Central limit theorems for smoothed extreme value estimates of point processes boundaries

Stéphane Girard, Ludovic Menneteau

Central limit theorems for smoothed extreme value estimates of Poisson point processes boundaries

Stéphane Girard[1] & Ludovic Menneteau[2]

[1] SMS/LMC, Université Grenoble 1,

BP 53, 38041 Grenoble Cedex 9, France.

Stephane.Girard@imag.fr

[2] Département de Mathématiques, Université Montpellier 2,

Place Eugène Bataillon, 34095 Montpellier Cedex 5, France.

mennet@math.univ-montp2.fr

Short title: Smoothed estimates of boundaries

Abstract

In this paper, we give sufficient conditions to establish central limit theorems for boundary estimates of Poisson point processes. The considered estimates are obtained by smoothing some bias corrected extreme values of the point process. We show how the smoothing leads Gaussian asymptotic distributions and therefore pointwise confidence intervals. Some new unidimensional and multidimensional examples are provided.

Keywords: Functional estimate, Central limit theorem, Extreme values, Poisson process, Boundary estimation.

AMS Subject Classification: Primary 60G70; Secondary 62M30, 62G05, 62G20.

1 Introduction

Many proposals are given in the literature for estimating a set S given a finite random set N of points drawn from the interior. This problem of support estimation arises in classification (Hardy & Rasson (1982)), clustering problems (Hartigan (1975)), discriminant analysis (Baufays & Rasson (1985)), and outliers detection (Devroye & Wise (1980)). Applications are also found in image analysis. For instance, the segmentation problem can be considered under the support estimation point of view, where the support is a connex bounded set in $\mathbb{R}^2$. We also point out some applications in econometrics (e.g. Deprins, *et al* (1984)). In such cases, the unknown support can be written

$$S = \{(x, y) : x \in E \ ; \ 0 \leq \ y \ \leq \ f(x)\}, \tag{1}$$

where f is an unknown function and E an arbitrary set. The set S is often called a boundary fragment, see Korostelev & Tsybakov (1993), Chapter 3. Then, the problem reduces to estimating f, sometimes called the production frontier (see for instance Härdle *et al* (1995a)). The data consist of pair (X, Y) where X represents the input, possibly multidimensional (labor, energy or capital), used to produce an output Y in a given firm. In such a framework, the value $f(x)$ can be interpreted as the maximum level of output which is attainable for the level of input x. Korostelev *et al* (1995) suppose f to be increasing and concave, from economical considerations, which suggests an adapted estimator, called the DEA (Data Envelopment Analysis) estimator. Its asymptotic distribution is established by Gijbels *et al* (1999).

Here, N is a Poisson point process, with observed points belonging to a subset S defined as in (1) where f is an unknown function which needs not to be monotone. An early paper was written by Geffroy (1964) for independent identically distributed observations from a density. The proposed estimator is a kind of histogram based on the extreme values of the sample. This work was extended in two main directions.

(a) On the one hand, piecewise polynomials were introduced and their optimality in an asymptotic minimax sense is proved under weak assumptions on the rate of decrease β of the density towards 0 and on the number q of continuous derivatives of f by Korostelev &

Tsybakov (1993) and by Härdle *et al* (1995b). The asymptotic distribution is established by Hall *et al* (1998). Extreme values methods are proposed by Hall *et al* (1997) and by Gijbels & Peng (1999) to estimate the parameter β.

(b) On the other hand, different propositions for smoothing Geffroy's estimate were made. Girard & Jacob (2001, 2003a, 2003b) introduced estimates based on kernel regressions and on projection methods. In the same spirit, Gardes (2002) proposed a Faber-Shauder estimate. In each case, the consistency and the limit distribution of the estimator are established.

Finally, the work of Mammen & Tsybakov (1995) offers a general framework for comparing the estimates of type (a) or (b). The optimal rates of convergence are derived for estimates of boundaries which have a smooth parametrisation. The existence of estimates reaching these optimal rates of convergence is proved by the minimization of contrast criteria over classes of functions.

Here, we introduce new estimates of type (b). The considered estimates are obtained by smoothing the bias corrected extreme values of the Poisson process (see Menneteau (2003a) for related work in the iid setting).

This approach offers several advantages. First, the bias correction allows to overcome the classical limitation due to the fact that the data lie below the boundary. Second, the smoothing permits to obtain Gaussian asymptotic distributions. Therefore, it is straightforward to obtain pointwise confidence intervals for $f(x)$ all the more so as our estimates benefit from explicit forms and are easy to compute. Finally, let us note that our estimates offer new features compared to those quoted in (b): i) They are not dedicated to unidimensional boundary estimation problems since there is no restriction on the set E in (1), ii) the bias correction is different and thus, iii) the intensity measure of the point process can be more general, iv) the smoothing is achieved with more general weight functions allowing v) better speeds of convergence than the previous estimates quoted in (b).

2 The boundary estimate

Let $(E, \mathcal{E}, \nu)$ be a probability space and $f : (E, \mathcal{E}) \to (\mathbb{R}^+, \mathcal{B}(\mathbb{R}^+))$ a measurable function, where $\mathcal{B}(\mathbb{R})$ is the Borel σ-algebra on $\mathbb{R}$. Consider $S = \{(x, y) \in E \times \mathbb{R}, 0 \leq y \leq f(x)\}$ and the sequence of Poisson point processes

$$N_n = \left\{ N_n(D) : D \in \mathcal{E} \otimes \mathcal{B}\left(\mathbb{R}^+\right) \right\}, \quad n \geq 1,$$

with intensity measure

$$nc(\nu \otimes \lambda)\, \mathbf{1}_S, \tag{2}$$

where $c > 0$, and λ is the Lebesgue measure on $\mathbb{R}^+$. Let $\{(X_{n,i}, Y_{n,i}),\ 1 \leq i \leq N_n(S)\}$ be the set of points associated to the point process. Our aim is then to estimate S via an estimation of f. Let $k_n \uparrow \infty$ and denote by $\{I_{n,r} :\ 1 \leq r \leq k_n\}$ a measurable partition of E. For all $1 \leq r \leq k_n$, note $\nu_{n,r} = \nu(I_{n,r})$,

$$D_{n,r} = \{(x, y) : x \in I_{n,r},\ 0 \leq y \leq f(x)\}$$

the cell of S built on $I_{n,r}$ and $N_{n,r} = N_n(D_{n,r})$. We introduce the extreme values

$$Y^*_{n,r} = \max\{Y_{n,i} : (X_{n,i}, Y_{n,i}) \in D_{n,r}\},$$

if $N_{n,r} \neq 0$ and $Y^*_{n,r} = 0$ otherwise. In the following, the convention $0 \times \infty = 0$ is adopted. For $x \in E$, our estimator of $f(x)$ is

$$\widehat{f}_n(x) = \sum_{r=1}^{k_n} \nu_{n,r} \kappa_{n,r}(x) \left(1 + \frac{1}{N_{n,r}}\right) Y^*_{n,r}, \tag{3}$$

where $\kappa_{n,r} : E \to \mathbb{R}$ is a weighting function determining the nature of the smoothing introduced in the estimate. In the next section, some general conditions are imposed on $\kappa_{n,r}$ and examples are provided in Section 5. It is well-known that $Y^*_{n,r}$ is an estimator of the maximum of f on $I_{n,r}$ with negative bias. The use of the random variable $(1 + N^{-1}_{n,r})Y^*_{n,r}$ allows to reduce this bias. This bias correction is motivated by the remark that, conditionally on $N_{n,r}$, $Y^*_{n,r}$ has approximatively the same distribution as the maximum of $N_{n,r}$ independent random variables uniformly distributed on $[0, \min\{f(x) : x \in I_{n,r}\}]$ (see Lemma 1 ii) below). Therefore, $\widehat{f}_n$ appears as a linear combination of extreme value estimates of sampled values of f. The asymptotic properties of $\widehat{f}_n$ are established

in Section 3, and proved in Section 4. Illustrations are presented in Section 5 with general kernel estimates including Parzen-Rosenblatt and Dirichlet kernels.

3 Main results

Define $\nu_n = \min\{\nu_{n,r},\ 1 \leq r \leq k_n\}$ and

$$\kappa_n(x) = \left(\sum_{r=1}^{k_n} \kappa_{n,r}^2(x) \right)^{1/2}, \ x \in E.$$

Let $m = \sup\{\alpha > 0 : \nu(\{f < \alpha\}) = 0\}$ and $M = \inf\{\alpha > 0 : \nu(\{f > \alpha\}) = 0\}$ be the ν-essential infimum and supremum of f on E. Similarly, for all $1 \leq r \leq k_n$, let

$$m_{n,r} = \sup\{\alpha > 0 : \nu(\{f < \alpha\} \cap I_{n,r}) = 0\},$$

$$M_{n,r} = \inf\{\alpha > 0 : \nu(\{f > \alpha\} \cap I_{n,r}) = 0\}$$

and

$$f_{n,r} = \nu_{n,r}^{-1} \int_{I_{n,r}} f \, d\nu$$

be respectively the ν-essential infimum, the ν-essential supremum and the mean value of f on $I_{n,r}$ and define the ν-essential oscillation of f on $I_{n,r}$ by

$$\Delta_n = \max\{M_{n,r} - m_{n,r},\ 1 \leq r \leq k_n\}.$$

Let us highlight that, in most applications (see Section 5), E is a subset of $\mathbb{R}^d$, ν is absolutely continuous with respect to Lebesgue measure and f is continuous. Hence, all essential infima and essential suprema considered below reduce to the classical minima and maxima.

Finally, set $w_{n,r}(x) = \kappa_{n,r}(x) / \kappa_n(x)$. We consider the following series of assumptions:

(H.1) $k_n \uparrow \infty$ and $n\nu_n \to \infty$ as $n \to \infty$.

(H.2) $0 < m \leq M < +\infty$ and

$$\delta_n := \max_{1 \leq r \leq k_n} \nu_{n,r}(M_{n,r} - m_{n,r}) = o(1/n) \text{ as } n \to \infty.$$

There exists $F \subset E$ such that:

(H.3) For each $(x_1, ..., x_p) \subset F$, there exists a covariance matrix $\Sigma_{(x_1,...,x_p)} = [\sigma(x_i, x_j)]_{1 \leq i,j \leq p}$ in $\mathbb{R}^p$ such that for all $1 \leq i,\ j \leq p$,

$$\sum_{r=1}^{k_n} w_{n,r}(x_i) w_{n,r}(x_j) \to \ \sigma(x_i, x_j) \text{ as } n \to \infty.$$

(H.4) For all $x \in F$,

$$\max_{1 \leq r \leq k_n} |w_{n,r}(x)| \to 0 \text{ as } n \to \infty.$$

(H.5) For all $x \in F$,

$$\left| \sum_{r=1}^{k_n} \nu_{n,r} \kappa_{n,r}(x) f_{n,r} - f(x) \right| = o\left(\frac{\kappa_n(x)}{n} \right) \text{ as } n \to \infty.$$

(H.6) For all $x \in F$,

$$\sum_{r=1}^{k_n} |w_{n,r}(x)| \max\left((n\delta_n)^2, n\nu_n \exp\left(-mcn\nu_n\right), \Delta_n \right) \to 0 \text{ as } n \to \infty.$$

Before proceeding, let us comment on the assumptions. (H.1)–(H.4) are devoted to the control of the centered estimator $\widehat{f_n}(x) - \mathbb{E}(\widehat{f_n}(x))$. Assumption (H.1) imposes that the mean number of points in each $D_{n,r}$ goes to infinity. (H.2) requires the unknown function f to be bounded away from 0. It also imposes that the mean number of points in $D_{n,r}$ above $m_{n,r}$ converges to 0. Note that (H.1) and (H.2) force the ν-essential oscillation of f on $I_{n,r}$ to converge uniformly to 0: $\Delta_n \to 0$ as $n \to \infty$. (H.3) is devoted to the multivariate aspects of the limit theorems. (H.4) imposes to the weight functions $\kappa_{n,r}(x)$ in the linear combination (3) to be approximatively of the same order. This is a natural condition to obtain an asymptotic Gaussian behavior. These assumptions are easy to verify in practice since they involve either $f(x)$ or $\kappa_{n,r}(x)$ without mixing these two quantities. Assumptions (H.5) and (H.6) are devoted to the control of the bias term $\mathbb{E}(\widehat{f_n}(x)) - f(x)$. They prevent it to be too important with respect to the variance of the estimate (which will reveal to be of order $\kappa_n(x)/n$). Consequently, these two assumptions involve both the unknown function $f(x)$ and the weight functions $\kappa_{n,r}(x)$. Finally, (H.6) can be looked at as a stronger version of (H.2).

