Extreme values and kernel estimates of point processes boundaries

Stéphane Girard[(1)] & Pierre Jacob[(2)]

[(1)] SMS/LMC, Université Grenoble 1, BP 53, 38041 Grenoble cedex 9, France.

`Stephane.Girard@imag.fr`

[(2)] EPS/I3M, Université Montpellier 2, place Eugène Bataillon,

34095 Montpellier cedex 5, France.

`jacob@math.univ-montp2.fr`

Abstract

We present a method for estimating the edge of a two-dimensional bounded set, given a finite random set of points drawn from the interior. The estimator is based both on a Parzen-Rosenblatt kernel and extreme values of point processes. We give conditions for various kinds of convergence and asymptotic normality. We propose a method of reducing the negative bias and edge effects, illustrated by some simulations.

Keywords: Kernel estimates, Extreme values, Poisson process, Shape estimation.

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

1 Introduction

We address the problem of estimating a bounded set S of $\mathbb{R}^2$ given a finite random set N of points drawn from the interior. This kind of problem arises in various frameworks such as classification [15], image processing [19] or econometrics problems [5]. A lot of different solutions were proposed since [7] and [22] depending on the properties of the observed random set N and of the unknown set S. In this paper, we focus on the special case where $S = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq 1 \; ; \; 0 \leq y \leq f(x)\}$, with f an unknown function. Thus, the estimation of the subset S reduces to the estimation of the function f. This problem arises for instance in econometrics where the function f is called the production frontier. This is the case in [14] where data consist of pairs (X_i, Y_i), X_i representing the input (labor, energy or capital) used to produce an output Y_i in a given firm i. 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.

Most papers on support estimation use to consider the random set of point N appearing under the frontier f as a n-sample. However, in practice, the *number* as well as the position of the points is random, so we do prefer for a long time to deal with point processes. Cox processes are known to provide a high level of generality among the point processes on a plane. However, after conditioning the intensity, the realization of a Cox process is merely the one of a Poisson point process, so what is really observed is a Poisson point process. Moreover in most applications such as medical imaging f delimits a frontier between two zones. A contrasting substance is spread on the whole domain, for instance the brain. The magnetic resonance imaging only displays the bleeding. So the healthy part acts as a mask. Inversely, but similarly, when investigating the retina, the patient does not detect the small luminous spots pointed on a destroyed area. In such cases, there is no way to consider the remaining observed points as a random sample. In fact, such truncated empirical point processes are no longer random samples but binomial point processes (see [21]). In fact, even the nature is unable to obtain a random sample on S in this way! It turns out that binomial point processes are well approximated by Poisson processes. Moreover, truncated Poisson point processes are still Poisson point processes and the same is true for general Cox processes. Naturally, as our point of view is not prevailing, we have to preserve the possibility of comparing our results with those of authors dealing with random samples. So in place of a uniform n-empirical process on S with the distribution $\lambda/\lambda(S)$, we consider a Poisson point process with the intensity $n\lambda/\lambda(S)$, where λ denotes the Lebesgue measure. The intensities of the two processes are obviously equal. Finally, we claim that we are able to deduce for samples similar results by means of Poisson approximations. But it not so simple to achieve, and we prefer to defer this work to a further paper.

In the wide range of nonparametric functional estimators [2], piecewise polynomials have been especially studied [10, 18, 19] and their asymptotic optimality is established under different regularity assumptions on f. See [13, 16, 20] for other cases. Estimators of f based upon orthogonal series appear in [1, 17]. In the case of Haar and C^1 bases, extreme values estimates are defined and studied in [8, 9, 12] and reveal better properties than those of [17]. In the same spirit, a Faber-Shauder estimate is proposed in [6]. Estimating f can also been considered as a regression problem $Y_i = f(X_i) + \varepsilon_i$ with negative noise ε_i. In this context, local polynomial estimates are introduced [11].

Here a kernel method is proposed in order to obtain estimates $\hat{f}_n$ which are at least continuous. We point out a term of systematic bias which permits a real improvement of the estimate at finite distance. We give limit laws with explicit speed of convergence σ_n for $\sigma_n^{-1}(\hat{f}_n - \mathrm{E}\,\hat{f}_n)$ and even for $\sigma_n^{-1}(\hat{f}_n - f)$ after reducing the bias. Section 2 is devoted to the definition of the estimator and basic properties of extreme values. Section 3 deals with *ad hoc* adaptation of Bochner approximation results. In Section 4, we give the main results of convergence: mean square uniform convergence and almost complete uniform convergence. We prove, in Section 5, the asymptotic normality of the estimator, when centered to its mathematical expectation. Section 6 is devoted to some bias reductions, allowing in certain cases asymptotic normality for an estimator, when centered to the function f. We also present a technique for avoiding edge effects. In Section 7 a simulation gives an idea of the improvements carried off by these modifications. Section 8 is dedicated to comparison of kernel estimates with the other propositions found in the literature.

2 Definition and basic properties

For all $n > 0$, let N be a Poisson point process with mean measure $nc\lambda$, where λ denotes the Lebesgue measure on a subset S of $\mathbb{R}^2$ defined as follows:

$$S = \{(x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq 1 \; ; \; 0 \leq y \leq f(x)\}. \tag{2.1}$$

The normalization parameter c is defined by $c = 1/\lambda(S)$ such that $\mathrm{E}(N(S)) = n$. We assume that on $[0,1]$, f is a bounded measurable function, strictly positive and α-Lipschitz, $0 < \alpha \leq 1$, with Lipschitz multiplicative constant L_f and that f vanishes elsewhere. We denote by m (and M) the lower (and the upper) bound of f on $[0,1]$. Given (h_n) a sequence of positive real numbers such that $h_n \to 0$ when $n \to \infty$, the function f is approximated by the convolution:

$$g_n(x) = \int_{\mathbb{R}} K_n(x-y) f(y) dy, \quad x \in [0,1], \tag{2.2}$$

where K_n is given by

$$K_n(t) = \frac{1}{h_n} K\left(\frac{t}{h_n}\right), \quad t \in \mathbb{R},$$

and K is a bounded positive Parzen-Rosenblatt kernel *i.e.* verifying:

$$\forall x \in \mathbb{R},\ 0 \leq K(x) \leq \sup_{\mathbb{R}} K < +\infty, \quad \int_{\mathbb{R}} K(t)dt = 1, \quad \lim_{|x| \to \infty} xK(x) = 0.$$

