应用概率统计 第 34 卷
第 5 期 2018 年 10 月

Chinese Journal of Applied Probability and Statistics
Oct., 2018, Vol. 34, No. 5, pp. 450-462
doi: 10.3969/j.issn.1001-4268.2018.05.002

Asymptotic Properties of Nonparametric Intensity Estimation for Replicated Spatial Point Patterns *

ZHAO Jin*
(*Department of Mathematics, Nanjing University, Nanjing, 210093, China*)
ZHOU Xiuqing
(*School of Mathematical Sciences, Nanjing Normal University, Nanjing, 210023, China*)

Abstract: A kernel-type nonparametric estimator of the intensity function for inhomogeneous spatial point patterns with replicated data is proposed. Asymptotic expansion of the mean square error is derived and the rate of convergence of the integrated square error is also investigated. Two methods, least-square and composite likelihood cross-validation, for selecting the bandwidth are described. The performance of the two procedures are illustrated using simulation data.

Keywords: kernel estimator; intensity function; replicated spatial point patterns; asymptotic properties; composite likelihood

2010 Mathematics Subject Classification: 62M30

Citation: ZHAO J, ZHOU X Q. Asymptotic properties of nonparametric intensity estimation for replicated spatial point patterns [J]. Chinese J Appl Probab Statist, 2018, 34(5): 450–462.

§1. Introduction

Spatial point patterns arise in a wide variety of applications including forestry, biology, anatomy, ecology, seismology, epidemiology and geography. Methods for the analysis of stationary point process data are now well established in many textbooks, see [1–3]. In recent years, inhomogeneous spatial point processes have been investigated by many authors. A brief survey can be found in [4,5]. For further references about inhomogeneous spatial point processes, see [6–11].

A key interest in inhomogeneous spatial point patterns is the first-order intensity function $\lambda(s)$ (See Section 2 for its definition), $s \in D$ and D is the observed domain of interest. Let $\lambda(s;\theta)$ be the intensity function depending on a vector of unknown parameters θ. The unknown parameters θ can be estimated by the composite likelihood approach [12]. The

*The project was supported by the National Natural Science Foundation of China (Grant No. 11201235).
*Corresponding author, E-mail: zhaojin@nju.edu.cn.
Received February 23, 2016. Revised September 11, 2017.

composite likelihood estimator is defined as the maximum of $\sum_{s\in N\cap D} \ln \lambda(s;\theta) - \int_D \lambda(s;\theta)\mathrm{d}s$, where N denote the spatial point process that is observed over a domain of interest D. It is also referred to as the Poisson maximum likelihood estimator, since it coincides with the true maximum likelihood when the underlying process is an inhomogeneous Poisson process. Schoenberg [13] proved that composite likelihood estimator is consistent under some mild conditions. Recently, Guan and Loh [14] proved the asymptotic normality of the composite likelihood estimator in general cases. However, in practice, a parametric model for $\lambda(s)$ is not always forthcoming. Then the nonparametric approaches can be considered. Diggle [15] proposed a kernel estimator for $\lambda(s)$ with a single realization of the underlying process, and the expression for the mean squared error of the kernel estimation was derived. Unfortunately, the asymptotic properties of the nonparametric estimation are difficult to be derived, where only one realization is available.

Since the previous application of spatial statistics was mainly in the fields of forestry and plant biology, where only a single realization is available. Analysis for replicated point patterns are still under development. However, with the relatively recent advances in microscopy and technology, particularly in biological sciences, clinical neuroanatomy, scientists are finding that their data may consist of replicated spatial point patterns. See [16–21].

The propose of this paper is to investigate the nonparametric estimator of the intensity function for replicated spatial point patterns. A kernel-type estimator is proposed. The asymptotic expansion of the pointwise mean squared error (MSE) is given, then the mean integrated square error (MISE) can be expressed and the optimal rate of convergence is given, similarly as the classic result of the kernel density estimation. Furthermore, we also study the rate of convergence of the integrated squared error (ISE), which is not always investigated in the study of the kernel estimation. As in any nonparametric smoothing applications, a key factor affecting the accuracy of the estimator is the choice of the bandwidth. An inappropriate value of bandwidth may lead to an estimator with a large bias or variance or both. Thus it is important to develop a data-driven procedure which can be used to automatically and objectively select the bandwidth. In this article two methods are proposed. One is a familiar least-squares cross-validation type of procedure, which has been often applied for bandwidth selection in nonparametric smoothing estimator. The other is the composite likelihood cross-validation, which recently was introduced to select the bandwidth for the estimator of the pair correlation function. See [22]. The efficacy of the two methods is studied by Monte Carlo simulation. The study suggests that the composite likelihood cross-validation preforms better in the case of strong clustering,

conversely, the least-squares cross-validation performs better in the case of weak clustering.

