Centrum voor Wiskunde en Informatica
REPORT RAPPORT

Consistent Estimation of the Intensity Function of a Cyclic Poisson Process

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

Probability, Networks and Algorithms (PNA)

PNA-R9915 December 31, 1999

Report PNA-R9915
ISSN 1386-3711

CWI
P.O. Box 94079
1090 GB Amsterdam
The Netherlands

Consistent Estimation of the Intensity Function of a Cyclic Poisson Process

Roelof Helmers
CWI
P.O. Box 94079, 1090 GB Amsterdam, The Netherlands

I Wayan Mangku
Department of Mathematics, Bogor Agricultural University,
Jl. Raya Pajajaran Bogor, 16144 Indonesia

Ričardas Zitikis
Department of Statistics, University of Manitoba,
Winnipeg, Canada R3T 2N2
and
Laboratory for Research in Statistics and Probability,
Carleton University, Ottawa, Canada K1S 5B6

ABSTRACT
We construct and investigate a consistent kernel-type nonparametric estimator of the intensity function of a cyclic Poisson process when the period is unknown. We assume that only a single realization of the Poisson process is observed in a bounded window. In particular, we prove that the proposed estimator is weakly and strongly consistent when the size of the window expands.

1991 Mathematics Subject Classification: 60G55, 62G05, 62G20.
Keywords and Phrases: Poisson process, cyclic Poisson process, periodic Poisson process, point process, cyclic intensity function, periodic intensity function, nonparametric estimation, consistency, rate of consistency.
Note: Work carried out under project PNA3.2. The second author was supported by a cooperation project between The Netherlands and Indonesia on "Applied Mathematics and Computational Methods". The third author was partially supported by the NSERC Canada Collaborative Grant at the Laboratory in Research in Statistics and Probability, Ottawa, as well as by The Netherlands Organization for Scientific Research (NWO). This paper is submitted for publication.

1. Introduction

Let X be a Poisson point process in $\mathbf{R}$ with (unknown) locally integrable intensity function λ which is assumed throughout to be periodic or, in other words, cyclic with (unknown) period $\tau \in \mathbf{R}^+$, that is

$$\lambda(s + k\tau) = \lambda(s) \tag{1.1}$$

for all $s \in \mathbf{R}$ and $k \in \mathbf{Z}$. Furthermore, let $W_1,\ W_2, \ldots$ be a sequence of intervals of $\mathbf{R}$, called windows, such that the size or the Lebesgue measure $|W_n|$ of W_n is finite for each fixed $n \in \mathbf{N}$, but

$$|W_n| \to \infty, \tag{1.2}$$

as $n \to \infty$. (In order to make the paper shorter, from now on we are to suppress "$n \to \infty$" whenever confusion is unlikely.)

Suppose now that, for some $\omega \in \Omega$, a single realization $X(\omega)$ of the cyclic Poisson process X is observed, though only within a bounded interval, called 'window' $W \subset \mathbf{R}$. Our goal in this paper is to construct a consistent non-parametric estimator of the intensity function λ at a given point $s \in \mathbf{R}$

from a single realization $X(\omega)$ of the Poisson process X observed in $W := W_n$. The requirement $s \in W_n$ can be dropped when we know the period τ.

If it is not stated otherwise, we assume throughout that s is a Lebesgue point of λ. This assumption appears to be a mild one since, due to the local integrability of λ, the set of all Lebesgue points of λ is dense in $\mathbf{R}$. Despite the latter observation, however, right after Theorem 1.2 below we discuss possible results without assuming that $s \in \mathbf{R}$ is a Lebesgue point of λ.

In order to give the definition of our estimator of $\lambda(s)$, we need to introduce further notations. Let $\hat{\tau}_n$ be any consistent estimator of the period τ, that is,

$$\hat{\tau}_n \xrightarrow{p} \tau.$$

For example, one may use the estimators constructed by Helmers and Mangku [3] or Vere-Jones [4].

Furthermore, let $K : \mathbf{R} \to \mathbf{R}$ be a function, called kernel, satisfying assumptions:

(K.1) K is a probability density function,

(K.2) K is bounded,

(K.3) K has support in $[-1, 1]$.

With the above introduced notations, we now define the estimator of $\lambda(s)$ as

$$\hat{\lambda}_{n,K}(s) := \frac{\hat{\tau}_n}{|W_n|} \sum_{k=-\infty}^{\infty} \frac{1}{h_n} \int_{W_n} K\left(\frac{x-(s+k\hat{\tau}_n)}{h_n}\right) X(dx), \tag{1.3}$$

where h_n is a sequence of positive real numbers converging to 0, that is,

$$h_n \downarrow 0. \tag{1.4}$$

In order to have that $\hat{\lambda}_{n,K}(s)$ is a consistent estimator of $\lambda(s)$, we need to impose another assumption on the kernel K, that is

(K.4) K has only a finite number of discontinuities.

Theorem 1.1 *Let the intensity function λ be periodic and locally integrable, and let the kernel K satisfy assumptions (K.1)–(K.4). Furthermore, let the bandwidth h_n be such that (1.4) holds true, and*

$$h_n|W_n| \to \infty. \tag{1.5}$$

If

$$|W_n||\hat{\tau}_n - \tau|/h_n \xrightarrow{p} 0, \tag{1.6}$$

then

$$\hat{\lambda}_{n,K}(s) \xrightarrow{p} \lambda(s), \tag{1.7}$$

provided s is a Lebesgue point of λ. In other words, $\hat{\lambda}_{n,K}(s)$ is a consistent estimator of $\lambda(s)$.

We note that, the assumption (K.4) can be weakened into assumption (K.4*) below. Therefore, in the next section, we give proof of Theorems 1.1 and 1.2 under the weaker assumption (K.4*). This assumption will allow us to control the fluctuations of the function

$$x \mapsto K\left(\frac{x-(s+k\hat{\tau}_n)}{h_n}\right)$$

depending on the fluctuations of $\hat{\tau}_n$ around τ. In particular, it will exclude functions K like

$$K_0 := \frac{1}{2}\mathbf{I}_{[-1,1]\setminus\mathcal{Q}},$$

where $\mathcal{Q}$ stands for the set of all rational numbers, and $\mathbf{I}_A$ denotes the indicator function of the set A. A more detail discussion on the necessity of excluding functions like K_0, which satisfies condition (K.1) - (K.3), is given in the Appendix.

Before the condition (K.4*), we note that, according to the Lusin's theorem, the measurability of function K (which is implicitly assumed by (K.1)) implies that

(L) For any $\alpha > 0$, there exists a compact set A_α and a continuous function $K_\alpha : \mathbf{R} \to \mathbf{R}$ such that the Lebesgue measure of the set $[-1,1] \setminus A_\alpha$ does not exceed α, and $|K(u) - K_\alpha(u)| \leq \alpha$ for all $u \in A_\alpha$.

