Statistical Inference for Stochastic Processes **2:** 227–251, 1999.

Histograms and Associated Point Processes

PIERRE JACOB[1] and PAULO EDUARDO OLIVEIRA[2]
[1]*Lab. de Probab. et Stat., Université de Montpellier II, 34095 Montpellier Cedex 05, France, e-mail: jacob@stat.math.univ-montp2.fr*
[2]*Dep. Matemática, Universidade de Coimbra, 3000 Coimbra, Portugal, e-mail: paulo@mat.uc.pt*

Abstract. Nonparametric inference for point processes is discussed by way histograms, which provide a nice tool for the analysis of on-line data. The construction of histograms depends on a sequence of partitions, which we take to be nonembedded. This is quite natural in what regards applications, but presents some theoretical problems. In another direction, we drop the usual independence assumption on the sample, replacing it by an association assumption. Under this setting, we study the convergence of the histogram, in probability and almost surely which, under association, depends on conditions on the covariance structure. In the final section we prove that the finite dimensional distributions converge in distribution to a Gaussian centered vector with a specified covariance. The main tool of analysis is a decomposition of second order moment measures.

AMS Mathematics Subject Classifications (1991): 62G05, 62G20.

Key words: point process, Radon–Nikodym derivative, estimation.

1. Introduction

Nonparametric inference for point processes has been developed by using methods similar to those employed in classical functional estimation, where the estimators are either histograms or kernel estimators. Although the kernel approach has become increasingly popular as it produces smooth estimators, the use of histograms still proves efficient in many situations. In addition, some recent variations on the classical histogram help improve the convergence rates of such an estimator (see Beirlant et al. [2]). Histograms have been used in estimation in several models depending on point processes. Some examples include regression, as in Bensaïd [3], Palm distributions, as in Karr [23–25] or Niéré [29], mean local distributions of composed random measures, as in Mendes Lopes [26] or Saleh [36, 37], or density estimation, as in Ellis [11]. These references are not an account of the existing literature, but rather a mention of examples illustrating each problem. For a more complete list of publications on these subjects the interested reader is referred to one of the following monographs: Bosq [6], Bosq and Lecoutre [8], Bosq and Nguyen [7] or Karr [23]. All the above-mentioned problems produce results which exhibit a similarity. This similarity is due to the fact that these problems may be addressed in a unified way by defining a convenient general framework, reducing the estimation of the functions in each case to the estimation of a Radon–

Nikodym derivative of the means of two given random measures. Some examples as to how this framework may include some of the problems referred to above will be given later. This general framework has been used in Bensaïd and Fabre [4], Ellis [11], Ferrieux [13, 14], Jacob and Mendes Lopes [17], Jacob, Oliveira [19–21] and Roussas [33–35]. Articles [19] and [21] are concerned with histograms, while the others study kernel type estimators. Papers [11] and [33–35] used a somewhat narrower framework by imposing some special properties on the random measures, namely, assuming one of them to be almost surely fixed. Jacob and Mendes Lopes [17] deals with absolutely continuous random measures, thus reducing the problem to an analysis of the random densities involved. Half of the articles cited study estimation based on an independent sampling of the point process. Some results for dependent sampling have been obtained by Bensaïd and Fabre [4] where the kernel estimator is constructed under strong mixing. Suppressing the independence assumption, Roussas [33] and, more recently Ferrieux [13, 14] considered kernel estimators based on associated samples. Roussas [34, 35] also studied kernel estimates for associated random fields.

Here we will be concerned with histograms based associated compound point processes. These models provide interesting examples for illustrative purposes. The use of histograms relies on the choice of a sequence of partitions of the base space, which typically is constructed by splitting some of the sets of a partition to obtain the next one. This procedure produces embedded partitions which are convenient as they allow the use of martingale tools for proving the required convergences. This was used by the authors in [19]. However, this procedure is quite unnatural from an applications point of view. For such cases, it is customary to require that the sets in each partition are of same size, with respect to some reference measure. This requirement, together with the embedding procedure, produces sets which decrease quite fast. This fact may mean that the results thus obtained are of limited interest, as the number of new observations needed to change to the next partition would be very large. Nonembedded partitions have been used, for example, in Abou-Jaoudé [1], Grenander [15] or Karr [23]. The conditions used typically link the number of sets in each partition to the moments of the unknown distribution, as it is done in Karr [23]. These authors [21] gave another solution to this problem, using the same general framework as is done here, but for independent samples. The conditions imposed depend only on the distribution or only on the sizes of the sets. As this seems a more natural procedure to apply, the results in [21] will be the base for the extension discussed here to associated samples.

2. Preliminaries

In order to define the framework more precisely let $\mathbf{S}$ be a complete, separable and locally compact metric space; let $\mathcal{B}$ be the ring of relatively compact Borel subsets of $\mathbf{S}$; and let $\mathcal{M}$ be the space of nonnegative Radon measures on $\mathbf{S}$. A random measure is any function defined on some probability space with values in

$\mathcal{M}$ measurable with respect to the σ-algebra induced by the topology of vague convergence (we refer the reader to Daley and Vere-Jones [10], Kallenberg [22] or Karr [23] for basic properties on random measures). In what follows ξ and η are random measures which are supposed to be integrable, that is, the set functions $\mu(B) = \mathbb{E}\eta(B)$ and $\nu(B) = \mathbb{E}\xi(B)$ define elements of $\mathcal{M}$, and these mean measures satisfy the absolute continuity relation $\mu \ll \nu$. As it will be evident, we will be interested in estimating a version of the Radon–Nikodym derivative $\mathrm{d}\mu/\mathrm{d}\nu$. We will denote by $\mathbf{I}_A$ the indicator function of the set A.

We now indicate how some of the estimation problems mentioned above may be included in the present framework. In each setting, we will be interested in the interpretations of the Radon–Nikodym derivative $\mathrm{d}\mu/\mathrm{d}\nu$.

- (Ellis [11]) Density estimation: let ν be a measure on $\mathbf{S}$ and take $\xi = \nu$ a.s., $\eta = \delta_X$, where X is a random variable with absolutely continuous distribution with respect to ν. Then $\mathrm{d}\mu/\mathrm{d}\nu$ is the density of X with respect to ν.
- Regression: suppose Y is an almost surely nonnegative real random variable and X is a random variable on $\mathbf{S}$. Then, if $\xi = \delta_X$ and $\eta = Y\delta_X$, the conditional expectation $\mathbb{E}(Y|X = s)$ is a version of $\mathrm{d}\mu/\mathrm{d}\nu$.
- Thinning: suppose $\xi = \sum_{i=1}^{N} \delta_{X_i}$, where the X_n, $n \in \mathbb{N}$, are random variables on $\mathbf{S}$, α_n, $n \in \mathbb{N}$, are Bernoulli variables, conditionally independent given the sequence X_n, $n \in \mathbb{N}$, with parameters $p(X_n)$, and put $\eta = \sum_{i=1}^{N} \alpha_i \delta_{X_i}$. Then $p = \mathrm{d}\mu/\mathrm{d}\nu$ is the thinning function giving the probability of suppressing each point.
- Marked point processes: let $\zeta = \sum_{i=1}^{N} \delta_{(X_i,T_i)}$ be a point process on $\mathbf{S}\times \mathrm{T}$ such that the margin $\xi = \sum_{i=1}^{N} \delta_{X_i}$ is itself a point process. If $B \subset \mathrm{T}$ is measurable, choosing $\alpha_n = \mathbf{I}_B(T_n)$, and $\eta = \sum_{i=1}^{N} \alpha_i \delta_{X_i}$, we have

  $$\mathbb{E}\zeta(A \times B) = \int_A \frac{\mathrm{d}\mu}{\mathrm{d}\nu}(s)\, \mathbb{E}\zeta(\mathrm{d}s \times \mathbb{R}).$$

  Thus $\mathrm{d}\mu/\mathrm{d}\nu$ is the marking function.
- Cluster point processes: suppose $\zeta = \sum_{i=1}^{N} \sum_{j=1}^{N_i} \delta_{(X_i,Y_{i,j})}$ is a point process on $\mathbf{S}\times\mathbf{S}$ such that $\sum_{i=1}^{N} \sum_{j=1}^{N_i} \delta_{Y_{i,j}}$ is also a point process (for which it suffices to assume that, for example, N and N_n, $n \in \mathbb{N}$ are almost surely finite). The process $\xi = \sum_{i=1}^{N} \delta_{X_i}$ identifies the cluster centers and the processes $\zeta_{X_i} = \sum_{i=1}^{N_i} \delta_{Y_{i,j}}$ identify the points. The distribution of ζ may be characterized by a Markovian kernel of distributions $(\pi_x,\ x \in \mathbf{S})$ with means $(\mu_x,\ x \in \mathbf{S})$ such that, conditionally on $\xi = \sum_{i=1}^{N} \delta_{x_i}$, $(\zeta_{x_1}, \dots, \zeta_{x_n})$, it has distribution $\pi_{x_1} \otimes \cdots \otimes \pi_{x_n}$. Defining $\eta(A) = \zeta(A \times B)$, with $B \in \mathcal{B}$ fixed, we have

  $$\frac{\mathrm{d}\mu}{\mathrm{d}\nu}(x) = \mu_x(B)$$

ν-almost everywhere.

- Markovian shifts: this is a special case of the previous example, when $N_i = 1$ a.s., $i \geqslant 1$. In reference to at the previous example, the conclusion is that $(Y_1, \dots, Y_n)$ has distribution $\mu_{x_1} \otimes \cdots \otimes \mu_{x_n}$ (we replaced the double index of the Y variables by a single one as, for each i fixed, there is only one such variable). Then it would follow that

$$\frac{\mathrm{d}\mu}{\mathrm{d}\nu}(x) = \mu_x(B) = \mathbb{P}(Y \in B | X = x).$$

So, as illustrated by the above examples, we will be concerned with the estimation of $\mathrm{d}\mu/\mathrm{d}\nu$, based on a sample $((\xi_1, \eta_1), \dots, (\xi_n, \eta_n))$ of the random pair (ξ, η). As already mentioned, we suppose the pairs (ξ_i, η_i), $i = 1, \dots, n$, to be associated: given $n \in \mathbb{N}$ and any two coordinatewise nondecreasing functions f, g defined on $\mathcal{M}^{2n}$, for which the covariance below exists, we have

$$\mathrm{Cov}(f(\xi_1, \eta_1, \dots, \xi_n, \eta_n),\ g(\xi_1, \eta_1, \dots, \xi_n, \eta_n)) \geqslant 0.$$

(For $\zeta_1, \zeta_2 \in \mathcal{M}$, we say that $\zeta_1 \leqslant \zeta_2$ if $\zeta_2 - \zeta_1 \in \mathcal{M}$). For basic results on association, we refer the reader to Newman [27], and for association of random measures to Burton and Waymire [9] or Evans [12]. An account of the relevant results pertinent to our purposes may be found in Ferrieux [13, 14].