The rest of this paper is organized as follows. Section 2 gives the kernel-type nonparametric estimator and studied its asymptotic properties. Section 3 introduces the least-squares and composite likelihood cross-validation procedures to select the bandwidth. Simulation study of the two procedures are presented in Section 4. Real data analysis is given in Section 5, which suggests that our nonparametric estimation is consistent with [16]'s and [3]'s conclusions. Proofs are given in Section 6. Section 7 gives the conclusion.

§2. The Estimator and Its Asymptotic Properties

Consider a d-dimensional spatial point process N. For any Borel set $D \subset R^d$, let $|D|$ denote the volume of D, and $N(D)$ denote the number of events of N in D. Let ds be an infinitesimal region containing s. Following [3], we define the first-order and second-order intensity functions of N as

$$\lambda(s) = \lim_{|ds|\to 0} \frac{\mathsf{E}[N(ds)]}{|ds|} \qquad \text{and} \qquad \lambda_2(s,t) = \lim_{|ds|,|dt|\to 0} \frac{\mathsf{E}[N(ds)N(dt)]}{|ds||dt|}$$

respectively.

Let $N_1, N_2, \ldots, N_n$ be independent and identically distributed realizations of the underlying process N. For a pattern N_i, there are $N_i(D)$ events $\{s_{i1}, s_{i2}, \ldots, s_{iN_i(D)}\}$ in a fixed region D of finite volume $|D|$. A kernel-type estimator of $\lambda(s)$ has been proposed in [15], where there is only a single realization of N. When replicated spatial point patterns are available, it is natural to consider the following kernel-type estimator:

$$\widehat{\lambda}_{nh}(s) = \frac{1}{n}\sum_{i=1}^{n} \lambda_{nh}^{(i)}(s), \tag{1}$$

where

$$\lambda_{nh}^{(i)} = \sum_{s_{ij}\in N_i\cap D} \frac{k_n(s - s_{ij})}{\omega_h(s)}. \tag{2}$$

In (2), $k_h(s) = (1/h^d)k(s/h)$ is a d-dimensional kernel function and $\omega_h(s)$ is an edge correction form, e.g., $\omega_h(s) = \int_D k_h(s-t)\mathrm{d}t$, as in [15].

We now derive the asymptotic properties of the kernel-type estimator as $n \to \infty$. For simplicity, let $\omega_h(s) \equiv 1$. Firstly, the pointwise mean square error admits the following bias and variance decomposition:

$$\mathrm{MSE}(s) = \mathsf{E}[\widehat{\lambda}_{nh}(s) - \lambda(s)]^2 = [\mathsf{E}\widehat{\lambda}_{nh}(s) - \lambda(s)]^2 + \mathsf{Var}\,[\widehat{\lambda}_{nh}(s)].$$

An approximation of the MSE is given by the following theorem.

Theorem 1 Suppose that k is a bounded kernel function with a bounded support and $\int xk(x)\mathrm{d}x = 0$. If λ has the continuous second derivation at an interior point s of the domain D, and $\lambda_2(s,t)$ has the continuous first derivation at (s,s), then

$$\mathsf{E}\widehat{\lambda}_{nh}(s) - \lambda(s) = \frac{h^2}{2}\int k(x)x^{\mathsf{T}}\nabla^2\lambda(s)x\mathrm{d}x + o(h^2) \tag{3}$$

and

$$\mathsf{Var}\,[\widehat{\lambda}_{nh}(s)] = \frac{\lambda(s)}{nh^d}\int k^2(x)\mathrm{d}x + o\Big(\frac{1}{nh^d}\Big). \tag{4}$$

Furthermore, a global measure can be obtain by using a integrated square error

$$\mathrm{ISE} = \int_{D'}[\widehat{\lambda}_{nh}(s) - \lambda(s)]^2\mathrm{d}s,$$

where D' is an arbitrary region we concerned and $D' \subset D$. To avoid the edge effect, let $D'_{\oplus\epsilon} = \{s : s \in \epsilon B(s'), s' \in D'\} \subset D$ for some $\epsilon > 0$, where $\oplus$ denote the Minkowski addition and $B(s)$ is a unit ball in R^d centered at s. Similarly, the mean integrated square error can be expressed by

$$\begin{aligned}\mathrm{MISE} &= \mathsf{E}\int_{D'}[\widehat{\lambda}_{nh}(s) - \lambda(s)]^2\mathrm{d}s \\ &= \frac{h^4}{4}\int_{D'}\Big[\int k(x)x^{\mathsf{T}}\nabla^2\lambda(s)x\mathrm{d}x\Big]^2\mathrm{d}s + \frac{1}{nh^d}\int_{D'}\lambda(s)\mathrm{d}s\int k^2(x)\mathrm{d}x + o\Big(\frac{1}{nh^d} + h^4\Big).\end{aligned}$$

By choosing $h = O(n^{-1/(4+d)})$, the MISE can achieves the optimal rate of convergence $n^{-4/(4+d)}$. Here we also investigate the rate of convergence of the ISE, which is given by the following theorem.