By slightly strengthening assumption (L), we can exclude all the functions K like K_0 and, consequently, prove consistency of the estimator $\hat{\lambda}_{n,K}(s)$ under assumption (K.4*), that is

(K.4*) For any $\alpha > 0$, there exists a finite collection of disjoint compact intervals B_1,..., B_{M_α} and a continuous function $K_\alpha : \mathbf{R} \to \mathbf{R}$ such that the Lebesgue measure of the set $[-1,1] \setminus \cup_{i=1}^{M_\alpha} B_i$ does not exceed α, and $|K(u) - K_\alpha(u)| \leq \alpha$ for all $u \in \cup_{i=1}^{M_\alpha} B_i$.

Note, that by taking $A_\alpha = \cup_{i=1}^{M_\alpha} B_i$, we immediately obtain that any kernel satisfying assumption (K.4*) also satisfies (L). However, assumption (K.4*) still covers all the kernel functions K of statistical relevance that we can think of. For example, any kernel K whose all discontinuity points can, for any fixed $\alpha > 0$, be covered by a finite collection of open intervals of total size not exceeding α obviously satisfies assumption (K.4*).

We are now to discuss possible affects of the estimator $\hat{\tau}_n$ on $\hat{\lambda}_{n,K}(s)$. Recall first that the Poisson process X is observed only in the window W_n. Using the available for us information in W_n, we construct an estimator $\hat{\tau}_n$ of τ (cf., for example, Helmers and Mangku [3] and Vere-Jones [4]. Let us furthermore assume that

$$|W_n|^\gamma |\hat{\tau}_n - \tau| \xrightarrow{p} 0, \tag{1.8}$$

as $n \to \infty$, for some $\gamma \geq 0$. Then, assumption (1.6) holds true if, for example,

$$|W_n|/h_n \leq |W_n|^\gamma \tag{1.9}$$

for all sufficiently large n. We can find h_n converging to 0 and satisfying (1.9) if $\gamma > 1$. If, however, $\gamma \leq 1$, then in order to find h_n converging to 0 and satisfying (1.9), we have to replace the window W_n in the definition (1.3) of $\hat{\lambda}_{n,K}(s)$, as well as in Theorem 1.1, by a smaller window $W_{0,n} \subset W_n$ of size

$$|W_{0,n}| \sim |W_n|^{\gamma - \rho} \tag{1.10}$$

for some (no matter how small) $\rho > 0$. Indeed, if the estimator $\hat{\tau}_n$ does not converge to τ sufficiently fast, then the estimator $\hat{\lambda}_{n,K}(s)$ may not be consistent due to the slow convergence of $\hat{\tau}_n$ to τ. Therefore, by using the smaller window $W_{0,n}$ in the definition of $\hat{\lambda}_{n,K}(s)$, even though X is observed and the estimator $\hat{\tau}_n$ is constructed in the bigger window W_n, we can reduce the accumulated error by $\hat{\tau}_n$ and make the estimator $\hat{\lambda}_{n,K}(s)$ converge to $\lambda(s)$. Let us look, for example, at two estimators constructed by Helmers and Mangku [3] which are shown therein to satisfy (1.8) for any $\gamma \in [0, 1/4)$ and $\gamma \in [0, 1)$. Therefore, with either of the two estimators of Helmers and Mangku [3] that are constructed there using the observations of X in the window W_n, we now use, according to the above discussion, smaller windows $W_{0,n}$ of the sizes $|W_n|^\gamma$ for any fixed $\gamma \in [0, 1/4)$ and $\gamma \in [0, 1)$, respectively, to construct $\hat{\lambda}_{n,K}(s)$ as in definition (1.3).

Under, naturally, stronger assumptions than those of Theorem 1.1, we also have the complete convergence of the estimator $\hat{\lambda}_{n,K}(s)$ which, in turn, gives a rate of consistency of the estimator $\hat{\lambda}_{n,K}(s)$. Throughout this paper, we use $\stackrel{c}{\to}$ to denote complete convergence.

Theorem 1.2 *Let the intensity function λ be periodic and locally integrable, and let the kernel K satisfy assumptions (K.1)–(K.4). Furthermore, let the bandwidth h_n be such that (1.4) holds true, and*

$$\sum_{n=1}^{\infty} \exp\big\{ -\epsilon\sqrt{|W_n| h_n}\big\} < \infty \tag{1.11}$$

for any $\epsilon > 0$. If

$$|W_n||\hat{\tau}_n - \tau|/h_n \stackrel{c}{\to} 0, \tag{1.12}$$

then

$$\hat{\lambda}_{n,K}(s) \stackrel{c}{\to} \lambda(s), \tag{1.13}$$

provided s is a Lebesgue point of λ.

One may naturally want to know where the estimator $\hat{\lambda}_{n,K}(s)$ converges when it is not assumed that s is a Lebesgue point. A careful inspection of the proof (given in the next section) of Theorem 1.1 shows, for example, that under the assumption

$$\frac{1}{h}\int_{-h}^{h} \lambda(s+x)dx = O(1), \quad h \to 0,$$

the estimator $\hat{\lambda}_{n,K}(s)$ estimates

$$\lambda^*(s) := \lim_{h\to 0}\int_{-1}^{1} K(x)\lambda(s+xh)dx, \tag{1.14}$$

provided that the limit in (1.14) exists. For example, if the left- and right-hand limits $\lambda(s-)$ and $\lambda(s-)$ of λ at s exist, then

$$\lambda^*(s) = \lambda(s-)\int_{-1}^{0} K(x)dx + \lambda(s+)\int_{0}^{1} K(x)dx.$$

Consequently, if we assume that the function K is symmetric, then, due to the fact that K is a probability density function by assumption (K.1), we have the following representation

$$\lambda^*(s) = \frac{1}{2}\{\lambda(s-) + \lambda(s+)\}.$$

In turn, if s is a continuity point of λ, then the latter representation implies the following one

$$\lambda^*(s) = \lambda(s), \tag{1.15}$$

as it should be expected. Let us note in passing that if λ is known to be either right- or left-continuous, then we also have equality (1.15), provided that K has "one-sided" supports $[0, 1]$ and $[-1, 0]$, respectively.

In a follow-up paper, we compute the bias, variance, and mean squared error (MSE) of the estimator $\hat{\lambda}_{n,K}(s)$.