Note that K^2 and K^3 are Lebesgue-integrable. In the sequel, we introduce extra hypothesis on K when necessary. Consider (k_n) a sequence of integers increasing to infinity and divide S into

k_n cells $D_{n,r}$ with:

$$D_{n,r} = \{ \ (x,y) \in S \mid x \in I_{n,r} \ \}, \qquad I_{n,r} = \left[\frac{r-1}{k_n}, \frac{r}{k_n} \right[, \quad r = 1, \ldots, k_n.$$

The convolution (2.2) is discretized on the $\{I_{n,r}\}$ subdivision of $[0,1]$:

$$f_n(x) = \frac{1}{k_n} \sum_{r=1}^{k_n} K_n(x - x_r) f(x_r), \quad x \in [0,1],$$

where x_r is the center of $I_{n,r}$. The values $f(x_r)$ of the function on the subdivision are estimated through $X^\star_{n,r}$ the supremum of the second coordinate of the points of the truncated process $N(. \cap D_{n,r})$. The considered estimator can be written as:

$$\hat{f}_n(x) = \frac{1}{k_n} \sum_{r=1}^{k_n} K_n(x - x_r) X^\star_{n,r}, \quad x \in [0,1]. \tag{2.3}$$

Formally, this estimator is very similar to the estimators based on expansion of f on L^2 bases [8, 9] although it is obtained by a different principle. Besides, combining the uniform kernel $K(t) = \mathbf{1}_{[-1/2,1/2]}(t)$ with the bandwidth $h_n = 1/k_n$ yields Geffroy's estimate:

$$\hat{f}^G_n(x) = \sum_{r=1}^{k_n} \mathbf{1}_{I_{n,r}}(x) X^\star_{n,r}, \quad x \in [0,1], \tag{2.4}$$

which is piecewise constant on the $\{I_{n,r}\}$ subdivision of $[0,1]$. At the opposite, here we focus on smooth estimators obtained by considering smooth kernels in (2.3). More precisely, we examine systematically the convergence properties of the estimator in two main situations:

(A) K is β-Lipschitz on $\mathbb{R}$, $0 < \beta \leq 1$, with Lipschitz multiplicative constant L_K, $x \to x^2 K(x)$ is integrable, $k_n = o(n)$, $h_n k_n^\alpha \to \infty$, and $h_n^{1+\beta} k_n^\beta \to \infty$ when $n \to \infty$.

(B) K has a compact support, a bounded first derivative and is piecewise C^2, $k_n = o(n)$ and $h_n k_n \to \infty$ when $n \to \infty$.

Of course, Geffroy's estimate does not fulfil these conditions. Some stochastic convergences will require extra conditions on the (k_n) sequence:

(C) $k_n = o\left(n/\ln n\right)$ and $n = o\left(k_n^{1+\alpha}\right)$.

Throughout this paper, we write:

$$\lambda(D_{n,r}) = \lambda_{n,r}, \ \min_{x \in I_{n,r}} f(x) = m_{n,r}, \ \max_{x \in I_{n,r}} f(x) = M_{n,r}.$$

The cumulative distribution function of $X^\star_{n,r}$ is easily calculated on $[0, m_{n,r}]$, after noticing that, for every measurable $B \subset S$, $P(N(B) = 0) = \exp\left(-nc\lambda(B)\right)$:

$$F_{n,r}(x) = P(X^\star_{n,r} \leq x) = \exp\left(\frac{nc}{k_n}(x - k_n\lambda_{n,r})\right), \ \ x \in [0, m_{n,r}]. \tag{2.5}$$

Of course, $F_{n,r}(x) = 0$ if $x < 0$ and $F_{n,r}(x) = 1$ if $x > M_{n,r}$. For $x \in [m_{n,r}, M_{n,r}]$, $F_{n,r}(x)$ is unknown, but $1 - F_{n,r}(m_{n,r})$ can be controlled through regularity conditions made on f. Finally, (2.5) and this control provide precise expansions for the first moments of $X^\star_{n,r}$. We quote that useful results in the following lemma.

Lemma 1 *Assume* **(C)** *is verified. Then,*

(i) $\max_r \left| E(X^\star_{n,r}) - k_n\lambda_{n,r} + \dfrac{k_n}{nc} \right| = O\left(\dfrac{n}{k_n^{1+2\alpha}}\right)$,

(ii) $\max_r \left| Var(X^\star_{n,r}) - \dfrac{k_n^2}{n^2c^2} \right| = O\left(\dfrac{1}{k_n^{2\alpha}}\right)$,

(iii) $\max_r E\left(\left|X^\star_{n,r} - E(X^\star_{n,r})\right|^3\right) = O\left(\dfrac{k_n^3}{n^3}\right)$.

We shall also need a lemma on the Parzen-Rosenblatt kernel.

Lemma 2 *Let $a \neq 0$. For any probability sequence (P_n), we have*

$$\int K(u)K\left(u + \frac{a}{h_n}\right) P_n(du) = o\left(h_n\right). \tag{2.6}$$

Proof of all lemmas are postponed to the Appendix.

Corollary 1 *For all $x \neq y$,*

$$\frac{1}{h_n k_n} \sum_{r=1}^{k_n} K\left(\frac{x - x_r}{h_n}\right) K\left(\frac{y - x_r}{h_n}\right) = o(1).$$

This result is deduced from Lemma 2 with $a = y - x$ and $P_n = \frac{1}{k_n} \sum_{r=1}^{k_n} \delta_{\frac{x-x_r}{h_n}}$.

3 Bias convergence

We first give conditions on the sequences (h_n) and (k_n) to obtain the local uniform convergence of f_n to f, that is, the uniform convergence on every compact subset C of $]0,1[$. Of course, since f is not continuous at 0 and 1, we cannot obtain uniform convergence on the whole compact $[0,1]$. We note in the sequel:

$$\|g\|_\infty^C = \sup_{x \in C} |g(x)|,$$

for all function $g : [0,1] \to \mathbb{R}$. The triangular inequality

$$\|f - f_n\|_\infty^C \leq \|f_n - g_n\|_\infty + \|g_n - f\|_\infty^C,$$

shows the two contributions to the bias. The first term, studied in Lemma 3, is a consequence of the discretization of (2.2). The second term is studied in Lemma 4. It appears in various other kernel estimates such as regression or density estimates.

Lemma 3

(i) *Under* **(A)**, $\|f_n - g_n\|_\infty = O\left(\frac{1}{h_n k_n^\alpha}\right) + O\left(\frac{1}{h_n^{1+\beta} k_n^\beta}\right)$.

