Centrum voor Wiskunde en Informatica

REPORT_RAPPORT_

PNA

Probability, Networks and Algorithms

Probability, Networks and Algorithms

A non-parametric estimator for the doubly-periodic
Poisson intensity function

R. Helmers, I.W. Mangku, R.Z. Zitikis

Centrum voor Wiskunde en Informatica (CWI) is the national research institute for Mathematics and Computer Science. It is sponsored by the Netherlands Organisation for Scientific Research (NWO).
CWI is a founding member of ERCIM, the European Research Consortium for Informatics and Mathematics.

CWI's research has a theme-oriented structure and is grouped into four clusters. Listed below are the names of the clusters and in parentheses their acronyms.

Probability, Networks and Algorithms (PNA)

Software Engineering (SEN)

Modelling, Analysis and Simulation (MAS)

Information Systems (INS)

A non-parametric estimator for the doubly-periodic Poisson intensity function

ABSTRACT

In a series of papers, J. Garrido and Y. Lu have proposed and investigated a doubly-periodic Poisson model, and then applied it to analyze hurricane data. The authors have suggested several parametric models for the underlying intensity function. In the present paper we construct and analyze a non-parametric estimator for the doubly-periodic intensity function. Assuming that only a single realization of the process is available in a bounded window, we show that the estimator is consistent and asymptotically normal when the window expands indefinitely. In addition we calculate the asymptotic bias and variance of the estimator, and in this way gain helpful information for optimizing the performance of the estimator.

A non-parametric estimator for the doubly-periodic Poisson intensity function

Roelof Helmers[1]

Centre for Mathematics and Computer Science (CWI), P.O. Box 94079, 1090 GB, Amsterdam, The Netherlands. E-mail: R.Helmers@cwi.nl

I Wayan Mangku[2]

Department of Mathematics, Bogor Agricultural University, Jl. Meranti, Kampus IPB Darmaga, Bogor 16680, Indonesia.

Ričardas Zitikis[3]

Department of Statistical and Actuarial Sciences, University of Western Ontario, London, Ontario, Canada N6A 5B7. E-mail: zitikis@stats.uwo.ca

Abstract. In a series of papers, J. Garrido and Y. Lu have proposed and investigated a doubly-periodic Poisson model, and then applied it to analyze hurricane data. The authors have suggested several parametric models for the underlying intensity function. In the present paper we construct and analyze a *non-parametric* estimator for the doubly-periodic intensity function. Assuming that only a single realization of the process is available in a bounded window, we show that the estimator is consistent and asymptotically normal when the window expands indefinitely. In addition, we calculate the asymptotic bias and variance of the estimator, and in this way gain helpful information for optimizing the performance of the estimator.

Keywords and phrases: Poisson process, doubly-periodic Poisson process, periodic intensity function, non-parametric estimation, consistency, asymptotic normality, bias, variance, mean-squared error.

1 Introduction and the model

Non-homogeneous and, in particular, *periodic* non-homogeneous Poisson processes have been used to model random phenomena in areas such as communications, hydrology, meteorology, insurance, reliability, seismology; we refer to the books by Lewis (1972), Cox and Lewis (1978), Cox and Isham (1980), Diggle (1983), Karr (1986), Daley and Vere-Jones (1988), Cressie (1991), Kingman (1993), Reiss (1993), Snyder and Miller (1995), Kutoyants (1998), as well as to the references therein.

Recently, a far reaching research has been conducted by Garrido and Lu (2004), Lu (2005), Lu and Garrido (2005, 2006), where the authors analyze El Niňo and La Niña related data using the therein proposed doubly-periodic Poisson model. The model is described by an underlying intensity function λ that can be expressed as the product of two periodic intensity functions that model short-term (e.g., one year) and long-term (e.g., five years) fluctuations. Specifically, let the (minimal) period of the short-term intensity function λ_S be $\tau > 0$, and the (minimal) period of the long-term intensity function λ_L be $T = K\tau$, where $K \geq 1$ is the number of short-term periods inside one long-term period. Throughout the paper we assume that τ and K (and thus T) are known. This is not an unreasonable assumption given the underlying example that has inspired the Garrido-Lu model. Specifically, Lu (2005) notes that "from a graphical analysis of the data set, we conclude that a long-term period of five years and a short-term period of one year reasonably describe the Atlantic hurricanes".

The Garrido-Lu model is a special case of the general periodic Poisson model for which non-parametric estimators and statistical inferential results have been developed by Helmers, Mangku and Zitikis (2003, 2005). However, the more refined structure of the intensity function λ as described above has allowed Garrido and Lu (2004), Lu (2005), Lu and Garrido (2005, 2006) to achieve new depths in the area. The authors have also suggested several parametric families for the intensity functions λ_S and λ_L. Nevertheless, the successful research by J. Garrido and Y. Lu also shows that choosing well fitting parametric families is not a simple task. In view of this, in the present paper we suggest to have a look at the *non-parametric* approach to the problem, which supplements the aforementioned research by J. Garrido and Y. Lu with new insights into the problem and possible ways to tackle it. We next define the problem rigorously.

Let X be a Poisson point process on the real line with (unknown) locally integrable

intensity function λ. We assume that λ is periodic (or cyclic) with period $T > 0$; that is, we have the equation $\lambda(s) = \lambda(s + kT)$ for all real s and all integer k. We assume that $\lambda = \lambda_S \lambda_L$, where λ_S has a known period $\tau > 0$ and λ_L has the period $T = K\tau > 0$ with a known integer $K \geq 1$. As to the long-term intensity function λ_L, we additionally assume the following step-wise structure (cf. Lu, 2005; Lu and Garrido, 2005)

$$\lambda_L(s) = \sum_{k=1}^{T/\tau} a_k \mathbf{1}\{s \in W_k^0\} \ \text{ for all } \ s \in [0, T),$$