2. Proof of Theorems 1.1 and 1.2

To give a better insight into our proof of Theorems 1.1 and 1.2, we are now to describe the idea behind the construction of the estimator $\hat{\lambda}_{n,K}(s)$. To start with, we note that since there is available only one realization of the Poisson process X, we have to collect necessary information about the (unknown) value of $\lambda(s)$ from different places of the window W_n. For this reason, assumption (1.1) plays a crucial role and leads to the following string of (approximate) equations

$$\begin{aligned}
\lambda(s) &= \frac{1}{N_n}\sum_{k=-\infty}^{\infty} \lambda(s+k\tau)\mathbf{I}\{s+k\tau \in W_n\} \\
&\approx \frac{1}{N_n}\sum_{k=-\infty}^{\infty} \frac{1}{|B_{h_n}(s+k\tau)|}\int_{B_{h_n}(s+k\tau)\cap W_n} \lambda(x)dx \\
&= \frac{1}{N_n}\sum_{k=-\infty}^{\infty} \frac{1}{2h_n}\mu(B_{h_n}(s+k\tau)\cap W_n) \\
&\approx \frac{1}{N_n}\sum_{k=-\infty}^{\infty} \frac{1}{2h_n}X(B_{h_n}(s+k\tau)\cap W_n) \\
&\approx \frac{\tau}{|W_n|}\sum_{k=-\infty}^{\infty} \frac{1}{2h_n}X(B_{h_n}(s+k\tau)\cap W_n), \qquad (2.1)
\end{aligned}$$

where

$$N_n = \#\{k : \; s+k\tau \in W_n\},$$

h_n denotes a sequence of positive numbers converging to 0, $B_h(x)$ stands for the interval $[x-h, x+h]$, and μ denote the measure defined as

$$\mu(A) := \mathbf{E}X(A) = \int_A \lambda(x)dx, \quad A \in \mathcal{B}(\mathbf{R}).$$

[We note that in order to make the first $\approx$ in (2.1) work, we have assumed that s is a Lebesgue point of λ and h_n converges to 0.] Thus, from (2.1) we conclude that

$$\lambda_n(s) := \frac{\tau}{|W_n|}\sum_{k=-\infty}^{\infty} \frac{1}{2h_n}X(B_{h_n}(s+k\tau)\cap W_n), \qquad (2.2)$$

is an estimator of $\lambda(s)$, provided that the period τ is known.

Remark 2.1 The idea described in (2.1) and (2.2) of constructing an estimator for $\lambda(s)$ resembles that of Helmers and Zitikis [2] where we obtained in a similar fashion a non-parametric estimator for an intensity function which, in addition to the periodic trend, also has a polynomial trend. In Helmers and Zitikis [2], just like when constructing the estimator $\lambda_n(s)$ in (2.2), the period τ is supposed to be known.

The estimator $\lambda_n(s)$ of (2.2) can be modified in order to cover intensity functions with unknown periods as well. Namely, let $\hat{\tau}_n$ be a consistent estimator of τ. For example, one can think about the estimators of Helmers and Mangku [3], or Vere-Jones [4], or any other estimator of the period τ. Then, we modify the estimator (2.2) by replacing the unknown period τ by its estimator $\hat{\tau}_n$ and obtain the following estimator

$$\hat{\lambda}_n(s) = \frac{\hat{\tau}_n}{|W_n|}\sum_{k=-\infty}^{\infty} \frac{1}{2h_n}X\left(B_{h_n}(s+k\hat{\tau}_n)\cap W_n\right)$$

of $\lambda(s)$. Note that the estimator $\hat{\lambda}_n(s)$ can be rewritten as

$$\hat{\lambda}_n(s) = \frac{\hat{\tau}_n}{|W_n|} \sum_{k=-\infty}^{\infty} \frac{1}{h_n} \int_{W_n} \frac{1}{2} \mathbf{I}_{[-1,1]} \left(B_{h_n}(s + k\hat{\tau}_n) \right) X(dx). \tag{2.3}$$

By replacing the function $2^{-1}\mathbf{I}_{[-1,1]}(\cdot)$ in (2.3) by the general kernel K, we immediately arrive at the estimator introduced in (1.3).

We are now to prove Theorems 1.1 and 1.2. These theorems are a consequence of the basic probabilistic tool, which is given in Theorem 2.1.

Theorem 2.1 *Let the intensity function λ be periodic and locally integrable, and let the kernel K satisfy assumptions (K.1)–(K.3) and (K.4*). Furthermore, let the bandwidth h_n be such that (1.4) holds true. Then, for every $\epsilon > 0$, there exists a (small) $\beta := \beta(\epsilon) > 0$ and a (large) $n(\epsilon)$ such that the bound*

$$\mathbf{P}\{|\hat{\lambda}_{n,K}(s) - \lambda(s)| \geq \epsilon\} \leq c \exp\{-\epsilon\sqrt{|W_n| h_n}\} + \mathbf{P}\{|W_n| \, |\hat{\tau}_n - \tau| \geq \beta h_n\}, \tag{2.4}$$

holds true for all $n \geq n(\epsilon)$, provided s is a Lebesgue point of the intensity function λ.

To prove Theorem 2.1, we need the following three lemmas.

Lemma 2.1 *Let the intensity function λ be periodic and locally integrable, and let the kernel K satisfy assumptions (K.1)–(K.3). Furthermore, let the bandwidth h_n be such that (1.4) holds true. Then*

$$\frac{1}{N_n} \sum_{k=-\infty}^{\infty} \frac{1}{h_n} \int_{W_n} K\left(\frac{x - (s + k\tau)}{h_n} \right) \mu(dx) \to \lambda(s), \tag{2.5}$$

provided s is a Lebesgue point of λ.

Proof: Obviously,

$$\begin{aligned}
&\frac{1}{N_n} \sum_{k=-\infty}^{\infty} \frac{1}{h_n} \int_{W_n} K\left(\frac{x - (s + k\tau)}{h_n} \right) \mu(dx) \\
&= \frac{1}{N_n} \sum_{k=-\infty}^{\infty} \frac{1}{h_n} \int_{\mathbf{R}} K\left(\frac{x - (s + k\tau)}{h_n} \right) \lambda(x) I(x \in W_n) dx \\
&= \frac{1}{N_n h_n} \int_{\mathbf{R}} K\left(\frac{x}{h_n} \right) \sum_{k=-\infty}^{\infty} \lambda(x + s + k\tau) I(x + s + k\tau \in W_n) dx.
\end{aligned} \tag{2.6}$$

Since λ is periodic with period τ, we have $\lambda(x + s + k\tau) = \lambda(x + s)$. Furthermore, it is obvious that

$$\sum_{k=-\infty}^{\infty} I(x + s + k\tau \in W_n) \in [N_n - 1, N_n + 1]. \tag{2.7}$$

Consequently, the r.h.s of (2.6) converges to $\lambda(s)$ when $n \to \infty$, provided that

$$\frac{1}{h_n} \int_{\mathbf{R}} K\left(\frac{x}{h_n} \right) \lambda(x + s) dx \to \lambda(s). \tag{2.8}$$

Note that

$$\frac{1}{h_n}\int_{\mathbf{R}} K\left(\frac{x}{h_n}\right)\lambda(s)dx = \lambda(s)\int_{\mathbf{R}} K\left(x\right)dx = \lambda(s), \tag{2.9}$$

where we used the assumption that K is a probability density function. Consequently, statement (2.8) follows if

$$\frac{1}{h_n}\int_{\mathbf{R}} K\left(\frac{x}{h_n}\right)\left\{\lambda(x+s)-\lambda(s)\right\}dx \to 0, \tag{2.10}$$

when $n\to\infty$. The latter statement obviously follows from the assumptions that K is bounded and with support in $[-1,1]$, and that s is a Lebesgue point of λ. This completes the proof of Lemma 2.1. □