We note that the density estimation case and the regression case mentioned above are not meaningful for the associated sampling. In fact, it is easily checked that, whenever a point process has a fixed number of independent points, it cannot be associated with itself. Thus, it is impossible to construct a sequence of associated point processes with that same distribution. To check this, suppose $\xi = \delta_X$. Then it is easily seen that $\mathrm{Cov}(\xi, \xi) = \mathbb{E}\delta_{(X,X)} - \mathbb{P}_X \otimes \mathbb{P}_X$. More generally, if $\xi = \sum_{i=1}^n \delta_{X_i}$, for some independent random elements X_i with distributions $\mathbb{P}_{X_i}$, not necessarily equal, then

$$\mathrm{Cov}(\xi, \xi) = \sum_{i=1}^n \left(\mathbb{E}\delta_{(X_i,X_i)} - \mathbb{P}_{X_i} \otimes \mathbb{P}_{X_i}\right).$$

As $\mathbb{E}\delta_{(X_i,X_i)}$ is a measure on $\mathbf{S}\times\mathbf{S}$ with support included in the diagonal and $\mathbb{P}_{X_i} \otimes \mathbb{P}_{X_i}$ is not supported by the diagonal (except in degenerate cases), we actually have a signed measure.

It should also be noted that it is not clear whether there is any connection between $X_1, \dots, X_n$ being associated and $\delta_{X_1}, \dots, \delta_{X_n}$ being associated. This implies that there is probably no overlap with the work of Ellis [11] or Roussas [33–35].

In order to define the histograms to be employed we need a sequence of partitions. For reasons that will be explained later we will take Π_k, $k \in \mathbb{N}$, to be a sequence of partitions of a fixed compact set $B \subset \mathbf{S}$, instead of partitions of the entire space. On the following I represents a set belonging to some partition Π_k. We impose the following assumptions:

(P1) for each $k \in \mathbb{N}$, $\Pi_k \subset \mathcal{B}$;
(P2) for each $k \in \mathbb{N}$, Π_k is finite;
(P3) $\theta_k = \sup\{\operatorname{diam}(I)\colon\ I \in \Pi_k\} \longrightarrow 0$;
(P4) for each $k \in \mathbb{N}$ and $I \in \Pi_k$, $\nu(I) > 0$;
(P5) $\max_{I\in\Pi_k} \nu(I) \longrightarrow 0$.

Note that **(P4)** and **(P5)** introduce assumptions which are relative to the measure ν. In some cases we need that **(P4)** and **(P5)** be satisfied with respect to some other reference measure λ, meaning that we require $\lambda(I) > 0$, for every $I \in \Pi_k$, $k \in \mathbb{N}$, and $\max_{I\in\Pi_k} \lambda(I) \longrightarrow 0$. The correct indication of this measure is of importance when coupled with conditions **(M1)** and **(M2)**, to be introduced later, where there exists a measure playing a role of reference. We need these two reference measures to be identical.

Before we proceed with the introduction of further assumptions, we may define an approximation to (a suitable version of) $\mathrm{d}\mu/\mathrm{d}\nu$. Given $s \in B$, we denote by $I_k(s)$ the unique set of Π_k containing the point s, and, for each $k \in \mathbb{N}$, define the function

$$g_k(s) = \sum_{I\in\Pi_k} \frac{\mu(I)}{\nu(I)} \mathbf{I}_I(s) = \frac{\mu(I_k(s))}{\nu(I_k(s))}.$$

In the case of embedded partitions, the convergence of g_k to some version of $\mathrm{d}\mu/\mathrm{d}\nu$ is just a martingale result, which, however, is no longer available in our setting. As is well known, if there exists a continuous version f of the Radon–Nikodym derivative $\mathrm{d}\mu/\mathrm{d}\nu$, and if the sequence of partitions Π_k, $k \in \mathbb{N}$, satisfies **(P1)–(P4)**, the convergence

$$\sup_{s\in B} |f(s) - g_k(s)| \longrightarrow 0$$

holds. The fact that everything is happening within a compact set is crucial to the proof of this uniform convergence. That is why we only consider partitions of a fixed compact set B.

Based on the sample $((\xi_1, \eta_1), \ldots, (\xi_n, \eta_n))$, define

$$\overline{\xi}_n = \frac{1}{n} \sum_{i=1}^{n} \xi_i \quad \text{and} \quad \overline{\eta}_n = \frac{1}{n} \sum_{i=1}^{n} \eta_i. \tag{1}$$

The histogram estimator of f is then

$$f_n(s) = \sum_{I\in\Pi_k} \frac{\overline{\eta}_n(I)}{\overline{\xi}_n(I)} \mathbf{I}_I(s) = \frac{\overline{\eta}_n(I_k(s))}{\overline{\xi}_n(I_k(s))} \tag{2}$$

(as usual, we define $f_n(s)$ as zero whenever the denominator vanishes), where the dependence of k on n is to be made precise later. The convergence of f_n to some version of $\mathrm{d}\mu/\mathrm{d}\nu$ follows from the convergence of $f_n - g_k$ to zero. This latter convergence was obtained, in the independent case, via a martingale result concerning product measures of the type $\mathbb{E}\zeta_1 \otimes \zeta_2$, where $\zeta_1, \zeta_2 \in \{\xi, \eta\}$ (see Lemma 3.1 in [19]). Again, this was a consequence of the embedding of the partitions, no longer available in the present framework. To circumvent this difficulty, we consider an assumption concerning a decomposition of measures on the product space $\mathbf{S}\times\mathbf{S}$, as is done in [21]. We will say that a measure m on $\mathbf{S}\times\mathbf{S}$ satisfies condition **(M)** with respect to the measure ν on $\mathbf{S}$ if $m = m_1 + m_2$ where m_2 is a measure on Δ, the diagonal of $\mathbf{S}\times\mathbf{S}$, and m_1 is a measure on $\mathbf{S}\times\mathbf{S} \setminus \Delta$, such that

(**M1**) $m_1 \ll \nu \otimes \nu$ and there exists a version γ_1 of the Radon–Nikodym derivative $\mathrm{d}m_1/\mathrm{d}\nu \otimes \nu$ which is bounded;

(**M2**) $m_2 \ll \nu^*$, where ν^* is the measure on Δ defined by lifting ν, that is, such that $\nu^*(A^*) = \nu(A)$ with $A^* = \{(s,s) : s \in A\}$, and there exits a continuous version γ_2 of the Radon–Nikodym derivative $\mathrm{d}m_2/\mathrm{d}\nu^*$.

Then the following result, which will play the role of the above-mentioned martingale lemma in the independent case, holds.

THEOREM 2.1 [21]. *Suppose m is a measure on $\mathbf{S}\times\mathbf{S}$ that satisfies condition* **(M)** *with respect to ν and suppose the sequence of partitions Π_k, $k \in \mathbb{N}$, satisfies* **(P1)–(P5)**. *Then*

$$\sum_{I\in\Pi_k} \frac{m(I \times I)}{\nu(I)} \mathbf{I}_I(s) \longrightarrow \gamma_2(s,s)$$

uniformly on B.

Proof. Using the decomposition included in **(M)** we have two terms to examine, corresponding to m_1 and m_2. Regarding the first term,

$$\begin{aligned}\sum_{I\in\Pi_k} \frac{m_1(I \times I)}{\nu(I)}\mathbf{I}_I(s) &= \sum_{I\in\Pi_k} \left(\frac{1}{\nu(I)} \int_{I\times I} \gamma_1 \,\mathrm{d}\nu \otimes \nu \right) \mathbf{I}_I(s) \\ &\leqslant \sup_{s,t\in B} |\gamma_1(s,t)| \sum_{I\in\Pi_k} \nu(I)\, \mathbf{I}_I(s) \\ &\leqslant \sup_{s,t\in B} |\gamma_1(s,t)| \max_{I\in\Pi_k} \nu(I) \longrightarrow 0.\end{aligned}$$

As for the second term

$$\begin{aligned}\sum_{I\in\Pi_k} \frac{m_2(I \times I)}{\nu(I)}\mathbf{I}_I(s) &= \sum_{I\in\Pi_k} \frac{m_2(I^*)}{\nu^*(I^*)}\, \mathbf{I}_I(s) \\ &= \sum_{I\in\Pi_k} \left(\frac{1}{\nu^*(I^*)} \int_{I^*} \gamma_2 \,\mathrm{d}\nu^* \right) \mathbf{I}_I(s)\end{aligned}$$

and the uniform convergence of this expression to $\gamma_2(s,s)$ is just another version of the result giving the already mentioned convergence of the sequence g_k, $k \in \mathbb{N}$. □

Note that **(M)** must be defined with respect to some measure. If we do not mention any such measure, it will be understood that the measure is ν. As it was stated after the introduction of conditions **(P1)–(P5)** what will be important is that the reference measure is the same in both cases. Then, the convergence stated in Theorem 2 still holds with the obvious modification on the definition of γ_2, becoming the Radon–Nikodym derivative of m_2 with respect to the lifting of the reference measure used.

We conclude this section by quoting a useful result, which makes possible the separation of the variables in the expression f_n.