Our first result states the multivariate central limit theorem for $\widehat{f_n}(x)$.

Theorem 1 *Under assumptions* (H.1)-(H.6) , *and for all* $(x_1, ..., x_p) \subset F$,

$$\left\{ \frac{nc}{\kappa_n(x_j)} \left(\widehat{f}_n(x_j) - f(x_j) \right) : 1 \leq j \leq p \right\} \underset{\mathcal{D}}{\to} N\left(0, \Sigma_{(x_1,...,x_p)}\right),$$

where c *is defined in* (2) , $\underset{\mathcal{D}}{\to}$ *denotes the convergence in distribution and* $N\left(0, \Sigma_{(x_1,...,x_p)}\right)$ *is the centered Gaussian distribution in* $\mathbb{R}^p$, *with covariance matrix* $\Sigma_{(x_1,...,x_p)}$.

In practice, c is not known and has to be estimated. In this aim, we introduce $\widehat{c}_n = N_n(S)/(n\widehat{a}_n)$, where

$$\widehat{a}_n = \sum_{r=1}^{k_n} \nu_{n,r} \left(1 + \frac{1}{N_{n,r}}\right) Y^*_{n,r}$$

is an estimator of $a = \int_E f d\nu$. We then have the following corollary:

Corollary 1 *Theorem 1 holds when* c *is replaced by* $\widehat{c}_n$.

For all $x \in F$, this leads to an explicit asymptotic γth confidence interval for $f(x)$:

$$\left[\sum_{r=1}^{k_n} \nu_{n,r} \left(\kappa_{n,r}(x) - z_\gamma \frac{\kappa_n(x)}{N_n(S)} \right) \left(1 + \frac{1}{N_{n,r}}\right) Y^*_{n,r}, \sum_{r=1}^{k_n} \nu_{n,r} \left(\kappa_{n,r}(x) + z_\gamma \frac{\kappa_n(x)}{N_n(S)} \right) \left(1 + \frac{1}{N_{n,r}}\right) Y^*_{n,r} \right],$$

where z_γ is the $(\gamma+1)/2$th quantile of the $N(0,1)$ distribution. Note that the computation of this interval does not require a bootstrap procedure as for instance in Hall *et al* (1998).

Remark. In the case where the measure ν is unknown, it is natural to introduce the boundary estimate:

$$\mathring{f}_n(x) = \sum_{r=1}^{k_n} \mathring{\nu}_{n,r} \kappa_{n,r}(x) \left(1 + \frac{1}{N_{n,r}}\right) Y^*_{n,r}, \tag{4}$$

where $\mathring{\nu}_{n,r}$ is an estimator of $\nu_{n,r}$. If no prior information is available on ν, one can use the following non-parametric estimate:

$$\mathring{\nu}_{n,r} = \frac{N_{n,r}}{ncY^*_{n,r}(1 + N_{n,r}^{-1})},$$

leading to

$$\mathring{f}_n(x) = \sum_{r=1}^{k_n} \kappa_{n,r}(x) \frac{N_{n,r}}{nc},$$

which has been first introduced by Jacob & Suquet (1995) with a particular choice of the weighting function $\kappa_{n,r}$ and when ν is the Lebesgue measure. Let us note that Theorem 1 does

not hold for $\overset{\circ}{f}_n$ since it converges slower than $\hat{f}_n$, see Jacob & Suquet (1995), Theorem 7. If ν is assumed to belong to a parametric family, another versions of (4) can be used, leading to semi-parametric estimates of f.

4 Proofs of the main results

The proofs are built as follows. First, we establish a multivariate central limit theorem for the finite dimensional projection of the centered process

$$\frac{nc}{\kappa_n(x)}\left(\widehat{f}_n(x) - \mathbb{E}\left(\widehat{f}_n(x)\right)\right), \quad x \in F$$

(see Proposition 1 below). To this aim, by the general framework of the appendix (Theorem 3) it is sufficient to control the centered moments of

$$\xi_{n,r} = \left(1 + \frac{1}{N_{n,r}}\right) Z^*_{n,r} = \left(1 + \frac{1}{N_{n,r}}\right) nc\nu_{n,r} Y^*_{n,r}.$$

This is achieved in Lemma 2. In a second time, we establish that the bias term

$$\frac{nc}{\kappa_n(x)}\left(\mathbb{E}\left(\widehat{f}_n(x)\right) - f(x)\right)$$

vanishes when $n \uparrow \infty$ (see Proposition 2). Finally, we prove in Lemma 3 that c can be replaced by $\widehat{c}_n$ in the multivariate central limit theorem. Before that, we introduce some new notations and definitions needed for our proofs. For all $1 \leq r \leq k_n$, each cell $D_{n,r}$ can be splitted as $D_{n,r} = \widetilde{D}_{n,r} \cup D^-_{n,r} \cup D^+_{n,r}$ where

- $\widetilde{D}_{n,r} = \{(x,y) \in I_{n,r} \times [0, m_{n,r}],\ f(x) < m_{n,r}\}$,
- $D^-_{n,r} = \{(x,y) \in D_{n,r},\ 0 \leq y \leq m_{n,r}\} = (I_{n,r} \times [0, m_{n,r}]) \setminus \widetilde{D}_{n,r}$,
- $D^+_{n,r} = \{(x,y) : x \in I_{n,r},\ m_{n,r} < y \leq f(x)\}$.

Moreover, for all $1 \leq r \leq k_n$, set

- $\lambda_{n,r} = nc\nu_{n,r}m_{n,r}$,

- $\mu_{n,r} = nc\nu_{n,r}f_{n,r}$,
- $N^-_{n,r} = N_n\left(D^-_{n,r}\right)$, $N^+_{n,r} = N_n\left(D^+_{n,r}\right)$,
- $Z^-_{n,r} = nc\nu_{n,r}\max\{Y_{n,i} : (X_{n,i}, Y_{n,i}) \in D^-_{n,r}\}$, if $N^-_{n,r} \neq 0$, and $Z^-_{n,r} = 0$ otherwise,

  $Z^+_{n,r} = nc\nu_{n,r}\max\{Y_{n,i} : (X_{n,i}, Y_{n,i}) \in D^+_{n,r}\}$, if $N^+_{n,r} \neq 0$, and $Z^+_{n,r} = 0$ otherwise,
- $\xi^-_{n,r} = \left(1 + \frac{1}{N^-_{n,r}}\right) Z^-_{n,r}$.

Some technical results are collected in Lemma 1. The second of them is the key tool for proving the following ones. It states that, conditionally on $N^-_{n,r}$, $Z^-_{n,r}$ has the same distribution as the maximum of $N^-_{n,r}$ independent random variables uniformly distributed on $[0, \lambda_{n,r}]$. This motivates the bias correction in (3).

Lemma 1 *Under assumptions* (H.1) *and* (H.2) *we have*

i) $\max\limits_{1\leq r\leq k_n} \mathbb{P}\left(N^+_{n,r} > 0\right) = O\left(n\delta_n\right) = o\left(1\right)$.

ii) For all $1 \leq r \leq k_n$, *and any* $t \in [0, \lambda_{n,r}]$, $\mathbb{P}\left(Z^-_{n,r} \leq t \mid N^-_{n,r}\right) = \left(\frac{t}{\lambda_{n,r}}\right)^{N^-_{n,r}}$.

iii) For all $1 \leq r \leq k_n$, *and any* $t \in [0, \lambda_{n,r}]$, $\mathbb{P}\left(Z^-_{n,r} \leq t\right) = \exp(t - \lambda_{n,r})$.

iv) For all $1 \leq r \leq k_n$, $\mathbb{E}\left(Z^-_{n,r}\right) = \lambda_{n,r} - \left(1 - e^{-\lambda_{n,r}}\right)$.

v) For all $1 \leq r \leq k_n$, $\mathbb{V}\left(Z^-_{n,r}\right) = 1 - 2\lambda_{n,r}e^{-\lambda_{n,r}} - e^{-2\lambda_{n,r}}$.

vi) For all $1 \leq r \leq k_n$, $\mathbb{E}\left(\frac{Z^-_{n,r}}{N^-_{n,r}}\right) = 1 - e^{-\lambda_{n,r}}\left(1 + \lambda_{n,r}\right)$.

vii) For all $\ell \geq 1$ *and* $1 \leq r \leq k_n$, $\mathbb{E}\left(\left(\frac{Z^-_{n,r}}{N^-_{n,r}}\right)^\ell\right) \leq \ell!$.

viii) $\max\limits_{1\leq r\leq k_n} \mathbb{V}\left(\frac{Z^-_{n,r}}{N^-_{n,r}}\right) = o(1)$.

ix) For all $\ell \geq 1$ *and* $1 \leq r \leq k_n$, $\mathbb{E}\left(\left|Z^-_{n,r} - \mathbb{E}\left(Z^-_{n,r}\right)\right|^\ell\right) \leq 1 + \ell!$.

x) For all $\ell \geq 1$, $\max\limits_{1\leq r\leq k_n} \mathbb{E}\left(|\xi_{n,r} - \xi^-_{n,r}|^\ell\right) = O(n\delta_n)$.

xi) $\max\limits_{1\leq r\leq k_n} \left|\mathbb{E}\left(Z^*_{n,r}\right) - \mu_{n,r} + 1\right| = O\left(\max\left(\exp\left(-mnc\nu_n\right); \left(n\delta_n\right)^2\right)\right)$.