Theorem 2 Suppose that k is a bounded kernel function with a bounded support and $\int xk(x)\mathrm{d}x = 0$. If λ has the continuous second derivation in D', and $\lambda_2(s,t)$ has the continuous first derivation in $D' \times D'$. $N(D)$ is bounded, i.e., $\mathsf{P}[N(D) > C] = 0$ for some constant $C > 0$, then

$$\sqrt{\mathrm{ISE}} = \sqrt{\int_{D'}[\widehat{\lambda}_{nh}(s) - \lambda(s)]^2\mathrm{d}s} = O\Big(\sqrt{\frac{\ln n}{nh^d}}\Big) + O\Big(\frac{1}{\sqrt{nh^d}} + h^2\Big) \qquad \text{a.s.}$$

For $h = O((\ln n/n)^{1/(4+d)})$, $\mathrm{ISE} = (\ln n/n)^{4/(4+d)}$ a.s.

§3. Bandwidth Selection

It is well known that the choice of bandwidth affects the performance of the kernel estimator greatly, and it is much more crucial than the choice of the kernel. The theoretical results in Section 2 confirm the intuition that increasing h implies increasing bias

an decreasing variance while increasing number of observations implies smaller variance. However the theoretical results are not helpful from a practical point of view when it comes to selecting the bandwidth. There are quite a few methods for selecting the bandwidth (e.g. [23]). Among these, an often used approach is the least-square cross validation. We propose a familiar procedure to select h as the minimizer of the following criterion:

$$\mathrm{LSCV}(h) = \int_D \widehat{\lambda}^2_{nh}(s)\mathrm{d}s - \frac{2}{n}\sum_{i=1}^{n}\sum_{s_{ij}\in N_i\cap D}\widehat{\lambda}_{n\backslash\{i\}h}(s_{ij}), \tag{5}$$

where $\widehat{\lambda}_{n\backslash\{i\}h}(s) = (n-1)^{-1}\sum\limits_{j\neq i}\lambda^{(j)}_{nh}(s)$, i.e., a kernel-type estimator based on the data with N_i omitted. It is easy to verify that $n^{-1}\sum\limits_{i=1}^{n}\sum\limits_{s_{ij}\in N_i\cap D}\widehat{\lambda}_{n\backslash\{i\}h}(s_{ij})$ is an unbiased estimator of $\int_D \widehat{\lambda}_{nh}(s)\lambda(s)\mathrm{d}s$. In fact, we have

$$\mathsf{E}\int_D \widehat{\lambda}_{nh}(s)\lambda(s)\mathrm{d}s = \int_D\int_D k_h(s-t)\lambda(s)\lambda(t)\mathrm{d}s\mathrm{d}t$$

and

$$\frac{1}{n}\sum_{i=1}^{n}\sum_{s_{ij}\in N_i\cap D}\widehat{\lambda}_{n\backslash\{i\}h}(s_{ij}) = \frac{1}{n(n-1)}\sum_{i\neq i'}\sum_{j,k}k_h(s_{ij}-s_{i'k}).$$

Another simple approach is the likelihood cross validation [24]. Recently, Guan [22] modified this approach to select the bandwidth for the estimator of the pair correlation function. Here a full likelihood function is not available, a natural alternative is to use the composite likelihood, i.e., $\sum\limits_{s\in N\cap D}\ln\lambda(s) - \int_D\lambda(s)\mathrm{d}s$. Thus composite likelihood cross validation defines the bandwidth as the maximum of

$$\mathrm{CLCV}(h) = \frac{1}{n}\sum_{i=1}^{n}\sum_{s_{ij}\in N_i\cap D}\ln\widehat{\lambda}_{n\backslash\{i\}h}(s_{ij}) - \int_D\widehat{\lambda}_{nh}(s)\mathrm{d}s. \tag{6}$$

§4. Simulation Study

The proposed bandwidth selection methods are applied to simulation data generated by a non-homogeneous Poisson process (NHPP) and a inhomogeneous Neyman-Scott process (see [25]). For the NHPP, the intensity function $\lambda(s_1, s_2) = \alpha\exp(-s_1 - s_2)$, we considered $\alpha = 20, 40, 60$. For the inhomogeneous Neyman-Scott process, the parents points form a HPP with intensity $\beta > 0$. The clusters are independent NHPP with intensity functions

$$\lambda(s_1, s_2; \sigma) = 4k_\sigma(s_1 - c_1, s_2 - c_2)\exp(-s_1 - s_2),$$

where $k_\sigma(s_1, s_2) = (2\pi\sigma^2)^{-1}\exp[-(s_1^2 + s_2^2)/(2\sigma^2)]$, and (c_1, c_2) is the location of parents. Let $\beta = 10, 20, 40$, for each value of β we consider $\sigma = 0.1, 0.4$, which correspond to respectively strong and weak clustering. In each case, we simulated $n = 100, 200, 400, 800$ replicated realizations on a unit square.

