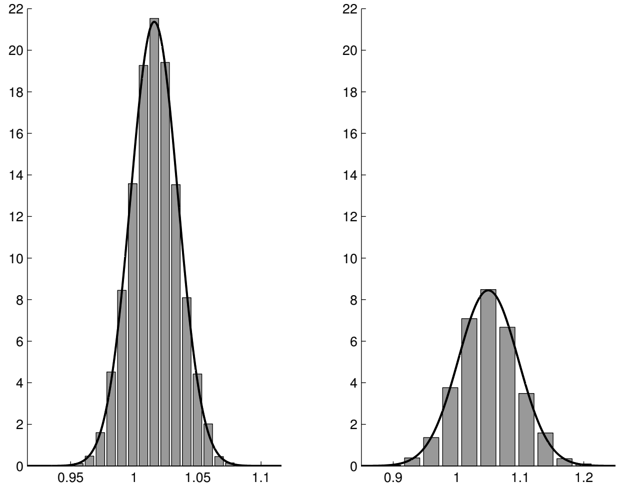

Figure 4: Histograms of $\frac{1}{2}(\hat{f}_n(0.0) + \hat{f}_n(0.5))$ for the bandwidths $h_1 = 0.10$ (left) and $h_2 = 0.01$ (right), together with the theoretical densities of the normal distribution for the data of Example 2.

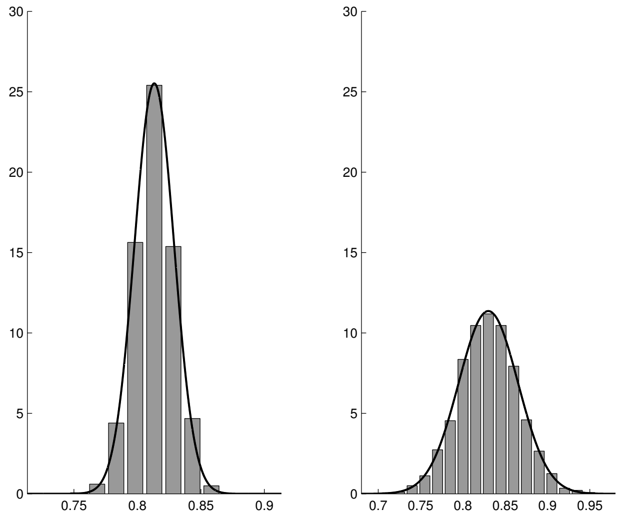

Figure 5: Histograms of $\frac{1}{3}(\hat{f}_n(-1.0) + \hat{f}_n(-0.5) + \hat{f}_n(0.0))$ for the bandwidths $h_1 = 0.10$ (left) and $h_2 = 0.01$ (right), together with the theoretical densities of the normal distribution for the data of Example 2.

5. Conclusions

In the paper, the kernel type density estimator $\hat{f}_n$ is considered. The underlying random field is m-dependent but the observation domain can be irregular. Nearly infill sampling scheme is supposed. Based on the CLT of Park, Kim, Park and Hwang [6] the joint asymptotic normality of $\hat{f}_1(x_1), \dots, \hat{f}_n(x_r)$ is obtained. The asymptotic covariance matrix is unusual in the sense that it is a combination of the covariance matrices in the continuous and the discrete parameter cases. Numerical evidence supports our results.

References

[1] Fazekas, I. and Chuprunov, A. (2004), A central limit theorem for random fields. *Acta Mathematica Academiae Paedagogicae Nyiregyhaziensis,* **20** (1), 93–104, www.emis.de/journals/AMAPN.

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

[3] Karácsony, Zs. and Filzmoser, P. (2010), Asymptotic normality of kernel type regression estimators for random fields. *Journal of Statistical Planning and Inference*, **140**, 872–886.

[4] Lahiri, S.N. (1999), Asymptotic distribution of the empirical spatial cumulative distribution function predictor and prediction bands based on a subsampling method. *Probab. Theory Related Fields*, **114** (1), 55–84.

[5] Lahiri, S.N., Kaiser, M.S., Cressie, N. and Hsu, N.J. (1999), Prediction of spatial cumulative distribution functions using subsampling. *J. Amer. Statist. Assoc.* **94** (445), 86–110.

[6] Park, B.U., Kim, T.Y., Park, T.-S., Hwang, S.Y. (2009), Practically Applicable Central Limit Theorem for Spatial Statistics. *Math. Geosci.* **41**, 555–569.