(ii) *Under* **(B)**, $\|f_n - g_n\|_\infty = O\left(\frac{1}{h_n^2 k_n^2}\right) + O\left(\frac{1}{k_n^\alpha}\right)$.

The function f is uniformly continuous on $[0,1]$ as soon as it is continuous on the same compact interval and the Bochner Lemma entails that $\|g_n - f\|_\infty^C \to 0$ as $n \to \infty$. The following lemma precises this result by providing the rates of the convergence of $\|g_n - f\|_\infty^C$ in different situations.

Lemma 4

(i) *If* $x \to x^2 K(x)$ *is integrable, then* $\|g_n - f\|_\infty^C = O\left(h_n^{\frac{2\alpha}{\alpha+2}}\right)$.

(ii) *If* K *has a compact support then* $\|g_n - f\|_\infty^C = O\left(h_n^\alpha\right)$.

Let us note that in situation **(B)**, $1/k_n^\alpha = o(h_n^\alpha)$. Thus, as a simple consequence of Lemma 3 and Lemma 4, we get:

Proposition 1

(i) *Under* **(A)**: $\|f_n - f\|_\infty^C = O\left(\frac{1}{h_n k_n^\alpha}\right) + O\left(\frac{1}{h_n^{1+\beta} k_n^\beta}\right) + O\left(h_n^{\frac{2\alpha}{\alpha+2}}\right)$.

(ii) *Under* **(B)**: $\|f_n - f\|_\infty^C = O\left(\frac{1}{h_n^2 k_n^2}\right) + O\left(h_n^\alpha\right)$.

In either case, f_n *converges uniformly locally to* f.

Applying Proposition 1 to the function $\mathbf{1}_{[0,1]}$ leads to the following corollary which will reveal useful in the following.

Corollary 2 *Under the conditions of Proposition 1,*

$$\lim_{n\to\infty} \left\| \frac{1}{k_n} \sum_{r=1}^{k_n} K_n(. - x_r) - 1 \right\|_\infty^C = 0.$$

4 Estimate convergences

This section is devoted to the study of the stochastic convergence of $\hat{f}_n$ to f. We establish sufficient conditions for mean square local uniform convergence and almost complete local uniform convergence.

4.1 Mean square local uniform convergence

In this paragraph, we give sufficient conditions for

$$\sup_{x\in C} \mathrm{E}\left[(\hat{f}_n(x) - f(x))^2\right] \to 0 \text{ as } n \to \infty,$$

where C is compact subset of $]0,1[$. The well-known expansion

$$\mathrm{E}\left[(\hat{f}_n(x) - f(x))^2\right] = \left[\mathrm{E}(\hat{f}_n(x)) - f(x)\right]^2 + \mathrm{Var}(\hat{f}_n(x))$$

allows one to consider the bias term and the variance term separately. The two following lemmas are devoted to the bias which splits in turn as

$$\left\|\mathrm{E}(\hat{f}_n) - f\right\|_\infty^C \leq \left\|\mathrm{E}(\hat{f}_n) - f_n\right\|_\infty^C + \|f_n - f\|_\infty^C.$$

Lemma 5 *Suppose* (**C**) *is verified. Under* (**A**) *or* (**B**)*:* $\left\|E(\hat{f}_n) - f_n\right\|_\infty^C = O\left(k_n/n\right)$.

As a consequence of Lemma 5 and Proposition 1, we obtain the behavior of the bias:

Lemma 6 *Suppose* (**C**) *is verified.*

(i) *Under* (**A**)*:*

$$\left\|E(\hat{f}_n) - f\right\|_\infty^C = O\left(\frac{k_n}{n}\right) + O\left(\frac{1}{h_n k_n^\alpha}\right) + O\left(\frac{1}{h_n^{1+\beta} k_n^\beta}\right) + O\left(h_n^{\frac{2\alpha}{\alpha+2}}\right).$$

(ii) *Under* (**B**)*:*

$$\left\|E(\hat{f}_n) - f\right\|_\infty^C = O\left(\frac{k_n}{n}\right) + O\left(\frac{1}{h_n^2 k_n^2}\right) + O\left(h_n^\alpha\right). \tag{4.1}$$

To conclude, it remains to consider the variance term.

Lemma 7 *Suppose* (**C**) *is verified. Under* (**A**) *or* (**B**)*:*

$$\lim_{n\to\infty} \left\|\frac{Var(\hat{f}_n)}{\sigma_n^2} - \sigma^2\right\|_\infty^C = 0,$$

where $\sigma_n = \frac{k_n^{1/2}}{n h_n^{1/2}}$ *and* $\sigma = \frac{\|K\|_2}{c}$.

In situation **(B)**, $\sigma_n = o(k_n/n)$, and therefore the variance of the estimator is small with respect to the bias. In both situations, as a consequence of Lemma 6 and Lemma 7, we get:

Theorem 1 *Suppose* **(C)** *is verified.*

(i) *Under* **(A)***:*

$$\left\| E(\hat{f}_n - f)^2 \right\|_\infty^C = O\left(\frac{k_n^2}{n^2}\right) + O\left(\frac{1}{h_n^2 k_n^{2\alpha}}\right) + O\left(\frac{1}{h_n^{2+2\beta} k_n^{2\beta}}\right) + O\left(h_n^{\frac{4\alpha}{\alpha+2}}\right) + O\left(\frac{k_n}{n^2 h_n}\right).$$

(ii) *Under* **(B)***:*

$$\left\| E(\hat{f}_n - f)^2 \right\|_\infty^C = O\left(\frac{k_n^2}{n^2}\right) + O\left(\frac{1}{h_n^4 k_n^4}\right) + O\left(h_n^{2\alpha}\right).$$

In either case, the mean square local uniform convergence of $\hat{f}_n$ to f follows.

In situation **(B)**, choosing $k_n = n^{\frac{\alpha+2}{3\alpha+2}}$ and $h_n = n^{-\frac{2}{3\alpha+2}}$ yields

$$\left\| \mathrm{E}(\hat{f}_n - f)^2 \right\|_\infty^C = O\left(n^{-\frac{4\alpha}{3\alpha+2}}\right),$$

and thus, we obtain the following bound for the L_1 norm:

$$\mathrm{E}\left(\| \hat{f}_n - f \|_1\right) = \mathrm{E}\left(\int_0^1 \left|\hat{f}_n(x) - f(x)\right| dx\right) \leq \left(\left\| \mathrm{E}(\hat{f}_n - f)^2 \right\|_\infty^C\right)^{1/2} = O\left(n^{-\frac{\alpha}{1+\frac{3}{2}\alpha}}\right). \quad (4.2)$$

As a comparison, the minimax rate in the n-sample case is $n^{-\frac{\alpha}{1+\alpha}}$ and is reached by Geffroy's estimate. A bias reduction method will be introduced in Section 6 in order to ameliorate the bound (4.2).