Proof. i) is straightforward.

ii) First, note that, since $(\nu \otimes \lambda)(\widetilde{D}_{n,r}) = 0$, then for all $A \in \mathcal{B}\left(\mathbb{R}^+\right)$,

$$N_n\left(D^-_{n,r} \cap \left(I_{n,r} \times A\right)\right) = N_n\left(I_{n,r} \times \left([0, m_{n,r}] \cap A\right)\right) \text{ a.s.}$$

which is a Poisson random variable with mean $nc\nu_{n,r}\lambda([0,m_{n,r}]\cap A)$. Second, set $t\in[0,\lambda_{n,r}]$, and define $t_{n,r}=t/(nc\nu_{n,r})$. Then, for all $q\geq 1$,

$$\begin{aligned}
&\mathbb{P}\left(Z^-_{n,r}\leq t \mid N^-_{n,r}=q\right)\\
=&\frac{\mathbb{P}\left(N_n\left(D^-_{n,r}\cap(I_{n,r}\times[0,t_{n,r}])\right)=q\right)\ \mathbb{P}\left(N_n\left(D^-_{n,r}\cap(I_{n,r}\times(t_{n,r},+\infty))\right)=0\right)}{\mathbb{P}\left(N_n\left(D^-_{n,r}\right)=q\right)}\\
=&\left(\frac{t}{\lambda_{n,r}}\right)^q. \qquad (5)
\end{aligned}$$

Noticing that (5) is obvious when $q=0$ gives the result.

iii)–v) are deduced from ii) by easy calculations.

vi) It follows from ii) that

$$\begin{aligned}
\mathbb{E}\left(\frac{Z^-_{n,r}}{N^-_{n,r}}\right)&=\mathbb{E}\left(\frac{\mathbf{1}_{\left\{N^-_{n,r}>0\right\}}}{N^-_{n,r}}\mathbb{E}\left(Z^-_{n,r}\mid N^-_{n,r}\right)\right)=\lambda_{n,r}\mathbb{E}\left(\frac{\mathbf{1}_{\left\{N^-_{n,r}>0\right\}}}{N^-_{n,r}+1}\right)=\sum_{q=1}^{\infty}\frac{\lambda_{n,r}^{q+1}}{(q+1)!}e^{-\lambda_{n,r}}\\
&=1-e^{-\lambda_{n,r}}\left(1+\lambda_{n,r}\right).
\end{aligned}$$

vii) We have,

$$\begin{aligned}
\mathbb{E}\left(\left(\frac{Z^-_{n,r}}{N^-_{n,r}}\right)^{\ell}\right)&=\mathbb{E}\left(\left(\frac{1}{N^-_{n,r}}\right)^{\ell}\mathbb{E}\left(\left(Z^-_{n,r}\right)^{\ell}\mid N^-_{n,r}\right)\right)=\lambda_{n,r}^{\ell}\mathbb{E}\left(\frac{\mathbf{1}_{\left\{N^-_{n,r}>0\right\}}}{\left(N^-_{n,r}\right)^{\ell-1}\left(\ell+N^-_{n,r}\right)}\right)\\
&=\ell!\sum_{q=1}^{\infty}\frac{(q+\ell-1)!}{\ell!q^{\ell-1}q!}\frac{\lambda_{n,r}^{q+\ell}}{(q+\ell)!}e^{-\lambda_{n,r}} \qquad (6)\\
&\leq\ell!,
\end{aligned}$$

since for all ℓ and q, $(q+\ell-1)!\leq\ell!q^{\ell-1}q!$.

viii) By (6) with $\ell=2$,

$$\begin{aligned}
\mathbb{E}\left(\left(\frac{Z^-_{n,r}}{N^-_{n,r}}\right)^{2}\right)&=\sum_{q=1}^{\infty}\left(1+\frac{1}{q}\right)\frac{\lambda_{n,r}^{q+2}}{(q+2)!}e^{-\lambda_{n,r}}\\
&=1-e^{-\lambda_{n,r}}\left(1+\lambda_{n,r}+\frac{\lambda_{n,r}^2}{2}\right)+\sum_{q=1}^{\infty}\frac{1}{q}\frac{\lambda_{n,r}^{q+2}}{(q+2)!}e^{-\lambda_{n,r}}.
\end{aligned}$$

Now, since,

$$\max_{1\leq r\leq k_n}e^{-\lambda_{n,r}}\left(1+\lambda_{n,r}+\frac{\lambda_{n,r}^2}{2}\right)\leq\max_{\lambda\geq mnc\nu_n}e^{-\lambda}\left(1+\lambda+\frac{\lambda^2}{2}\right)=o\left(1\right),$$

and for all $Q\geq 1$,

$$\begin{aligned}
\max_{1\leq r\leq k_n}\sum_{q=1}^{\infty}\frac{1}{q}\frac{\lambda_{n,r}^{q+2}}{(q+2)!}e^{-\lambda_{n,r}}&\leq\max_{1\leq r\leq k_n}\sum_{q=1}^{Q}\frac{\lambda_{n,r}^{q+2}}{(q+2)!}e^{-\lambda_{n,r}}+\frac{1}{Q}\max_{1\leq r\leq k_n}\sum_{q=1}^{\infty}\frac{\lambda_{n,r}^{q+2}}{(q+2)!}e^{-\lambda_{n,r}}\\
&\leq o\left(1\right)+\frac{1}{Q},
\end{aligned}$$

it follows that

$$\max_{1\leq r\leq k_n}\left|\mathbb{E}\left(\left(\frac{Z_{n,r}^-}{N_{n,r}^-}\right)^2\right)-1\right|=o(1). \tag{7}$$

Collecting (7) and vi) gives the result.

ix) Note that

$$\mathbb{E}\left(\left|Z_{n,r}^- - \mathbb{E}\left(Z_{n,r}^-\right)\right|^\ell\right)=\ell\int_0^{+\infty} t^{\ell-1}\mathbb{P}\left(\left|Z_{n,r}^- - \mathbb{E}\left(Z_{n,r}^-\right)\right|>t\right)dt.$$

Moreover, by iii) and iv),

$$\begin{aligned}\mathbb{P}\left(\left|Z_{n,r}^- - \mathbb{E}\left(Z_{n,r}^-\right)\right|>t\right) &= \mathbb{P}\left(Z_{n,r}^- > \mathbb{E}\left(Z_{n,r}^-\right)+t\right)+\mathbb{P}\left(Z_{n,r}^- < \mathbb{E}\left(Z_{n,r}^-\right)-t\right)\\ &= \left[1-\exp\left(t-\lambda_{n,r}+\mathbb{E}\left(Z_{n,r}^-\right)\right)\right]\mathbf{1}_{\left[0,\lambda_{n,r}-\mathbb{E}\left(Z_{n,r}^-\right)\right]}(t)\\ &+\exp\left(-t-\lambda_{n,r}+\mathbb{E}\left(Z_{n,r}^-\right)\right)\mathbf{1}_{\left[0,\mathbb{E}\left(Z_{n,r}^-\right)\right]}(t)\\ &\leq \mathbf{1}_{[0,1]}(t)+\exp\left(e^{-\lambda_{n,r}}-1\right)e^{-t}.\end{aligned}$$

Hence,

$$\begin{aligned}\mathbb{E}\left(\left|Z_{n,r}^- - \mathbb{E}\left(Z_{n,r}^-\right)\right|^\ell\right) &\leq \ell\int_0^1 t^{\ell-1}dt+\exp\left(e^{-\lambda_{n,r}}-1\right)\ell\int_0^{+\infty}t^{\ell-1}e^{-t}dt\\ &\leq 1+\exp\left(e^{-\lambda_{n,r}}-1\right)\ell!.\end{aligned}$$

x) Since $Z_{n,r}^* = \left(Z_{n,r}^+ - Z_{n,r}^-\right)\mathbf{1}_{\left\{N_{n,r}^+>0\right\}}+Z_{n,r}^-$, we get

$$\begin{aligned}\xi_{n,r}-\xi_{n,r}^- &= \left(1+\frac{1}{N_{n,r}}\right)\left(Z_{n,r}^+ - Z_{n,r}^-\right)\mathbf{1}_{\left\{N_{n,r}^+>0\right\}}+Z_{n,r}^-\left(\frac{1}{N_{n,r}}-\frac{1}{N_{n,r}^-}\right)\\ &= \left(1+\frac{1}{N_{n,r}}\right)\left(Z_{n,r}^+ - \lambda_{n,r}\right)\mathbf{1}_{\left\{N_{n,r}^+>0\right\}}+\left(1+\frac{1}{N_{n,r}}\right)\left(\lambda_{n,r}-\mathbb{E}\left(Z_{n,r}^-\right)\right)\mathbf{1}_{\left\{N_{n,r}^+>0\right\}}\\ &+\left(1+\frac{1}{N_{n,r}}\right)\left(\mathbb{E}\left(Z_{n,r}^-\right)-Z_{n,r}^-\right)\mathbf{1}_{\left\{N_{n,r}^+>0\right\}}+Z_{n,r}^-\left(\frac{1}{N_{n,r}}-\frac{1}{N_{n,r}^-}\right)\\ &:= \gamma_{n,r,1}+\gamma_{n,r,2}+\gamma_{n,r,3}+\gamma_{n,r,4}.\end{aligned}$$

Hence, by the triangle inequality,

$$\max_{1\leq r\leq k_n}\mathbb{E}\left(\left|\xi_{n,r}-\xi_{n,r}^-\right|^\ell\right)^{1/\ell}\leq\sum_{j=1}^4\max_{1\leq r\leq k_n}\mathbb{E}\left(\left|\gamma_{n,r,j}\right|^\ell\right)^{1/\ell}. \tag{8}$$

First,

$$\max_{1\leq r\leq k_n}\mathbb{E}\left(\left|\gamma_{n,r,1}\right|^\ell\right)^{1/\ell}\leq 2nc\max_{1\leq r\leq k_n}\nu_{n,r}\left(M_{n,r}-m_{n,r}\right)\mathbb{P}\left(N_{n,r}^+>0\right)=o\left(n\delta_n\right). \tag{9}$$

Second, by iv):

$$\max_{1\leq r\leq k_n}\mathbb{E}\left(\left|\gamma_{n,r,2}\right|^\ell\right)^{1/\ell}\leq 2\max_{1\leq r\leq k_n}\mathbb{P}\left(N_{n,r}^+>0\right)=O\left(n\delta_n\right). \tag{10}$$

Third, the independence of $N^+_{n,r}$ and $Z^-_{n,r}$ yields with ix):

$$\max_{1\le r\le k_n}\mathbb{E}\left(|\gamma_{n,r,3}|^\ell\right)^{1/\ell}\le 2\left(1+\ell!\right)^{1/\ell}\max_{1\le r\le k_n}\mathbb{P}\left(N^+_{n,r}>0\right)=O\left(n\delta_n\right). \tag{11}$$

Finally, since $|\gamma_{n,r,4}|\le\lambda_{n,r}(N^-_{n,r}+1)^{-1}\mathbf{1}_{\left\{N^+_{n,r}>0\right\}}$ and taking into account that

$$\mathbb{E}\left(\left(N^-_{n,r}+1\right)^{-\ell}\right)=\ell!\lambda^{-\ell}_{n,r}\sum_{q=0}^{\infty}\frac{(q+\ell)!}{(q+1)^\ell\,\ell!q!}\frac{\lambda^{q+\ell}_{n,r}}{(q+\ell)!}e^{-\lambda_{n,r}}\le\ell!\lambda^{-\ell}_{n,r}, \tag{12}$$

(where we used the fact that for all ℓ and q, $(q+\ell)!\le(q+1)^\ell\,\ell!q!$), we get

$$\begin{aligned}\max_{1\le r\le k_n}\mathbb{E}\left(|\gamma_{n,r,4}|^\ell\right)^{1/\ell}&\le\max_{1\le r\le k_n}\lambda_{n,r}\mathbb{E}\left(\left(N^-_{n,r}+1\right)^{-\ell}\right)^{1/\ell}\mathbb{P}\left(N^+_{n,r}>0\right)\\&\le\left(\ell!\right)^{1/\ell}\max_{1\le r\le k_n}\mathbb{P}\left(N^+_{n,r}>0\right)=O\left(n\delta_n\right).\end{aligned} \tag{13}$$

The result is a consequence of $(8)-(13)$.

xi) Note that

$$\begin{aligned}\mathbb{E}\left(Y^*_{n,r}\right)&=\int_0^{M_{n,r}}\mathbb{P}\left(Y^*_{n,r}\ge u\right)du\\&=\int_0^{m_{n,r}}\left(1-\mathbb{P}\left(N_n\left(D_{n,r}\cap\left(I_{n,r}\times[u,M_{n,r}]\right)\right)\right)\right)du+\int_{m_{n,r}}^{M_{n,r}}\mathbb{P}\left(Y^*_{n,r}>u\right)du\\&=\int_0^{m_{n,r}}\left(1-\exp\left(-nc\nu_{n,r}f_{n,r}+nc\nu_{n,r}u\right)\right)du+\int_{m_{n,r}}^{M_{n,r}}\mathbb{P}\left(Y^*_{n,r}>u\right)du\\&=m_{n,r}-\left(nc\nu_{n,r}\right)^{-1}\exp\left(-nc\nu_{n,r}\left(f_{n,r}-m_{n,r}\right)\right)+\left(nc\nu_{n,r}\right)^{-1}\exp\left(-nc\nu_{n,r}f_{n,r}\right)\\&+\int_{m_{n,r}}^{M_{n,r}}\mathbb{P}\left(Y^*_{n,r}>u\right)du.\end{aligned}$$

Hence,

$$\begin{aligned}&\max_{1\le r\le k_n}\left|\mathbb{E}\left(Z^*_{n,r}\right)-\mu_{n,r}+1\right|\\&\le\max_{1\le r\le k_n}\left|nc\nu_{n,r}\left(f_{n,r}-m_{n,r}\right)-1+\exp\left(-nc\nu_{n,r}\left(f_{n,r}-m_{n,r}\right)\right)\right|\\&+\exp\left(-nc\nu_nm\right)+nc\max_{1\le r\le k_n}\nu_{n,r}\mathbb{P}\left(Y^*_{n,r}>m_{n,r}\right)\left(M_{n,r}-m_{n,r}\right)\\&\le\max_{0\le t\le cn\delta_n}\left(e^{-t}+t-1\right)+\exp\left(-nc\nu_nm\right)+O\left(\left(n\delta_n\right)^2\right)\\&=O\left(\max\left(\exp\left(-mnc\nu_n\right);\left(n\delta_n\right)^2\right)\right).\end{aligned}$$

■

In the next lemma we give an uniform upper bound on the centered moments of $(\xi_{n,r})$ and an exact uniform control of the variances and expectations.

