Spatial Statistics 18 (2016) 24–39

Contents lists available at ScienceDirect

Spatial Statistics

journal homepage: www.elsevier.com/locate/spasta

Standard and robust intensity parameter estimation for stationary determinantal point processes

Christophe A.N. Biscio [a], Jean-François Coeurjolly [b,*]

[a] *Department of Mathematical Sciences, Aalborg University, Denmark*

[b] *University Grenoble Alpes, France*

ARTICLE INFO

Article history:
Received 24 November 2015
Accepted 21 April 2016
Available online 4 May 2016

Keywords:
Repulsive point processes
Statistical inference
Robust statistics
Sample quantiles
Brillinger mixing

ABSTRACT

This work is concerned with the estimation of the intensity parameter of a stationary determinantal point process. We consider the standard estimator, corresponding to the number of observed points per unit volume and a recently introduced median-based estimator more robust to outliers. The consistency and asymptotic normality of estimators are obtained under mild assumptions on the determinantal point process. We illustrate the efficiency of the procedures in a simulation study.

1. Introduction

Spatial point patterns are datasets containing the random locations of some event of interest which arise in many scientific fields such as biology, epidemiology, seismology and hydrology. Spatial point processes are the stochastic models generating such data. We refer to Stoyan et al. (1995), Illian et al. (2008) or Møller and Waagepetersen (2004) for an overview on spatial point processes. The Poisson point process is the reference process to model random locations of points without interaction. Many alternative models such as Cox point processes (including Neyman–Scott processes, shot noise Cox processes, log-Gaussian Cox processes) or Gibbs point processes allow us to introduce clustering effects or to produce regular patterns (see again e.g. Møller and Waagepetersen, 2004 or Illian et al., 2008). First introduced by Macchi (1975), the interesting class of determinantal point processes

* Corresponding author.
E-mail addresses: christophe@math.aau.dk (C.A.N. Biscio), Jean-Francois.Coeurjolly@upmf-grenoble.fr (J.-F. Coeurjolly).

http://dx.doi.org/10.1016/j.spasta.2016.04.007
2211-6753/

has been revisited recently by Lavancier et al. (2015) in a statistical context. Such processes are in particular designed to model repulsive point patterns.

In this paper, we focus on stationary point processes, that is on point processes with distribution invariant by translation, and on first order characteristics for such processes, that is on the intensity parameter denoted by λ. The nonnegative real parameter λ measures the mean number of points per unit volume and is needed for the estimation of second-order characteristics of point processes such as the pair correlation function or the Ripley's K-function, see for instance Møller and Waagepetersen (2004). Thus, the estimation of λ has been the subject of a large literature (see e.g. Illian et al., 2008). Asymptotic properties for estimators of λ are non trivial and may be particularly challenging to obtain for some models such as the class of Gibbs point processes.

In this paper, we investigate the theoretical and practical properties of two different estimators of λ for the class of stationary determinantal point processes. The first estimator is the standard one, corresponding to the number of observed points divided by the volume of the observation domain. The second one is a median-based estimator recently proposed by Coeurjolly (in press) to handle outliers such as extra points or missing points. The form of these two estimators is not novel and follows the aforementioned references. Asymptotic properties for these two estimators have been established under general conditions on the underlying point process. However, these conditions have been checked mainly for Cox processes. We propose two contributions. First, we provide conditions on the kernel C, defining a determinantal point process (see Section 2 for details), which ensure that the standard and the median-based estimators are consistent and satisfy a central limit theorem. Second, we investigate the finite-sample size properties of the proposed procedures through a simulation study, where, in particular, we evaluate the ability of the estimators to be robust to outliers.

The rest of the paper is organized as follows. A short background on stationary determinantal point processes is presented in Section 2. Section 3.1 focuses on the standard estimator and details asymptotic properties for this estimator as well as an estimator of its asymptotic variance. Section 3.2 deals with the median-based estimator. Finally, we conduct a simulation study in Section 4 to compare these estimators in different scenarios. Proofs of the main results are postponed to Appendix.

2. Stationary determinantal point processes

2.1. Background and definition

For $d \geq 1$, let $\mathbf{X}$ be a spatial point process defined on $\mathbb{R}^d$, which we see as a random locally finite subset of $\mathbb{R}^d$. Let $\mathcal{B}(\mathbb{R}^d)$ denote the class of bounded Borel sets in $\mathbb{R}^d$. For $u = (u^1, \ldots, u^d)^\top \in \mathbb{R}^d$ and for $A, B \in \mathcal{B}(\mathbb{R}^d)$, we denote by $|u| = \max_{i=1,\ldots,d} |u^i|$ and by $d(A, B)$ the minimal distance between A and B.

For any $W \in \mathcal{B}(\mathbb{R}^d)$, we denote by $|W|$ its Lebesgue measure, by $N(\mathbf{X} \cap W)$ the number of points in $\mathbf{X} \cap W$ and a realization of $\mathbf{X} \cap W$ is of the form $\mathbf{x} = \{x_1, \ldots, x_m\} \subset W$ for some nonnegative finite integer m. We consider simple point processes which means that two points of the process never coincide almost surely. For further details about point processes, we refer to Daley and Vere-Jones (2003, 2008) and Møller and Waagepetersen (2004).

The factorial moment measures are quantities of special interest for point processes. For any integer $l \geq 1$, $\mathbf{X}$ is said to have an lth order factorial moment measure $\alpha^{(l)}$ if for all nonnegative measurable functions h defined on $\mathbb{R}^{dl}$,

$$\mathrm{E} \sum_{u_1,\ldots,u_l \in \mathbf{X}}^{\neq} h(u_1, \ldots, u_l) = \int_{\mathbb{R}^{dl}} h(u_1, \ldots, u_l)\, \alpha^{(l)}(\mathrm{d}u_1 \times \cdots \times \mathrm{d}u_l) \tag{1}$$

where the sign $\neq$ over the summation means that $u_1, \ldots, u_l$ are pairwise distinct. If $\alpha^{(l)}$ admits a density with respect to the Lebesgue measure on $\mathbb{R}^{dl}$, this density is called the lth order product density of $\mathbf{X}$ and is denoted by ρ_l. Note that $\rho_1 = \lambda$ and that for the homogeneous Poisson point process $\rho_l(u_1, \ldots, u_l) = \lambda^l$. We assume from now on, that λ is a positive real number.