4.2 Almost complete local uniform convergence

We shall give sufficient conditions for the convergence of the series

$$\forall \varepsilon > 0, \qquad \sum_{n=1}^{+\infty} P\left(\left\{\left\| \hat{f}_n - f \right\|_\infty^C > \varepsilon\right\}\right) < +\infty.$$

Theorem 2 *Suppose $k_n = o(n/\ln n)$. Under* **(A)** *or* **(B)**, *$\hat{f}_n$ is almost completely locally uniformly convergent to f.*

Proof : Let C be a compact subset of $]0,1[$ and $\varepsilon > 0$. From Proposition 1, f_n converges uniformly to f on C. It remains to consider $\left\|\hat{f}_n - f_n\right\|_\infty^C$. For $x \in C$, we have:

$$\begin{aligned}
\left|\hat{f}_n(x) - f_n(x)\right| &\leq \frac{1}{k_n}\sum_{r=1}^{k_n} K_n(x - x_r) \max_r \left|X_{n,r}^\star - f(x_r)\right| \\
&\leq \left(1 + \left\|\frac{1}{k_n}\sum_{r=1}^{k_n} K_n(. - x_r) - 1\right\|_\infty^C\right) \max_r \left|X_{n,r}^\star - f(x_r)\right| \\
&\leq (1 + o(1)) \max_r \left|X_{n,r}^\star - f(x_r)\right|,
\end{aligned}$$

with Corollary 2. Now, since f is continuous on $[0,1]$, $M_{n,r} - m_{n,r} < \varepsilon/2$ uniformly in r, for n large enough, and therefore

$$\left\{\max_r \left|X_{n,r}^\star - f(x_r)\right| > \varepsilon\right\} \subset \bigcup_r \left\{f(x_r) - X_{n,r}^\star > \varepsilon\right\} \subset \bigcup_r \left\{X_{n,r}^\star < m_{n,r} - \varepsilon/2\right\}.$$

As a consequence,

$$P\left(\left\{\max_r \left|X_{n,r}^\star - f(x_r)\right| > \varepsilon\right\}\right) \leq \sum_{r=1}^{k_n} F_{n,r}(m_{n,r} - \varepsilon/2),$$

where $F_{n,r}$ is given by (2.5). Then, the inequality

$$P\left(\left\{\max_r \left|X_{n,r}^\star - f(x_r)\right| > \varepsilon\right\}\right) \leq k_n \exp\left(-\frac{nc\varepsilon}{2k_n}\right)$$

entails the convergence of the series with $k_n = o(n/\ln n)$. ■

5 Asymptotic distributions

In Theorem 3, we give the limiting distribution of the random variable $\hat{f}_n(x)$ for a fixed $x \in C$, a compact subset of $]0,1[$. In Theorem 4, we study the asymptotic distribution of the random vector obtained by evaluating $\hat{f}_n$ in several distinct points of C.

Theorem 3 *Suppose* **(C)** *is verified. Under* **(A)** *or* **(B)**, $s_n(x) = \sigma_n^{-1}(\hat{f}_n(x) - E(\hat{f}_n(x)))$ *converges in distribution to a centered Gaussian variable with variance* σ^2, *for all* $x \in C$.

Proof : Let $x \in C$ be fixed. Introducing the k_n independent random variables

$$Y_{n,r}(x) = \frac{n}{k_n^{3/2} h_n^{1/2}} K\left(\frac{x - x_r}{h_n}\right)(X_{n,r}^\star - \mathrm{E}\, X_{n,r}^\star),$$

the quantity $s_n(x)$ can be rewritten as

$$s_n(x) = \sum_{r=1}^{k_n} Y_{n,r}(x).$$

Our goal is to prove that the Lyapounov condition

$$\lim_{n\to\infty} \sum_{r=1}^{k_n} \frac{\mathrm{E}\left(|Y_{n,r}(x)|^3\right)}{\mathrm{Var}^{3/2}(s_n(x))} = 0$$

holds under condition **(A)**, **(C)** or **(B)**, **(C)**. Remark first that

$$\mathrm{Var}(s_n(x)) = \mathrm{Var}(\hat{f}_n(x))/\sigma_n^2 \to \sigma^2,$$

as $n \to \infty$ with Lemma 7. Second, we have

$$\begin{aligned}
\sum_{r=1}^{k_n} \mathrm{E}\left(|Y_{n,r}(x)|^3\right) &\le \frac{n^3}{h_n^{3/2} k_n^{9/2}} \sum_{r=1}^{k_n} K^3\left(\frac{x - x_r}{h_n}\right) \max_r \mathrm{E}\left(\left|X_{n,r}^\star - E(X_{n,r}^\star)\right|^3\right) \\
&\le \left(\frac{1}{k_n h_n} \sum_{r=1}^{k_n} K^3\left(\frac{x - x_r}{h_n}\right)\right) O\left(\frac{1}{k_n^{1/2} h_n^{1/2}}\right),
\end{aligned}$$

with Lemma 1. Then,

$$\begin{aligned}
\frac{1}{h_n k_n} \sum_{r=1}^{k_n} K^3\left(\frac{x - x_r}{h_n}\right) &\le \int_{\mathbb{R}} K^3(u)du + \left\| \frac{1}{h_n k_n} \sum_{r=1}^{k_n} K^3\left(\frac{\cdot - x_r}{h_n}\right) - \int_{\mathbb{R}} K^3(u)du \right\|_\infty^C \\
&\le \int_{\mathbb{R}} K^3(u)du + o(1),
\end{aligned}$$

with Corollary 2 applied to the kernel $K^3/\int K^3(u)du$. As a conclusion,

$$\sum_{r=1}^{k_n} \frac{\mathrm{E}\left(|Y_{n,r}(x)|^3\right)}{\mathrm{Var}^{3/2}(s_n(x))} = O\left(\frac{1}{h_n^{1/2} k_n^{1/2}}\right),$$

and the result follows. ∎

Theorem 4 *Let $(y_1,\dots,y_q)$ be distinct points in C and denote $\mathbb{I}_q$ the identity matrix of size q. Under the conditions of Theorem 3, the random vector $(s_n(y_j),\ j=1,\dots,q)$ converges in distribution to a centered Gaussian vector of $\mathbb{R}^q$ with covariance matrix $\sigma^2\mathbb{I}_q$.*