Denote

$$D_n := \frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}\int_{W_n} K\left(\frac{x-(s+k\tau)}{h_n}\right)X(dx) - \frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}\int_{W_n} K\left(\frac{x-(s+k\tau)}{h_n}\right)\mu(dx)$$

Lemma 2.2 *Let the intensity function λ be periodic and locally integrable, and let the kernel K satisfy assumptions (K.1)–(K.3). Furthermore, let the bandwidth h_n be such that (1.4) holds true. Then there is a (large) constant n_1 such that for any constant $c_1>0$ there exists another one $c_2>0$ such that*

$$\mathbf{P}\{|D_n|\geq c_1\epsilon\} \leq c_2\exp\big\{-\epsilon\sqrt{|W_n|h_n}\big\}, \tag{2.11}$$

for every $\epsilon>0$ and all $n\geq n_1$, provided s is a Lebesgue point of λ.

Proof: For every $t>0$, we have that

$$\mathbf{P}\{|D_n|\geq c_1\epsilon\} \leq \exp\{-c_1\epsilon t\}\big(\mathbf{E}\exp\{tD_n\}+\mathbf{E}\exp\{-tD_n\}\big). \tag{2.12}$$

To make our further considerations more transparent, we denote

$$\xi_k := \int_{W_n} K\left(\frac{x-(s+k\tau)}{h_n}\right)X(dx)$$

and then rewrite D_n as

$$D_n = \frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}\{\xi_k-\mathbf{E}\xi_k\}.$$

Since $h_n\downarrow 0$, the random variables ξ_k, $k=1,2,...$ are independent for all sufficiently large n (depending on the period τ). Thus, for sufficiently large n, we obtain

$$\mathbf{E}\exp\{\pm tD_n\} = \prod_{k=-\infty}^{\infty}\mathbf{E}\exp\left\{\pm\frac{t}{N_nh_n}(\xi_k-\mathbf{E}\xi_k)\right\}. \tag{2.13}$$

Using the well known formula for the Laplace transform of the Poisson process, we obtain that

$$\mathbf{E}\exp\left\{\pm\frac{t}{N_nh_n}\xi_k\right\} = \exp\left\{\int_{W_n}(e^{K^*(x)}-1)\lambda(x)dx\right\},$$

where we used the notation

$$K^*(x) := \pm \frac{t}{N_n h_n} K\left(\frac{x-(s+k\tau)}{h_n}\right)$$

Consequently, for every factor on the r.h.s. of (2.13) we have the following formula

$$\mathbf{E}\exp\left\{\pm\frac{t}{N_n h_n}\{\xi_k - \mathbf{E}\xi_k\}\right\} = \exp\left\{\int_{W_n} (e^{K^*(x)} - 1 - K^*(x))\lambda(x)dx\right\}. \tag{2.14}$$

Since $|\exp\{x\} - 1 - x|$ does not exceed $x^2\exp\{|x|\}$, we obtain from (2.14) that

$$\mathbf{E}\exp\left\{\pm\frac{t}{N_n h_n}\{\xi_k - \mathbf{E}\xi_k\}\right\} \le \exp\left\{\int_{W_n} |K^*(x)|^2 e^{|K^*(x)|}\lambda(x)dx\right\}. \tag{2.15}$$

We now make the following choice

$$t := \frac{1}{c_1}\sqrt{N_n h_n}. \tag{2.16}$$

Using the assumption that K is bounded and has support in the interval $[-1,1]$, we obtain from (2.15) with (2.16) that

$$\mathbf{E}\exp\left\{\pm\frac{t}{N_n h_n}\{\xi_k - \mathbf{E}\xi_k\}\right\} \le \exp\left\{c\frac{1}{N_n h_n}\mu\left(B_{h_n}(s+k\tau)\cap W_n\right)\right\}, \tag{2.17}$$

for a constant c that does not depend on n. Applying bound (2.17) on the r.h.s. of (2.13), we obtain

$$\mathbf{E}\exp\{\pm t D_n\} \le \exp\left\{c\frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}\mu\left(B_{h_n}(s+k\tau)\cap W_n\right)\right\}. \tag{2.18}$$

Furthermore, we note that the quantity $\mu\left(B_{h_n}(s+k\tau)\cap W_n\right)$ obviously equals to

$$\int_{B_{h_n}(0)} \lambda(s+k\tau+x) I(s+k\tau+x\in W_n)dx.$$

Consequently, using the periodicity of λ and (2.7) on the r.h.s. of (2.18), we obtain that

$$\mathbf{E}\exp\{\pm t D_n\} \le \exp\left\{c\frac{1}{h_n}\int_{B_{h_n}(0)}\lambda(s+x)dx\right\}. \tag{2.19}$$

Since s is a Lebesgue point of λ, we have that

$$\frac{1}{2h_n}\int_{B_{h_n}(0)}\lambda(s+x)dx \to \lambda(s),$$

when $n\to\infty$. Thus,

$$\lim_{n\to\infty}\mathbf{E}\exp\{\pm t D_n\} \le c < \infty. \tag{2.20}$$

Bound (2.20), when applied on the r.h.s. of (2.12), implies that

$$\mathbf{P}\{|D_n| \ge \epsilon\} \le \exp\left\{-\epsilon\sqrt{N_n h_n}\right\}, \tag{2.21}$$

due to our choice of t as in (2.16). Lemma 2.2 is therefore proved. □

We denote

$$\Lambda_n := \frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}\int_{W_n} K\left(\frac{x-(s+k\tau)}{h_n}\right)X(dx) - \frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}\int_{W_n} K\left(\frac{x-(s+k\hat{\tau}_n)}{h_n}\right)X(dx). \quad (2.22)$$

Lemma 2.3 *Let the intensity function λ be periodic and locally integrable, and let the kernel K satisfy assumptions (K.1)–(K.3) and (K.4*). Furthermore, let the bandwidth h_n be such that (1.4) holds true. Then, for every $\epsilon > 0$, there exists a (small) $\beta := \beta(\epsilon) > 0$ and a (large) $n(\epsilon) \in \mathbf{N}$ such that the bound*

$$\mathbf{P}\{|\Lambda_n| \geq \epsilon\} \leq c\exp\left\{-\epsilon\sqrt{|W_n|h_n}\right\} + \mathbf{P}\left\{|W_n||\hat{\tau}_n - \tau| \geq \beta h_n\right\}, \quad (2.23)$$

holds true for all $n \geq n(\epsilon)$, provided s is a Lebesgue point of λ.