where $\mathbf{1}\{\cdot\}$ is the indicator function, $W_k^0 = [(k-1)\tau, k\tau)$, and $a_k \in (0, \infty)$ are unknown 'amplitudes'. The periodic extension of $\lambda_L(s)$ to the entire real line is accomplished by the equation $\lambda_L(s) = \lambda_L(s + kT)$ for all integer k. The amplitudes a_k are unknown, but we find it convenient to assume – which does not restrict the generality – that one of them, say a_1, is equal to 1. Indeed, writing the equation $\lambda(s) = \lambda_S^*(s) \sum_{k=1}^{K} a_k^* \mathbf{1}\{s \in W_k^0\}$ with $\lambda_S^*(s) = a_1 \lambda_S(s)$ and $a_k^* = a_k/a_1$, we obviously have the property $a_1^* = 1$. Skipping the stars from the above equations, we have (cf. eq. (11) in Lu and Garrido, 2005, p. 23)

$$\lambda(s) = \lambda_S(s) \sum_{k=1}^{T/\tau} a_k \mathbf{1}\{s \in W_k^0\} \ \text{ for all } \ s \in [0, T), \tag{1.1}$$

with the desired property $a_1 = 1$. The equation $\lambda(s) = \lambda(s + kT)$ for all integer k extends the above definition of $\lambda(s)$ to the entire real line. Hence, in view of $a_1 = 1$, the earlier made assumption about the minimality of the periods, and thus in particular of T, implies that $a_k \neq 1$ for every $2 \leq k \leq K$. We conclude the current paragraph with the note that, if desired, one can extend the above doubly-period model of λ to products of finite number of periodic intensity functions; in the present paper, however, we restrict ourselves to only the Garrido-Lu doubly-periodic model, as it has a solid practical justification.

The rest of the paper is organized as follows. In Section 2 we construct a consistent estimator for the intensity function $\lambda(s)$. In Section 3 we investigate statistical properties of the estimator, including its consistency, unbiasedness and asymptotic normality. Section 4 provides a number of auxiliary results that we later use in Section 5, which contains proofs of the statistical properties formulated in Section 3.

2 Constructing an estimator

Suppose that at time n we have a single realization $\{X(t; \omega) : t \in W_n\}$ of a Poisson process X in an interval $W_n \subset [0, \infty)$ usually called 'window', where ω is an element of the sample space. Let the intensity function of X be the above specified doubly-periodic λ. We want to construct a non-parametric estimator for $\lambda(s)$ at any given point s on the real line using the realization $\{X(t; \omega) : t \in W_n\}$. We assume that the windows $W_1, W_2, \cdots \subset \mathbf{R}$ are intervals of finite length, containing the origin $t = 0$, and such that their length $|W_n|$ increases indefinitely when $n \to \infty$. It is natural to expect that, in practice, many situations involve nested windows $W_n \subseteq W_{n+1}$, which means that we do not lose information, only accumulate it as the time n progresses.

Given that both τ and T are known, equation (1.1) implies that in order to construct an estimator for $\lambda(s)$ (cf. equation (2.5) below), we need to estimate the amplitudes a_k as well as the short-term intensity function λ_S. We start with estimating the amplitudes. Using equation (1.1) and the law of large numbers, we have that

$$a_k = \frac{\int_{W_k^0} \lambda(s)ds}{\int_{W_1^0} \lambda_S(s)ds} \approx a_{k,n}, \tag{2.2}$$

where

$$a_{k,n} := \frac{X(W_{k,n})}{X(W_{1,n})} \quad \text{and} \quad W_{k,n} = \bigcup_{j=-\infty}^{\infty} \left\{ \left(W_k^0 + jT \right) \cap W_n \right\}.$$

The ratio $a_{k,n}$ may not always be well defined as the numerator and/or denominator can be zero. Hence, we define an estimator of a_k as follows:

$$\widehat{a}_{k,n} = a_{k,n} \mathbf{1}\{a_{k,n} \in (0, \infty)\}.$$

Plugging $\widehat{a}_{k,n}$ instead of a_k in the definition of the long-term intensity function λ_L, we have $\lambda(s) \approx \lambda_n(s)$, where

$$\lambda_n(s) = \sum_{k=1}^{T/\tau} \widehat{a}_{k,n} \lambda_S(s_k) \mathbf{1}\left\{ s \in W_k^0 \right\} \tag{2.3}$$

with the notation $s_k = s - (k-1)\tau \in [0, \tau)$. Note that $\lambda_n(s)$ is neither an empirical estimator nor a population function; it is just an auxiliary function. To make this function an estimator, we estimate the short-term intensity λ_S, keeping in mind that we need to estimate it at the point $s_k \in [0, \tau)$ only, as the right-hand side of equation (2.3) suggests.

To this end, we aggregate information about $\lambda_S(s_k)$ from $[0, \tau)$ and also from other regions of the window W_n, as suggested by the equation $\lambda_S(s_k) = \lambda_S(s_k + j\tau)$. This way we obtain the first equation below:

$$\lambda_S(s_k) = \frac{\sum_{j=-\infty}^{\infty} \lambda_S(s_k + j\tau)\mathbf{1}\left\{s_k + j\tau \in W_n\right\}}{\sum_{j=-\infty}^{\infty} \mathbf{1}\left\{s_k + j\tau \in W_n\right\}}$$

$$\approx \frac{\tau}{|W_n|} \sum_{j=-\infty}^{\infty} \lambda_S(s_k + j\tau)\mathbf{1}\left\{s_k + j\tau \in W_n\right\}, \tag{2.4}$$

where the second (approximate) equation holds since $\sum_{j=-\infty}^{\infty} \mathbf{1}\left\{s_k + j\tau \in W_n\right\}$ is approximately $|W_n|/\tau$, which is the number of short-term periods in the window W_n. Next we replace $\lambda_S(s_k + j\tau)$ on the right-hand side of equation (2.4) by $\lambda(s_k + j\tau)$, which we can estimate from data. For this, we first rewrite the sum on the right-hand side of equation (2.4) by separating its short- and long-term related summands as follows:

$$\lambda_S(s_k) \approx \frac{\tau}{|W_n|} \sum_{j=-\infty}^{\infty} \sum_{m=1}^{T/\tau} \lambda_S(s_k + (m-1)\tau + jT)\mathbf{1}\left\{s_k + (m-1)\tau + jT \in W_n\right\}.$$