Proof : Our goal is to prove that, $\forall (u_1,\dots,u_q)\in\mathbb{R}^q$, the random variable

$$\tilde{s}_n = \sum_{i=1}^{q} u_i s_n(y_i)$$

converges in distribution to a centered Gaussian variable with variance $\| u \|_2^2\sigma^2$. A straightforward calculation yields

$$\tilde{s}_n = \sum_{r=1}^{k_n} \tilde{Y}_{n,r},$$

where we have defined

$$\tilde{Y}_{n,r} = \frac{n}{k_n^{3/2}h_n^{1/2}} \sum_{i=1}^{q} u_i K\left(\frac{y_i-x_r}{h_n}\right)(X^\star_{n,r} - \mathrm{E}\, X^\star_{n,r}).$$

We use a chain of arguments similar to the ones in Theorem 3 proof. First, the variance of $\tilde{Y}_{n,r}$ is evaluated with Lemma 1**(ii)**:

$$\begin{aligned}
\mathrm{Var}(\tilde{Y}_{n,r}) &= \frac{n^2}{h_n k_n^3}\left(\sum_{i=1}^{q} u_i K\left(\frac{y_i-x_r}{h_n}\right)\right)^2 \mathrm{Var}(X^\star_{n,r}) \\
&= \frac{1}{c^2}\frac{1}{h_n k_n}\left(\sum_{i=1}^{q} u_i K\left(\frac{y_i-x_r}{h_n}\right)\right)^2 \left(1+O\left(\frac{n^2}{k_n^{2\alpha+2}}\right)\right) \\
&\sim \frac{1}{c^2}\frac{1}{h_n k_n}\left(\sum_{i=1}^{q} u_i K\left(\frac{y_i-x_r}{h_n}\right)\right)^2 .
\end{aligned}$$

Then, the variance of $\tilde{s}_n$ can be expanded as

$$\mathrm{Var}(\tilde{s}_n) \sim \sum_{i=1}^{q}\sum_{r=1}^{k_n} \frac{u_i^2}{c^2}\frac{1}{h_n k_n} K^2\left(\frac{y_i-x_r}{h_n}\right) + \sum_{i\neq j}\frac{u_i u_j}{c^2}\frac{1}{h_n k_n}\sum_{r=1}^{k_n} K\left(\frac{y_i-x_r}{h_n}\right) K\left(\frac{y_j-x_r}{h_n}\right).$$

Corollary 2 provides the limit of the first term and, from Corollary 1, the second term goes to 0 when n goes to infinity. As a partial conclusion, $\mathrm{Var}(\tilde{s}_n)\to \| u \|_2^2\sigma^2$ when $n\to\infty$. Now, we have

$$\mathrm{E}\left(\left|\tilde{Y}_{n,r}\right|^3\right) = \frac{n^3}{h_n^{3/2}k_n^{9/2}}\left|\sum_{i=1}^{q} u_i K\left(\frac{y_i-x_r}{h_n}\right)\right|^3 \mathrm{E}\left(\left|X^\star_{n,r} - E(X^\star_{n,r})\right|^3\right).$$

Then, Lemma 1**(iii)** entails

$$\sum_{r=1}^{k_n} \mathrm{E}\left(\left|\tilde{Y}_{n,r}\right|^3\right) = O\left(\frac{1}{h_n^{3/2}k_n^{3/2}}\right)\sum_{r=1}^{k_n}\left|\sum_{i=1}^{q} u_i K\left(\frac{y_i - x_r}{h_n}\right)\right|^3,$$

and remarking that

$$\left|\sum_{i=1}^{q} u_i K\left(\frac{y_i - x_r}{h_n}\right)\right|^3 \le \|K\|_\infty \| u \|_1 \left(\sum_{i=1}^{q} u_i K\left(\frac{y_i - x_r}{h_n}\right)\right)^2$$

shows finally that

$$\sum_{r=1}^{k_n} \mathrm{E}\left(\left|\tilde{Y}_{n,r}\right|^3\right) = O\left(\frac{1}{h_n^{1/2}k_n^{1/2}}\right),$$

and the conclusion follows. ■

6 Bias reduction

It is worth noticing that, in Section 5, the negative bias of $\hat{f}_n$ is too large to obtain a limit distribution for $(\hat{f}_n - f)$. We introduce a corrected estimator $\tilde{f}_n$ sharp enough to obtain a limiting distribution for $(\hat{f}_n - f)$ under conditions **(B)** and **(C)**.

It is clear, in view of Lemma 1, that $X^\star_{n,r}$ is an estimator of $k_n\lambda_{n,r}$ with a negative bias asymptotically equivalent to $-k_n/(nc)$. To reduce this bias, we introduce the random variable defined by

$$Z_n = \frac{1}{n-k_n}\sum_{r=1}^{k_n} X^\star_{n,r}.$$

Lemma 1 implies that, under **(C)**,

$$\mathrm{E}(Z_n) = \frac{k_n}{nc} + O\left(\frac{1}{k_n^{2\alpha}}\right). \tag{6.1}$$

This suggests to consider the estimator

$$\tilde{f}_n(x) = \frac{1}{k_n}\sum_{r=1}^{k_n} K_n(x - x_r)\left(X^\star_{n,r} + Z_n\right), \quad x \in [0,1].$$

A more precise version of Theorem 3 can be given in situation **(B)** at the expense of additional conditions. To this end, we need a preliminary lemma providing the bias of the new estimator $\tilde{f}_n$.

Lemma 8 *Under* **(B)**, **(C)**:

$$\left\| E(\tilde{f}_n) - f \right\|_\infty^C = O\left(\frac{1}{h_n^2 k_n^2}\right) + O\left(h_n^\alpha\right).$$

The bias of the new estimator $\tilde{f}_n(x)$ is asymptotically lower than the bias of $\hat{f}_n(x)$ since the k_n/n term of (4.1) is cancelled in Lemma 8. Let us also note that the variance of $\tilde{f}_n(x)$ is bounded above by the variance of $\hat{f}_n(x)$: Since

$$\operatorname{Var} \tilde{f}_n(x) \leq 2 \operatorname{Var} \hat{f}_n(x) + 2\left(\frac{1}{k_n}\sum_{r=1}^{k_n} K_n(x - x_r)\right)^2 \operatorname{Var} Z_n, \tag{6.2}$$

it follows from Lemma 1, Lemma 7 and Corollary 2 that

$$\frac{\operatorname{Var} \tilde{f}_n(x)}{\operatorname{Var} \hat{f}_n(x)} \leq 2 + O\left(\frac{\operatorname{Var} Z_n}{\sigma_n^2}\right) = 2 + O\left(\frac{k_n \operatorname{Var} X_{n,1}^\star}{n^2 \sigma_n^2}\right) = 2 + O\left(\frac{h_n k_n^2}{n^2}\right) = 2 + o\left(1\right). \tag{6.3}$$

These remarks allow to give the asymptotic distribution of $(\tilde{f}_n(x) - f(x))$.