To employ the kernel-type nonparametric estimator, one needs to choose the kernel function and the bandwidth. It is well known that the choice of kernel functions is not very important. Here we use the uniform kernel function, i.e., $k(s_1, s_2) = (1/4)I_{[-1,1]\times[-1,1]}(s_1, s_2)$, where $I_A(\cdot)$ denotes the indicator function of the set A. The bandwidth were selected by minimizing (5) and maximizing (6). We compare the two methods by two criterions: l_2 distance and l_∞ distance, i.e., $l_2 = \sqrt{\int_D [\widehat{\lambda}_{nh}(s) - \lambda(s)]^2 \mathrm{d}s}$ and $l_\infty = \sup_D |\widehat{\lambda}_{nh}(s) - \lambda(s)|$, where D is the observation region.

Tables 1 – 3 list the simulation results for the inhomogeneous Poisson processes, while Tables 4 – 6 show the results for the inhomogeneous Neyman-Scott processes. In all cases, the l_2 and l_∞ distances decrease as the sample size increases. From Tables 1 – 3, the least-square cross validation has smaller l_2 distances, but the l_∞ distances are larger. For the cluster process case in Tables 4 – 6, the composite likelihood cross validation performs better when $\sigma = 0.1$, which correspond to strong clustering. However, in case of $\sigma = 0.4$, it performs worse than least-square cross validation. As a result, we recommend the use of composite likelihood cross validation in case of strong clustering and use least-square cross validation in case of weak clustering.

Table 1 Cross-validation procedures in NHPP case ($\alpha = 20$)

	Least-square		Composite likelihood	
n	l_2	l_∞	l_2	l_∞
100	2.369	4.400	2.266	4.041
200	2.090	3.754	2.116	3.726
400	1.809	3.448	1.973	3.288
800	1.708	2.807	1.854	2.580

Table 2 Cross-validation procedures in NHPP case ($\alpha = 40$)

	Least-square		Composite likelihood	
n	l_2	l_∞	l_2	l_∞
100	2.519	5.589	2.746	4.949
200	2.077	4.475	2.529	4.537
400	1.951	3.488	2.200	3.316
800	1.808	2.930	2.104	2.621

Table 3 Cross-validation procedures in NHPP case ($\alpha = 60$)

	Least-square		Composite likelihood	
n	l_2	l_∞	l_2	l_∞
100	4.066	10.766	4.351	7.867
200	3.197	6.356	3.364	6.087
400	2.785	5.610	2.976	5.414
800	2.534	4.980	2.640	4.655

Table 4 Cross-validation procedures in the Poisson cluster case ($\beta = 10$)

		Least-square		Composite likelihood	
n	σ	l_2	l_∞	l_2	l_∞
100	0.1	2.767	5.073	2.417	4.140
	0.4	3.091	7.783	3.540	10.863
200	0.1	2.353	3.633	2.023	3.566
	0.4	2.821	6.949	2.987	9.006
400	0.1	1.969	2.748	1.923	2.587
	0.4	2.551	5.501	2.625	6.176
800	0.1	1.742	1.994	1.634	1.790
	0.4	1.995	3.951	2.106	4.252

Table 5 Cross-validation procedures in the Poisson cluster case ($\beta = 20$)

		Least-square		Composite likelihood	
n	σ	l_2	l_∞	l_2	l_∞
100	0.1	4.574	7.098	4.353	6.887
	0.4	5.158	9.409	5.454	9.497
200	0.1	4.214	6.102	4.014	6.011
	0.4	4.878	8.016	5.132	8.085
400	0.1	3.790	4.890	3.435	4.677
	0.4	4.241	6.487	4.488	6.753
800	0.1	2.764	3.538	2.430	3.329
	0.4	3.246	5.015	3.486	5.348

§5. Application

As an example, we consider the data shown in Figure 1, which was originally analyzed by Diggle et al.[16] These data consist of 12 point patterns, each of which concerns the locations of pyramidal neurons in the cingulate cortex of human brains. Post-morterm

Table 6 Cross-validation procedures in the Poisson cluster case ($\beta = 40$)

		Least-square		Composite likelihood	
n	σ	l_2	l_∞	l_2	l_∞
100	0.1	6.367	8.932	6.128	8.704
	0.4	6.890	10.933	7.109	11.135
200	0.1	5.879	7.360	5.540	7.186
	0.4	6.341	9.131	6.602	9.670
400	0.1	5.220	6.144	4.976	5.882
	0.4	5.764	7.878	5.903	8.010
800	0.1	4.182	5.094	3.958	4.857
	0.4	4.534	6.281	4.808	6.397

slices of the same area in each of the subject's brains were obtained. The locations of the neurons were identified through digitization and scaled to the unit square. The 12 point patterns are presumed to be normal control cases. Diggle [3] used the estimated K-function to show that the normal group was not distinguishable from the complete spatial randomness (CSR).