Since $s_k \in [0, \tau)$, equation (1.1) implies that $\lambda_S(s_k + (m-1)\tau + jT)$ equals $a_m^{-1}\lambda(s_k + (m-1)\tau + jT)$. Next we replace the amplitude a_m by its estimator $\widehat{a}_{m,n}$ in the latter quantity. These arguments lead to the approximation

$$\lambda_S(s_k) \approx \frac{\tau}{|W_n|} \sum_{m=1}^{T/\tau} \frac{1}{\widehat{a}_{m,n}} \sum_{j=-\infty}^{\infty} \lambda(s_k + (m-1)\tau + jT)\mathbf{1}\left\{s_k + (m-1)\tau + jT \in W_n\right\}.$$

Replacing $\lambda(s_k + (m-1)\tau + jT)$ by its empirical estimator, we arrive at an estimator of $\lambda_S(s_k)$:

$$\widehat{\lambda}_{S,n}(s_k) = \frac{\tau}{|W_n|} \sum_{m=1}^{T/\tau} \frac{1}{\widehat{a}_{m,n}} \sum_{j=-\infty}^{\infty} \frac{1}{2h_n} X\left(\left(s_k + (m-1)\tau + jT + [-h_n, h_n]\right) \cap W_n\right),$$

where $h_n \downarrow 0$ when $n \to \infty$. (The choice of h_n depends on how fast $|W_n|$ tends to infinity.) Plugging in $\widehat{\lambda}_{S,n}(s_k)$ instead of $\lambda_S(s_k)$ on the right-hand side of equation (2.3), we arrive at an estimator for $\lambda(s)$ at any $s \in [0, T)$:

$$\widehat{\lambda}_n(s) = \sum_{k=1}^{T/\tau} \widehat{a}_{k,n} \widehat{\lambda}_{S,n}(s_k)\mathbf{1}\left\{s \in W_k^0\right\}. \tag{2.5}$$

In the next section we investigate statistical properties of the estimator $\widehat{\lambda}_n(s)$, such as consistency, bias, mean-squared error, asymptotic normality.

As we have already noted above, the construction of the estimator $\widehat{\lambda}_n(s)$ assumes that the periods τ and T are known. Despite the reasonableness, for example, of the choices $\tau = 1$ and $T = 5$ in the aforementioned hurricane-related context, we can of course find situations where it is desirable to dispense with the assumption that the periods τ and T are known. Finding good estimators for the periods, however, is a challenging task. For hints and references on the topic, we refer to Vere-Jones (1982) for a periodogram-type estimator, and to Mangku (2001), Helmers and Mangku (2003), Bebbington and Zitikis (2004) for several non-parametric estimators. Naturally, when estimators of τ and T have been found, plugging them in on the right-hand side of equation (2.5) produces a desired modification of the underlying intensity function. This, of course, introduces further technical complexities, as seen from Helmers, Mangku and Zitikis (2003, 2005) where the case of the classical 'singly' periodic Poisson process is investigated.

3 Statistical properties

The intensity function λ is periodic, and its period T, by assumption, is known. Hence, when estimating the intensity function, we can and thus do restrict ourselves to only those s that are in the interval $[0, T)$. We also assume throughout the paper that s is a Lebesgue point of λ_S, which is a weak assumption meaning that $(2\delta)^{-1} \int_{-\delta}^{\delta} |\lambda_S(s+x) - \lambda_S(s)| dx \to 0$ when $\delta \downarrow 0$. Indeed, if the intensity function λ_S is locally integrable (which is minor requirement), then the set of all Lebesgue points of λ_S is dense in the real line.

Theorem 3.1 *If $h_n|W_n| \to \infty$, then $\widehat{\lambda}_n(s)$ is a (weakly) consistent estimator of $\lambda(s)$.*

To achieve asymptotic unbiasedness of the estimator $\widehat{\lambda}_n(s)$, we need to slightly modify it. Namely, we introduce (small) $\epsilon_n > 0$ such that $a_k \in (\epsilon_n, \epsilon_n^{-1})$ for every $1 \leq k \leq K$ and for all sufficiently large n. The only (minor) restriction that we impose on $\epsilon_n > 0$ is the existence of $L \geq 1$ such that $\epsilon_n|W_n|^L \to \infty$ when $n \to \infty$. Hence, if we think about the window W_n as the interval $[0, n]$, then can simply choose $\epsilon_n = 1/n$. Of course, if we know (e.g., from historical data) a sufficiently small $\epsilon > 0$ such that all the amplitudes a_k are in the interval $(\epsilon, \epsilon^{-1})$, then we can take $\epsilon_n = \epsilon$. Hence, assuming that we have decided on the choice of ϵ_n, our new estimator of a_k is

$$\widetilde{a}_{k,n} = \frac{X(W_{k,n})}{X(W_{1,n})} \mathbf{1}\left\{ \frac{X(W_{k,n})}{X(W_{1,n})} \in (\epsilon_n, \epsilon_n^{-1}) \right\}.$$

In turn, the estimator of $\lambda(s)$ becomes

$$\widetilde{\lambda}_n(s) = \sum_{k=1}^{T/\tau} \widetilde{a}_{k,n} \, \widetilde{\lambda}_{S,n}(s_k) \mathbf{1}\left\{s \in W_k^0\right\},$$

where

$$\widetilde{\lambda}_{S,n}(s_k) = \frac{\tau}{|W_n|} \sum_{m=1}^{T/\tau} \frac{1}{\widetilde{a}_{m,n}} \sum_{j=-\infty}^{\infty} \frac{1}{2h_n} X\Big(\big(s_k + (m-1)\tau + jT + [-h_n, h_n]\big) \cap W_n\Big).$$

Theorem 3.2 *Let* $h_n^4 |W_n| \to \infty$, *and let the second derivative* $\lambda_S''(s)$ *be finite. Then*

$$\mathbf{E}[\widetilde{\lambda}_n(s)] = \lambda(s) + \frac{1}{6}\lambda_S''(s)\lambda_L(s)h_n^2 + o(h_n^2). \tag{3.6}$$

Statement (3.6) suggests choosing h_n as fast converging to 0 as possible, as long as $h_n^4 |W_n| \to \infty$ holds. However, the variance of the estimator has to be taken into account as well, which is the topic of the next theorem. We shall see later in this section that the optimal (i.e., minimizing the mean-squared error) bandwidth is $h_n = \mathcal{O}(|W_n|^{-1/5})$.