Lemma 2 *Under assumptions* (H.1) *and* (H.2) *we have*

i) $\limsup_{n\to\infty} \max_{\ell\geq 2} \frac{1}{6^\ell \ell!} \max_{1\leq r\leq k_n} \mathbb{E}\left(\left|\xi_{n,r} - \mathbb{E}(\xi_{n,r})\right|^\ell\right) < 1.$

ii) $\max_{1\leq r\leq k_n} \left|\mathbb{V}\left(\xi_{n,r}\right) - 1\right| = o\left(1\right).$

iii) $\max_{1\leq r\leq k_n} \left|\mathbb{E}(\xi_{n,r} - \mu_{n,r})\right| = O\left(\max\left(\left(n\delta_n\right)^2, n\nu_n \exp\left(-mcn\nu_n\right), \Delta_n\right)\right).$

Proof. i) It is easy to see that

$$\max_{1\leq r\leq k_n} \mathbb{E}\left(\left|\xi_{n,r} - \mathbb{E}(\xi_{n,r})\right|^\ell\right) \leq 3^\ell \max\left\{ \begin{array}{c} \max_{1\leq r\leq k_n} \mathbb{E}\left(\left|Z^-_{n,r} - \mathbb{E}\left(Z^-_{n,r}\right)\right|^\ell\right), \\ 2^\ell \max_{1\leq r\leq k_n} \mathbb{E}\left(\left|\frac{Z^-_{n,r}}{N^-_{n,r}}\right|^\ell\right),\ 2^\ell \max_{1\leq r\leq k_n} \mathbb{E}\left(\left|\xi_{n,r} - \xi^-_{n,r}\right|^\ell\right) \end{array} \right\}$$

and the result follows from Lemma 1 vii), ix) and x).

ii) Introduce for the sake of simplicity $\gamma_{n,r} = \xi_{n,r} - Z^-_{n,r} = (\xi_{n,r} - \xi^-_{n,r}) + (Z^-_{n,r}/N^-_{n,r})$. This yields

$$\mathbb{V}\left(\xi_{n,r}\right) = \mathbb{V}\left(Z^-_{n,r}\right) + 2\mathbb{C}ov\left(Z^-_{n,r}, \gamma_{n,r}\right) + \mathbb{V}\left(\gamma_{n,r}\right)$$

and we thus have,

$$\max_{1\leq r\leq k_n} \left|\mathbb{V}\left(\xi_{n,r}\right) - 1\right| \leq \max_{1\leq r\leq k_n} \left|\mathbb{V}\left(Z^-_{n,r}\right) - 1\right| + \max_{1\leq r\leq k_n} \mathbb{V}\left(\gamma_{n,r}\right) + 2\left[\max_{1\leq r\leq k_n} \mathbb{V}\left(\gamma_{n,r}\right)\mathbb{V}\left(Z^-_{n,r}\right)\right]^{1/2}.$$

Lemma 1 v) shows that

$$\max_{1\leq r\leq k_n} \left|\mathbb{V}\left(Z^-_{n,r}\right) - 1\right| = \max_{1\leq r\leq k_n} \left|2\lambda_{n,r}e^{-\lambda_{n,r}} + e^{-2\lambda_{n,r}}\right| = o\left(1\right).$$

Besides,

$$\max_{1\leq r\leq k_n} \mathbb{V}\left(\gamma_{n,r}\right) \leq 2 \max_{1\leq r\leq k_n} \mathbb{V}\left(\frac{Z^-_{n,r}}{N^-_{n,r}}\right) + 2 \max_{1\leq r\leq k_n} \mathbb{E}\left(\left|\xi_{n,r} - \xi^-_{n,r}\right|^2\right) = o(1)$$

by Lemma 1 viii) and x) in the particular case where $\ell = 2$.

iii) First, by the triangle inequality and Lemma 1 (vi),

$$\begin{aligned} \max_{1\leq r\leq k_n} \left|\mathbb{E}\left(\frac{Z^*_{n,r}}{N_{n,r}}\right) - 1\right| &\leq \max_{1\leq r\leq k_n} \left|\mathbb{E}\left(\left(\frac{Z^+_{n,r}}{N_{n,r}} - 1\right)\mathbf{1}_{\{N^+_{n,r}>0\}}\right)\right| \\ &\quad + \max_{1\leq r\leq k_n} \left|\mathbb{E}\left(\left(\frac{Z^-_{n,r}}{N^-_{n,r}} - 1\right)\mathbf{1}_{\{N^+_{n,r}=0\}}\right)\right| \\ &\leq \max_{1\leq r\leq k_n} \left|\mathbb{E}\left(\left(\frac{Z^+_{n,r} - \mathbb{E}(N_{n,r})}{N_{n,r}}\right)\mathbf{1}_{\{N^+_{n,r}>0\}}\right)\right| \\ &\quad + \max_{1\leq r\leq k_n} \left|\mathbb{E}\left(\left(\frac{N_{n,r} - \mathbb{E}(N_{n,r})}{N_{n,r}}\right)\mathbf{1}_{\{N^+_{n,r}>0\}}\right)\right| + O\left(n\nu_n \exp\left(-mcn\nu_n\right)\right). \end{aligned} \tag{14}$$

Now, since $\mathbb{E}(N_{n,r}) = \mu_{n,r}$, we get using Lemma 1 (ii) and (12)

$$\begin{aligned}\max_{1\le r\le k_n}\left|\mathbb{E}\left(\left(\frac{Z^+_{n,r}-\mathbb{E}(N_{n,r})}{N_{n,r}}\right)\mathbf{1}_{\{N^+_{n,r}>0\}}\right)\right| &\le nc\delta_n \max_{1\le r\le k_n}\mathbb{E}\left(\left(N^-_{n,r}+1\right)^{-1}\right)\mathbb{P}\left(N^+_{n,r}>0\right)\\ &= o\left((nc\delta_n)^2\right). \end{aligned} \tag{15}$$

Moreover,

$$\begin{aligned}\max_{1\le r\le k_n}\left|\mathbb{E}\left(\left(\frac{N_{n,r}-\mathbb{E}(N_{n,r})}{N_{n,r}}\right)\mathbf{1}_{\{N^+_{n,r}>0\}}\right)\right| &\le \max_{1\le r\le k_n}\left|\mathbb{E}\left(\left(\frac{N^+_{n,r}-\mathbb{E}(N^+_{n,r})}{N_{n,r}}\right)\mathbf{1}_{\{N^+_{n,r}>0\}}\right)\right|\\ &\quad+\max_{1\le r\le k_n}\left|\mathbb{E}\left(\left(\frac{N^-_{n,r}-\mathbb{E}(N^-_{n,r})}{N_{n,r}}\right)\mathbf{1}_{\{N^+_{n,r}>0\}}\right)\right| \end{aligned} \tag{16}$$

and since for all large n, $\max_{1\le r\le k_n}\mathbb{E}(N^+_{n,r}) \le nc\delta_n < 1$, we get eventually, using (12) again,

$$\begin{aligned}\max_{1\le r\le k_n}\left|\mathbb{E}\left(\left(\frac{N^+_{n,r}-\mathbb{E}(N^+_{n,r})}{N_{n,r}}\right)\mathbf{1}_{\{N^+_{n,r}>0\}}\right)\right| &= \max_{1\le r\le k_n}\mathbb{E}\left(\left(\frac{N^+_{n,r}-\mathbb{E}(N^+_{n,r})}{N_{n,r}}\right)\mathbf{1}_{\{N^+_{n,r}>0\}}\right)\\ &\le \max_{1\le r\le k_n}\mathbb{E}\left(\left(\frac{N^+_{n,r}-\mathbb{E}(N^+_{n,r})}{N^-_{n,r}+1}\right)\mathbf{1}_{\{N^+_{n,r}>0\}}\right)\\ &\le \max_{1\le r\le k_n}\mathbb{E}\left(\left(N^-_{n,r}+1\right)^{-1}\right)\mathbb{E}\left(N^+_{n,r}\right)\\ &\le \max_{1\le r\le k_n}\lambda_{n,r}^{-1}nc\nu_{n,r}\left(M_{n,r}-m_{n,r}\right)\\ &= O\left(\Delta_n\right). \end{aligned} \tag{17}$$

Finally, since, for all $r \le k_n$,

$$\begin{aligned}&\left|\mathbb{E}\left(\left(\frac{N^-_{n,r}-\mathbb{E}(N^-_{n,r})}{N_{n,r}}\right)\mathbf{1}_{\{N^+_{n,r}>0\}}\right)\right|\\ &\le\left|\mathbb{E}\left(\frac{N^-_{n,r}-\mathbb{E}(N^-_{n,r})}{N^-_{n,r}+1}\right)\right|\mathbb{P}\left(N^+_{n,r}=1\right)+\sum_{j\ge 2}\mathbb{E}\left(\frac{\left|N^-_{n,r}-\mathbb{E}(N^-_{n,r})\right|}{N^-_{n,r}+j}\right)\mathbb{P}\left(N^+_{n,r}=j\right)\\ &\le\left|\mathbb{E}\left(\frac{N^-_{n,r}-\mathbb{E}(N^-_{n,r})}{N^-_{n,r}+1}\right)\right|\mathbb{P}\left(N^+_{n,r}=1\right)+\mathbb{E}\left(\frac{\left|N^-_{n,r}-\mathbb{E}(N^-_{n,r})\right|}{N^-_{n,r}+1}\right)\mathbb{P}\left(N^+_{n,r}\ge 2\right)\\ &\le\left|\mathbb{E}\left(\frac{N^-_{n,r}-\mathbb{E}(N^-_{n,r})}{N^-_{n,r}+1}\right)\right|\mathbb{P}\left(N^+_{n,r}=1\right)+\left(\mathbb{V}\left(N^-_{n,r}\right)\mathbb{E}\left(\left(N^-_{n,r}+1\right)^{-2}\right)\right)^{1/2}\mathbb{P}\left(N^+_{n,r}\ge 2\right)\\ &\le\left|\mathbb{E}\left(\frac{N^-_{n,r}-\mathbb{E}(N^-_{n,r})}{N^-_{n,r}+1}\right)\right|\mathbb{P}\left(N^+_{n,r}=1\right)+o\left(\mathbb{P}\left(N^+_{n,r}\ge 2\right)\right) \end{aligned} \tag{18}$$