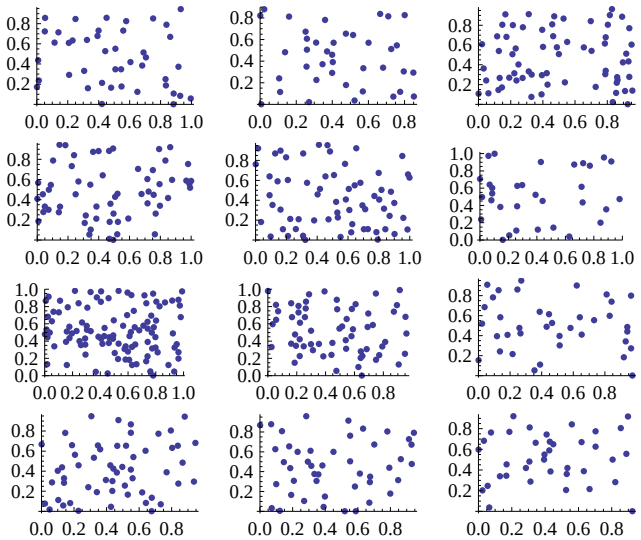

Figure 1 Locations of pyramidal neurons from 12 normal subjects

We use the kernel-type method to estimate the intensity function of the 12 replicated point patterns. We choose the Gaussian kernel function, i.e., $k(s_1, s_2) = (2\pi)^{-1}\mathrm{e}^{-(s_1^2+s_2^2)/2}$. The bandwidth selection approach is the least-square cross validation. Figure 2 shows the surface of the kernel estimator $\widehat{\lambda}(s_1, s_2)$. The fluctuation of the function surface is mild.

In fact,

$$\sup_{s_1\in(0,1),s_2\in(0,1)} |\widehat{\lambda}(s_1,s_2)-\widetilde{\lambda}| = 4.490, \qquad \text{where } \widetilde{\lambda}=\frac{\sum N_i(D)}{12}\approx 54.4,$$

which is the maximum likelihood estimator of the intensity when the underlying process is a HPP. The value is small relative to the intensity. This suggests that we have no reason to reject the CSR of the 12 normal subjects. Again, our result is consistent with [16]'s and [3]'s conclusions.

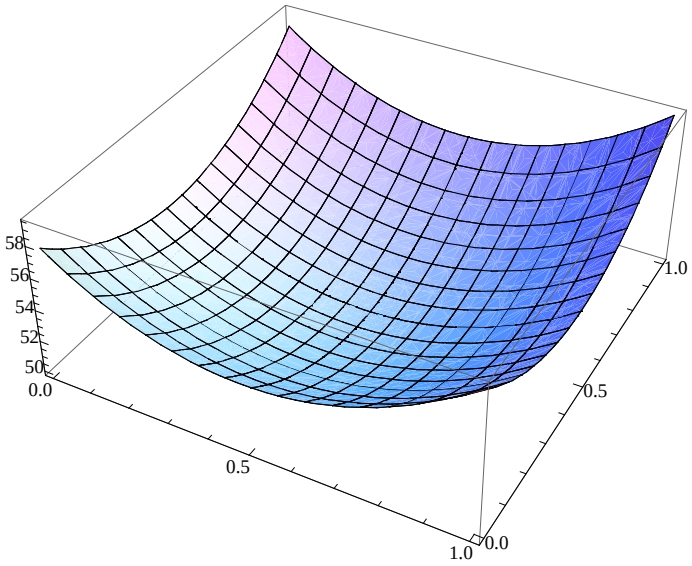

Figure 2 Function surface of the kernel estimator

§6. Proofs

Proof of Theorem 1 By the Campbell theorem, we have

$$\mathsf{E}\widehat{\lambda}_{nh}(s)-\lambda(s)=\mathsf{E}\lambda_{nh}^{(1)}(s)-\lambda(s)=\int_D k_h(s-x)\lambda(s)\mathrm{d}s-\lambda(s).$$