Theorem 3.3 *If* $h_n |W_n| \to \infty$, *then*

$$\mathbf{Var}[\widetilde{\lambda}_n(s)] = \frac{\tau\lambda(s)\lambda_L(s)}{2h_n|W_n|}\left(\frac{1}{T/\tau}\sum_{m=1}^{T/\tau}\frac{1}{a_m}\right) + o\left(\frac{1}{h_n|W_n|}\right). \tag{3.7}$$

Combining statements (3.6) and (3.7), and also noting that $\lambda(s) = \lambda_S(s)\lambda_L(s)$, we see that the mean-squared error of the estimator $\widetilde{\lambda}_n(s)$ is minimal with respect to h_n when

$$h_n = \frac{c_0^{1/5}}{|W_n|^{1/5}} \quad \text{with} \quad c_0 = \frac{9\tau\lambda_S(s)}{2(\lambda_S''(s))^2}\left(\frac{1}{T/\tau}\sum_{i=1}^{T/\tau}\frac{1}{a_i}\right). \tag{3.8}$$

The average on the right-hand side of the definition of c_0 is the harmonic mean of the amplitudes a_i. If all the amplitudes are equal to 1, then the harmonic mean is equal to 1 and thus statement (3.7) reduces to a special case of statement (3.4) in Helmers, Mangku and Zitikis (2005).

Corollary 3.1 *Let* $h_n^4 |W_n| \to \infty$, *and let* λ_S *be twice differentiable at* $s \in [0,T)$. *If, in addition to the above, we have that* $h_n^5 |W_n| = O(1)$, *then*

$$\sqrt{\frac{h_n|W_n|}{d_0(s)}}\left(\widetilde{\lambda}_n(s) - \lambda(s) - \frac{1}{6}\lambda_S''(s)\lambda_L(s)h_n^2\right) \to_d \mathcal{N}(0,1),$$

where

$$d_0(s) = \frac{\tau\lambda(s)\lambda_L(s)}{2}\left(\frac{1}{T/\tau}\sum_{m=1}^{T/\tau}\frac{1}{a_m}\right).$$

The corollary can be used for constructing confidence intervals for, and test hypotheses about, $\lambda(s)$. It is useful to note in this regard that under the conditions of Corollary 3.1 and with the optimal h_n given by equations (3.8), we have that

$$\sqrt{\frac{h_n|W_n|}{d_0(s)}} \left(\widetilde{\lambda}_n(s) - \lambda(s)\right) \to_d \mathcal{N}(\mu, 1) \tag{3.9}$$

with $\mu = 1/2$. If, however, h_n is such that $h_n^5|W_n| = o(1)$, then the limiting distribution becomes standard normal $\mathcal{N}(0, 1)$, that is, statement (3.9) holds with $\mu = 0$.

4 Auxiliary results

Based on the construction of the estimator $\widehat{\lambda}_n(s)$, we derive its asymptotic properties by establishing appropriate convergence rates of $\widehat{a}_{k,n}$ to a_k and $\widehat{\lambda}_{S,n}(s_k)$ to $\lambda_S(s_k)$ when $n \to \infty$. Furthermore, since the estimator $\widehat{\lambda}_{S,n}(s_k)$ contains randomness in both $\widehat{a}_{m,n}$ and $X(\dots)$, we find it convenient to separate them using the auxiliary function

$$\lambda_{S,n}(z) = \frac{\tau}{|W_n|} \sum_{m=1}^{T/\tau} \frac{1}{a_m} \sum_{j=-\infty}^{\infty} \frac{1}{2h_n} X\Big(\big(z + (m-1)\tau + jT + [-h_n, h_n]\big) \cap W_n \Big) \tag{4.10}$$

defined on the interval $[0, \tau)$. Specifically, in the proofs below we frequently use the decomposition

$$\widehat{\lambda}_n(s) = \sum_{k=1}^{T/\tau} (\widehat{a}_{k,n} - a_k)\, \widehat{\lambda}_{S,n}(s_k)\, \mathbf{1}\{s \in W_k^0\} + \sum_{k=1}^{T/\tau} a_k \left(\widehat{\lambda}_{S,n}(s_k) - \lambda_{S,n}(s_k) \right) \mathbf{1}\{s \in W_k^0\}$$

$$+ \sum_{k=1}^{T/\tau} a_k\, \lambda_{S,n}(s_k)\mathbf{1}\{s \in W_k^0\} \tag{4.11}$$

(or an analogous one with 'hats' replaced by 'tildes') and show that the first two sums on the right-hand side of equation (4.11) are asymptotically smaller than the third one. To accomplish this, we need a rate of convergence of $\widehat{a}_{k,n}$ to a_k; hence, the next lemma.

Lemma 4.1 *For every $\gamma < 1/2$ and $1 \le k \le K$, we have $\widehat{a}_{k,n} = a_k + o_{\mathbf{P}}(|W_n|^{-\gamma})$.*

Proof. We only need to show that $a_{k,n} = a_k + o_{\mathbf{P}}(|W_n|^{-\gamma})$. Write $a_{k,n}$ as the ratio ξ_k/ξ_1 with the notation $\xi_k = KX(W_{k,n})/|W_n|$, and then express a_k as the ratio of $a_k\theta$ and $a_1\theta$ (recall that $a_1 = 1$) with $\theta = \tau^{-1} \int_0^\tau \lambda_S(s)ds$. Hence, $a_{k,n} = a_k + o_{\mathbf{P}}(|W_n|^{-\gamma})$

holds provided that $|W_n|^\gamma (\xi_k - a_k\theta) \to_{\mathbf{P}} 0$. The latter statement, in turn, follows from $|W_n|^\gamma(\xi_k - \mathbf{E}[\xi_k]) \to_{\mathbf{P}} 0$ and $|W_n|^\gamma(\mathbf{E}[\xi_k] - a_k\theta) \to 0$. The former statement follows from the latter. Indeed, using the Chebyshev inequality together with the fact that the mean and the variance of $X(W_{k,n})$ coincide, we have the bound

$$\mathbf{P}\left[|W_n|^\gamma|\xi_k - \mathbf{E}[\xi_k]| > \epsilon\right] \le \frac{K}{\epsilon^2 |W_n|^{1-2\gamma}}\left(a_k\theta + \mathcal{O}\left(\frac{1}{|W_n|^\gamma}\right)\right). \tag{4.12}$$