LEMMA 2.2 [18]. *Let X and Y be non-negative integrable random variables. Then, for* $\varepsilon > 0$ *small enough,*

$$\left\{\left|\frac{X}{Y} - \frac{\mathbb{E}(X)}{\mathbb{E}(Y)}\right| > \varepsilon\right\} \subset$$
$$\subset \left\{\left|\frac{X}{\mathbb{E}(X)} - 1\right| > \frac{\varepsilon}{4}\frac{\mathbb{E}(Y)}{\mathbb{E}(X)}\right\} \cup \left\{\left|\frac{Y}{\mathbb{E}(Y)} - 1\right| > \frac{\varepsilon}{4}\frac{\mathbb{E}(Y)}{\mathbb{E}(X)}\right\}.$$

Using this Lemma, it follows that, for $\varepsilon > 0$ small enough,

$$\begin{aligned}\{|f_n(s) - g_k(s)| > \varepsilon\} &= \left\{\left|\frac{\overline{\eta}_n(I_k(s))}{\overline{\xi}_n(I_k(s))} - \frac{\mu(I_k(s))}{\nu(I_k(s))}\right| > \varepsilon\right\} \subset \\ &\subset \left\{\left|\overline{\eta}_n(I_k(s)) - \mu(I_k(s))\right| > \frac{\varepsilon}{4}\nu(I_k(s))\right\} \cup \\ &\cup \left\{\left|\overline{\xi}_n(I_k(s)) - \nu(I_k(s))\right| > \frac{\varepsilon}{4}\frac{\nu^2(I_k(s))}{\mu(I_k(s))}\right\}. \end{aligned} \tag{3}$$

3. Convergence of the Estimator

Having introduced all the definitions and preliminary results needed, we may now investigate the convergence of the estimator f_n. We begin with the convergence in probability, for which we state two versions. The second version extends to an almost complete result which we will not state here for reasons that will be explained later. In order to be more explicit about the dependence between the different indices used, we will denote the set involved by $I_{k(n)}$ to emphasize the dependence of k on n, the size of the sample.

THEOREM 3.1. *Let $B \in \mathcal{B}$ be compact and let f be a version of $\mathrm{d}\mu/\mathrm{d}\nu$ continuous on B. Suppose the sequence of partitions Π_k, $k \in \mathbb{N}$, satisfies conditions* **(P1)–(P5)** *and that there exist measures $m^{\xi,\xi}$ and $m^{\eta,\eta}$ such that, for every $n \in \mathbb{N}$,*

$$\frac{1}{n}\sum_{i,j=1}^{n} \mathrm{Cov}(\xi_i, \xi_j) \leqslant m^{\xi,\xi} \quad \text{and} \quad \frac{1}{n}\sum_{i,j=1}^{n} \mathrm{Cov}(\eta_i, \eta_j) \leqslant m^{\eta,\eta} \tag{4}$$

with $m^{\xi,\xi}$ and $m^{\eta,\eta}$ both satisfying **(M)** *with respect to ν and*

$$n \min_{I \in \Pi_{k(n)}} \nu(I) \longrightarrow +\infty. \tag{5}$$

Then, for every $s \in B$, $f_n(s)$ converges in probability to $f(s)$.

Proof. After separation of variables by using (3), we apply Chebyshev's inequality. The term corresponding to η leads to

$$\begin{aligned}
&\mathbb{P}\left(\left|\overline{\eta}_n(I_{k(n)}(s)) - \mu(I_{k(n)}(S))\right| > \frac{\varepsilon\,\nu(I_{k(n)}(s))}{4}\right) \\
&\quad\leqslant \frac{16}{\varepsilon^2\, n\, \nu^2(I_{k(n)}(s))}\frac{1}{n}\sum_{i,j=1}^{n} \mathrm{Cov}(\eta_i(I_{k(n)}(s)), \eta_j(I_{k(n)}(s))) \\
&\quad\leqslant \frac{16}{\varepsilon^2\, n\, \nu(I_{k(n)}(s))}\frac{m_1^{\eta,\eta}(I_{k(n)}(s)\times I_{k(n)}(s)) + m_2^{\eta,\eta}(I_{k(n)}(s)\times I_{k(n)}(s))}{\nu(I_{k(n)}(s))},
\end{aligned} \tag{6}$$

and this last expression converges to zero according to (5) and Theorem 2. The other term is treated analogously after separation of variables. □

Note that in the preceding result, association implies that the covariance measures introduced are really measures and not just signed measures. We may relax (4) by requiring only that the covariances involved to be bounded on B. This will mean a slower decrease rate of measures of the sets.

COROLLARY 3.2. *Let $B \in \mathcal{B}$ and let f be a version of $\mathrm{d}\mu/\mathrm{d}\nu$ continuous on B. Suppose there exist constants $c_1,\ c_2 > 0$ such that*

$$\frac{1}{n}\sum_{i,j=1}^{n} \mathrm{Cov}\left(\eta_i(B),\ \eta_j(B)\right) \leqslant c_1, \tag{7}$$

$$\frac{1}{n}\sum_{i,j=1}^{n} \mathrm{Cov}\left(\xi_i(B),\ \xi_j(B)\right) \leqslant c_2. \tag{8}$$

If

$$n^{1/2} \min_{I \in \Pi_{k(n)}} \nu(I) \longrightarrow +\infty, \tag{9}$$

then $f_n(s)$ converges in probability to $f(s)$ ν-almost everywhere in B.

Proof. As in the proof of the theorem, we begin by applying Chebyshev's inequality to find the upper bound in the middle line of (6). The sets $I_{k(n)}(s)$ are, by definition of the partitions Π_k, subsets of B so, by association, this upper bound is still bounded above by

$$\frac{16}{\varepsilon^2\, n\, \nu^2(I_{k(n)}(s))}\frac{1}{n}\sum_{i,j=1}^{n}\mathrm{Cov}(\eta_i(B),\eta_j(B)),$$

which converges to zero according to (7) and (9). □

Note that condition (7), for the case $\eta = \delta_X$, is rewritten as

$$\frac{1}{n}\sum_{i,j=1}^{n}[\mathbb{P}(X_i\in I_{k(n)}, X_j\in I_{k(n)}) - \mathbb{P}(X_i\in I_{k(n)})\,\mathbb{P}(X_j\in I_{k(n)})]\leqslant c_1.$$

This kind of sum appears in other situations as well when studying association. In fact, a general condition for tightness of empirical processes in $L^2[0,1]$ is the uniform convergence of these expressions, as proved in Oliveira and Suquet [30, 31]. The discussion of the same problem, but in the space $D[0,1]$, also depends on these expressions, as is done in Yu [40] and Shao and Yu [38].

The method used for proving Corollary 3.2 may be extended, requiring the existence of higher order moments, to derive an almost complete result. We would then be lead to use moment inequalities for sums of associated variables by Birkel [5]. These would require a quite slow convergence rate of the sets used at each step and, further, this convergence rate should be well tuned with the decrease rate of the covariance structure of the sequences $\xi_n(B)$, $\eta_n(B)$, $n\in\mathbb{N}$. Thus, we would have conditions with the same drawbacks as those already mentioned linking the size of the sets to the moments of the unknown distribution, which we are trying to avoid here. Another method to derive the almost complete convergence is based on exponential inequalities. One such inequality for associated random variables appeared in the literature while this after this article was submitted (see Ioannides and Roussas [16]). This inequality really provide the means for an almost complete result, but the conditions it requires are of a different sort and much stronger than those we have been assuming in this article. Namely, for the use of Ioannides and Roussas's exponential inequality it would be necessary to assume that the point processes were uniformly bounded, at least on the compact set B. In this article, we have been using only moment conditions on the point processes. So, we choose no to include an almost complete result and prove only an almost sure result. Instead of using separation of variables based on Lemma 2.2, the crucial step towards an almost sure theorem is to observe that we do not change the partition each time a new observation is added to our sample, that is, we go on using the same sets until the number of observations increases enough to justify the use of the next partition. This is what is implicitly included in conditions such us (5) or (9). We will not investigate the difference $f_n - g_k$, but rather we rewrite $f_n(s)$ as

$$f_n(s) = \frac{\mu(I_{k(n)}(s))}{\nu(I_{k(n)}(s))} \frac{\overline{\eta}_n(I_{k(n)}(s))/\mu(I_{k(n)}(s))}{\overline{\xi}_n(I_{k(n)}(s))/\nu(I_{k(n)}(s))}.$$

So, in order to prove the almost sure convergence of $f_n(s)$, it is enough to prove that both expressions $\overline{\eta}_n(I_{k(n)}(s))/\mu(I_{k(n)}(s))$ and $\overline{\xi}_n(I_{k(n)}(s))/\nu(I_{k(n)}(s))$ converge almost surely to 1. We will suppress the argument s where confusion does not arise. For the almost sure convergence, we need to identify where we really change from one partition to the next one. For very small values of n, the construction of the histogram estimator f_n uses sets belonging to Π_1. As the sample size n increases, that will mean we eventually will use, for constructing $f_n(s)$, sets from Π_2. Define $t_1 = 1$ and t_2 the first n for which we use, for the construction of f_n, sets of Π_2. As n continues to increase, we will eventually base the construction of $f_n(s)$ in sets belonging to Π_k. We define t_k as the sample size for which we use, for the first times, sets from the partition Π_k.

THEOREM 3.3. *Let $B \in \mathcal{B}$ be compact and let f be a version of* $\mathrm{d}\mu/\mathrm{d}\nu$ *continuous and bounded away from zero on B. Suppose the sequence of partitions Π_k, $k \in \mathbb{N}$, satisfies* **(P1)–(P5)**, *that there exist measures $m^{\xi,\xi}$ and $m^{\eta,\eta}$, such that, for every $n \in \mathbb{N}$,*

$$\frac{1}{n}\sum_{i,j=1}^{n} \operatorname{Cov}(\xi_i, \xi_j) \leqslant m^{\xi,\xi} \quad \textit{and} \quad \frac{1}{n}\sum_{i,j=1}^{n} \operatorname{Cov}(\eta_i, \eta_j) \leqslant m^{\eta,\eta}$$

with $m^{\xi,\xi}$ and $m^{\eta,\eta}$ both satisfying **(M)**, *and*

$$\frac{t_{k+1}}{t_k} \tag{10}$$

being bounded and

$$\sum_{k=1}^{\infty} \frac{1}{t_k \min_{I\in\Pi_k} \nu(I)} < \infty. \tag{11}$$

Then, for every $s \in B$, $f_n(s)$ converges almost surely to $f(s)$.