with

$$\mathbb{E}\left(\frac{N_{n,r}^{-}-\mathbb{E}\left(N_{n,r}^{-}\right)}{N_{n,r}^{-}+1}\right) = \sum_{q=0}^{\infty}\left(q-\lambda_{n,r}\right)\frac{\lambda_{n,r}^{q}}{(q+1)!}e^{-\lambda_{n,r}}$$
$$= \lambda_{n,r}^{-1}\sum_{q=0}^{\infty}\left(q+1-1\right)\frac{\lambda_{n,r}^{q+1}}{(q+1)!}e^{-\lambda_{n,r}} - \sum_{q=0}^{\infty}\frac{\lambda_{n,r}^{q+1}}{(q+1)!}e^{-\lambda_{n,r}}$$
$$= \lambda_{n,r}^{-1}\sum_{q=1}^{\infty}\left(q-1\right)\frac{\lambda_{n,r}^{q}}{q!}e^{-\lambda_{n,r}} - \sum_{q=1}^{\infty}\frac{\lambda_{n,r}^{q}}{q!}e^{-\lambda_{n,r}}$$
$$= \lambda_{n,r}^{-1}\left(\lambda_{n,r}-1+e^{-\lambda_{n,r}}\right) - \left(1-e^{-\lambda_{n,r}}\right)$$
$$= -\lambda_{n,r}^{-1}+e^{-\lambda_{n,r}}\left(1+\lambda_{n,r}^{-1}\right), \tag{19}$$

and

$$\mathbb{P}\left(N_{n,r}^{+}\geq 2\right) \leq \sum_{q=2}^{\infty}\frac{\left(nc\delta_n\right)^{q}}{q!}e^{-nc\delta_n} \leq \left(nc\delta_n\right)^{2} \tag{20}$$

we get collecting $(14)-(20)$ that

$$\max_{1\leq r\leq k_n}\left|\mathbb{E}\left(\frac{Z_{n,r}^{*}}{N_{n,r}}\right)-1\right| = O\left(\max\left(\left(n\delta_n\right)^{2},\Delta_n\right)\right).$$

As a conclusion, Lemma 1 (xi) yields:

$$\max_{1\leq r\leq k_n}\left|\mathbb{E}(\xi_{n,r}-\mu_{n,r})\right| \leq \max_{1\leq r\leq k_n}\left|\mathbb{E}\left(Z_{n,r}^{*}\right)-\mu_{n,r}+1\right| + \max_{1\leq r\leq k_n}\left|\mathbb{E}\left(\frac{Z_{n,r}^{*}}{N_{n,r}}\right)-1\right|$$
$$= O\left(\max\left(\left(n\delta_n\right)^{2},n\nu_n\exp\left(-mcn\nu_n\right),\Delta_n\right)\right).$$

■

Proposition 1 *Under assumptions* (H.1) – (H.4) *and for all* $(x_1,...,x_p)\subset F$,

$$\left\{\frac{nc}{\kappa_n\left(x_j\right)}\left(\widehat{f}_n\left(x_j\right)-\mathbb{E}\left(\widehat{f}_n\left(x_j\right)\right)\right) : 1\leq j\leq p\right\} \underset{\mathcal{D}}{\rightarrow} N\left(0,\Sigma_{(x_1,...,x_p)}\right).$$

Proof. The proof is based on Theorem 3 in the Appendix: For all $1\leq r\leq k_n$, set $\zeta_{n,r}=\xi_{n,r}-\mathbb{E}(\xi_{n,r})$ and $w_{n,r}={}^{t}\left(w_{n,r}\left(x_1\right),...,w_{n,r}\left(x_p\right)\right)$. It is easily seen that

$$\left\{\frac{nc}{\kappa_n\left(x_j\right)}\left(\widehat{f}_n\left(x_j\right)-\mathbb{E}\left(\widehat{f}_n\left(x_j\right)\right)\right) : 1\leq j\leq p\right\} = \sum_{r=1}^{k_n} w_{n,r}\zeta_{n,r},$$

where the $\left(\zeta_{n,r}\right)_{1\leq r\leq k_n}$ are independent. Then (H.3), (H.4) and Lemma 2 i), ii) show that the assumptions of Theorem 3 are satisfied. ■

Proposition 2 *Under assumptions* (H.1), (H.2), (H.5), (H.6), *we have for all* $x \in F$,

$$\frac{nc}{\kappa_n(x)}\left(\mathbb{E}\left(\widehat{f}_n(x)\right) - f(x)\right) \to 0 \text{ as } n \to \infty.$$

Proof. For all $x \in F$, we get, by the triangle inequality and assumption (H.5),

$$\begin{aligned}
\frac{nc}{\kappa_n(x)}\left|\mathbb{E}\left(\widehat{f}_n(x)\right) - f(x)\right| &= \left|\sum_{r=1}^{k_n} w_{n,r}(x)\,\mathbb{E}(\xi_{n,r}) - \frac{nc}{\kappa_n(x)} f(x)\right| \\
&\leq \left|\sum_{r=1}^{k_n} w_{n,r}(x)\left(\mathbb{E}(\xi_{n,r}) - \mu_{n,r}\right)\right| + \left|\sum_{r=1}^{k_n} w_{n,r}(x)\,\mu_{n,r} - nc\frac{f(x)}{\kappa_n(x)}\right| \\
&\leq \left(\sum_{r=1}^{k_n} |w_{n,r}(x)|\right) \max_{1\leq r\leq k_n} |\mathbb{E}(\xi_{n,r}) - \mu_{n,r}| + o(1).
\end{aligned}$$

Lemma 2 iii) and condition (H.6) give the result. ■

Theorem 1 is a straightforward consequence of Proposition 1 and Proposition 2. The following lemma shows that $\widehat{c}_n$ converges to c almost surely. In particular, it implies that c can be replaced by $\widehat{c}_n$ in the above theorem.

Lemma 3 *Under assumptions* (H.1) *and* (H.2), *for all* $\delta > 0$, *there exist* $\alpha_\delta > 0$ *and* $n_\delta > 0$ *such that* $\forall n \geq n_\delta$, $\mathbb{P}(|\widehat{c}_n - c| \geq \delta) \leq 3\exp(-n\alpha_\delta)$.

Proof. We have

$$|\widehat{c}_n - c| \leq \frac{1}{n\widehat{a}_n}|N_n(S) - nac| + ac\left|\frac{1}{\widehat{a}_n} - \frac{1}{a}\right|.$$

Let $\delta > 0$ and $\eta_\delta = \min(a/2,\ a\delta/(4c))$. Then,

$$|\widehat{c}_n - c| \leq |\widehat{c}_n - c|\mathbf{1}_{\{|\widehat{a}_n - a| > \eta_\delta\}} + \left(\frac{2}{na}|N_n(S) - nac| + \frac{2c\eta_\delta}{a}\right)\mathbf{1}_{\{|\widehat{a}_n - a| \leq \eta_\delta\}},$$

and therefore

$$\begin{aligned}
\mathbb{P}(|\widehat{c}_n - c| \geq \delta) &\leq \mathbb{P}\left(\frac{1}{na}|N_n(S) - nac| \geq \frac{1}{2}\left(\delta - \frac{2c\eta_\delta}{a}\right)\right) + \mathbb{P}(|\widehat{a}_n - a| > \eta_\delta) \\
&\leq \mathbb{P}\left(\frac{N_n(S)}{na} \notin \left]c - \frac{\delta}{4},\ c + \frac{\delta}{4}\right[\right) + \mathbb{P}(|\widehat{a}_n - a| > \eta_\delta). \qquad (21)
\end{aligned}$$

Let us consider the first term of (21). Since $N_n(S)$ has a Poisson distribution with mean nac, it can be expanded as $N_n(S) = \sum_{k=1}^{n} \pi_k$, where the π_k are independent Poisson random variables with mean ac. Introducing $\Lambda_\pi(s) = \log \mathbb{E}(e^{s\pi}) = ac(e^s - 1)$ and denoting

$$\Lambda_\pi^*(t) = \sup_{s\in\mathbb{R}}(st - \Lambda_\pi(s)) = \begin{cases} t\log(t/ac) - t + ac & \text{if} \quad t \geq 0 \\ +\infty & \text{if} \quad t < 0, \end{cases}$$

Cramer's theorem (see Dembo & Zeitouni (1991), Theorem 2.2.3) yields

$$\limsup_{n\to\infty}\frac{1}{n}\log\mathbb{P}\left(\frac{N_n(S)}{na}\notin\left]c-\frac{\delta}{4},\ c+\frac{\delta}{4}\right[\right)\leq-\inf\left\{\Lambda_\pi^*(t),\ t\notin\left]c-\frac{\delta}{4},\ c+\frac{\delta}{4}\right[\right\}<0.$$

Consequently, there exists $\alpha'_\delta>0$ such that

$$\forall n\geq 1,\mathbb{P}\left(\frac{N_n(S)}{na}\notin\left]c-\frac{\delta}{4},\ c+\frac{\delta}{4}\right[\right)\leq\exp\left(-n\alpha'_\delta\right). \tag{22}$$

Consider now the second term of (21) and observe that

$$\mathbb{E}(\widehat{a}_n)-a=\frac{1}{nc}\sum_{r=1}^{k_n}(\mathbb{E}(\xi_{n,r})-\mu_{n,r})$$

Lemma 2 iii) implies that

$$|\mathbb{E}(\widehat{a}_n)-a|=\frac{k_n}{n}\left(\max\left(\Delta_n,(n\delta_n)^2,n\nu_n\exp\left(-mnc\nu_n\right)\right)\right)$$

which converges to 0 under (H.1) and (H.2). Therefore, there exists $n_\delta>0$ such that

$$\forall n\geq n_\delta,\ \mathbb{P}(|\widehat{a}_n-a|\geq\eta_\delta)\leq\mathbb{P}(|\widehat{a}_n-\mathbb{E}(\widehat{a}_n)|\geq\eta_\delta/2)=\mathbb{P}\left(\left|\sum_{r=1}^{k_n}((\xi_{n,r}-\mathbb{E}(\xi_{n,r}))\right|\geq\eta_\delta nc/2\right),$$

and in view of Lemma 2 i), applying Bernstein's inequality (see Shorack & Wellner (1986), p. 855) yields that for some constants C_1 and C_2,

$$\forall n\geq n_\delta,\ \mathbb{P}(|\widehat{a}_n-a|\geq\eta_\delta)\leq 2\exp\left(-\frac{\eta_\delta^2c^2n^2}{C_1k_n+C_2\eta_\delta cn}\right)\leq 2\exp\left(-\frac{\eta_\delta^2c^2}{C_1+C_2\eta_\delta c}n\right). \tag{23}$$

Defining $\alpha_\delta=\min(\alpha'_\delta,\eta_\delta^2c^2/(C_1+C_2\eta_\delta c))$ and collecting (21)–(23) give the result. ■

Proof of Corollary 1. It remains to verify that the difference

$$D_n(x)=\frac{n}{\kappa_n(x)}(\widehat{c}_n-c)(\widehat{f}_n(x)-f(x))$$

can be neglected in the central limit theorem. In this aim, let $(x_1,\ldots,x_p)\subset F$. For all $\eta>0$ and $\delta>0$, we have

$$\mathbb{P}\left(\max_{1\leq j\leq p}|D_n(x_j)|\geq\eta\right)\leq\mathbb{P}(|\widehat{c}_n-c|\geq\delta)+\mathbb{P}\left(\max_{1\leq j\leq p}\frac{nc}{\kappa_n(x_j)}|\widehat{f}_n(x_j)-f(x_j)|\geq\eta c/\delta\right),$$

where the first term converges to 0 as $n\to\infty$ in view of Lemma 3. Thus,

$$\begin{aligned}\limsup_{n\to\infty}\mathbb{P}\left(\max_{1\leq j\leq p}|D_n(x_j)|\geq\eta\right)&\leq\limsup_{n\to\infty}\mathbb{P}\left(\max_{1\leq j\leq p}\frac{nc}{\kappa_n(x_j)}|\widehat{f}_n(x_j)-f(x_j)|\geq\eta c/\delta\right)\\&=\mathbb{P}\left(\max_{1\leq j\leq p}|G_j|\geq\eta c/\delta\right),\end{aligned} \tag{24}$$

where $(G_1, \ldots, G_p)$ follows the distribution $N(0, \Sigma_{(x_1,\ldots,x_p)})$ under the conditions of Theorem 1. Letting $\delta \to 0$ in (24) yields

$$\limsup_{n\to\infty} \mathbb{P}\left(\max_{1\leq j\leq p} |D_n(x_j)| \geq \eta\right) = 0,$$

and therefore $\{D_n(x_j),\ 1 \leq j \leq p\} \xrightarrow{\mathbb{P}} 0$. ■

5 Applications

We first introduce a general class of kernel estimators which will be shown to satisfy our main result given in Theorem 1. Then, we focus on the particular cases of Parzen-Rosenblatt and Dirichlet kernels.