Proof: Fix any $\alpha > 0$ and denote

$$A_\alpha := \bigcup_{i=1}^{M_\alpha} B_i \subset [-1,1],$$

where B_1,..., B_{M_α} are compact disjoint intervals defined in assumption (K.4*). Furthermore, using the (continuous) function K_α of assumption (K.4*) and the Weierstrass's theorem, we get that there exists a Lipschitz function L_α such that $|K(u) - L_\alpha(u)| \leq \alpha$ for all $u \in A_\alpha$. Now, we decompose both K on the right-hand side of (2.22) as follows

$$\begin{aligned} K(u) =& \{K(u) - L_\alpha(u)\}\mathbf{I}_{A_\alpha^c}(u) \\ &+ \{K(u) - L_\alpha(u)\}\mathbf{I}_{A_\alpha}(u) \\ &+ L_\alpha(u). \end{aligned} \quad (2.24)$$

Since K and L_α are bounded, we easily see that the quantity

$$\left|\frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}\int_{W_n}(K-L_\alpha)\left(\frac{x-(s+k\tau)}{h_n}\right)\mathbf{I}_{A_\alpha^c}\left(\frac{x-(s+k\tau)}{h_n}\right)X(dx) - \frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}\int_{W_n}(K-L_\alpha)\left(\frac{x-(s+k\hat{\tau}_n)}{h_n}\right)\mathbf{I}_{A_\alpha^c}\left(\frac{x-(s+k\hat{\tau}_n)}{h_n}\right)X(dx)\right|$$

does not exceed the sum of the following two quantities

$$\Lambda_{n,1} := c(K,L_\alpha)\frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}X\left(\{s+k\tau+h_nA_\alpha^c\}\cap W_n\right),$$

$$\Lambda_{n,2} := c(K,L_\alpha)\frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}X\left(\{s+k\hat{\tau}_n+h_nA_\alpha^c\}\cap W_n\right),$$

where $c(K, L_\alpha)$ denotes a constant depending only on $\sup\{|K(u)| : u \in [-1,1]\}$ and $\sup\{|L_\alpha(u)| : u \in [-1,1]\}$.

The quantity

$$\left|\frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}\int_{W_n}(K-L_\alpha)\left(\frac{x-(s+k\tau)}{h_n}\right)\mathbf{I}_{A_\alpha}\left(\frac{x-(s+k\tau)}{h_n}\right)X(dx)\right.$$
$$\left.-\frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}\int_{W_n}(K-L_\alpha)\left(\frac{x-(s+k\hat{\tau}_n)}{h_n}\right)\mathbf{I}_{A_\alpha}\left(\frac{x-(s+k\hat{\tau}_n)}{h_n}\right)X(dx)\right|$$

does not exceed the sum of the following two quantities

$$\Lambda_{n,3} := \alpha\frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}X\left(\{s+k\tau+h_n[-1,1]\}\cap W_n\right),$$
$$\Lambda_{n,4} := \alpha\frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}X\left(\{s+k\hat{\tau}_n+h_n[-1,1]\}\cap W_n\right).$$

Next, without loss of generality we assume that the support of the Lipschitz function L_α is in the interval $[-1,1]$. Using this fact, we obtain that

$$|L_\alpha(u)-L_\alpha(v)| \le c(L_\alpha)|u-v|\left(\mathbf{I}\{u\in[-1,1]\}+\mathbf{I}\{v\in[-1,1]\}\right)$$

for all $u, v \in [-1,1]$. Consequently, the quantity

$$\left|\frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}\int_{W_n}L_\alpha\left(\frac{x-(s+k\tau)}{h_n}\right)X(dx)-\frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}\int_{W_n}L_\alpha\left(\frac{x-(s+k\hat{\tau}_n)}{h_n}\right)X(dx)\right|$$

does not exceed the sum of the following two quantities

$$\Lambda_{n,5} := c(L_\alpha)\frac{1}{N_n}\sum_{k=-\infty}^{\infty}\left|\frac{k(\hat{\tau}_n-\tau)}{h_n}\right|\frac{1}{h_n}X\left(\{s+k\tau+h_n[-1,1]\}\cap W_n\right),$$
$$\Lambda_{n,6} := c(L_\alpha)\frac{1}{N_n}\sum_{k=-\infty}^{\infty}\left|\frac{k(\hat{\tau}_n-\tau)}{h_n}\right|\frac{1}{h_n}X\left(\{s+k\hat{\tau}_n+h_n[-1,1]\}\cap W_n\right).$$

Taking above obtained bounds together we see that the probability that $\Lambda_n \ge \epsilon$ does not exceed the probability that $\Lambda_{n,1}+\cdots+\Lambda_{n,6}\ge\epsilon$. This observation, in turn, implies that, for any $\beta > 0$,

$$\mathbf{P}\{\Lambda_n \ge \epsilon\} \le \mathbf{P}\{\Lambda_{n,1}+\cdots+\Lambda_{n,6}\ge\epsilon,\ |W_n|\ |\hat{\tau}_n-\tau|\le\beta h_n\}+\mathbf{P}\{|W_n|\ |\hat{\tau}_n-\tau|\ge\beta h_n\}. \quad (2.25)$$

We are now to estimate $\Lambda_{n,1}$,..., $\Lambda_{n,6}$ under the restriction $|W_n|\ |\hat{\tau}_n-\tau|\le\beta h_n$. Let start with the observation that even though $\Lambda_{n,1}$,..., $\Lambda_{n,6}$ are infinite sums, in each sum there is only a finite number of non-zero summands. As an example, let us first give a close look at $\Lambda_{n,1}$. Since $X(\emptyset)=0$, we have that

$$X\left(\{s+k\tau+h_nA_\alpha^c\}\cap W_n\right)=0$$

when, for example,

$$\{s+k\tau+h_n[-1,1]\}\cap W_n=\emptyset.$$

The latter statement is, obviously, equivalent to the following one

$$\left\{k+h_n\left[-\frac{1}{\tau},\frac{1}{\tau}\right]\right\}\cap\left\{\frac{1}{\tau}(W_n-s)\right\}=\emptyset. \quad (2.26)$$

Since $s \in W_n$ by assumption, the set $\tau^{-1}(W_n - s)$ contains 0. Therefore, (2.26) holds true if, for example,

$$|k| \geq \left|\frac{1}{\tau}(W_n - s)\right| + \frac{1}{\tau}h_n = \frac{1}{\tau}\{|W_n| + h_n\}. \tag{2.27}$$

We are now to apply similar reasoning to the quantity $\Lambda_{n,2}$. Namely, using the restriction $|W_n|\,|\hat{\tau}_n - \tau| \leq \beta h_n$, we get that

$$X\left(\{s + k\hat{\tau}_n + h_n A_\alpha^c\} \cap W_n\right) = 0$$

if, for example,

$$\left\{s + k\tau + h_n \frac{k}{|W_n|}[-\beta, \beta] + h_n[-1, 1]\right\} \cap W_n = \emptyset.$$