Proof. We shall show that, under the assumptions of the theorem $\overline{\eta}_n(I_{k(n)})/\mu(I_{k(n)})$ converges to 1 a.s. The term corresponding to ξ is treated analogously. The proof will follow the classical method: first we show the convergence along the subsequence defined by the indices t_k, $k \in \mathbb{N}$, and then establish bounds for the difference between these subsequences and the remaining terms of the sequence.

The first step reduces to an application of Chebyshev's inequality, as follows:

$$\begin{aligned}
&\mathbb{P}\left(\left|\frac{\overline{\eta}_{t_k}(I_k)}{\nu(I_k)}-1\right|>\varepsilon\right)\\
&\quad=\mathbb{P}\left(\left|\sum_{i=1}^{t_k}(\eta_i(I_k)-\mu(I_k))\right|>\varepsilon t_k\mu(I_k)\right)\\
&\quad\leqslant\frac{1}{\varepsilon^2 t_k^2\mu(I_k)}\sum_{i,j=1}^{t_k}\operatorname{Cov}(\eta_i(I_k),\eta_j(I_k))\\
&\quad\leqslant\frac{1}{\varepsilon^2}\frac{1}{t_k\nu(I_k)}\frac{\nu^2(I_k)}{\mu^2(I_k)}\frac{m_1^{\eta,\eta}(I_k\times I_k)+m_2^{\eta,\eta}(I_k\times I_k)}{\nu(I_k)}
\end{aligned}$$

and this defines a convergent series, according to (11) and Theorem 2.1.

Suppose now that $n\in[t_k,t_{k+1})$. According to the definition of t_k, it follows that $I_{k(n)}=I_k$, so that

$$\begin{aligned}
&\frac{\overline{\eta}_n(I_{k(n)})}{\mu(I_{k(n)})}-\frac{\overline{\eta}_{t_k}(I_k)}{\mu(I_k)}\\
&\quad=\sum_{i=1}^{t_k}\left(\frac{1}{n}-\frac{1}{t_k}\right)\frac{\eta_i(I_k)-\mu(I_k)}{\mu(I_k)}+\frac{1}{n}\sum_{i=t_k+1}^{n}\frac{\eta_i(I_k)-\mu(I_k)}{\mu(I_k)}.
\end{aligned}\tag{12}$$

The first term equals $(t_k/n-1)(\overline{\eta}_{t_k}(I_k)/\mu(I_k)-1)$. As $t_k\leqslant n$, the first factor is bounded, and the other factor in this last expression converges almost surely to 0, as proved in the first step. As for the second term in (12), we have, by using the generalization of the Kolmogorov inequality for associated variables proved by Newman and Wright [28],

$$\begin{aligned}
&\mathbb{P}\left(\max_{t_k\leqslant n<t_{k+1}}\frac{1}{n}\left|\sum_{i=t_k+1}^{n}\frac{\eta_i(I_k)-\mu(I_k)}{\mu(I_k)}\right|>\varepsilon\right)\\
&\quad\leqslant\mathbb{P}\left(\max_{t_k\leqslant n<t_{k+1}}\left|\sum_{i=t_k+1}^{n}[\eta_i(I_k)-\mu(I_k)]\right|>\varepsilon t_k\mu(I_k)\right)\\
&\quad\leqslant\frac{2}{\varepsilon^2 t_k^2\mu^2(_k)}\sum_{i,j=t_k}^{t_{k+1}}\operatorname{Cov}(\eta_i(I_k),\eta_j(I_k))\\
&\quad\leqslant\frac{2}{\varepsilon^2}\frac{t_{k+1}}{t_k}\frac{1}{t_k\nu(I_k)}\frac{\nu^2(I_k)}{\mu^2(I_k)}\frac{m_1^{\eta,\eta}(I_k\times I_k)+m_2^{\eta,\eta}(I_k\times I_k)}{\nu(I_k)},
\end{aligned}$$

which defines a convergent series according to (10), (11) and Theorem 2.1. So the second term in (12) also converges almost surely to zero, and this concludes the proof. □

4. Finite-Dimensional Distributions

We now investigate the finite-dimensional asymptotics of $f_n - g_k$, properly normalized. As in Jacob and Oliveira [21], in this section we will suppose that ν is absolutely continuous with respect to some fixed nonatomic measure λ on **S**, with Radon–Nikodym derivative f_ν continuous on the compact set B, and that the sets in each partition have equal λ measure. Denote by h_n the λ measure of each set in $\Pi_{k(n)}$. Obviously, μ will also be absolutely continuous with respect to λ and we will denote by f_μ a version of the Radon–Nikodym derivative $\mathrm{d}\mu/\mathrm{d}\lambda$ which we will suppose also to be continuous on B. Further, we will suppose that both f_ν and f_μ are bounded away from zero on B. Let us fix $s_1, \dots, s_r \in B$ and denote by $I_{n,1}, \dots, I_{n,r}$ the sets in partition $\Pi_{k(n)}$ containing each one of the given points. To prove the convergence in distribution of the finite-dimensional distributions, we will need a weak form of weak stationarity on the sample, expressed by the conditions to be imposed on the decomposition of the covariance measures (13). The proof is based on the method used in the proof of Theorem 9 in Oliveira and Suquet [32], consisting in approximating the sums involved by the sums of suitably defined blocks and showing that we may reason as if these blocks were independent. For this latter part, the main tool is the inequality proved in Theorem 16 in Newman [27], regarding the characteristic functions of associated random vectors. Before we proceed with the result regarding the finite-dimensional distributions of the estimator, we state a lemma which is a suitable version of the inequality just referred to.

LEMMA 4.1. *Let Y_n, $n \in \mathbb{N}$, be associated random variables, let $r \in \mathbb{N}$ and let $\alpha_0, \dots, \alpha_r \in \mathbb{R}$. For each $n \in \mathbb{N}$, define*

$$X_n = \sum_{k=0}^{r} \alpha_k Y_{k+n} \quad \text{and} \quad \overline{X}_n = \sum_{k=0}^{r} |\alpha_k|\, Y_{k+n}.$$

Then, for every $u_1, \dots, u_r \in \mathbb{R}$,

$$\left| \mathbb{E} e^{i \sum_{j=1}^{m} u_j X_j} - \prod_{j=1}^{m} \mathbb{E} e^{i u_j X_j} \right| \leqslant 2 \sum_{k \neq l} \left| u_k u_l \operatorname{Cov}(\overline{X}_k, \overline{X}_l) \right|.$$

Proof. For each $n \in \mathbb{N}$, define $f_n(y_1, y_2, \dots) = \sum_{k=0}^{r} \alpha_k y_{k+n}$ and $\overline{f}_n(y_1, y_2, \dots) = \sum_{k=0}^{r} |\alpha_k|\, y_{k+n}$. Then $f_n(y_1, y_2, \dots) + \overline{f}_n(y_1, y_2, \dots) = \sum_{k=0}^{r} (\alpha_k + |\alpha_k|) y_{k+n}$ and $\overline{f}_n(y_1, y_2, \dots) - f_n(y_1, y_2, \dots) = \sum_{k=0}^{r} (|\alpha_k| - \alpha_k) y_{k+n}$, both are coordinatewise increasing, as the coefficients of these linear combinations are non-negative. Thus, we may apply Theorem 16 of Newman [27], which yields the conclusion of this lemma. □

For each $j, k \in \mathbb{N}$, let us introduce the measures

$$\theta_{j,k} = \frac{1}{k} \sum_{l,l'=(j-1)k+1}^{jk} \mathrm{Cov}(\zeta_{1,l}, \zeta_{2,l'}), \tag{13}$$

where $\zeta_{1,l} = \xi_l$ or $\zeta_{1,l} = \eta_l$ for every $l \in \mathbb{N}$, and analogously for $\zeta_{2,l}$. Decomposition **(M)** defines measures which we will denote by $m_{1,j,k}^{\zeta_1,\zeta_2}$ and $m_{2,j,k}^{\zeta_1,\zeta_2}$, where $\zeta_1, \zeta_2 \in \{\xi, \eta\}$, and analogously for the corresponding Radon–Nikodym derivatives.

In the course of the proof of the next theorem and lemmas we need to assume that the sequence t_k, $k \in \mathbb{N}$, is such that the differences $t_{k+1} - t_k$ are strictly increasing. As this is true, at least for some subsequence, we will assume in the sequel that this property is satisfied.

THEOREM 4.2. *Suppose the sequence of partitions Π_k, $k \in \mathbb{N}$, satisfies* **(P1)–(P5)** *with respect to λ, and*

$$n\,h_n \longrightarrow +\infty, \tag{14}$$

$$\frac{h_{n+1}}{h_n} \longrightarrow 1. \tag{15}$$

For $k \in \mathbb{N}$, let m be the largest integer less than or equal to n/k. For each choice of $\zeta_1, \zeta_2 \in \{\xi, \eta\}$, suppose that the measures $\theta_{j,k}$ satisfy condition **(M)** *with respect to λ and the Radon–Nikodym derivatives defined there satisfy*

$$\sup_{j,k,n\in\mathbb{N},\, jk\leqslant n}\ \sup_{x\in B} \left|\gamma_{1,j,k}^{\zeta_1,\zeta_2}(x)\right| \leqslant c_0 < \infty, \tag{16}$$

$$\lim_{m\to+\infty} \frac{1}{m} \sum_{j=1}^{m} \gamma_{2,j,k}^{\zeta_1,\zeta_2} = g_{2,k}^{\zeta_1,\zeta_2} \quad \textit{uniformly on } B, \tag{17}$$

$$\lim_{k\to+\infty} g_{2,k}^{\zeta_1,\zeta_2} = g_2^{\zeta_1,\zeta_2} \quad \textit{uniformly on } B, \tag{18}$$

for some functions $g_{2,k}^{\zeta_1,\zeta_2}$ and $g_2^{\zeta_1,\zeta_2}$ continuous on B. Suppose further that for every sequence $I_n \in \cup_{k=1}^{\infty}\Pi_k$ decreasing to a discrete set and every constant $C > 0$,

$$\int_{\{\zeta_2^2(I_n)>Cnh_n\}} \frac{1}{h_n}\zeta_1^2(I_n)\,\mathrm{d}\mathbb{P} \longrightarrow 0 \tag{19}$$

for every choice $\zeta_1, \zeta_2 \in \{\xi, \eta\}$. Then, the random vector

$$n^{1/2} h_n^{-1/2}(\overline{\eta}_n(I_{n,1}) - \mu(I_{n,1}), \ldots, \overline{\eta}_n(I_{n,r}) - \mu(I_{n,r}),$$
$$\overline{\xi}_n(I_{n,1}) - \nu(I_{n,1}), \ldots, \overline{\xi}_n(I_{n,r}) - \nu(I_{n,r})) \tag{20}$$

converges in distribution to a centered Gaussian random vector with covariance matrix

$$\Gamma = \begin{bmatrix} g_2^{\eta,\eta}(s_1,s_1) & \cdots & 0 & g_2^{\xi,\eta}(s_1,s_1) & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & g_2^{\eta,\eta}(s_r,s_r) & 0 & \cdots & g_2^{\xi,\eta}(s_r,s_r) \\ g_2^{\xi,\eta}(s_1,s_1) & \cdots & 0 & g_2^{\xi,\xi}(s_1,s_1) & \cdots & 0 \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & g_2^{\xi,\eta}(s_r,s_r) & 0 & \cdots & g_2^{\xi,\xi}(s_r,s_r) \end{bmatrix}.$$