The right-hand side of bound (4.12) converges to 0 since $\gamma < 1/2$ and $|W_n| \to \infty$. Consequently, we complete the proof of the lemma by showing that $|W_n|^\gamma(\mathbf{E}[\xi_k] - a_k\theta) \to 0$. In fact, we next prove (and use later) a stronger statement, namely, $|W_n|(\mathbf{E}[\xi_k] - a_k\theta) \to 0$, which we start proving with the equations

$$\mathbf{E}[\xi_k] = \frac{K}{|W_n|} \sum_{j=-\infty}^{\infty} \mathbf{E}\big[X\left([(k-1+jK)\tau, (k+jK)\tau] \cap W_n\right)\big]$$

$$= \frac{K}{|W_n|} \sum_{j=-\infty}^{\infty} \int_{W_k^0} \lambda(x+jT)\mathbf{1}\{x+jT \in W_n\}dx.$$

Note that $\lambda(x+jT)$ is equal to $\lambda(x)$, which is $a_k\lambda_S(x)$ for $x \in W_k^0$. Furthermore, uniformly in x, the sum $\sum_{j=-\infty}^{\infty} \mathbf{1}\{x+jT \in W_n\}$ is equal to $|W_n|/T + \mathcal{O}(1)$. Consequently,

$$\mathbf{E}[\xi_k] = \frac{T}{|W_n|}\left(\frac{1}{\tau}\int_{W_k^0} a_k\lambda_S(x)dx\right)\left(\frac{|W_n|}{T} + \mathcal{O}(1)\right),$$

which implies $|W_n|(\mathbf{E}[\xi_k] - a_k\theta) \to 0$ and completes the proof of Lemma 4.1. $\qquad\square$

Lemma 4.2 *Let $h_n|W_n| \to \infty$, and let $z \in [0, \tau)$ be a Lebesgue point of λ_S. Then*

$$\mathbf{E}[\lambda_{S,n}(z)] = \lambda_S(z) + o(1) \tag{4.13}$$

and

$$\mathbf{Var}[\lambda_{S,n}(z)] = \frac{\tau\lambda_S(z)}{2h_n|W_n|}\left(\frac{1}{T/\tau}\sum_{m=1}^{T/\tau}\frac{1}{a_m}\right) + o\left(\frac{1}{h_n|W_n|}\right). \tag{4.14}$$

Proof. We start with statement (4.13). The expectation $\mathbf{E}[\lambda_{S,n}(z)]$ is equal to

$$\frac{\tau}{|W_n|}\sum_{m=1}^{T/\tau}\frac{1}{a_m 2h_n}\int_{-h_n}^{h_n}\lambda(x+z+(m-1)\tau)\sum_{j=-\infty}^{\infty}\mathbf{1}\{x+z+(m-1)\tau+jT \in W_n\}dx,$$

where we also used the fact that the period of λ is T. Uniformly in x, the inner sum (with respect to j) is $|W_n|/T + \mathcal{O}(1)$. Furthermore, since $z \in [0, \tau)$ is a Lebesgue point

9

of λ_S, so is the point $z + (m-1)\tau$. Hence, $(2h_n)^{-1} \int_{-h_n}^{h_n} \lambda(x + z + (m-1)\tau)dx$ equals $\lambda(z + (m-1)\tau) + o(1)$, which is $a_m \lambda_S(z) + o(1)$ since $z + (m-1)\tau \in W_m^0$. Hence,

$$\mathbf{E}[\lambda_{S,n}(z)] = \frac{\tau}{|W_n|} \sum_{m=1}^{T/\tau} \frac{1}{a_m} \left(a_m \lambda_S(z) + o(1)\right) \left(\frac{|W_n|}{T} + \mathcal{O}(1)\right),$$

which is $\lambda_S(z)+o(1)$. This completes the proof of statement (4.13). The proof of statement (4.14) is similar. Assuming that n is sufficiently large so that h_n is small and the summands of $\lambda_{S,n}(s)$ are independent, the variance $\mathbf{Var}[\lambda_{S,n}(s)]$ can be written as

$$\frac{\tau^2}{2h_n|W_n|^2} \sum_{m=1}^{T/\tau} \frac{1}{a_m^2 2h_n} \int_{-h_n}^{h_n} \lambda(x + z + (m-1)\tau) \sum_{j=-\infty}^{\infty} \mathbf{1}\{x + z + (m-1)\tau + jT \in W_n\}dx.$$

Using analogous arguments to those in the proof of statement (4.13), we have that

$$\mathbf{Var}[\lambda_{S,n}(s)] = \frac{\tau^2}{2h_n|W_n|^2} \sum_{m=1}^{T/\tau} \frac{1}{a_m^2} \left(a_m \lambda_S(z) + o(1)\right) \left(\frac{|W_n|}{T} + \mathcal{O}(1)\right),$$

which completes the proof of statement (4.14). $\qquad\square$

Lemma 4.3 *Let $h_n^2|W_n| \to \infty$, and let λ_S be twice differentiable at $z \in [0, \tau)$. Then*

$$\mathbf{E}[\lambda_{S,n}(z)] = \lambda_S(z) + \frac{1}{6}\lambda_S''(z)h_n^2 + o(h_n^2).$$

Proof. The Taylor expansion gives the equation

$$\frac{1}{2h_n} \int_{-h_n}^{h_n} \lambda(x + z + (m-1)\tau)dx = \lambda(z + (m-1)\tau) + \frac{1}{6}\lambda''(z + (m-1)\tau)h_n^2 + o(h_n^2).$$