Theorem 5 *If* **(B)** *holds,* $n = o\left(k_n^{1/2} h_n^{-1/2-\alpha}\right)$, $n = o\left(k_n^{5/2} h_n^{3/2}\right)$ *and* $k_n = o\left(n / \ln n\right)$*, then* $t_n(x) = \sigma_n^{-1}(\tilde{f}_n(x) - f(x))$ *converges in distribution to a centered Gaussian variable with variance* σ^2*, for all* x *in a compact subset of* $]0,1[$.

Proof : Consider the expansion

$$\begin{aligned} t_n(x) &= \sigma_n^{-1}(\tilde{f}_n(x) - \mathrm{E}(\tilde{f}_n(x))) + \sigma_n^{-1}(\mathrm{E}(\tilde{f}_n(x)) - f(x)) \\ &= \sigma_n^{-1}(\hat{f}_n(x) - \mathrm{E}(\hat{f}_n(x))) + \frac{\sigma_n^{-1}}{k_n}\sum_{r=1}^{k_n} K_n(x - x_r)(Z_n - \mathrm{E}(Z_n)) + \sigma_n^{-1}(\mathrm{E}(\tilde{f}_n(x)) - f(x)). \end{aligned}$$

In view of Theorem 3, the first term converges in distribution to a centered Gaussian variable with variance σ^2. The second term is centered and its variance converges to zero from (6.2) and (6.3). Therefore, the second term converges to 0 in probability. The third term is controlled with Lemma 8:

$$\sigma_n^{-1}\left\|\mathrm{E}(\tilde{f}_n) - f\right\|_\infty^C = O\left(\frac{n}{h_n^{3/2}k_n^{5/2}}\right) + O\left(\frac{nh_n^{1/2+\alpha}}{k_n^{1/2}}\right) \to 0,$$

and the conclusion follows. ■

The uniform mean square distance between $\tilde{f}_n$ and f is derived from Lemma 7 and Lemma 8:

$$\left\|\mathrm{E}(\tilde{f}_n - f)^2\right\|_\infty^C = O\left(\frac{k_n}{n^2 h_n}\right) + O\left(\frac{1}{h_n^4 k_n^4}\right) + O\left(h_n^{2\alpha}\right). \tag{6.4}$$

Possible choices are $k_n = n^{\frac{4+2\alpha}{4+5\alpha}}$ and $h_n = n^{-\frac{4}{4+5\alpha}}$ leading to

$$\left\|\mathrm{E}(\tilde{f}_n - f)^2\right\|_\infty^C = O\left(n^{-\frac{8\alpha}{4+5\alpha}}\right),$$

and thus,

$$\mathrm{E}\left(\| \tilde{f}_n - f \|_1\right) = O\left(n^{-\frac{\alpha}{1+\frac{5}{4}\alpha}}\right), \tag{6.5}$$

which is a significant improvement of (4.2). It is well-known that non-parametric estimators based on Parzen-Rosenblatt kernels suffer from a lack of performance on the boundaries of the estimation interval. To overcome this limitation, symmetrization techniques have been developed [4]. The application of such a method to $\tilde{f}_n(x)$ yields the following estimator:

$$\check{f}_n(x) = \frac{1}{k_n}\sum_{r=1}^{k_n}\left(K_n(x - x_r) + K_n(x + x_r) + K_n(x + x_r - 2)\right)\left(X_{n,r}^\star + Z_n\right), \quad x \in [0,1].$$

The convergence properties of $\hat{f}_n$ and $\tilde{f}_n$ on the compact subsets of $]0,1[$ can be extended to $\check{f}_n$ on the whole interval $[0,1]$ without difficulties.

7 Illustration on simulations

We consider a Poisson process on a set S defined as in (2.1) with

$$f(x) = \left| \begin{array}{ll} 1 + \exp\left(-60(x - 1/4)^2\right) & \text{if } 0 \leq x \leq 1/3, \\ 1 + \exp\left(-5/12\right) & \text{if } 1/3 < u \leq 2/3, \\ -6\exp\left(-5/12\right)x + 1 + 5\exp\left(-5/12\right) & \text{if } 2/3 < u \leq 5/6, \\ 6x - 4 & \text{if } 5/6 < u \leq 1. \end{array} \right. ,$$

with $n = 300$. Note that f is continuous but is not derivable at $x = 1/3$, $x = 2/3$ and $x = 5/6$. We use a subdivision of $[0, 1]$ into $k_n = 30$ intervals leading to about 10 simulated points in each interval. The following kernel is chosen

$$K(t) = \cos^2(\pi t/2)\mathbf{1}_{[-1,1]}(t),$$

with associated window $h_n = 2/k_n$, which corresponds to a smoothing over 4 intervals. The automatic choice of h_n and k_n is under investigation. A possible solution would be to use a plug-in procedure. The idea consists in selecting the sequences minimizing a more precise version of (6.4) including the multiplicative constants.

The experiment involves several steps:

– First, $m = 100$ replications of the Poisson process are simulated.

– For each of the m previous set of points, the three kernel estimators $\hat{f}_n$, $\tilde{f}_n$, $\check{f}_n$ and Geffroy's estimator $\hat{f}_n^G$ are computed.

– For each of four estimators, the m associated L_1 distances to f are evaluated on a grid.

– Finally, for each of the four estimators, the best situation (ie the estimation corresponding to the smallest L_1 error) is represented on Figure 1 and the worst situation (ie the estimation corresponding to the largest L_1 error) is represented on Figure 2.

It appears on Figure 1(a), that even in the best case, $\hat{f}_n$ suffers from a negative bias and boundary problems. Figure 1(b) illustrates the efficiency of the bias correction proposed in Section 6. Boundary effects disappear on Figure 1(c) with the symmetrization procedure described in Section 6. As a comparison, Geffroy's estimate is represented on Figure 1(d). The negative bias appears clearly on the $[1/3, 2/3]$ interval. These conclusions also hold for the worst cases (Figure 2).