The proof of this theorem follows several steps taking care of the approximations needed to handle the dependence of the variables. In order to improve readability, we will present this proof divided into four lemmas, followed by a final step, presented as the proof of the theorem itself, gathering all the partial results. Before embarking in the proof of the lemmas we give a brief description of the step that is accomplished in each one of the following lemmas. Lemma 4.3 shows that we only need to treat those values of n which are multiples of the fixed integer k. The variable introduced in the previous lemma is decomposed into the sum of several dependent variables. The usual coupling technique replaces these variables with independent ones with the same distributions. Lemmas 4.4 and 4.5 justify the use of this coupling by controlling the difference of the respective characteristic functions. This control has to be accomplished in two steps due to the nature of the variables treated. Finally, Lemma 4.6 shows that, after coupling, the Lindeberg condition is satisfied, so the Central Limit Theorem holds.

In the course of the proof we will need some notation which will be used throughout the lemmas. Let $c_1, \ldots, c_r, d_1, \ldots, d_r \in \mathbb{R}$ be fixed and, for each $n \in \mathbb{N}$, $i = 1, \ldots, n$, $q = 1, \ldots, r$, define the random variables

$$T_{n,i}^q = \frac{1}{\sqrt{h_n}}[c_q(\xi_i(I_{n,q}) - \nu(I_{n,q})) + d_q(\eta_i(I_{n,q}) - \mu(I_{n,q}))]$$

and

$$T_n^q = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} T_{n,i}^q, \qquad Z_{n,i} = \sum_{q=1}^{r} T_{n,i}^q, \qquad Z_n = \frac{1}{\sqrt{n}}\sum_{i=1}^{n} Z_{n,i} = \sum_{q=1}^{r} T_n^q.$$

For each $j = 1, \ldots, m$ and $q = 1, \ldots, r$ define

$$Y_{n,j}^q = \frac{1}{\sqrt{k}}\sum_{l=(j-1)k+1}^{jk} T_{n,l}^q.$$

Then

$$T_{mk}^q = \frac{1}{\sqrt{mk}}\sum_{i=1}^{mk} T_{mk,i}^q = \frac{1}{\sqrt{m}}\sum_{j=1}^{m} Y_{mk,j}^q.$$

The variable Z_n is the linear combination of the coordinates of (20), required for the application of the Cramér-Wold Theorem, while the variables $Y^q_{n,j}$ correspond to the blocks in which the sums are decomposed.

LEMMA 4.3. *Suppose the assumptions of Theorem 4.2 are satisfied and let k be fixed. Then, at least one of the following convergences hold:*

$$\lim_{n\to+\infty} \left|\mathbb{E}e^{\mathrm{i}uZ_n} - \mathbb{E}e^{\mathrm{i}uZ_{mk}}\right| = 0 \tag{21}$$

or

$$\lim_{n\to+\infty} \left|\mathbb{E}e^{\mathrm{i}uZ_n} - \mathbb{E}e^{\mathrm{i}uZ_{(m+1)k}}\right| = 0. \tag{22}$$

Proof. For fixed k and large enough n, there is at most one change of partition between the sample sizes mk and $(m+1)k$. Suppose for the moment there are no changes of partitions, or, if there is one corresponding to the sample size $t_l \in [mk, (m+1)k)$, then $mk \leqslant n < t_l$. In this case, we approximate Z_n by Z_{mk}.

$$\begin{aligned}\left|\mathbb{E}e^{\mathrm{i}uZ_n} - \mathbb{E}e^{\mathrm{i}uZ_{mk}}\right| &\leqslant \mathbb{E}[|u|\,|Z_n - Z_{mk}|]\\ &\leqslant |u|\,\mathrm{Var}^{1/2}(Z_n - Z_{mk})\\ &\leqslant |u|\left[\frac{1}{\sqrt{n}}\mathrm{Var}^{1/2}\left(\sum_{i=1}^{mk}(Z_{n,i} - Z_{mk,i})\right) + \right.\\ &\quad + \left(\frac{1}{\sqrt{mk}} - \frac{1}{\sqrt{n}}\right)\mathrm{Var}^{1/2}\left(\sum_{i=1}^{mk} Z_{mk,i}\right) + \\ &\quad \left. + \frac{1}{\sqrt{n}}\mathrm{Var}^{1/2}\left(\sum_{i=mk+1}^{n} Z_{n,i}\right)\right].\end{aligned} \tag{23}$$

We now prove that this sum converges to zero. The square of the first term is

$$\begin{aligned}&\frac{1}{n}\mathrm{Var}\left(\sum_{i=1}^{mk}(Z_{n,i} - Z_{mk,i})\right)\\ &\quad = \frac{1}{n}\sum_{i,j=1}^{mk}\big[\mathrm{Cov}(Z_{n,i}, Z_{n,j}) - \mathrm{Cov}(Z_{n,i}, Z_{mk,j}) -\\ &\quad - \mathrm{Cov}(Z_{mk,i}, Z_{n,j}) + \mathrm{Cov}(Z_{mk,i}, Z_{mk,j})\big].\end{aligned} \tag{24}$$

Expanding the first of these terms, we find

$$\begin{aligned}&\frac{1}{n\,h_n}\sum_{q,q'=1}^{r}\sum_{i,j=1}^{mk}[c_q c_{q'}\,\mathrm{Cov}(\xi_i(I_{n,q}), \xi_j(I_{n,q'})) +\\ &\quad + c_q d_{q'}\,\mathrm{Cov}(\xi_i(I_{n,q}), \eta_j(I_{n,q'})) + d_q c_{q'}\,\mathrm{Cov}(\eta_i(I_{n,q}), \xi_j(I_{n,q'})) +\\ &\quad + d_q d_{q'}\,\mathrm{Cov}(\eta_i(I_{n,q}), \eta_j(I_{n,q'}))].\end{aligned} \tag{25}$$

Using the decomposition **(M)** the first term of this last expansion equals

$$\frac{mk\,h_{mk}}{n\,h_n}\frac{m^{\xi,\xi}_{1,1,mk}(I_{n,q}\times I_{n,q})}{h_{mk}}+\frac{mk}{n}\frac{m^{\xi,\xi}_{2,1,mk}(I_{n,q}\times I_{n,q})}{h_{mk}}.$$

Now,

$$\frac{m^{\xi,\xi}_{1,1,mk}(I_{n,q}\times I_{n,q})}{h_{mk}}\leqslant c_0\frac{\lambda(I_{n,q})\lambda(I_{n,q'})}{h_n}\longrightarrow 0$$

by the assumptions on the partitions. The second term on the decomposition equals 0 if $q\neq q'$, as in this case the set $I_{n,q}\times I_{n,q'}$ does not intersect the diagonal of the product space. When $q=q'$, we find, by (18),

$$\frac{m^{\xi,\xi}_{2,1,mk}(I_{n,q}\times I_{n,q})}{h_{mk}}=\frac{m^{\xi,\xi}_{2,1,mk}(I^*_{n,q})}{\lambda^*(I^*_{n,q})}\longrightarrow g_2^{\xi,\xi}(s_q,s_q).$$

The remaining terms in (25) are treated analogously. Thus, remembering that $mk/n\to 1$, we get,

$$\lim_{n\to+\infty}\frac{1}{n}\sum_{i,j=1}^{mk}\mathrm{Cov}(Z_{n,i},Z_{n,j})$$
$$=\sum_{q=1}^{r}\Big(c_q^2\,g_2^{\xi,\xi}(s_q,s_q)+2c_qd_q\,g_2^{\xi,\eta}(s_q,s_q)+d_q^2\,g_2^{\eta,\eta}(s_q,s_q)\Big)$$

(note that we should consider two terms corresponding to $g_2^{\xi,\eta}$ and $g_2^{\eta,\xi}$, but as we only need their values on the diagonal, these terms coincide). The fourth term in (24) is analogous to the one just discussed. The second and third are slightly different, requiring the use of the sequence $t_l,\ l\in\mathbb{N}$. In fact,

$$\frac{1}{n}\sum_{i,j=1}^{mk}\mathrm{Cov}(Z_{n,i},Z_{mk,j})$$
$$=\frac{1}{n\sqrt{h_nh_{mk}}}\sum_{q,q'=1}^{r}\sum_{i,j=1}^{mk}[c_qc_{q'}\,\mathrm{Cov}(\xi_i(I_{n,q}),\xi_j(I_{mk,q'}))+$$
$$+c_qd_{q'}\,\mathrm{Cov}(\xi_i(I_{n,q}),\eta_j(I_{mk,q'}))+d_qc_{q'}\,\mathrm{Cov}(\eta_i(I_{n,q}),\xi_j(I_{mk,q'}))+$$
$$+d_qd_{q'}\,\mathrm{Cov}(\eta_i(I_{n,q}),\eta_j(I_{mk,q'}))].$$

We supposed that there was no change of partition between mk and $(m+1)k$ or that $mk\leqslant n<t_l<(m+1)k$. In either case, it follows that $I_{mk,q'}=I_{n,q'}$, so the convergence of this expression to

$$\sum_{q=1}^{r}\Big(c_q^2\,g_2^{\xi,\xi}(s_q,s_q)+2c_qd_q\,g_2^{\xi,\eta}(s_q,s_q)+d_q^2\,g_2^{\eta,\eta}(s_q,s_q)\Big)$$

follows as in the discussion of the first term in (24).