5.1 General kernel estimates

Consider an unbiaised version of Geffroy's estimator:

$$\overline{f}_n(x) = \sum_{r=1}^{k_n} \mathbf{1}_{I_{n,r}}(x)\left(1 + N_{n,r}^{-1}\right) Y_{n,r}^*.$$

In order to smooth this estimator, a sequence $K_n : E \times E \to \mathbb{R}$, of general smoothing kernels is introduced. Conditions on this sequence will be imposed later. The general kernel estimate is defined by

$$\begin{aligned}\widehat{f}_n(x) &= \int_E K_n(x,t)\,\overline{f}_n(t)\,\nu(dt) \\ &= \sum_{r=1}^{k_n}\left(\int_{I_{n,r}} K_n(x,t)\,\nu(dt)\right)\left(1 + N_{n,r}^{-1}\right) Y_{n,r}^*. \end{aligned} \tag{25}$$

It appears that (25) is a particular case of (3) with $\kappa_{n,r}(x) = \nu_{n,r}^{-1}\int_{I_{n,r}} K_n(x,t)\,\nu(dt)$. In the case where the calculation of this mean value is computationally expensive, it can be approximated by $K_n(x, x_{n,r})$ for some $x_{n,r} \in I_{n,r}$, leading to the simplified estimate

$$\widetilde{f}_n(x) = \sum_{r=1}^{k_n} \nu_{n,r}\ K_n(x, x_{n,r})\ \left(1 + N_{n,r}^{-1}\right) Y_{n,r}^*, \tag{26}$$

which is still a particular case of (3) with $\kappa_{n,r}(x) = K_n(x, x_{n,r})$.

In order to introduce the assumptions needed on K_n, we set, for all $x \in E$,

$$\Gamma_n(x) = \max_{r \leq k_n} \sup \{K_n(x,t) - K_n(x,s) : (s,t) \in I_{n,r} \times I_{n,r}\}$$

and

$$\Psi_n(x) = \left| \int_E K_n(x,t) f(t) \nu(dt) - f(x) \right|.$$

For the sake of simplicity, assume that, for all $n \geq 1$, the partitions $\{I_{n,r} : 1 \leq r \leq k_n\}$ are such that $\nu_{n,r} = k_n^{-1}$ for all $r \leq k_n$. Finally, for all function $g : E \to \mathbb{R}$, we note

$$\|g\|_1 = \int_E |g(t)| \nu(dt), \quad \|g\|_2 = \left(\int_E g(t)^2 \nu(dt) \right)^{1/2} \text{ and } \|g\|_E = \sup_{t \in E} |g(t)|.$$

In this context, the general assumptions (H.1)–(H.6) can be simplified as:

(H′.1) $k_n \uparrow \infty$ and $n^{-1} k_n \log(n) \to 0$ as $n \to \infty$.

(H.2) $0 < m \leq M < +\infty$ and $n k_n^{-1} \Delta_n \to 0$ as $n \to \infty$.

(K.0) For all $n \geq 1$, $\int_{E \times E} |K_n(x,t)| \nu(dx) \nu(dt) < \infty$.

(K.1) For all $(x_1, x_2) \in F \times F$,

$$\Gamma_n(x_1) \|K_n(x_2, \,.\,)\|_1 = o\left(\|K_n(x_1, \,.\,)\|_2 \|K_n(x_2, \,.\,)\|_2 \right) \text{ as } n \to \infty.$$

(K.2) For all $(x_1, x_2) \in F \times F$,

$$\langle K_n(x_1, \,.\,), K_n(x_2, \,.\,) \rangle_2 \left(\|K_n(x_1, \,.\,)\|_2 \|K_n(x_2, \,.\,)\|_2 \right)^{-1} \to \sigma(x_1, x_2) \text{ as } n \to \infty.$$

(K.3) For all $x \in F$,

$$k_n^{-1/2} \|K_n(x, \,.\,)\|_2^{-1} \|K_n(x, \,.\,)\|_E \to 0 \text{ as } n \to \infty.$$

(K.4) For all $x \in F$,

$$n k_n^{-1/2} \|K_n(x, \,.\,)\|_2^{-1} \max\left(\Psi_n(x); \Delta_n \|K_n(x, \,.\,)\|_1 \right) \to 0 \text{ as } n \to \infty.$$

(K.5) For all $x \in F$,

$$n k_n^{-1/2} \|K_n(x, \,.\,)\|_2^{-1} \left(\sum_{r=1}^{k_n} \int_{I_{n,r}} \left(K_n(x,t) - K_n(x, x_{n,r}) \right) \nu(dt) \right) \to 0 \text{ as } n \to \infty.$$

The results established in Section 3 yield:

Theorem 2 *a) Under* (H′.1) , (H.2) , (K.0) − (K.4) , *and for all* $(x_1, ..., x_p) \subset F$,

$$\left\{nck_n^{-1/2} \|K_n(x_j, \ .\,)\|_2^{-1} \left(\widehat{f}_n(x_j) - f(x_j)\right) : 1 \leq j \leq p\right\} \underset{\mathcal{D}}{\to} N\left(0, \Sigma_{(x_1,...,x_p)}\right). \tag{27}$$

b) If, moreover, (K.5) *holds, then for all* $(x_1, ..., x_p) \subset F$,

$$\left\{nck_n^{-1/2} \|K_n(x_j, \ .\,)\|_2^{-1} \left(\widetilde{f}_n(x_j) - f(x_j)\right) : 1 \leq j \leq p\right\} \underset{\mathcal{D}}{\to} N\left(0, \Sigma_{(x_1,...,x_p)}\right). \tag{28}$$

c) (27) *and* (28) *also hold when* c *is replaced by* $\hat{c}_n$.

Proof. a) For all $x \in F$, we just verify (H.1) − (H.6) for $\kappa_{n,r}(x) := k_n \int_{I_{n,r}} K_n(x,t)\,\nu(dt)$.

(H.1) and (H.2) hold trivialy. Moreover, by (K.1),

$$\begin{aligned}
\sum_{r=1}^{k_n} \kappa_{n,r}(x_1)\,\kappa_{n,r}(x_2) &= k_n^2 \sum_{r=1}^{k_n} \int_{I_{n,r}} \int_{I_{n,r}} K_n(x_1,s) K_n(x_2,t) \nu(ds)\,\nu(dt) \\
&= k_n \langle K_n(x_1, \ .\,), K_n(x_2, \ .\,)\rangle_2 \\
&+ k_n^2 \sum_{r=1}^{k_n} \int_{I_{n,r}} \int_{I_{n,r}} K_n(x_2,t)\left(K_n(x_1,s) - K_n(x_1,t)\right)\nu(dt)\,\nu(ds) \\
&= k_n \langle K_n(x_1, \ .\,), K_n(x_2, \ .\,)\rangle_2 + k_n O\left(\sum_{r=1}^{k_n} \Gamma_n(x_1) \int_{I_{n,r}} |K_n(x_2,t)|\,\nu(dt)\right) \\
&= k_n \left[\langle K_n(x_1, \ .\,), K_n(x_2, \ .\,)\rangle_2 + o\left(\|K_n(x_1, \ .\,)\|_2 \|K_n(x_2, \ .\,)\|_2\right)\right].
\end{aligned}$$

Hence,

$$\kappa_n(x) = k_n^{1/2} \|K_n(x, \ .\,)\|_2 \left(1 + o(1)\right), \tag{29}$$

and (K.2) leads to

$$\sum_{r=1}^{k_n} w_{n,r}(x_1)\, w_{n,r}(x_2) = \sigma(x_1,x_2) + o(1),$$

which is (H.3) . Now, (29) entails for all large n,

$$\begin{aligned}
\max_{1\leq r\leq k_n} |w_{n,r}(x)| &\leq 2k_n^{-1/2} \|K_n(x, \ .\,)\|_2^{-1} \ k_n \max_{1\leq r\leq k_n} \left|\int_{I_{n,r}} K_n(x,t)\nu(dt)\right| \\
&\leq 2k_n^{-1/2} \|K_n(x, \ .\,)\|_2^{-1} \ \|K_n(x, \ .\,)\|_E \\
&\to 0 \text{ as } n \to \infty \qquad \text{by (K.3)},
\end{aligned}$$

i.e. (H.4) holds. In order to show (H.5), note that using (29) again in combination with Fubini Theorem (which holds by (K.0)) and the triangle inequality yield

$$\begin{aligned}
&\frac{n}{\kappa_n(x)}\left|\sum_{r=1}^{k_n}\nu_{n,r}\kappa_{n,r}(x)f_{n,r}-f(x)\right| \\
&=\frac{n}{\kappa_n(x)}\left|\sum_{r=1}^{k_n}k_n\int_{I_{n,r}\times I_{n,r}}K_n(x,t)f(s)\nu(dt)\nu(ds)-f(x)\right| \\
&\leq\frac{n}{\kappa_n(x)}\left|\sum_{r=1}^{k_n}k_n\int_{I_{n,r}\times I_{n,r}}K_n(x,t)(f(s)-f(t))\nu(dt)\nu(ds)\right|+\frac{n}{\kappa_n(x)}\Psi_n(x) \\
&\leq 2nk_n^{-1/2}\|K_n(x,\,.\,)\|_2^{-1}\left(\Delta_n\sum_{r=1}^{k_n}\left|\int_{I_{n,r}}K_n(x,t)\nu(dt)\right|+\Psi_n(x)\right) \\
&\leq 2nk_n^{-1/2}\|K_n(x,\,.\,)\|_2^{-1}\left(\Delta_n\|K_n(x,\,.\,)\|_1+\Psi_n(x)\right) \\
&\to 0 \text{ as } n\to\infty \qquad \text{by (K.4)}.
\end{aligned}$$

Finally, we show that (H.6) holds. Since $\max\left((n\delta_n)^2,\Delta_n\right)=o\left(nk_n^{-1}\Delta_n\right)$, it follows that

$$\begin{aligned}
&\sum_{r=1}^{k_n}|w_{n,r}(x)|\max\left((n\delta_n)^2,\Delta_n\right) \\
&\leq 2k_n^{-1/2}\|K_n(x,\,.\,)\|_2^{-1}\sum_{r=1}^{k_n}k_n\left|\int_{I_{n,r}}K_n(x,t)\nu(dt)\right|o\left(nk_n^{-1}\Delta_n\right) \\
&=o\left(nk_n^{-1/2}\|K_n(x,\,.\,)\|_2^{-1}\|K_n(x,\,.\,)\|_1\Delta_n\right)=o(1) \quad \text{by (K.4)}
\end{aligned}$$

and, since, by (H′.1), $n\exp\left(-mcnk_n^{-1}\right)\to 0$,

$$\begin{aligned}
\sum_{r=1}^{k_n}|w_{n,r}(x)|nk_n^{-1}\exp\left(-mcnk_n^{-1}\right)&\leq n\exp\left(-mcnk_n^{-1}\right)\max_{1\leq r\leq k_n}|w_{n,r}(x)| \\
&=o\left(\max_{1\leq r\leq k_n}|w_{n,r}(x)|\right)=o(1) \quad \text{by (H.4)}.
\end{aligned}$$

b) For all $x\in F$, it is easy to see that

$$\begin{aligned}
&\frac{n}{\kappa_n(x)}\left|\widehat{f}_n(x)-\widetilde{f}_n(x)\right| \\
&\leq 4Mnk_n^{-1/2}\|K_n(x,\,.\,)\|_2^{-1}\left(\sum_{r=1}^{k_n}\int_{I_{n,r}}(K_n(x,t)-K_n(x,x_{n,r}))\nu(dt)\right) \\
&=o(1) \quad \text{by (K.5)},
\end{aligned}$$

which, combined with (a), give the intended result by standard arguments.

c) is straightforward by Corollary 1. ■

Two illustrations of this result are now provided. See Menneteau (2003b) for other applications.