Using the assumption that k has a bounded support and s is an interior point of D, we have for h small enough,

$$\mathsf{E}\widehat{\lambda}_{nh}(s)-\lambda(s)=\int k(x)[\lambda(s-hx)-\lambda(s)]\mathrm{d}s=\frac{h^2}{2}\int k(x)x^{\mathsf{T}}\nabla^2\lambda(s-\xi)x\mathrm{d}x,$$

where ξ lies between 0 and hx. By the Lebesgue convergence theorem, we have

$$\mathsf{E}\widehat{\lambda}_{nh}(s)-\lambda(s)=\frac{h^2}{2}\int k(x)x^{\mathsf{T}}\nabla^2\lambda(s)x\mathrm{d}x+o(h^2).$$

Next,

$$\mathsf{Var}\,[\widehat{\lambda}_{nh}(s)] = \frac{1}{n}\mathsf{Var}\,[\lambda^{(1)}_{nh}(s)].$$

Note that $\mathsf{E}\lambda^{(1)}_{nh}(s) = O(1)$, we have

$$\begin{aligned}
&\mathsf{Var}\,[\widehat{\lambda}_{nh}(s)]\\
&= \frac{1}{n}\mathsf{E}\Big[\sum_{s_{1j}\in N_1\cap D} k_h(s-s_{1j})\Big]^2 + O\Big(\frac{1}{n}\Big)\\
&= \frac{1}{n}\Big[\int_D\int_D k_h(s-x)k_h(s-y)\lambda_2(x,y)\mathrm{d}x\mathrm{d}y + \int_D k_h^2(s-x)\lambda(x)\mathrm{d}x\Big] + O\Big(\frac{1}{n}\Big)\\
&= \frac{1}{n}\Big[\int\int k(x)k(y)\lambda_2(s-hx,s-hy)\mathrm{d}x\mathrm{d}y + \frac{1}{h^d}\int k^2(x)\lambda(s-hx)\mathrm{d}x\Big] + O\Big(\frac{1}{n}\Big).
\end{aligned}$$

Using the assumption on $\lambda(s)$ and $\lambda_2(s,t)$, we have

$$\int\int k(x)k(y)\lambda_2(s-hx,s-hy)\mathrm{d}x\mathrm{d}y = \lambda_2(s,s) + o(1)$$

and

$$\int k^2(x)\lambda(s-hx)\mathrm{d}x = \lambda(s)\int k^2(x)\mathrm{d}x + o(1),$$

provided that $h\to 0$. Hence equation (4) follows and this completes the proof. □

To prove Theorem 2, we need the following [26] exponential inequality.

Lemma 3 Let $\{S_n\}_{n\geqslant 1}$ be a martingale based on the independent r.v. $X_1, X_2, \ldots$ Suppose that exist r.v. $T_n = \psi_n(X_1, X_2, \ldots, X_n)$ and constant T_o and c_n, such that

$$T_{n-1} \leqslant S_n \leqslant T_{n-1} + c_n, \qquad n = 1, 2, \ldots$$

Then for all $n \geqslant 1$,

$$\mathsf{P}(|S_n - \mathsf{E}S_1| \geqslant t) \leqslant 2\exp\Big(-2t^2/\sum_{i=1}^n c_i^2\Big).$$

Proof of Theorem 2 By the triangle and Jensen's inequality, we have

$$\begin{aligned}
\mathsf{E}\sqrt{\int_{D'}[\widehat{\lambda}_{nh}(s)-\lambda(s)]^2\mathrm{d}s} &\leqslant \mathsf{E}\sqrt{\int_{D'}[\widehat{\lambda}_{nh}(s)-\mathsf{E}\widehat{\lambda}_{nh}(s)]^2\mathrm{d}s} + \sqrt{\int_{D'}[\mathsf{E}\widehat{\lambda}_{nh}(s)-\lambda(s)]^2\mathrm{d}s}\\
&\leqslant \sqrt{\int_{D'}\mathsf{Var}\,\widehat{\lambda}_{nh}(s)\mathrm{d}s} + \sqrt{\int_{D'}[\mathsf{E}\widehat{\lambda}_{nh}(s)-\lambda(s)]^2\mathrm{d}s}\\
&= O\Big(\frac{1}{\sqrt{nh^d}} + h^2\Big).
\end{aligned}$$

Thus, we need only to show that

$$\sqrt{\int_{D'}[\widehat{\lambda}_{nh}(s)-\lambda(s)]^2\mathrm{d}s} - \mathsf{E}\sqrt{\int_{D'}[\widehat{\lambda}_{nh}(s)-\lambda(s)]^2\mathrm{d}s} = O\Big(\sqrt{\frac{\ln n}{nh^d}}\Big) \qquad \text{a.s.}$$