The latter statement is equivalent to the following one

$$\left\{k\left(1 + h_n\left[-\frac{\beta}{\tau|W_n|}, \frac{\beta}{\tau|W_n|}\right]\right) + h_n\left[-\frac{1}{\tau}, \frac{1}{\tau}\right]\right\} \cap \left\{\frac{1}{\tau}(W_n - s)\right\} = \emptyset. \tag{2.28}$$

Obviously, (2.28) holds true, for example, for all k such that

$$|k| \geq \frac{1}{\tau}(|W_n| + h_n)/\left(1 - h_n\frac{\beta}{|W_n|\tau}\right). \tag{2.29}$$

Due to assumptions $h_n \to 0$ and $|W_n| \to \infty$, we have that, for all sufficiently large n, both bounds (2.27) and (2.29) hold true for all k such that, for example,

$$|k| \geq \frac{2}{\tau}|W_n|. \tag{2.30}$$

The above presented consideration actually prove that the summands of $\Lambda_{n,3}$,..., $\Lambda_{n,6}$ are also 0 for all k such that (2.30) holds true. Consequently, when estimating $\Lambda_{n,1}$,..., $\Lambda_{n,6}$ we can always restrict ourselves to the summands with k such that $|k| \leq 2\tau^{-1}|W_n|$ only. This immediately implies the following bounds

$$\begin{aligned}
&\Lambda_{n,1}, \Lambda_{n,2} \leq c(K, L_\alpha)\Lambda_n^*,\\
&\Lambda_{n,3}, \Lambda_{n,4} \leq \alpha\Lambda_n^{**},\\
&\Lambda_{n,5}, \Lambda_{n,6} \leq c(L_\alpha)\Lambda_n^{**},
\end{aligned} \tag{2.31}$$

where

$$\Lambda_n^* := \frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}X\left(\left\{s + k\tau + h_n\left[-\frac{2\beta}{\tau}, \frac{2\beta}{\tau}\right] + h_n A_\alpha^c\right\} \cap W_n\right),$$

$$\Lambda_n^{**} := \frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}X\left(\left\{s + k\tau + h_n\left[-1 - \frac{2\beta}{\tau}, 1 + \frac{2\beta}{\tau}\right]\right\} \cap W_n\right).$$

Consequently, we have proved the following bound

$$\mathbf{P}\{|\Lambda_n| \geq \epsilon\} \leq \mathbf{P}\left\{c(K, L_\alpha)\Lambda_n^* + \{\alpha + \beta c(L_\alpha)\}\Lambda_n^{**} \geq \epsilon\right\} + \mathbf{P}\left\{|W_n|\,|\hat{\tau}_n - \tau| \geq \beta h_n\right\}.$$

The latter bound shows that the proof of Lemma 2.3 is completed if we show that

$$\mathbf{P}\left\{c(K, L_\alpha)\Lambda_n^* + \{\alpha + \beta c(L_\alpha)\}\Lambda_n^{**} \geq \epsilon\right\} \leq c\exp\left\{-\epsilon\sqrt{|W_n|h_n}\right\}. \tag{2.32}$$

The left-hand side of bound (2.32) does not exceed

$$\mathbf{P}\left\{c(K,L_\alpha)|\Lambda_n^* - \mathbf{E}\Lambda_n^*| + \{\alpha+\beta c(L_\alpha)\}|\Lambda_n^{**} - \mathbf{E}\Lambda_n^{**}| \geq c_\epsilon\right\}, \tag{2.33}$$

where

$$c_\epsilon := \epsilon - c(K,L_\alpha)\mathbf{E}\Lambda_n^* - \{\alpha+\beta c(L_\alpha)\}\mathbf{E}\Lambda_n^{**}.$$

We now want to show that the parameters α and β can be chosen in such a way that, for example,

$$c_\epsilon \geq \frac{\epsilon}{2} \tag{2.34}$$

when n is sufficiently large. To start with, we note that $\mathbf{E}\Lambda_n^{**}$ can be rewritten in the following way

$$2\left(1+\frac{2\beta}{\tau}\right)\frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n^*}\int_{W_n}\frac{1}{2}\mathbf{I}_{[-1,1]}\left(\frac{x-(s+k\tau)}{h_n^*}\right)\mu(dx), \tag{2.35}$$

where $h_n^* := (1+2\beta/\tau)h_n$. Using Lemma 2.1 with $K = 2^{-1}\mathbf{I}_{[-1,1]}$, we immediately obtain that the quantity of (2.35) converges to $2(1+2\beta/\tau)\lambda(s)$ when $n\to\infty$, and so does $\mathbf{E}\Lambda_n^{**}$. This implies that by choosing $\alpha>0$ and $\beta>0$ sufficiently small, we can make the quantity $\{\alpha+\beta c(L_\alpha)\}\mathbf{E}\Lambda_n^{**}$ smaller than $\epsilon/4$ for all sufficiently large n. In view of this fact, we obtain the desired bound (2.34), provided that

$$c(K,L_\alpha)\mathbf{E}\Lambda_n^* \leq \frac{\epsilon}{4} \tag{2.36}$$

for all sufficiently large n. We are now to prove (2.36). Denote

$$\mathfrak{A} := \left[-\frac{2\beta}{\tau},\frac{2\beta}{\tau}\right] + A_\alpha^c$$

for notational simplicity. Then

$$\begin{aligned}\mathbf{E}\Lambda_n^* &= \frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}\mathbf{E}X\left(\{s+k\tau+h_n\mathfrak{A}\}\cap W_n\right)\\ &= \frac{1}{N_n}\sum_{k=-\infty}^{\infty}\frac{1}{h_n}\int_{h_n\mathfrak{A}}\lambda(x+s+k\tau)\mathbf{I}_{W_n}(x+s+k\tau)dx\\ &= \frac{1}{N_nh_n}\int_{h_n\mathfrak{A}}\lambda(x+s)\sum_{k=-\infty}^{\infty}\mathbf{I}_{W_n}(x+s+k\tau)dx\\ &\leq \frac{2}{h_n}\int_{h_n\mathfrak{A}}\lambda(x+s)dx\\ &\leq \frac{2}{h_n}\Big|\int_{h_n\mathfrak{A}}\{\lambda(x+s)-\lambda(s)\}dx\Big| + 2\lambda(s)|\mathfrak{A}|.\end{aligned} \tag{2.37}$$

Note that the first summand on the right-hand side of (2.37) converges to 0, due to the assumption that s is a Lebesgue point of λ. Thus, in order to achieve the desired bound (2.36) we have to demonstrate that by choosing sufficiently small parameters $\alpha>0$ and $\beta>0$ we can make the quantity $|\mathfrak{A}|$ as small as we want. Here, only here, we need to employ assumption (K.4*).