The right-hand side of the equations is equal to $a_m \lambda_S(z) + 6^{-1}a_m \lambda_S''(s)h_n^2 + o(h_n^2)$. Hence, proceeding as in the proof of Lemma 4.2, we have that

$$\mathbf{E}[\lambda_{S,n}(z)] = \frac{\tau}{|W_n|} \sum_{m=1}^{T/\tau} \frac{1}{a_m} \left(a_m \lambda_S(z) + \frac{1}{6}a_m \lambda_S''(z)h_n^2 + o(h_n^2)\right) \left(\frac{|W_n|}{T} + \mathcal{O}(1)\right)$$

$$= \lambda_S(z) + \frac{1}{6}\lambda_S''(z)h_n^2 + o(h_n^2) + \mathcal{O}\left(\frac{1}{|W_n|}\right).$$

Since $h_n^2|W_n| \to \infty$, the remainder term $\mathcal{O}(|W_n|^{-1})$ is $o(h_n^2)$. This finishes the proof of Lemma 4.3. $\qquad\square$

The classical Lindeberg theorem and Slutzky's arguments (cf., e.g., Serfling, 1980) together with Lemmas 4.2 and 4.3 imply the following corollary.

Corollary 4.1 *Let $h_n^2|W_n| \to \infty$, and let λ_S be twice differentiable at $z \in [0, \tau)$. If $h_n^5|W_n| = O(1)$, then*

$$\sqrt{\frac{h_n|W_n|}{D(z)}} \left(\lambda_{S,n}(z) - \lambda_S(z) - \frac{1}{6}\lambda_S''(z)h_n^2 \right) \to_d \mathcal{N}(0, 1),$$

where

$$D(z) = \frac{\tau \lambda_S(z)}{2} \left(\frac{1}{T/\tau} \sum_{m=1}^{T/\tau} \frac{1}{a_m} \right).$$

Corollary 4.1 is tailored to accommodate the optimal choice of h_n, which, by definition, minimizes the mean-squared error (MSE) of $\lambda_{S,n}(z)$. The MSE is the sum of the variance and the squared bias, which are $\mathbf{Var}[\lambda_{S,n}(z)]$ and $\mathbf{Bias}[\lambda_{S,n}(z)] = \mathbf{E}[\lambda_{S,n}(z)] - \lambda_S(z)$, respectively. Hence, Lemmas 4.2 and 4.3 imply that the optimal h_n is

$$h_n = \frac{d_S(z)^{1/5}}{|W_n|^{1/5}} \quad \text{with} \quad d_S(z) = \frac{9\tau \lambda_S(z)}{2(\lambda_S''(z))^2} \left(\frac{1}{T/\tau} \sum_{i=1}^{T/\tau} \frac{1}{a_i} \right).$$

Given this choice of the bandwidth h_n, we have the asymptotic normality

$$\sqrt{\frac{h_n|W_n|}{D(z)}} \left(\lambda_{S,n}(z) - \lambda_S(z) \right) \to_d \mathcal{N}(\mu, 1) \tag{4.15}$$

with $\mu = 1/2$. If, however, we choose $h_n^5|W_n| = o(1)$, then $(1/6)\lambda_S''(z)h_n^2$ can be dropped from the statement of Corollary 4.1; equivalently, we have statement (4.15) with $\mu = 0$.

5 Proofs

Proof of Theorem 3.1 To prove Theorem 3.1, it suffices to show (cf. equation (4.11)) that for every $1 \le k \le K$ and when $n \to \infty$, the following three statements hold: A) $(\widehat{a}_{k,n} - a_k)\widehat{\lambda}_{S,n}(s_k) \to_{\mathbf{P}} 0$, B) $\widehat{\lambda}_{S,n}(s_k) - \lambda_{S,n}(s_k) \to_{\mathbf{P}} 0$, and C) $\lambda_{S,n}(s_k) \to_{\mathbf{P}} \lambda_S(s_k)$. Lemma 4.2 implies statement C). Statement C) and Lemma 4.1 imply B). Statements B) and C) give $\widehat{\lambda}_{S,n}(s_k) = \mathcal{O}_{\mathbf{P}}(1)$, which reduces the proof of statement A) to showing that $\widehat{a}_{k,n} \to_{\mathbf{P}} a_k$. The latter follows from Lemma 4.1, completing the proof of Theorem 3.1. $\square$

Proof of Theorem 3.2 We start with equation (4.11) and replace all the 'hats' by 'tildes'. By Lemma 4.3, the expectation of the right-most sum in equation (4.11) equals the right-hand side of equation (3.6). Hence, we are left to check that the expected values of the first

two sums on the right-hand side of equation (4.11) are of the order $o(h_n^2)$. The expectation of the first sum is such if $\mathbf{E}[|\widetilde{a}_{k,n} - a_k|\widetilde{\lambda}_{S,n}(s_k)] = o(h_n^2)$. To prove this statement, we use the Cauchy-Schwarz inequality and reduce the problem to showing that $\mathbf{E}[(\widetilde{a}_{k,n} - a_k)^2] = o(h_n^4)$ and $\mathbf{E}[\widetilde{\lambda}_{S,n}^2(s_k)] = \mathcal{O}(1)$. To prove that the expectation $\mathbf{E}[(\widetilde{a}_{k,n} - a_k)^2]$ is of the order $o(h_n^4)$, we write the expectation as the sum of $a_k^2 \mathbf{P}[a_{k,n} \notin (\epsilon_n, \epsilon_n^{-1})]$ and $\mathbf{E}[(a_{k,n} - a_k)^2 \mathbf{1}\{a_{k,n} \in (\epsilon_n, \epsilon_n^{-1})\}]$. We need to show that the latter probability and expectation are of the order $o(h_n^4)$. We start with the probability and write (cf. the proof of Lemma 4.1 for detail; recall the notation of $\xi_k = KX(W_{k,n})/|W_n|$)

$$a_{k,n} = \frac{a_k\theta + \mathcal{O}(|W_n|^{-1}) + (\xi_k - \mathbf{E}[\xi_k])}{\theta + \mathcal{O}(|W_n|^{-1}) + (\xi_1 - \mathbf{E}[\xi_1])}. \tag{5.16}$$