Define $\psi_n(N_1, N_2, \ldots, N_n) = \sqrt{\int_{D'}[\widehat{\lambda}_{nh}(s) - \lambda(s)]^2 \mathrm{d}s}$, then, ψ_n is a symmetric function of $N_1, N_2, \cdots, N_n$. Let $S_k = \mathsf{E}(\psi_n \,|\, N_1, N_2, \ldots, N_k)$, $k = 1, 2, \ldots, n-1$ and $S_n = \psi_n$. Thus $S_1, S_2, \ldots, S_n$ is a martingale. Since

$$\inf_{N_k} S_k \leqslant S_k \leqslant \inf_{N_k} S_k + \gamma_k,$$

where $\gamma_k = \sup_{N_k} S_k - \inf_{N_k} S_k$. We now give a bound on γ_k. It is convenient to denote N_k in the supremum as N and in the infimum as N', where N and N' are i.i.d. realizations of the underlying process. We have

$$\begin{aligned}
\gamma_k =& \sup_{N,N'} \mathsf{E}(\psi_n \,|\, N_1, N_2, \ldots, N_{k-1}, N) - \mathsf{E}(\psi_n \,|\, N_1, N_2, \ldots, N_{k-1}, N') \\
=& \sup_{N,N'} \int [\psi_n(N_1, N_2, \ldots, N_{k-1}, N, N_{k+1}, N_{k+2}, \ldots, N_n) \\
& - \psi_n(N_1, N_2, \ldots, N_{k-1}, N', N_{k+1}, N_{k+2}, \ldots, N_n)] \mathrm{d}\mathsf{P}(N_{k+1}, N_{k+2}, \ldots, N_n) \\
\leqslant & \sup_{N_1, N_2, \ldots, N_{k-1}, N, N', N_{k+1}, N_{k+2}, \ldots, N_n} |\psi_n(N_1, N_2, \ldots, N_{k-1}, N, N_{k+1}, N_{k+2}, \ldots, N_n) \\
& - \psi_n(N_1, N_2, \ldots, N_{k-1}, N', N_{k+1}, N_{k+2}, \ldots, N_n)| \\
=& \sup_{N, N', N_2, N_3, \ldots, N_n} |\psi(N, N_2, N_3, \ldots, N_n) - \psi(N', N_2, N_3, \ldots, N_n)|.
\end{aligned}$$

The last equality is follows by the symmetry of ψ_n. From the triangle inequality,

$$\gamma_k \leqslant \frac{1}{n} \sup_{N,N'} \Big\{ \int_{D'} \Big[\sum_{x_j \in N \cap D} k_h(s - x_j) - \sum_{x'_j \in N \cap D} k_h(s - x'_j) \Big]^2 \mathrm{d}s \Big\}^{1/2} \leqslant \frac{C'}{nh^{d/2}},$$

where C' is a constant. Using Lemma 3, we have

$$\mathsf{P}(|S_n - \mathsf{E}S_1| \geqslant t) < 2 \exp\Big(-2t^2 \Big/ \sum_{k=1}^{n} \frac{C'^2}{n^2 h^d} \Big) = 2 \exp\Big(-\frac{2nh^d t^2}{C'^2} \Big).$$

By the Borel-Cantelli lemma, the theorem is proved. □

§7. Conclusions

Intensity function is an important characteristic for inhomogeneous spatial point processes. Recently, data of replicated spatial point patterns are given in the fields of biological sciences and clinical neuroanatomy. In this paper we investigate the nonparametric estimator of the intensity function for replicated spatial point patterns. A kernel-type estimator is proposed. The asymptotic expansion of the pointwise MSE and the MISE are

given, which are similar as the classic results of the kernel density estimation. Furthermore, we also study the rate of convergence of the ISE, which is not always presented in the study of the kernel estimation.

As a key factor in any nonparametric smoothing applications, we proposed two methods for the choice of the bandwidth, which are the familiar least-squares cross-validation type of procedure and composite likelihood cross-validation. The efficacy of the two methods is studied by Monte Carlo simulation. The study suggests that the composite likelihood cross-validation preforms better in the case of strong clustering and the least-square cross-validation performs better in case of weak clustering. Real data analysis also shows that our nonparametric estimator is useful.

References

[1] RIPLEY B D. *Statistical Inference for Spatial Processes* [M]. New York: Cambridge University Press, 1988.

[2] CRESSIE N A C. *Statistics for Spatial Data* [M]. New York: Wiley. 1993.

[3] DIGGLE P J. *Statistical Analysis of Spatial and Spatio-Temporal Point Patterns* [M]. 3rd ed. Boca Raton, FL: Chapman and Hall/CRC, 2013.

[4] MØLLER J, WAAGEPETERSEN R P. *Statistical Inference and Simulation for Spatial Point Processes* [M]. New York: Chapman and Hall/CRC, 2003.

[5] MØLLER J, WAAGEPETERSEN R P. Modern statistics for spatial point processes (with discussion) [J]. *Scand J Statist*, 2007, **34(4)**: 643–711.