5.2 Parzen kernel estimates

In the following, we take $E = [0,1]^d$ $(d \in \mathbb{N}^*)$, ν is the Lebesgue measure on E and $\{I_{n,r} : 1 \leq r \leq k_n\}$ an adjacent equidistant partition of E such that $I_{n,r} = \prod_{j=1}^{d} J_{n,r,j}$ where the $J_{n,r,j}$ are interval of $[0,1]$ of length $k_n^{-1/d}$, leading to $\nu_{n,r} = k_n^{-1}$ for all $1 \leq r \leq k_n$. Besides, we denote by $x_{n,r}$ the center of the cell $I_{n,r}$, $r = 1, \ldots, k_n$. The multivariate Parzen kernel estimate is then defined by the kernel

$$K_n(x,t) = \frac{1}{h_n^d} K^{PR}\left(\frac{x-t}{h_n}\right),$$

where $K^{PR} : \mathbb{R}^d \to \mathbb{R}^+$ is a 1-Lipschitzian Parzen-Rosenblatt kernel with compact support $\mathcal{K}$, and (h_n) is a sequence of positive real numbers tending to zero. It tunes the smoothing introduced by the kernel. For a review on non-parametric regression, see HÄRDLE (1990). We suppose that f is α-Lipschitzian $(0 < \alpha \leq 1)$, in particular,

$$\Delta_n = O\left(k_n^{-\alpha/d}\right). \tag{30}$$

Corollary 2 *Assume that (i)* $n^{-1}k_n \log(n) \to 0$, *(ii)* $h_n^d k_n \to \infty$ *and (iii)* $nk_n^{-1/2}h_n^{\alpha+d/2} \to 0$ *then for all* $(x_1, ..., x_p) \subset \overset{\circ}{E} = (0,1)^d$,

$$\left\{v_n c\left(\widehat{f}_n(x_j) - f(x_j)\right) : 1 \leq j \leq p\right\} \underset{\mathcal{D}}{\to} N\left(0, \left\|K^{PR}\right\|_2^2 I_p\right), \tag{31}$$

where I_p *is the identity matrix of* $\mathbb{R}^p$ *and* $v_n = nh_n^{d/2}k_n^{-1/2}$.

The choice $h_n = n^{-\frac{1}{\alpha+d}}$ *and* $k_n = n^{\frac{d}{\alpha+d}}u_n^2$ *lead to* $v_n = n^{\frac{\alpha}{\alpha+d}}u_n^{-1}$, *where* $u_n \to \infty$ *arbitrary slowly.*

Proof. (K.0) holds trivialy and assumption (i) gives (H′.1). To show (H.2), note that (30) entails

$$nk_n^{-1}\Delta_n = O\left(nk_n^{-(1+\alpha/d)}\right) \tag{32}$$

and thus, by (ii) and (iii),

$$nk_n^{-(1+\alpha/d)} = \left(nk_n^{-1/2}h_n^{\alpha+d/2}\right)\left(h_n^d k_n\right)^{-\left(\frac{\alpha}{d}+\frac{1}{2}\right)} = o(1).$$

Let us consider now (K.1) – (K.4). To this aim, set $x \in \overset{\circ}{E}$. For large enough n (i.e. such that $\mathcal{K} \subset h_n^{-1}(x - E)$),

$$\|K_n(x, .)\|_1 = \int_{h_n^{-1}(x-E)} K^{PR}(u)du = 1. \tag{33}$$

$$\|K_n(x,\ .\ )\|_2 = h_n^{-\frac{d}{2}} \left(\int_{h_n^{-1}(x-E)} \left(K^{PR}\right)^2 (u) du \right)^{1/2} = h_n^{-\frac{d}{2}} \left\| K^{PR} \right\|_2 . \tag{34}$$

$$\begin{aligned} \Psi_n(x) &= h_n^{-d} \left| \int_E K^{PR}\left(h_n^{-1}(x-t)\right)(f(t)-f(x))\, dt \right| \\ &= \left| \int_{\mathcal{K}} K^{PR}(u)\,(f(x-h_n u)-f(x))\, du \right| \\ &= O(h_n^{\alpha}). \end{aligned} \tag{35}$$

Moreover, since K^{PR} is 1-Lipschitzian,

$$\Gamma_n(x) = O\left(k_n^{-1/d} h_n^{-(d+1)}\right). \tag{36}$$

To check (K.1), take $(x_1, x_2) \in F \times F$, then (33), (34), (36) and (ii) entail

$$\begin{aligned} \Gamma_n(x_1)\,\|K_n(x_2,\ .\ )\|_1 &= O\left(k_n^{-1/d} h_n^{-(d+1)}\right) = \left(k_n h_n^d\right)^{-1/d} O\left(h_n^{-d}\right) \\ &= o\left(\|K_n(x_1,\ .\ )\|_2 \,\|K_n(x_2,\ .\ )\|_2\right), \end{aligned}$$

(K.2) follows from the fact that for $x_1 \neq x_2$, we eventually have,

$$\langle K_n(x_1,\ .\ ), K_n(x_2,\ .\ )\rangle_2 = h_n^{-d} \int_{h_n^{-1}(x_1-E)} K^{PR}(u)\, K^{PR}\left(u + h_n^{-1}(x_1 - x_2)\right) du = 0.$$

For (K.3), note that, for all $x \in F$,

$$k_n^{-1/2} \|K_n(x,\ .\ )\|_2^{-1}\ \|K_n(x,\ .\ )\|_E = k_n^{-1/2} h_n^{d/2} \left\| K^{PR} \right\|_2^{-1} h_n^{-d} \left\| K^{PR} \right\|_E = o(1)$$

with (ii). Finally, for all $x \in F$, (33) – (35), (ii) and (iii) entail

$$\begin{aligned} nk_n^{-1/2} \|K_n(x,\ .\ )\|_2^{-1}\ \max\left(\Psi_n(x);\ \Delta_n \|K_n(x,\ .\ )\|_1\right) &= O\left(nk_n^{-1/2} h_n^{d/2}\left(h_n^{\alpha} + k_n^{-\alpha/d}\right)\right) \\ &= O\left(nk_n^{-1/2} h_n^{\frac{d}{2}+\alpha}\right) = o(1). \end{aligned}$$

The end of the proof is straightforward. ■

From the asymptotical point of view, $\widehat{f}_n$ is better than $\widetilde{f}_n$ and than the estimator based on Parzen kernel proposed, in the unidimensional case, by Girard & Jacob (2001). When f is α-Lipschitzian, the speed of convergence of $\widehat{f}_n$ can be chosen arbitrarily close to the minimax speed $n^{-\frac{\alpha}{\alpha+d}}$ (see Härdle *et al* (1995b)). Let us also note that the regularity of $\widehat{f}_n$ and $\widetilde{f}_n$ is determined by the choice of the Parzen-Rosenblatt kernel.

5.3 Projection estimates: Dirichlet kernels

In the sequel (b_n) is a sequence of integers tending to infinity. Let $(e_j)_{j\in\mathbb{N}}$ be an orthonormal basis of $L^2(E,\nu)$. The expansion of f on this basis truncated to the b_n first terms is noted

$$f_n(x) = \sum_{j=0}^{b_n} a_j e_j(x), \qquad x \in E.$$

Each $a_j = \int_E e_j(t) f(t) \nu(dt)$ is then estimated by

$$\widehat{a}_{j,k_n} = \sum_{r=1}^{k_n} \left(\int_{I_{n,r}} e_j(t) \nu(dt) \right) (1 + N_{n,r}^{-1}) Y_{n,r}^*, \qquad 1 \le j \le b_n,$$

leading to an estimate $\widehat{f}_n(x)$ of $f_n(x)$ via:

$$\widehat{f}_n(x) = \sum_{\ell=0}^{b_n} \widehat{a}_{j,k_n} e_\ell(x) = \sum_{r=1}^{k_n} \left(\int_{I_{n,r}} K_n^D(x,t) \nu(dt) \right) (1 + N_{n,r}^{-1}) Y_{n,r}^* \tag{37}$$

where K_n^D the Dirichlet's kernel associated to the orthonormal basis $(e_j)_{j\in\mathbb{N}}$ defined by

$$K_n^D(x,t) = \sum_{j=0}^{b_n} e_j(x) e_j(t), \qquad (x,t) \in E^2. \tag{38}$$

It appears that (37) is a particular case of (25) with $K_n = K_n^D$. Of course, the sometimes more easy to handle estimates

$$\widetilde{f}_n(x) = \sum_{r=1}^{k_n} \nu_{n,r} K_n^D(x, x_{n,r}) (1 + N_{n,r}^{-1}) Y_{n,r}^*, \tag{39}$$

can also be defined. Below, we focus on the trigonometric basis on $E = [0,1]$, ν is the Lebesgue measure on E, $\{I_{n,r} : 1 \le r \le k_n\}$ is the equidistant partition of E and then $\nu_{n,r} = 1/k_n$ for all $1 \le r \le k_n$. This basis is defined for $x \in [0,1]$ by

$$e_0(x) = 1, \qquad e_{2k-1}(x) = \sqrt{2}\cos(2k\pi x), \qquad e_{2k}(x) = \sqrt{2}\sin(2k\pi x), \quad k \ge 1.$$

It is easily seen in that case that the Dirichlet kernel is

$$K_n^D(x,t) = \frac{\sin((1+b_n)\pi(x-t))}{\sin(\pi(x-t))} \quad \text{for } x \neq t$$

$$= 1 + b_n \quad \text{if } x = t.$$

In the following, we assume that f is C^2. In particular,

$$\Delta_n = O\left(k_n^{-1}\right). \tag{40}$$

Besides, we introduce the boundary conditions $f(0) = f(1)$ and $f'(0) = f'(1)$.