If both $|\xi_k - \mathbf{E}[\xi_k]|$ and $|\xi_1 - \mathbf{E}[\xi_1]|$ do not exceed a sufficiently small $\delta > 0$, then for all large n, we have from equation (5.16) that $a_{k,n} \in (\epsilon_n, \epsilon_n^{-1})$, since the latter open interval contains all the amplitudes a_k. Hence, $\mathbf{P}[a_{k,n} \notin (\epsilon_n, \epsilon_n^{-1})]$ is of the order $o(h_n^4)$ if both $\mathbf{P}[|\xi_k - \mathbf{E}[\xi_k]| \geq \delta]$ and $\mathbf{P}[|\xi_1 - \mathbf{E}[\xi_1]| \geq \delta]$ are of the same order, which follows from bound (4.12) with $\gamma = 0$ since $h_n^4|W_n| \to \infty$. Consequently, we are left to show that the expectation $\mathbf{E}[(a_{k,n} - a_k)^2 \mathbf{1}\{a_{k,n} \in (\epsilon_n, \epsilon_n^{-1})\}]$ is of the order $o(h_n^4)$, which reduces to proving the statements

$$\mathbf{E}\Big[(a_{k,n} - a_k)^2 \mathbf{1}\{a_{k,n} \in (\epsilon_n, \epsilon_n^{-1})\}\mathbf{1}\{|\xi_1 - \mathbf{E}[\xi_1]| \geq \theta/2\}\Big] = o(h_n^4)$$

and

$$\mathbf{E}\Big[(a_{k,n} - a_k)^2 \mathbf{1}\{|\xi_1 - \mathbf{E}[\xi_1]| \leq \theta/2\}\Big] = o(h_n^4). \tag{5.17}$$

The first statement holds since $\epsilon_n^{-2}\mathbf{P}[|\xi_1 - \mathbf{E}[\xi_1]| \geq \theta/2] = \mathcal{O}(\epsilon_n^{-2}|W_n|^{-L})$ for any $L \geq 1$, which follows from: the Markov inequality, the Poisson distribution of $X(W_{k,n})$, and the assumption $h_n^4|W_n| \to \infty$. To verify statement (5.17), we use equation (5.16) and have

$$|a_{k,n} - a_k|\mathbf{1}\{|\xi_1 - \mathbf{E}[\xi_1]| \leq \theta/2\} \leq \frac{\mathcal{O}(|W_n|^{-1}) + |\xi_k - \mathbf{E}[\xi_k]| + a_k|\xi_1 - \mathbf{E}[\xi_1]|}{\theta/2 + \mathcal{O}(|W_n|^{-1})}. \tag{5.18}$$

The second moment of the right-hand side of bound (5.18) is of the order $\mathcal{O}(|W_n|^{-1})$, which implies statement (5.17). Hence, the statement $\mathbf{E}[(\widetilde{a}_{k,n} - a_k)^2] = o(h_n^4)$ holds. To complete the proof of $\mathbf{E}[(\widetilde{a}_{k,n} - a_k)\widetilde{\lambda}_{S,n}(s_k)] = o(h_n^2)$, we still need to show that $\mathbf{E}[\widetilde{\lambda}_{S,n}^2(s_k)] = \mathcal{O}(1)$. To this end, we next prove that $\mathbf{E}[(\widetilde{\lambda}_{S,n}(s_k) - \lambda_{S,n}(s_k))^2] = \mathcal{O}(1)$ and $\mathbf{E}[\lambda_{S,n}^2(s_k)] = \mathcal{O}(1)$. The latter statement immediately follows from Lemma 4.2. To prove that the former statement holds, we use the Cauchy-Schwarz inequality and see that it follows

from $\mathbf{E}[(\widetilde{a}_{i,n}^{-1} - a_i^{-1})^2] = o(h_n^4)$. The latter statement can be established similarly to the eaerlier statement $\mathbf{E}[(\widetilde{a}_{k,n} - a_k)^2] = o(h_n^4)$, as the only difference between the two statements is the reversed roles of $X(W_{k,n})$ and $X(W_{1,n})$. With this note we conclude the proof that the expectation of the first sum on the right-hand side of equation (4.11) is of the order $o(h_n^2)$. To show that the expectation of the second sum is of the order $o(h_n^2)$, we need to verify the statement $\mathbf{E}[|\widetilde{\lambda}_{S,n}(s_k) - \lambda_{S,n}(s_k)|] = o(h_n^2)$, which follows from the Cauchy-Schwarz inequality and the already proved statement $\mathbf{E}[\lambda_{S,n}^2(s_k)] = \mathcal{O}(1)$. The proof of Theorem 3.2 is now complete. $\qquad\square$

Proof of Theorem 3.3 We start with equation (4.11) and replace all 'hats' by 'tildes'. By Lemma 4.2, the variance of the right-most sum in equation (4.11) is equal to the right-hand side of equation (3.7). Hence, Theorem 3.3 follows if $\mathbf{E}[(\widetilde{a}_{k,n} - a_k)^2\widetilde{\lambda}_{S,n}^2(s_k)]$ and $\mathbf{E}[(\widetilde{\lambda}_{S,n}(s_k) - \lambda_{S,n}(s_k))^2]$ are of the order $o(h_n^{-1}|W_n|^{-1})$ when $n \to \infty$. To verify the first of these conditions, we reduce it to showing that $\mathbf{E}[(\widetilde{a}_{k,n} - a_k)^4] = \mathcal{O}(|W_n|^{-2})$ and $\mathbf{E}[\widetilde{\lambda}_{S,n}^4(s_k)] = \mathcal{O}(1)$. Analogous arguments to those in the proof of Theorem 3.2 show, for example, that $\mathbf{E}[(\widetilde{a}_{k,n} - a_k)^4] = \mathcal{O}(|W_n|^{-2})$ holds if the fourth moment of the right-hand side of bound (5.18) is of the order $\mathcal{O}(|W_n|^{-2})$, which follows easily from the fact that the fourth moments of the variable $\xi_k - \mathbf{E}[\xi_k]$ is of the order $\mathcal{O}(|W_n|^{-2})$. In summary, we have that the expectation $\mathbf{E}[(\widetilde{a}_{k,n} - a_k)^2\widetilde{\lambda}_{S,n}^2(s_k)]$ is of the order $\mathcal{O}(|W_n|^{-1})$, and hence $o(h_n^{-1}|W_n|^{-1})$. We are left to prove $\mathbf{E}[(\widetilde{\lambda}_{S,n}(s_k) - \lambda_{S,n}(s_k))^2] = o(h_n^{-1}|W_n|^{-1})$. Using analogous arguments to those in the proof of Theorem 3.2, we have that $\mathbf{E}[(\widetilde{a}_{i,n}^{-1} - a_i^{-1})^4] = \mathcal{O}(|W_n|^{-2})$ and $\mathbf{E}[\lambda_{S,n}^4(s_k)] = \mathcal{O}(1)$. Combining these facts with the Cauchy-Schwarz inequality, we obtain that the expectation $\mathbf{E}[(\widetilde{\lambda}_{S,n}(s_k) - \lambda_{S,n}(s_k))^2]$ is of the order $\mathcal{O}(|W_n|^{-1})$, and hence $o(h_n^{-1}|W_n|^{-1})$. This completes the proof of Theorem 3.3. $\qquad\square$