So, adding up these terms, we finally get that

$$\frac{1}{\sqrt{n}} \operatorname{Var}\left(\sum_{i=1}^{mk}(Z_{n,i} - Z_{mk,i})\right) \longrightarrow 0.$$

We now proceed with the second term in (23). Again expanding its square, we find

$$\left(\frac{1}{\sqrt{mk}} - \frac{1}{\sqrt{n}}\right)^2 \frac{1}{h_{mk}} \sum_{i,j=1}^{mk} \operatorname{Cov}(Z_{mk,i}, Z_{mk,j})$$
$$= \left(1 - \frac{\sqrt{mk}}{\sqrt{n}}\right)^2 \frac{1}{mk\, h_{mk}} \sum_{q,q'=1}^{r} \sum_{i,j=1}^{mk} [c_q c_{q'} \operatorname{Cov}(\xi_i(I_{mk,q}), \xi_j(I_{mk,q'})) +$$
$$+ c_q d_{q'} \operatorname{Cov}(\xi_i(I_{mk,q}), \eta_j(I_{mk,q'})) + d_q c_{q'} \operatorname{Cov}(\eta_i(I_{mk,q}), \xi_j(I_{mk,q'})) +$$
$$+ d_q d_{q'} \operatorname{Cov}(\eta_i(I_{mk,q}), \eta_j(I_{mk,q'}))].$$

All the terms have now the same form as those in (25), so the above expression converges to zero, as $(1 - \sqrt{mk}/\sqrt{n})^2 \longrightarrow 0$.

Finally, we investigate the third term in (23). Expanding its square, we find

$$\frac{1}{n} \sum_{i,j=mk+1}^{n} \operatorname{Cov}(Z_{n,i}, Z_{n,j})$$
$$= \frac{1}{n\, h_n} \sum_{q,q'=1}^{r} \sum_{i,j=mk+1}^{n} [c_q c_{q'} \operatorname{Cov}(\xi_i(I_{n,q}), \xi_j(I_{n,q'})) +$$
$$+ c_q d_{q'} \operatorname{Cov}(\xi_i(I_{n,q}), \eta_j(I_{n,q'})) + d_q c_{q'} \operatorname{Cov}(\eta_i(I_{n,q}), \xi_j(I_{n,q'})) +$$
$$+ d_q d_{q'} \operatorname{Cov}(\eta_i(I_{n,q}), \eta_j(I_{n,q'}))].$$

However, all these terms converge to zero because of (15) and the nonnegativity of the covariances, due to association of the variables. So, we have finally proved that (21) holds.

It remains to check the case where $mk \leqslant t_l \leqslant n < (m+1)k$. Presently, we approximate the characteristic function of Z_n by that of $Z_{(m+1)k}$. The boundedness used in (23) is modified here as follows. In the first two terms of (23), just replace m by $m+1$. This does not affect the arguments used in the subsequent discussion, and $I_{n,q} = I_{(m+1)k,q}$. Thus, the first two terms in (23) converge to zero. The third term in (23) is replaced by

$$\frac{1}{\sqrt{n}} \operatorname{Var}^{1/2}\left(\sum_{i=n+1}^{(m+1)k} Z_{n,i}\right),$$

which converges to zero as was the case in the corresponding term in the analysis carried previously when $mk \leqslant n < t_l < (m+1)k$. So, in this case, we also have that (22) is satisfied. □

LEMMA 4.4. *Suppose the assumptions of Theorem 4.2 are satisfied and let k be fixed. Then,*

$$\lim_{n\to+\infty}\left|\mathbb{E}e^{\mathrm{i}uZ_{mk}}-\prod_{q=1}^{r}\mathbb{E}e^{\mathrm{i}uT^q_{mk}}\right|=0.$$

Proof. We now work with the difference between $Z_{mk}=\sum_q T^q_{mk}$ and the same sum if the variables $T^1_{mk},\dots,T^r_{mk}$ were independent. For each $n\in\mathbb{N}$, $i=1,\dots,n$, $q=1,\dots,r$ define the random variables

$$\overline{T}^q_{n,i}=\frac{1}{\sqrt{h_n}}[|c_q|(\xi_i(I_{n,q})-\nu(I_{n,q}))+|d_q|(\eta_i(I_{n,q})-\mu(I_{n,q}))]$$

and let

$$\overline{T}^q_n=\frac{1}{\sqrt{n}}\sum_{i=1}^{n}\overline{T}^q_{n,i}.$$

By Lemma 4.1,

$$\begin{aligned}&\left|\mathbb{E}e^{\mathrm{i}uZ_{mk}}-\prod_{q=1}^{r}\mathbb{E}e^{\mathrm{i}uT^q_{mk}}\right|\\&\quad=\left|\mathbb{E}e^{\mathrm{i}u\sum_q T^q_{mk}}-\prod_{q=1}^{r}\mathbb{E}e^{\mathrm{i}uT^q_{mk}}\right|\leqslant 2u^2\sum_{q\neq q'}\mathrm{Cov}(\overline{T}^q_{mk},\overline{T}^{q'}_{mk}).\end{aligned}$$

This expression converges to zero, as is easily seen by expanding one of the covariance terms,

$$\begin{aligned}\mathrm{Cov}(\overline{T}^q_{mk},\overline{T}^{q'}_{mk}) &= \frac{1}{mk}\sum_{i,j=1}^{mk}\mathrm{Cov}(\overline{T}^q_{mk,i},\overline{T}^{q'}_{mk,j})\\&= \frac{1}{mk\,h_{mk}}\sum_{i,j=1}^{mk}[|c_qc_{q'}|\,\mathrm{Cov}(\xi_i(I_{mk,q}),\xi_j(I_{mk,q'}))\ +\\&\quad+|c_qd_{q'}|\,\mathrm{Cov}(\xi_i(I_{mk,q}),\eta_j(I_{mk,q'}))+\\&\quad+|d_qc_{q'}|\,\mathrm{Cov}(\eta_i(I_{mk,q}),\xi_j(I_{mk,q'}))+\\&\quad+|d_qd_{q'}|\,\mathrm{Cov}(\eta_i(I_{mk,q}),\eta_j(I_{mk,q'}))].\end{aligned}$$

Using the decomposition **(M)**, the first term on the right-hand side above equals

$$\frac{m^{\xi,\xi}_{1,1,mk}(I_{mk,q}\times I_{mk,q'})}{h_{mk}}+\frac{m^{\xi,\xi}_{2,1,mk}(I_{mk,q}\times I_{mk,q'})}{h_{mk}}.$$

This expression converges to zero, as was seen above, by taking into account that $q \neq q'$, so that for large enough n, $I_{mk,q} \cap I_{mk,q'} = \emptyset$. □

LEMMA 4.5. *Suppose the assumptions of Theorem 4.2 are satisfied and let k be fixed. Then,*

$$\limsup_{n\to+\infty} \left| \prod_{q=1}^{r} \mathbb{E} e^{\mathrm{i}uT_{mk}^{q}} - \prod_{q=1}^{r}\prod_{j=1}^{m} \mathbb{E} e^{\mathrm{i}\frac{u}{\sqrt{m}} Y_{mk,j}^{q}} \right| \leqslant 2\,u^2 \sum_{q=1}^{r} (\overline{a}^q - \overline{a}_k^q), \tag{26}$$

where

$$\overline{a}^q = c_q^2\, g_2^{\xi,\xi}(s_q, s_q) + 2\left|c_q d_q\right|\; g_2^{\xi,\eta}(s_q, s_q) + d_q^2\, g_2^{\eta,\eta}(s_q, s_q)$$

and

$$\overline{a}_k^q = c_q^2\, g_{2,k}^{\xi,\xi}(s_q, s_q) + 2\left|c_q d_q\right|\; g_{2,k}^{\xi,\eta}(s_q, s_q) + d_q^2\, g_{2,k}^{\eta,\eta}(s_q, s_q).$$

Proof. The sums $T_{mk}^q = (1/\sqrt{m}) \sum_{j=1}^{m} Y_{mk,j}^q$ are now to be approximated by independent summands. With q fixed, we may reason as follows:

$$\begin{aligned} &\left| \prod_{q=1}^{r} \mathbb{E} e^{\mathrm{i}uT_{mk}^{q}} - \prod_{q=1}^{r}\prod_{j=1}^{m} \mathbb{E} e^{\mathrm{i}\frac{u}{\sqrt{m}} Y_{mk,j}^{q}} \right| \\ &\quad\leqslant \sum_{q=1}^{r} \left| \mathbb{E} e^{\mathrm{i}uT_{mk}^{q}} - \prod_{j=1}^{m} \mathbb{E} e^{\mathrm{i}\frac{u}{\sqrt{m}} Y_{mk,j}^{q}} \right|. \end{aligned}$$

For each $j = 1, \ldots, m$ and $q = 1, \ldots, r$ define

$$\overline{Y}_{n,j}^q = \frac{1}{\sqrt{k}} \sum_{l=(j-1)k+1}^{jk} \overline{T}_{n,l}^q.$$

Then, an application of Lemma 4.1 yields

$$\begin{aligned} &\left| \mathbb{E} e^{\mathrm{i}uT_{mk}^{q}} - \prod_{j=1}^{m} \mathbb{E} e^{\mathrm{i}\frac{u}{\sqrt{m}} Y_{mk,j}^{q}} \right| \\ &\quad= \left| \mathbb{E} e^{\mathrm{i}\frac{u}{\sqrt{m}} \sum_{j=1}^{m} Y_{mk,j}^{q}} - \prod_{j=1}^{m} \mathbb{E} e^{\mathrm{i}\frac{u}{\sqrt{m}} Y_{mk,j}^{q}} \right| \\ &\quad\leqslant 2u^2 \sum_{j\neq j'} \frac{1}{m} \operatorname{Cov}(\overline{Y}_{mk,j}^q, \overline{Y}_{mk,j'}^q) \\ &\quad= 2u^2 \left[\frac{1}{mk} \sum_{i,j=1}^{mk} \operatorname{Cov}(\overline{T}_{n,i}^q, \overline{T}_{n,j}^q) - \frac{1}{m} \sum_{j=1}^{m} \operatorname{Cov}(\overline{Y}_{mk,j}^q, \overline{Y}_{mk,j}^q) \right]. \end{aligned} \tag{27}$$