Corollary 3 *Assume that (i)* $n^{-1}k_n \log(n) \to 0$, *(ii)* $n^{-1}k_n^2 \to \infty$, *(iii)* $(b_n \log(b_n))^{-1} k_n \to \infty$ *and (iv)* $nk_n^{-1/2} b_n^{-3/2} \to 0$. *Then, for all* $(x_1, ..., x_p) \subset [0,1]$,

$$\left\{ v_n c \left(\widehat{f}_n(x_j) - f(x_j) \right) : 1 \leq j \leq p \right\} \underset{\mathcal{D}}{\to} N(0, I_p), \tag{41}$$

where $v_n = n(b_n k_n)^{-1/2}$. *The choice* $b_n = n^{\frac{1}{2}}$ *and* $k_n = n^{\frac{1}{2}} \log(n) u_n^2$ *leads to* $v_n = n^{\frac{1}{2}} \log(n)^{-1/2} u_n^{-1}$, *where* $u_n \to \infty$ *arbitrarily slowly.*

Proof. (K.0) holds trivialy. Assumptions (i) and (ii) give (H′.1) and (H.2). The following facts are well known (see e.g. Tolstov (1962))

$$\left\| K_n^D(x, .) \right\|_E = 1 + b_n, \quad \left\| K_n^D(x, .) \right\|_2 = (1 + b_n)^{1/2}, \quad \left\| K_n^D(x, .) \right\|_1 = O(\log(b_n)), \tag{42}$$

$$\langle K_n(x_1, .), K_n(x_2, .) \rangle_2 = K_n(x_1, x_2) = o(b_n) \text{ for } x_1 \neq x_2. \tag{43}$$

Since f is C^2, and taking into account of $f(0) = f(1)$ and $f'(0) = f'(1)$, a double integration by parts yields,

$$\max_{j \geq 0} j^2 \int_0^1 f(t) e_j(t) dt = O(1).$$

Hence,

$$\Psi_n(x) = \left| \sum_{j > b_n} \int_0^1 f(t) e_j(t) dt \, e_j(x) \right| = O\left(\sum_{j > b_n} j^{-2} \right) = O\left(b_n^{-1} \right). \tag{44}$$

Moreover, since $\max_{j \geq 1} j^{-1} \left\| e_j' \right\|_E = O(1)$, the Taylor formula gives

$$\begin{aligned} \Gamma_n(x) &\leq \sum_{j=0}^{b_n} |e_j(x)| \sup \{ e_j(t) - e_j(s) : (s,t) \in I_{n,r} \times I_{n,r} \} \\ &= O\left(k_n^{-1} \sum_{j=0}^{b_n} j \right) = O\left(k_n^{-1} b_n^2 \right). \end{aligned} \tag{45}$$

Finally, (42) – (45) together with (i)-(iv) imply (K.1) – (K.4), the proof being similar to the one of Corollary 2. ■

In this situation, both estimates $\widehat{f}_n$ and $\widetilde{f}_n$ are C^∞. From the asymptotical point of view, $\widehat{f}_n$ is better than $\widetilde{f}_n$ and than the estimator based on projections proposed by Girard & Jacob (2003b). Nevertheless, when f is C^2, the above estimates are suboptimal, since the minimax speed of convergence is $n^{-2/3}$ (see e.g. Hall *et al* (1998)). The use of expansions on wavelet bases should

lead to a better speed of convergence. This is part of our future work.

We refer to Girard & Menneteau (2002) for a brief comparison on simulations of some of the previous estimates. Let us also emphasize that in such finite sample situations, the quality of the estimation strongly depends on the choice of the hyper-parameters. The estimates of type (b) described in introduction and more generally the estimates (3) require the choice of two hyper-parameters: the number of extreme values (k_n) and a smoothing parameter (b_n or h_n). Similarly, the estimates of type (a) usually require to select two hyper-parameters: the rate of decrease of the density towards 0 (noted β in the introduction) and the number of continuous derivatives of f (noted q in the introduction). In our opinion, one of the main problems in both cases is now to define an adaptive method for choosing the hyper-parameters.

6 Appendix

We provide a general theorem about the central limit property of a sequence of random $\mathbb{R}^p$ valued vectors

$$\theta_n = \sum_{r=1}^{k_n} w_{n,r}\zeta_{n,r}, \qquad n \geq 1,$$

where $(w_{n,r})_{1\leq r\leq k_n} \subset \mathbb{R}^p$ and $(\zeta_{n,r})_{1\leq r\leq k_n}$ are random variables such that:

(A.1) $(\zeta_{n,r})_{1\leq r\leq k_n}$ are centered and independent random variables.

(A.2) $\max_{1\leq r\leq k_n} \left|\mathbb{E}\left(\zeta_{n,r}^2\right) - 1\right| \to 0.$

(A.3) There exists a covariance matrix Σ in $\mathbb{R}^p$ such that for all $\lambda \in \mathbb{R}^p$,

$$\sum_{r=1}^{k_n} \langle w_{n,r}, \lambda\rangle_{\mathbb{R}^p}^2 \to {}^t\lambda\Sigma\lambda.$$

(A.4) $\max_{1\leq r\leq k_n} \|w_{n,r}\|_{\mathbb{R}^p} = o(1).$

(A.5) $\limsup_{\alpha\to\infty} \limsup_{n\to\infty} \max_{1\leq r\leq k_n} \mathbb{E}\left(\zeta_{n,r}^2 \mathbf{1}_{\{|\zeta_{n,r}|>\alpha\}}\right) = 0.$

Theorem 3 *Under assumptions* (A.1) – (A.5), $\theta_n \underset{\mathcal{D}}{\to} N(0,\Sigma)$.

Proof. We have to show that, for all $\lambda \in \mathbb{R}^p$,

$$\langle \theta_n, \lambda \rangle_{\mathbb{R}^p} \underset{\mathcal{D}}{\to} N\left(0, {}^t\lambda\Sigma\lambda\right). \tag{46}$$

Now, by Lindeberg Theorem (see e.g : DUDLEY (1989) p. 248), it is easy to see that (46) holds whenever for all $\varepsilon > 0$,

$$\limsup_{n\to\infty} \sum_{r=1}^{k_n} \langle w_{n,r}, \lambda \rangle^2_{\mathbb{R}^p} \mathbb{E}\left(\zeta^2_{n,r} \mathbf{1}_{\left\{|\langle w_{n,r},\lambda\rangle_{\mathbb{R}^p}\zeta_{n,r}|>\varepsilon\right\}}\right) = 0. \tag{47}$$

Fix $\lambda \in \mathbb{R}^p$, $\varepsilon > 0$ and $\alpha > 0$. Using (A.4), we get for all n large enough and all $1 \le r \le k_n$ that $\mathbf{1}_{\left\{|\langle w_{n,r},\lambda\rangle_{\mathbb{R}^p}\zeta_{n,r}|>\varepsilon\right\}} \le \mathbf{1}_{\{|\zeta_{n,r}|>\alpha\}}$. Hence,

$$\begin{aligned}
&\limsup_{n\to\infty} \sum_{r=1}^{k_n} \langle w_{n,r}, \lambda \rangle^2_{\mathbb{R}^p} \mathbb{E}\left(\zeta^2_{n,r} \mathbf{1}_{\left\{|\langle w_{n,r},\lambda\rangle_{\mathbb{R}^p}\zeta_{n,r}|>\varepsilon\right\}}\right) \\
&\le \limsup_{n\to\infty} \left(\sum_{r=1}^{k_n} \langle w_{n,r}, \lambda \rangle^2_{\mathbb{R}^p}\right) \max_{1\le r\le k_n} \mathbb{E}\left(\zeta^2_{n,r}\mathbf{1}_{\{|\zeta_{n,r}|>\alpha\}}\right) \\
&\le^t \lambda\Sigma\lambda \limsup_{n\to\infty} \max_{1\le r\le k_n} \mathbb{E}\left(\zeta^2_{n,r}\mathbf{1}_{\{|\zeta_{n,r}|>\alpha\}}\right).
\end{aligned}$$

and we get the result by (A.5) when $\alpha \uparrow \infty$. ■

References

Baufays, P. and Rasson, J.P. (1985) A new geometric discriminant rule. *Computational Statistics Quarterly*, **2**, 15–30.

Dembo, A. and Zeitouni O. (1991) *Large deviations techniques and applications*, Jones and Bartlett.

Deprins, D., Simar, L. and Tulkens, H. (1984) Measuring Labor Efficiency in Post Offices. in *The Performance of Public Enterprises: Concepts and Measurements* by M. Marchand, P. Pestieau and H. Tulkens, North Holland ed, Amsterdam.

Devroye, L.P. and Wise, G.L. (1980) Detection of abnormal behavior via non parametric estimation of the support. *SIAM J. Applied Math.*, **38**, 448–480.

Dudley, R.M. (1989) *Real Analysis and Probability*, Wadsworth and Brooks / Cole.

Gardes, L. (2002) Estimating the support of a Poisson process via the Faber-Shauder basis and extreme values. *Publications de l'Institut de Statistique de l'Université de Paris*, **XXXXVI**, 43–72.

Geffroy, J. (1964) Sur un problème d'estimation géométrique. *Publications de l'Institut de Statistique de l'Université de Paris*, **XIII**, 191–200.

Gijbels, I. and Peng, L. (1999) Estimation of a support curve via order statistics. *Discussion Paper* **9905**, Institut de Statistique, Université Catholique de Louvain.

Gijbels, I., Mammen, E., Park, B.U. and Simar, L. (1999) On estimation of monotone and concave frontier functions. *Journal of the American Statistical Association*, **94**, 220–228.

Girard, S. and Jacob, P. (2003a) Extreme values and Haar series estimates of point processes boundaries. *Scandinavian Journal of Statistics*, **30**(2), 369–384.

Girard, S. and Jacob, P. (2003b) Projection estimates of point processes boundaries. *Journal of Statistical Planning and Inference*, **116**(1), 1–15.

Girard, S. and Jacob, P. (2001) Extreme values and kernel estimates of point processes boundaries. *Technical report ENSAM-INRA-UM2*, **01-02**.

Girard, S. and Menneteau, L. (2002) Limit theorems for extreme values estimates of point processes boundaries. *Technical report INRIA*, **RR-4366**.

Hall, P., Nussbaum, M. and Stern, S. E. (1997) On the estimation of a support curve of indeterminate sharpness. *J. Multivariate Anal.*, **62**, 204–232.

Hall, P., Park, B. U. and Stern, S. E. (1998) On polynomial estimators of frontiers and boundaries. *J. Multivariate Anal.*, **66**, 71–98.

Härdle, W. (1990) *Applied nonparametric regression*, Cambridge University Press, Cambridge.

Härdle, W., Hall, P. and Simar, L. (1995a) Iterated bootstrap with application to frontier models. *J. Productivity Anal.*, **6**, 63–76.

Härdle, W., Park, B. U. and Tsybakov, A. B. (1995b) Estimation of a non sharp support boundaries. *J. Multivariate Anal.*, **43**, 205–218.

Hardy, A. and Rasson, J. P. (1982) Une nouvelle approche des problèmes de classification automatique. *Statistique et Analyse des données*, **7**, 41–56.

Hartigan, J. A. (1975) *Clustering Algorithms.* Wiley, Chichester.

Jacob, P. and Suquet, C. (1995) Estimating the edge of a Poisson process by orthogonal series. *Journal of Statistical Planning and Inference*, **46**, 215–234.

Korostelev, A. P., Simar, L. and Tsybakov, A. B. (1995) Efficient estimation of monotone boundaries. *The Annals of Statistics*, **23**, 476–489.

Korostelev, A. P. and Tsybakov, A.B. (1993) Minimax theory of image reconstruction. in *Lecture Notes in Statistics*, **82**, Springer-Verlag, New York.

Mammen, E. and Tsybakov, A. B. (1995) Asymptotical minimax recovery of sets with smooth boundaries. *The Annals of Statistics*, **23**, 502–524.

Menneteau, L. (2003a) Multidimentional limit theorems for smoothed extreme value estimates of empirical point processes boundaries, submitted.

Menneteau, L. (2003b) Some limit theorems for piecewise constant kernel smoothed estimates of Poisson point process boundaries, submitted.

Shorack, G. R. and Wellner, J.A. (1986) *Empirical processes with applications to statistics*, Wiley, New York.

Tolstov, G.P. (1962) *Fourier Series*, Dover Publications, New York.

Acknowledgement

The authors are indebted to the associate editor and anonymous referees for their helpful comments and suggestions. They have contributed to a greatly improved presentation of the results of this paper.