Remark 2.2 If we do not assume (K.4*), then we only have (L). In this case, the set A_α^c can be so scattered over the interval $[-1,1]$ that the set $[-\beta,\beta]+A_\alpha^c$ may fill almost all interval $[-1,1]$ and thus the Lebesgue measure of $[-\beta,\beta]+A_\alpha^c$ may be close, for example, to that of $[-1,1]$ – the case which we definitely want to avoid by assuming (K.4*). □

By choosing the parameter $\beta > 0$ sufficiently small, we can achieve the situation when $\mathfrak{A}$ is a disjoint union of the sets $[-2\beta/\tau, 2\beta/\tau] + B_i$, $i = 1,..., M_\alpha$. Consequently,

$$|\mathfrak{A}| = \sum_{i=1}^{M_\alpha} \left| \left[-\frac{2\beta}{\tau}, \frac{2\beta}{\tau} \right] + B_i \right| = \sum_{i=1}^{M_\alpha} |B_i| + 2M_\alpha \frac{2\beta}{\tau} = |A_\alpha^c| + 2M_\alpha \frac{2\beta}{\tau}$$
$$\leq \alpha + 2M_\alpha \frac{2\beta}{\tau}. \tag{2.38}$$

Obviously, the right-hand side of (2.38) can be made as small as we want by taking $\alpha > 0$ and $\beta > 0$ sufficiently small. Thus, the desired bound (2.36) can indeed be achieved for all sufficiently large n. This, in turn, implies that, for all sufficiently large n, the quantity of (2.33) does not exceed

$$\mathbf{P}\left\{ c(K, L_\alpha)|\Lambda_n^* - \mathbf{E}\Lambda_n^*| + \{\alpha + \beta c(L_\alpha)\}|\Lambda_n^{**} - \mathbf{E}\Lambda_n^{**}| \geq \frac{\epsilon}{2} \right\}.$$

The latter quantity does not exceed the sum of $\mathbf{P}\{|\Lambda_n^* - \mathbf{E}\Lambda_n^*| \geq c_1^* \epsilon\}$ and $\mathbf{P}\{|\Lambda_n^{**} - \mathbf{E}\Lambda_n^{**}| \geq c_1^{**}\epsilon\}$, where $c_1^* > 0$ and $c_1^{**} > 0$ are some constants. Using Lemma 2.2 with the kernel $K := |\mathfrak{A}|^{-1}\mathbf{I}_{\mathfrak{A}}$ we obtain the bound

$$\mathbf{P}\{|\Lambda_n^* - \mathbf{E}\Lambda_n^*| \geq c_1^*\epsilon\} \leq c_2^* \exp\big\{ -\epsilon\sqrt{|W_n|h_n} \big\},$$

Furthermore, an application of Lemma 2.2 with the kernel $K := |\mathfrak{B}|^{-1}\mathbf{I}_{\mathfrak{B}}$, where

$$\mathfrak{B} := \left[-1 - \frac{2\beta}{\tau}, 1 + \frac{2\beta}{\tau} \right]$$

implies

$$\mathbf{P}\{|\Lambda_n^{**} - \mathbf{E}\Lambda_n^{**}| \geq c_1^{**}\epsilon\} \leq c_2^{**} \exp\big\{ -\epsilon\sqrt{|W_n|h_n} \big\},$$

Thus, the quantity of (2.33) does not exceed $c\exp\big\{-\epsilon\sqrt{|W_n|h_n}\big\}$, which completes the proof of bound (2.32) and, in turn, Lemma 2.3. □

Let us denote

$$\bar{\lambda}_{n,K}(s) := \frac{1}{N_n} \sum_{k=-\infty}^{\infty} \frac{1}{h_n} \int_{W_n} K\left(\frac{x - (s + k\hat{\tau}_n)}{h_n} \right) X(dx).$$

Then it is easy to see that Lemmas 2.1, 2.2 and 2.3 taken together imply that, for every $\epsilon > 0$, there exists a (small) $\beta := \beta(\epsilon) > 0$ and a (large) $n(\epsilon) \in \mathbf{N}$ such that the bound

$$\mathbf{P}\left\{ |\bar{\lambda}_{n,K}(s) - \lambda(s)| \geq \epsilon \right\} \leq c\exp\left\{ -\epsilon\sqrt{|W_n|h_n} \right\} + \mathbf{P}\left\{ |W_n|\ |\hat{\tau}_n - \tau| \geq \beta h_n \right\}, \tag{2.39}$$

holds true for all $n \geq n(\epsilon)$. In a little while we shall use this result to complete the proof of Theorem 2.1. Now, we proceed as follows.

Elementary algebra shows that

$$\mathbf{P}\left\{ |\hat{\lambda}_{n,K}(s) - \lambda(s)| \geq \epsilon \right\} \leq \mathbf{P}\left\{ \big(\big|\frac{\hat{\tau}_n N_n}{|W_n|} - 1\big| + 1\big)|\bar{\lambda}_{n,K}(s) - \lambda(s)| + \big|\frac{\hat{\tau}_n N_n}{|W_n|} - 1\big|\lambda(s) \geq \epsilon \right\}. \tag{2.40}$$

It is also easy to check that

$$\begin{aligned}\left|\frac{\hat{\tau}_n N_n}{|W_n|}-1\right| &\leq \frac{|\hat{\tau}_n-\tau|}{\tau}\left|\frac{\tau N_n}{|W_n|}-1\right|+\frac{|\hat{\tau}_n-\tau|}{\tau}+\left|\frac{\tau N_n}{|W_n|}-1\right| \\ &\leq \frac{|\hat{\tau}_n-\tau|}{\tau}\left(\frac{\tau}{|W_n|}+1\right)+\frac{\tau}{|W_n|}, \end{aligned} \tag{2.41}$$

where the second bound of (2.41) was obtained using $|\tau N_n - |W_n|| \leq \tau$. Since $|W_n|$ converges to ∞ by assumption, we can make the right-hand side of (2.41) as small as we want provided that we assume $|W_n|\ |\hat{\tau}_n - \tau| \leq \beta h_n$. Consequently, the right-hand side of (2.40) does not exceed

$$\mathbf{P}\left\{|\bar{\lambda}_{n,K}(s)-\lambda(s)| \geq \frac{\epsilon}{2}\right\}+\mathbf{P}\left\{|W_n|\ |\hat{\tau}_n-\tau| \geq \beta h_n\right\},$$

This fact together with (2.39) completes the proof of Theorem 2.1. □

3. Appendix: Discussion concerning assumptions (K.4) and (K.4*)

We are now to discuss the role of assumption (K.4*) in our considerations and in Theorem 1.1 in particular, and to give an explanation about the necessity to exclude kernel functions like K_0. Let we decompose K_0 as

$$K_0 = K_1 - K_2,$$

where

$$\begin{aligned} K_1 &:= \frac{1}{2}\mathbf{I}_{[-1,1]}, \\ K_2 &:= \frac{1}{2}\mathbf{I}_{[-1,1]\cap\mathcal{Q}}. \end{aligned}$$

Consequently, we have the following decomposition

$$\hat{\lambda}_{n,K_0}(s) = \hat{\lambda}_{n,K_1}(s) - \hat{\lambda}_{n,K_2}(s). \tag{3.1}$$