8 Comparison with other estimates

Let us emphasize that such comparisons are only relevant within a same framework, which excludes hypotheses such as the convexity or the monotonicity of f. Thus, the competitive methods to our kernel approach are essentially local polynomial estimates [11], piecewise polynomial estimates [18, 19, 10] and our projection estimate [8, 9].

– From the theoretical point of view, piecewise polynomial estimates benefit from the minimax optimality whereas the estimates proposed in this paper are suboptimal. In the class of continuous functions f having a Lipschitzian k-th derivative, the optimal rate of convergence is attained by minimizing, on each cell of a partition of $[0, 1]$, the measure of a domain with a polynomial edge of degree k. For instance, in the case of a α-Lipschitzian frontier, the minimax optimal rate for the L_1 norm is $n^{-\frac{\alpha}{1+\alpha}}$ and the corresponding rate is $n^{-\frac{\alpha}{5/4+\alpha}}$ for $\tilde{f}_n$ (see (6.5)). The difference of speed increases with α, but even if $\alpha = 1$ (which is the worst situation for us), one obtains "similar" rates of convergence, that is $n^{-1/2}$ and $n^{-4/9}$. In this sense, kernel estimates bring a significant improvement to projection estimates.

– From the practical point of view, all the previous estimates require the selection of two hyperparameters. In case of piecewise polynomial and local polynomial estimators, the construction of the estimate requires to select the degree of the polynomial function (which corresponds to k

in the piecewise polynomial framework) and a kind of bandwidth parameter. Nevertheless, our opinion is that kernel estimates are the most pleasant to use in practice for the following reasons. The computation of local and piecewise polynomial estimates requires to solve an optimization problem. For instance, the computation of piecewise polynomial estimates is not straightforward, at least for $k > 0$. When $k = 0$, piecewise polynomial estimates reduce to Geffroy's estimate, whose unsatisfying behavior on finite sample situations has been illustrated in Section 7. At the opposite, kernel and projection estimators enjoy explicit forms and are thus easily implementable. Besides, these methods yield smooth estimates whereas piecewise polynomial estimates are discontinuous whatever the regularity degree of f is. Finally, only kernel and projection estimates benefit from an explicit asymptotic distribution. This property allows to build pointwise confident intervals without costfly Monte-Carlo methods (see [10] for the case of piecewise polynomial estimates). In the local polynomial estimates situation, both the limiting distribution and the normalization sequences are not explicit making difficult the reduction of the asymptotic bias.

9 Appendix : proof of lemmas

Proof of Lemma 1. We give here the complete proof of **(i)** and a sketch of the proofs of **(ii)** and **(iii)** since the methods in use are similar.

(i) The mathematical expectation can be expanded in three terms:

$$\begin{aligned} \mathrm{E}(X^{\star}_{n,r} - k_n\lambda_{n,r}) = -k_n\lambda_{n,r}e^{-nc\lambda_{n,r}} &+ \int_0^{m_{n,r}} (x - k_n\lambda_{n,r})F'_{n,r}(x)dx \\ &+ \int_{m_{n,r}}^{M_{n,r}} (x - k_n\lambda_{n,r})F_{n,r}(dx). \end{aligned}$$

The first term of the sum is asymptotically negligible:

$$\max_r k_n\lambda_{n,r}e^{-nc\lambda_{n,r}} = o\left(n^{-s}\right), \tag{9.1}$$

for all $s > 0$, when $n \to \infty$. Using (2.5), the second term can be rewritten as

$$\begin{aligned}\int_0^{m_{n,r}} (x - k_n\lambda_{n,r}) F'_{n,r}(x) dx &= \frac{nc}{k_n}\int_0^{m_{n,r}} (x - k_n\lambda_{n,r}) \exp\left[\frac{nc}{k_n}(x - k_n\lambda_{n,r})\right] dx \\ &= -\frac{k_n}{nc}\int_{\frac{nc}{k_n}(k_n\lambda_{n,r} - m_{n,r})}^{nc\lambda_{n,r}} u \exp(-u) du.\end{aligned}$$

Let us note $\phi(u) = (u+1)\exp(-u)$ a primitive of $-u\exp(-u)$. We have $\phi(u) = 1 + O(u^2)$ when $u \to 0$ and $\phi(u) = o(u^{-s})$, $\forall s > 0$ when $u \to \infty$. Consequently, remarking that the upper bound goes to infinity, and that the lower bound goes to 0 under the assumption $n = o(k_n^{1+\alpha})$ yields

$$\max_r \left| \int_0^{m_{n,r}} (x - k_n\lambda_{n,r}) F'_{n,r}(x) dx + \frac{k_n}{nc} \right| = O\left(\frac{n}{k_n^{1+2\alpha}}\right). \tag{9.2}$$

The third term is bounded above by

$$\int_{m_{n,r}}^{M_{n,r}} (x - k_n\lambda_{n,r}) F_{n,r}(dx) \le (M_{n,r} - m_{n,r})\left[1 - \exp\left(\frac{nc}{k_n}(m_{n,r} - k_n\lambda_{n,r})\right)\right]$$

and thus

$$\max_r \left| \int_{m_{n,r}}^{M_{n,r}} (x - k_n\lambda_{n,r}) F_{n,r}(dx) \right| = O\left(\frac{n}{k_n^{1+2\alpha}}\right). \tag{9.3}$$

Collecting (9.1), (9.2) and (9.3) proves the result.

(ii) It is convenient to write the variance as

$$\mathrm{Var}(X^\star_{n,r}) = \mathrm{Var}(X^\star_{n,r} - k_n\lambda_{n,r}) = \mathrm{E}\left[(X^\star_{n,r} - k_n\lambda_{n,r})^2\right] - \mathrm{E}^2(X^\star_{n,r} - k_n\lambda_{n,r}).$$

With a method very similar to the one used to prove **(i)**, we obtain uniformly in r,

$$\mathrm{E}\left[(X^\star_{n,r} - k_n\lambda_{n,r})^2\right] = \frac{2k_n^2}{n^2c^2} + O\left(\frac{1}{k_n^{2\alpha}}\right).$$

Besides, **(i)** entails

$$\mathrm{E}^2(X^\star_{n,r} - k_n\lambda_{n,r}) = \frac{k_n^2}{n^2c^2} + O\left(\frac{1}{k_n^{2\alpha}}\right),$$

uniformly in r, and the conclusion follows.