Proof of Corollary 3.1. The corollary follows from equation (4.11) with 'hats' replaced by 'tildes' and the following three statements: 1) $\sqrt{h_n|W_n|}\ \mathbf{E}[|\widetilde{a}_{k,n} - a_k|\widetilde{\lambda}_{S,n}(s_k)] \to_{\mathbf{P}} 0$, which follows from the earlier proved (cf. the proof of Theorem 3.2) $\mathbf{E}[|\widetilde{a}_{k,n} - a_k|\widetilde{\lambda}_{S,n}(s_k)] = o(h_n^2)$ since $h_n^5|W_n| = O(1)$ by assumption; 2) $\sqrt{h_n|W_n|}\ \mathbf{E}[|\widetilde{\lambda}_{S,n}(s_k) - \lambda_{S,n}(s_k)|] \to_{\mathbf{P}} 0$, which follows from the earlier proved (cf. the proof of Theorem 3.2) $\mathbf{E}[|\widetilde{\lambda}_{S,n}(s_k) - \lambda_{S,n}(s_k)|] = o(h_n^2)$ since $h_n^5|W_n| = O(1)$; and 3)

$$\sqrt{\frac{h_n|W_n|}{d_0(s)}}\left(\sum_{k=1}^{T/\tau} a_k\,\lambda_{S,n}(s_k)\mathbf{1}\{s \in W_k^0\} - \lambda(s) - \frac{1}{6}\lambda_S''(s)\lambda_L(s)h_n^2\right) \to_d \mathcal{N}(0,1).$$

13

To verify statement 3), we note that $s \in [0, T)$ uniquely determines k, which we denote by κ. Hence, statement 3) is equivalent to

$$\sqrt{\frac{h_n|W_n|}{D(s_\kappa)}} \left(\lambda_{S,n}(s_\kappa) - \lambda_S(s_\kappa) - \frac{1}{6}\lambda_S''(s_\kappa)h_n^2 \right) \rightarrow_d \mathcal{N}(0, 1),$$

which follows from Corollary 4.1 with $z = s_\kappa$. The proof of Corollary 3.1 is complete. $\square$

Acknowledgments

References

Bebbington, M. and Zitikis, R. (2004). A robust heuristic estimator for the period of a Poisson intensity function. *Methodology and Computing in Applied Probability*, **6**, 441-462.

Cox, D.R. and Isham, V. (1980). *Point Processes.* Chapman and Hall, London.

Cox, D.R. and Lewis, P.A.W. (1978). *The Statistical Analysis of Series of Events.* Chapman and Hall, London.

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

Daley, D.J. and Vere-Jones, D. (1988). *Introduction to the Theory of Point Processes.* Springer, New York.

Diggle, P.J. (1983). *Statistical Analysis of Spatial Point Processes.* Academic Press, London.

Garrido, J. and Lu, Y. (2004). On double periodic non-homogeneous Poisson processes. *Bulletin of the Association of Swiss Actuaries*, **2**, 195-212.

Helmers, R. and Mangku, I.W. (2003). On estimating the period of a cyclic Poisson process. In: *Mathematical Statistics and Applications: Festschrift for Constance van Eeden,* 345-356, Inst. Math. Statist., Beachwood, OH.

Helmers, R., Mangku, I.W. and Zitikis, R. (2003). Consistent estimation of the intensity function of a cyclic Poisson process. *J. Multivariate Anal.* **84** (2003), 19-39.

Helmers, R., Mangku, I.W. and Zitikis, R. (2005). Statistical properties of a kernel-type estimator of the intensity function of a cyclic Poisson process. *J. Multivariate Anal.*, **92**, 1-23.

Karr, A.F. (1986). *Point Processes and their Statistical Inference.* Marcel Dekker, New York.

Kingman, J.F.C. (1993). *Poisson Processes*, Clarendon Press, Oxford.

Kutoyants, Y. A. (1998). *Statistical Inference for Spatial Poisson Processes.* Springer, New York.

Lewis, P.A.W. (editor), (1972). *Stochastic Point Processes.* Wiley, New York.

Lu, Y. (2005). *On Periodic and Markovian Non-homogeneous Poisson Processes and their Application in Risk Theory* (Ph.D. Thesis), Concordia University, Montreal.

Lu, Y. and Garrido, J. (2005). Doubly periodic non-homogeneous Poisson models for hurricane data. *Statistical Methodology*, **2**, 17-35.

Lu, Y. and Garrido, J. (2006). Regime-switching periodic non-homogeneous Poisson processes. *North American Actuarial Journal*, **10** (4), 235-248.

Mangku, I.W. (2001). *Estimating the Intensity of a Cyclic Poisson Process* (Ph.D. Thesis), University of Amsterdam, Amsterdam.

Reiss, R.-D. (1993). *A Course on Point Processes.* Springer, New York.

Serfling, R.J. (1980). *Approximation Theorems of Mathematical Statistics*. Wiley, New York.

Snyder, D.L. and Miller, M.I. (1995). *Random Point Processes in Time and Space.* Springer, New York.

Vere-Jones, D. (1982). On the estimation of frequency in point-process data. *J. Appl. Prob.* 19A, 383–394.