[6] BADDELEY A J, MØLLER J, WAAGEPETERSEN R. Non- and semi-parametric estimation of interaction in inhomogeneous point patterns [J]. *Statist Neerlandica*, 2000, **54(3)**: 329–350.

[7] BRIX A, SENOUSSI R, COUTERON P, et al. Assessing goodness of fit of spatially inhomogeneous Poisson processes [J]. *Biometrika*, 2001, **88(2)**: 487–497.

[8] WAAGEPETERSEN R, GUAN Y T. Two-step estimation for inhomogeneous spatial point processes [J]. *J R Stat Soc Ser B Stat Methodol*, 2009, **71(3)**: 685–702.

[9] GUAN Y T. On nonparametric variance estimation for second-order statistics of inhomogeneous spatial point processes with a known parametric intensity form [J]. *J Amer Statist Assoc*, 2009, **104(488)**: 1482–1491.

[10] YUE Y R, LOH J M. Bayesian semiparametric intensity estimation for inhomogeneous spatial point processes [J]. *Biometrics*, 2011, **67(3)**: 937–946.

[11] YUE Y R, LOH J M. Bayesian nonparametric estimation of pair correlation function for inhomogeneous spatial point processes [J]. *J Nonparametr Stat*, 2013, **25(2)**: 463–474.

[12] LINDSAY B G. Composite likelihood methods [J]. *Contemp Math*, 1988, **80**: 221–239.

[13] SCHOENBERG F P. Consistent parametric estimation of the intensity of a spatial – temporal point process [J]. *J Statist Plann Inference*, 2005, **128(1)**: 79–93.

[14] GUAN Y T, LOH J M. A thinned block bootstrap variance estimation procedure for inhomogeneous spatial point patterns [J]. *J Amer Statist Assoc*, 2007, **102(480)**: 1377–1386.

[15] DIGGLE P. A kernel method for smoothing point process data [J]. *J Roy Statist Soc Ser C Appl Statist*, 1985, **34(2)**: 138–147.

[16] DIGGLE P J, LANGE N, BENEŠ F M. Analysis of variance for replicated spatial point patterns in clinical neuroanatomy [J]. *J Amer Statist Assoc*, 1991, **86(415)**: 618–625.

[17] DIGGLE P J, MATEU J, CLOUGH H E. A comparison between parametric and non-parametric approaches to the analysis of replicated spatial point patterns [J]. *Adv Appl Probab*, 2000, **32(2)**: 331–343.

[18] BADDELEY A J, MOYEED R A, HOWARD C V, et al. Analysis of a three-dimensional point pattern with replication [J]. *J Roy Statist Soc Ser C Appl Statist*, 1993, **42(4)**: 641–668.

[19] MATEU J. Parametric procedures in the analysis of replicated pairwise interaction point patterns [J]. *Biom J*, 2001, **43(3)**: 375–394.

[20] BELL M L, GRUNWALD G K. Mixed models for the analysis of replicated spatial point patterns [J]. *Biostatistics*, 2004, **5(4)**: 633–648.

[21] MYLLYMÄKI M, PANOUTSOPOULOU I G, SÄRKKÄ A. Analysis of spatial structure of epidermal nerve entry point patterns based on replicated data [J]. *J Microsc*, 2012, **247(3)**: 228–239.

[22] GUAN Y T. A composite likelihood cross-validation approach in selecting bandwidth for the estimation of the pair correlation function [J]. *Scand J Statist*, 2007, **34(2)**: 336–346.

[23] Silverman B W. *Density Estimation for Statistics and Data Analysis* [M]. London: Chapman and Hall, 1986.

[24] DUIN R P W. On the choice of smoothing parameters for Parzen estimators of probability density functions [J]. *IEEE Trans Comput*, 1976, **25(11)**: 1175–1179.

[25] WAAGEPETERSEN R P. An estimating function approach to inference for inhomogeneous Neyman – Scott processes [J]. *Biometrics*, 2007, **63(1)**: 252–258.

[26] HOEFFDING W. Probability inequalities for sums of bounded random variables [J]. *J Amer Statist Assoc*, 1963, **58(301)**: 13–30.

重复数据空间点过程的非参数强度函数估计的渐进性质

赵 进 周秀轻

(南京大学数学系, 南京, 210093) (南京师范大学数学科学学院, 南京, 210023)

摘 要: 在重复数据的情况下, 我们给出了非齐次点过程强度函数的核估计. 并研究了该估计的均方差与积分均方差的渐进性质. 对于核估计的窗宽选择, 我们给出了两种不同的方法, 并用模拟的方法比较了这两种方法的效率.

关键词: 核估计; 强度函数; 重复空间点过程; 渐进性质; 复合似然

中图分类号: O212.7