Next,

$$\frac{1}{mk}\sum_{i,j=1}^{mk}\operatorname{Cov}(\overline{T}^q_{n,i},\overline{T}^q_{n,j})$$
$$=\frac{1}{mk\,h_{mk}}\sum_{i,j=1}^{mk}[c_q^2\operatorname{Cov}(\xi_i(I_{mk,q}),\xi_j(I_{mk,q}))+$$
$$+|c_qd_q|\operatorname{Cov}(\xi_i(I_{mk,q}),\eta_j(I_{mk,q}))+|d_qc_q|\operatorname{Cov}(\eta_i(I_{mk,q}),\xi_j(I_{mk,q}))+$$
$$+d_q^2\operatorname{Cov}(\eta_i(I_{mk,q}),\eta_j(I_{mk,q}))]$$

and this expression converges to $\overline{a}^q$, as was seen earlier. The remaining term in (22) is discussed as follows by using decomposition **(M)**:

$$\operatorname{Cov}(\overline{Y}^q_{mk,j},\overline{Y}_{mk,j})$$
$$=\frac{1}{k\,h_{mk}}\sum_{l,l'=(j-1)k+1}^{jk}[c_q^2\operatorname{Cov}(\xi_l(I_{mk,q}),\xi_{l'}(I_{mk,q}))+$$
$$+|c_qd_q|\operatorname{Cov}(\xi_l(I_{mk,q}),\eta_{l'}(I_{mk,q}))+|d_qc_q|\operatorname{Cov}(\eta_l(I_{mk,q}),\xi_{l'}(I_{mk,q}))+$$
$$+d_q^2\operatorname{Cov}(\eta_l(I_{mk,q}),\eta_{l'}(I_{mk,q}))].$$

Utilizing same arguments as above, this expression converges to

$$c_q^2\gamma^{\xi,\xi}_{2,j,k}(s_q,s_q)+2|c_qd_q|\gamma^{\xi,\eta}_{2,j,k}(s_q,s_q)+d_q^2\gamma^{\eta,\eta}_{2,j,k}(s_q,s_q),$$

so that, by using (17),

$$\lim_{n\to+\infty}\frac{1}{m}\sum_{j=1}^{m}\operatorname{Cov}(\overline{Y}^q_{mk,j},\overline{Y}^q_{mk,j})=\overline{a}^q_k.$$

Thus, by relation (27), the inequality (26) follows. □

LEMMA 4.6. *Suppose the assumptions of Theorem 4.2 are satisfied and let k and q be fixed. Then,*

$$\lim_{n\to+\infty}\frac{1}{\operatorname{Var}(\sum_{j=1}^m Y^q_{mk,j})}\int_{\left\{\left|Y^q_{mk,j}\right|>\varepsilon a^q_k\sqrt{m}\right\}}(Y^q_{mk,j})^2\,\mathrm{d}\mathbb{P}=0.$$

Proof. We now show that the Lindeberg conditions is satisfied by the triangular array $m^{-1/2}\,Y^q_{mk,j},\ j=1,\ldots,m$. So, supposing the variables to be independent, a Central Limit Theorem will follow.

Repeating the arguments in the previous lemmas, is easily shown that

$$\frac{1}{m}\operatorname{Var}\left(\sum_{j=1}^{m}Y^q_{mk,j}\right)\longrightarrow a^q_k$$
$$:=c_q^2\,g^{\xi,\xi}_{2,k}(s_q,s_q)+2c_qd_q\,g^{\xi,\eta}_{2,k}(s_q,s_q)+d_q^2\,g^{\eta,\eta}_{2,k}(s_q,s_q).$$

In such a case Lindeberg condition reduces to

$$\sum_{j=1}^{m} \int_{\left\{\left|Y^q_{mk,j}\right|>\varepsilon a^q_k\sqrt{m}\right\}} \frac{1}{m}(Y^q_{mk,j})^2 \,\mathrm{d}\mathbb{P} \longrightarrow 0.$$

The variables $Y^q_{mk,j}$ are defined as sums of the variables $T^q_{mk,l}$, $l = (j-1)k + 1, \ldots, jk$, so we may apply Lemma 4 of Utev [39], which gives an upper bound for the integral of the square of a sum in terms of the sum of the squares of the variables, to this last integral to find the upper bound

$$\begin{aligned}
&\sum_{j=1}^{m} \int_{\left\{\left|\sum_{l=(j-1)k+1}^{jk} T^q_{mk,l}\right|>\varepsilon a^q_k\sqrt{mk}\right\}} \frac{1}{mk}\left(\sum_{l=(j-1)k+1}^{jk} T^q_{mk,l}\right)^2 \mathrm{d}\mathbb{P} \\
&\quad\leqslant \frac{2}{m}\sum_{j=1}^{m}\sum_{j=(j-1)k+1}^{jk} \int_{\left\{\left|T^q_{mk,l}\right|>\frac{\varepsilon a^q_k}{2}\sqrt{\frac{m}{k}}\right\}} (T^q_{mk,l})^2 \,\mathrm{d}\mathbb{P} \\
&\quad= \frac{2}{m}\sum_{j=1}^{mk} \int_{\left\{\left|T^q_{mk,j}\right|>\frac{\varepsilon a^q_k}{2k}\sqrt{mk}\right\}} (T^q_{mk,j})^2 \,\mathrm{d}\mathbb{P}.
\end{aligned}$$

As k is fixed, the above sum has the same form as the one appearing in the proof of Theorem 4.1 in Jacob and Oliveira [21] (see also, Theorem 6.1 in [19]), which was proved to converge to zero on account of (19). □

Proof of Theorem 4.2. Now, in order to complete the proof of the theorem, we set $a^q = c_q^2\, g_2^{\xi,\xi}(s_q, s_q) + 2c_q d_q\, g_2^{\xi,\eta}(s_q, s_q) + d_q^2\, g_2^{\eta,\eta}(s_q, s_q)$, and have

$$\begin{aligned}
&\left|\mathbb{E}e^{\mathrm{i}uZ_{mk}} - e^{-\frac{u^2}{2}\sum_{q=1}^r a^q}\right| \\
&\quad\leqslant \left|\mathbb{E}e^{\mathrm{i}uZ_{mk}} - \prod_{q=1}^{r}\mathbb{E}e^{\mathrm{i}uT^q_{mk}}\right| + \sum_{q=1}^{r}\left|\mathbb{E}e^{\mathrm{i}uT^q_{mk}} - \prod_{j=1}^{m}\mathbb{E}e^{\frac{\mathrm{i}u}{\sqrt{m}}Y^q_{mk,j}}\right| + \\
&\qquad + \sum_{q=1}^{r}\left|\prod_{j=1}^{m}\mathbb{E}e^{\frac{\mathrm{i}u}{\sqrt{m}}Y^q_{mk,j}} - e^{-\frac{u^2}{2}a^q_k}\right| + \sum_{q=1}^{r}\left|e^{-\frac{u^2}{2}a^q_k} - e^{-\frac{u^2}{2}a^q}\right|. \qquad (28)
\end{aligned}$$

For the moment, suppose k is fixed. The first term in the last expression above converges to zero by Lemma 4.4. The third term converges to zero by Lemma 4.6. In fact, the product appearing in this term is the characteristic function of the vector $m^{-1/2}(Y^q_{mk,1}, \ldots, Y^q_{mk,m})$, supposing the coordinates to be independent. A shown in Lemma 4.6 this converges to a centered Gaussian vector with covariance matrix of the same form as Γ but with the $g_2^{\zeta_1,\zeta_2}$ replaced by $g_{2,k}^{\zeta_1,\zeta_2}$, so the convergence to zero of the third term in (28) follows. So, taking into account of Lemma 4.5, we have, for each $k \in \mathbb{N}$,

$$\limsup_{m\to+\infty} \left| \mathbb{E} e^{\mathrm{i}uZ_{mk}} - e^{-\frac{u^2}{2}\sum_{q=1}^r a^q} \right| \leqslant 2u^2 \sum_{q=1}^r (\overline{a}^q - \overline{a}_k^q) + \sum_{q=1}^r \left| e^{-\frac{u^2}{2}a_k^q} - e^{-\frac{u^2}{2}a^q} \right|.$$

Finally, letting $k \longrightarrow +\infty$, the expression on the right-hand side above converges to zero on account of (18). Thus, the convergence in distribution asserted in (20) is established. □

A discussion on condition (19) has been presented by the authors in [19], indicating that it is a reasonable one. It is fulfilled in Poisson processes and also some other point processes constructed from Poisson processes.

An application of the δ-method yields the convergence of the finite-dimensional distributions of the estimator f_n itself.