Note that the kernel K_1 satisfies all four assumptions (K.1)-(K.3), (K.4*). Therefore, by Theorem 2.1, we have the following bound

$$\mathbf{P}\left\{|\hat{\lambda}_{n,K_1}(s)-\lambda(s)| \geq \epsilon\right\} \leq c\exp\left\{-\epsilon\sqrt{|W_n|h_n}\right\}+\mathbf{P}\left\{|W_n|\ |\hat{\tau}_n-\tau| \geq \beta h_n\right\},$$

with the same parameters as in Theorem 2.1. We now easily see that if $h_n|W_n| \to \infty$ and $|W_n||\hat{\tau}_n - \tau|/h_n \xrightarrow{p} 0$, as $n \to \infty$, then $\hat{\lambda}_{n,K_1}(s)$ is a consistent estimator of $\lambda(s)$. In view of this fact and decomposition (3.1), the random variable $\hat{\lambda}_{n,K_0}(s)$ can be a consistent estimator of $\lambda(s)$ if and only if

$$\hat{\lambda}_{n,K_2}(s) \xrightarrow{p} 0, \tag{3.2}$$

as $n \to \infty$. Let us now look at $\hat{\lambda}_{n,K_2}(s)$ more closely. By the very definition, $\hat{\lambda}_{n,K_2}(s)$ has the following form

$$\hat{\lambda}_{n,K_2}(s) = \frac{\hat{\tau}_n}{|W_n|}\sum_{k=-\infty}^{\infty}\frac{1}{2h_n}X\left(\{s+k\hat{\tau}_n+h_n\mathcal{Q}\}\cap W_n\right).$$

If $\hat{\tau}_n$ were identically equal to τ, the expectation of the random variable

$$X\left(\{s+k\hat{\tau}_n+h_n\mathcal{Q}\}\cap W_n\right)\ \ [= X\left(\{s+k\tau+h_n\mathcal{Q}\}\cap W_n\right)]$$

would obviously be equal to 0, which, in turn, would be a strong evidence that the statement (3.2) holds true (in fact, one can easily verify that it is so under the assumption $\hat{\tau}_n \equiv \tau$). However, if $\hat{\tau}_n$ is a truly random estimator of τ, then the validity of statement (3.2) becomes highly questionable, provided that no additional information about $\hat{\tau}_n$ is available except that $|W_n|\ |\hat{\tau}_n - \tau|/h_n \xrightarrow{p} 0$, for example. To give a more rigorous justification of the latter claim, we note that statement (3.2) can be reduced to showing that, for ant $\epsilon > 0$ and $\beta > 0$,

$$\mathbf{P}\left\{\hat{\lambda}_{n,K_2}(s) \geq \epsilon,\ |W_n|\ |\hat{\tau}_n - \tau| \leq \beta h_n\right\} \to 0, \tag{3.3}$$

as $n \to \infty$. The "restriction" $|W_n|\ |\hat{\tau}_n - \tau| \leq \beta h_n$ in (3.3) actually says that what we really know about the estimator $\hat{\tau}_n$ is only the following confidence interval

$$\hat{\tau}_n \in \tau + \frac{\beta}{|W_n|} h_n[-1,1]. \tag{3.4}$$

With the notation of (3.4), we rewrite (3.3) more explicitly as

$$\mathbf{P}\left\{\frac{\hat{\tau}_n}{|W_n|}\sum_{k=-\infty}^{\infty}\frac{1}{2h_n}X\left(\{s + k\hat{\tau}_n + h_n\mathcal{Q}\}\cap W_n\right) \geq \epsilon,\ \hat{\tau}_n \in \tau + \frac{\beta}{|W_n|}h_n[-1,1]\right\} \to 0, \tag{3.5}$$

as $n \to \infty$. If we now use the only available for us information $\hat{\tau}_n \in \tau + \beta|W_n|^{-1}h_n[-1,1]$ to estimate the random variable $X(\{s + k\hat{\tau}_n + h_n\mathcal{Q}\}\cap W_n)$ in (3.5), we shall inevitably end up with the necessity of proving that

$$\mathbf{P}\{\hat{\lambda}^*_{n,K_2}(s) \geq \epsilon\} \to 0, \tag{3.6}$$

as $n \to 0$, where

$$\hat{\lambda}^*_{n,K_2}(s) := \frac{\hat{\tau}_n}{|W_n|}\sum_{k=-\infty}^{\infty}\frac{1}{2h_n}X\left(\{s + k\tau + k\beta|W_n|^{-1}h_n[-1,1] + h_n\mathcal{Q}\}\cap W_n\right).$$

But statement (3.6) appears to be impossible if $\lambda(s) > 0$. Indeed, since the interval $\beta|W_n|^{-1}h_n[-1,1]$ has a positive Lebesgue measure (and it does not matter how small it is), we have that the set $k\beta|W_n|^{-1}h_n[-1,1] + h_n\mathcal{Q}$ completely covers the interval $h_n[-1,1]$. This observation immediately implies that

$$\hat{\lambda}^*_{n,K_2}(s) \geq \hat{\lambda}_{n,K_1}(s).$$

But we have already noted above that $\hat{\lambda}_{n,K_1}(s)$ is a consistent estimator of $\lambda(s)$. Thus, $\hat{\lambda}^*_{n,K_2}(s)$ cannot converge in probability to 0, as $n \to \infty$, if $\lambda(s) > 0$.

The above given discussion indicates that without additional information about the relationship between X and $\hat{\tau}_n$ in the expression

$$X\left(\{s + k\hat{\tau}_n + h_n\mathcal{Q}\}\cap W_n\right),$$

it may be impossible to prove statements like (3.3) or (3.2). And we emphasise that, by not considering any specific estimator $\hat{\tau}_n$ in the present paper, we do not have more information about $\hat{\tau}_n$ except that $\hat{\tau}_n$ is a consistent estimator of τ and, possibly, a rate of consistency like $|W_n|\ |\hat{\tau}_n - \tau|/h_n \xrightarrow{p} 0$, as $n \to \infty$. However, it is important to call readers attention that no matter how attractive the problem of including the kernel K_0 into Theorem 1.1 could be from the mathematical point of view, it does not seem relevant from the statistical point of view at all. Indeed, as far as we understand, all the kernels K of statistical relevance satisfy assumptions (K.1)–(K.4), and are thus covered by Theorems 1.1.

References

1. Daley, D.J., and Vere-Jones, D. (1988). *Introduction to the Theory of Point Processes.* Springer, New York.
2. Helmers, R., and Zitikis, R. (1999). On estimation of Poisson intensity functions. *Ann. Inst. Stat. Math.* **51**, 2, 265-280.
3. Helmers, R., and Mangku, I.W. (2000). Nonparametric estimation of the period of a cyclic Poisson point process (part I): consistency. *CWI report,* in preparation.
4. Vere-Jones, D. (1982). On the estimation of frequency in point-process data. *J. Appl. Prob.* **19A**, 383-394.