THEOREM 4.7. *Suppose the conditions of Theorem 4.2 are satisfied. Then*

$$n^{1/2} h_n^{1/2} \left(\frac{\overline{\eta}_n(I_{n,1})}{\overline{\xi}_n(I_{n,1})} - \frac{\mu(I_{n,1})}{\nu(I_{n,1})}, \dots, \frac{\overline{\eta}_n(I_{n,r})}{\overline{\xi}_n(I_{n,r})} - \frac{\mu(I_{n,r})}{\nu(I_{n,r})} \right) \tag{29}$$

converges in distribution to a centered Gaussian random vector with diagonal covariance matrix Γ^* *with*

$$\gamma_{q,q}^* = \frac{g_2^{\xi,\xi}(s_q, s_q)}{f_\mu^2(s_q)} - \frac{2g_2^{\xi,\eta}(s_q, s_q) f_\mu(s_q)}{f_\nu^3(s_q)} + \frac{f_\mu^2(s_q) g_2^{\eta,\eta}(s_q, s_q)}{f_\nu^4(s_q)}, \quad q = 1, \dots, r.$$

Proof. Define the random vector

$$U_n = \left(\overline{\eta}_n(I_{n,1}), \dots, \overline{\eta}_n(I_{n,r}), \overline{\xi}_n(I_{n,1}), \dots, \overline{\xi}_n(I_{n,r}) \right)$$

and the real valued function φ on $\mathbb{R}^{2r}$ by

$$\varphi(y) = \sum_{q=1}^r b_q \frac{y_q}{y_{r+q}},$$

where $b_1, \dots, b_q$ are real numbers, so that $\sqrt{nh_n}\,(\varphi(U_n) - \varphi(\mathbb{E}U_n))$ is a linear combination of the coordinates in (29). Using the Taylor expansion, we find

$$\sqrt{n\,h_n}\,(\varphi(U_n) - \varphi(\mathbb{E}U_n)) = \sum_{q=1}^{2r} h_n \frac{\partial \varphi}{\partial y_q}(\mathbb{E}U_n) \sqrt{\frac{n}{h_n}} (U_{n,q} - \mathbb{E}U_{n,q}) + h_n \sqrt{\frac{n}{h_n}} \| U_n - \mathbb{E}U_n \| \varepsilon(\| U_n - \mathbb{E}U_n \|),$$

where ε is continuous and $\lim_{y\to 0}\varepsilon(y) = 0$. As $U_n \longrightarrow \mathbb{E}U_n$ in probability, by Theorem 3.1, and $n^{1/2}h_n^{-1/2}\| U_n - \mathbb{E}U_n \|$ converges in distribution, the last term above converges in probability to zero. Now consider the vector $u = (f_\mu(s_1), \ldots, f_\mu(s_r), f_\nu(s_1), \ldots, f_\nu(s_r))$ and rewrite the first term of the Taylor expansion as follows:

$$\sum_{q=1}^{2r} \frac{\partial\varphi}{\partial y_q}(u)\sqrt{\frac{n}{h_n}}(U_{n,q} - \mathbb{E}U_{n,q}) + \\ + \sum_{q=1}^{2r}\left[h_n \frac{\partial\varphi}{\partial y_q}(\mathbb{E}U_n) - \frac{\partial\varphi}{\partial y_q}(u)\right]\sqrt{\frac{n}{h_n}}(U_{n,q} - \mathbb{E}U_{n,q}).$$

Computing the derivatives, it is easily seen that

$$h_n \frac{\partial\varphi}{\partial y_q}(\mathbb{E}U_n) \longrightarrow \frac{\partial\varphi}{\partial y_q}(u),$$

so that the limiting distribution of $\sqrt{nh_n}\,(\varphi(U_n) - \varphi(\mathbb{E}U_n))$ is the same as that of

$$\sum_{q=1}^{2r} \frac{\partial\varphi}{\partial y_q}(u)\sqrt{\frac{n}{h_n}}(U_{n,q} - \mathbb{E}U_{n,q}),$$

which, in Theorem 4.2 was shown to be Gaussian. Its variance is easily shown to be

$$\sum_{q=1}^{2r} b_q^2 \left(\frac{g_2^{\xi,\xi}(s_q, s_q)}{f_\mu^2(s_q)} - \frac{2g_2^{\xi,\eta}(s_q, s_q) f_\mu(s_q)}{f_\nu^3(s_q)} + \frac{f_\mu^2(s_q) g_2^{\eta,\eta}(s_q, s_q)}{f_\nu^4(s_q, s_q)}\right),$$

as it follows by replacing c_q and d_q, in the computation of the variance in the Theorem 4.2 by $b_q / f_\mu(s_q)$ and $b_q f_\mu(s_q)/f_\nu^2(s_q)$, respectively. □

Acknowledgements

The authors would like to express their gratitude to the anonymous referee for his careful reading of the first version and help on the preparation of a better version of this article. P. E. Oliveira was supported by grant PRAXIS/2/2.1/MAT/19/94 from FCT and Centro de Matemática da Universidade de Coimbra.

References

1. Abou-Jaoudé, S.: La convergence L^1 et L^∞ de certains estimateurs d'une densité de probabilité, Thesis, University of Pierre et Marie Curie, Paris, 1977.
2. Beirlant, J., Berlinet, A. and Györfi, L.: On piecewise linear density estimators, Statist. Neerlandica, to appear.

3. Bensaïd, N.: Contribution à l'estimation et à la prévision non paramétrique d'un processus ponctuel multidimensionnel, Thesis, University of Paris VI, 1992.
4. Bensaïd, N. and Fabre, J.: Estimation par noyau d'une derivée de Radon–Nikodym sous des conditions de mélange, *Canad. J. Stat.* **28** (1998), 267–282.
5. Birkel, T.: Moment Bounds for associated sequences, *Ann. Probab.* **16** (1988), 1184–1193.
6. Bosq, D.: *Nonparametric Statistics for Stochastic Processes*, Lecture Notes in Statistics 110, Springer, 1996.
7. Bosq, D. and Nguyen, H. T.: *A Course in Stochastic Processes. Stochastic Models and Statistical Inference*, Kluwer Acad. Publ., Dordrecht, 1996.
8. Bosq, D. and Lecoutre, J. P.: Théorie de l'estimation fonctionelle, Economica, Paris, 1987.
9. Burton, R. M. and Waymire, E.: The central limit problem for infinitely divisible random measures, in E. Eberlein, M. S. Taqqu (eds), *Dependence in Probability and Statistics*, Birkhäuser, 1986, pp. 383–395.
10. Daley, D. J. and Vere-Jones, D.: An introduction to the theory of point processes, Springer-Verlag, New York, Berlin, 1988.
11. Ellis, S. P.: Density estimation for point processes, *Stoch. Proc. Appl.* **39** (1991), 345–358.
12. Evans, S.: Association and random measures, *Probab. Th. Related Fields* **86** (1990), 1–19.
13. Ferrieux, D.: Estimation de densités de mesures moyennes de processus ponctuels associés, Thesis, University of Montpellier II, 1996.
14. Ferrieux, D.: Estimation à noyau de densités moyennes de mesures aléatoires associées, Compte Rendus Acad. Sciences de Paris **326** (1998), Série I, 1131–1134.
15. Grenander, U.: *Abstract Inference*, Wiley, New York, 1981.
16. Ioannides, D. A. and Roussas, G. G.: Exponential inequality for associated random variables, *Stat. Probab. Lett.* **42** (1999), 423–431.
17. Jacob, P. and Mendes Lopes, N.: Un théorème central limite fonctionnel pour un estimateur d'un noyau de transition, *Port. Math.* **48** (1991), 217–231.
18. Jacob, P. and Niéré, L.: Contribution à l'estimation des lois de Palm d'une mesure aléatoire, *Pub. Inst. Stat. Univ. Paris* **35** (1990), 39–49.
19. Jacob, P. and Oliveira, P. E.: A general approach to nonparametric histogram estimation, *Statistics* **27** (1995), 73–92.
20. Jacob, P. and Oliveira, P. E.: Kernel estimators of general Radon–Nikodym derivatives, *Statistics* **30** (1997), 25–46.
21. Jacob, P. and Oliveira. P. E.: Using non-embedded partitions in histogram estimation, Preprint, Pré-publicações Dep. Matemática, University of Coimbra, 1997, pp. 97–25.
22. Kallenberg, O.: *Random Measures*, Academic Press, London, New York, 1983.
23. Karr, A.: *Point Processes and Their Statistical Inference*, Marcel Dekker, New York, 1986.
24. Karr, A.: Estimation of Palm measures of stationary point processes, *Probab. Th. Related Fields* **74** (1987), 55–69.
25. Karr, A.: Palm distributions of point processes and their applications to statistical inference. Statistical inference from stochastic processes, Contemp. Math. 87, Amer. Math. Soc., Providence, 1988.
26. Mendes Lopes, N.: Uniform consistency of an estimator class of mean local distribution of a random composite measure, *Pub. IRMA Lille* **31** (1988), I.
27. Newman, C.: Asymptotic independence and limit theorems for positively and negatively dependent random variables, in Y. L. Tong (ed.), *Inequalities in Statistics and Probability*, IMS Lecture Notes – Monograph Series, Vol. 5, 1984, pp. 127–140.
28. Newman, C. and Wright, A. L.: An invariance principle for certain dependent sequences, *Ann. Probab.* **9** (1981), 671–675.
29. Niéré, L.: Estimation des distributions de Palm, Thesis, University of Lille, 1987.
30. Oliveira, P. E. and Suquet, C.: $L^2[0, 1]$ weak invariance principle of the empirical process for dependent variables, in *Proc. XVèmes Rencontres Franco-Belges de Statisticiens (Ondelettes*

et Statistique), Lecture Notes in Statistics 103, Wavelets and Statistics, A. Antoniadis, G. Oppenheim (eds), 1995, pp. 331–344.

31. Oliveira, P. E. and Suquet, C.: Processus empirique sous dépendence dans L^2, *Compte Rendus Acad. Sciences de Paris* **321** (1995), Série I, 939–944.
32. Oliveira, P. E. and Suquet, C.: An $L^2[0, 1]$ invariance principle for LPQD random variables, *Port. Math.* **53** (1996), 367–379.
33. Roussas, G. G.: Kernel estimates under association: strong uniform consistency, *Stat. Probab. Lett.* **12** (1991), 393–403.
34. Roussas, G. G.: Curve estimation in random fields of associated processes, *J. Nonparam. Stat.* **2** (1993), 215–224.
35. Roussas, G. G.: Asymptotic normality of random fields of positively or negatively associated processes, *J Multiv. Analysis* **50** (1994), 152–173.
36. Saleh, S.: Mesure composite aléatoire, *Pub. Inst. Stat. Univ. Paris* **29** (1984), 59–67.
37. Saleh, S.: Estimation de la distribution locale moyenne d'une measure composite aléatoire, Comptes Rendus de l'Acad. Sciences Paris 301, Série 1, 545–548.
38. Shao, Q. M. and Yu, H.: Weak convergence for weighted empirical processes of dependent sequences, *Ann. Probab.* **24** (1996), 2052–2078.
39. Utev, S. A.: On the central limit theorem for φ-mixing arrays of random variables, *Th. Probab. Appl.* **35** (1990), 131–139.
40. Yu, H.: A Glivenko-Cantelli lemma and weak convergence for empirical processes of associated sequences, *Probab. Th. Related Fields* **95** (1993), 357–370.