(iii) The proof is similar. It requires the calculation of $\mathrm{E}(|X^\star_{n,r} - k_n\lambda_{n,r}|^3)$ and the use of **(i)** and **(ii)**. ■

Proof of Lemma 2. Let $\varepsilon > 0$ and split (2.6) into

$$\begin{aligned}\int K(u)K\left(u+\frac{a}{h_n}\right)P_n(du) &= \int_{|u|>\frac{|a|}{2h_n}} K(u)K\left(u+\frac{a}{h_n}\right)P_n(du)\\ &+ \int_{|u|\le\frac{|a|}{2h_n}} K(u)K\left(u+\frac{a}{h_n}\right)P_n(du),\end{aligned}$$

and consider the two terms separately.

- The first term is bounded above by

$$\begin{aligned}\int_{|u|>\frac{|a|}{2h_n}} K(u)K\left(u+\frac{a}{h_n}\right)P_n(du) &\le \|K\|_\infty \int_{|u|>\frac{|a|}{2h_n}} K(u)P_n(du) \\ &= \|K\|_\infty \int_{|u|>\frac{|a|}{2h_n}} |u|\,K(u)\frac{1}{|u|}P_n(du).\end{aligned} \tag{9.4}$$

  Since $uK(u)\to 0$ when $|u|\to\infty$, for n large enough $|uK(u)|<\varepsilon$ entailing

$$\int_{|u|>\frac{|a|}{2h_n}} K(u)K\left(u+\frac{a}{h_n}\right)P_n(du) \le \|K\|_\infty\varepsilon \int_{|u|>\frac{|a|}{2h_n}} \frac{1}{|u|}P_n(du) \le \frac{2\varepsilon h_n\|K\|_\infty}{|a|}.$$

  We have proved that $\forall\varepsilon>0$, for n large enough

$$\frac{1}{h_n}\int_{|u|>\frac{|a|}{2h_n}} K(u)K\left(u+\frac{a}{h_n}\right)P_n(du) \le \frac{2\varepsilon\|K\|_\infty}{|a|},$$

  or equivalently,

$$\int_{|u|>\frac{|a|}{2h_n}} K(u)K\left(u+\frac{a}{h_n}\right)P_n(du) = o\left(h_n\right).$$

- The second term is bounded above by

$$\begin{aligned}\int_{|u|\le\frac{|a|}{2h_n}} K(u)K\left(u+\frac{a}{h_n}\right)P_n(du) &\le \|K\|_\infty \int_{|u|\le\frac{|a|}{2h_n}} K\left(u+\frac{a}{h_n}\right)P_n(du)\\ &= \|K\|_\infty \int_{|v|>\frac{|a|}{2h_n}} K(v)P_n(dv),\end{aligned}$$

  with $v = u + a/h_n$, and the end of the proof is the same as for (9.4). ∎

Proof of Lemma 3. Taking into account that f vanishes out of $[0,1]$, we have

$$f_n(x) - g_n(x) = \frac{1}{k_n}\sum_{r=1}^{k_n} K_n(x-x_r)f(x_r) - \int_0^1 K_n(x-y)f(y)dy.$$

Let us define $\phi_{n,x}(y) = K_n(x-y)f(y)$ for $(x,y) \in \mathbb{R}^2$. With this notation,

$$f_n(x) - g_n(x) = \sum_{r=1}^{k_n} \int_{I_{n,r}} [\phi_{n,x}(x_r) - \phi_{n,x}(y)]\, dy = \frac{1}{k_n} \sum_{r=1}^{k_n} \int_{-1/2}^{1/2} \left[\phi_{n,x}(x_r) - \phi_{n,x}\left(x_r + \frac{u}{k_n}\right)\right] du. \tag{9.5}$$

and we have the following expansion:

$$\begin{aligned}
\phi_{n,x}(x_r) - \phi_{n,x}\left(x_r + \frac{u}{k_n}\right) &= K_n\left(x - x_r - \frac{u}{k_n}\right)\left(f(x_r) - f\left(x_r + \frac{u}{k_n}\right)\right) && (9.6)\\
&+ f(x_r)\left(K_n(x - x_r) - K_n\left(x - x_r - \frac{u}{k_n}\right)\right), && (9.7)
\end{aligned}$$

Now, since f is α-Lispchitz, (9.6) is uniformly bounded above by $\|K\|_\infty L_f/(h_n k_n^\alpha)$. The rest of the proof depends on the assumptions made on K:

(i) Under **(A)**, K is β-Lipschitzian and thus (9.7) is uniformly bounded above by $\|f\|_\infty L_K/(h_n^{1+\beta} k_n^\beta)$, and the conclusion follows.

(ii) Under **(B)**, since K has a compact support, the number of nonzero terms in (9.5) is $O(k_n h_n)$. Thus, the contribution of (9.6) is $O(1/k_n^\alpha)$. Two situations have to be considered for the term (9.7). If r is such that K has only a bounded first derivative at $x - x_r$, then (9.7) is uniformly bounded above by $\|f\|_\infty \|K'\|_\infty/(h_n^2 k_n)$. Remarking there are only a finite number of such terms in (9.5) shows that the contribution of (9.7) is $O\left(1/(h_n^2 k_n^2)\right)$. If r is such that K is C^2 at $x - x_r$, then a second order Taylor expansion yields

$$K_n(x - x_r) - K_n\left(x - x_r - \frac{u}{k_n}\right) = \frac{u}{k_n} K_n'(x - x_r) - \frac{u^2}{2k_n^2} K_n''\left(x - x_r + \theta_n(u)\frac{u}{k_n}\right),$$

with $\theta_n(u) \in]0,1[$. Replacing in (9.5), the first order term vanishes, and thus the contribution of (9.7) is bounded above by $\|f\|_\infty \|K_n''\|_\infty h_n/(24k_n^2)$. Since $\|K_n''\|_\infty = O\left(1/h_n^3\right)$ the result follows. ■

Proof of Lemma 4. For any compact subset $C \subset]0,1[$, there exist $0 < a < b < 1$ such that $C \subset [a,b]$. Let $x \in [a,b]$ and consider

$$g_n(x) - f(x) = \int_{\mathbb{R}} K_n(u)(f(x-u) - f(x))du. \tag{9.8}$$

Let (δ_n) be a positive sequence tending to 0. Then, (9.8) is bounded above by

$$\begin{aligned} |g_n(x) - f(x)| &\leq \sup_{|u|\leq\delta_n} |f(x-u) - f(x)| + \int_{|u|\geq\delta_n} K_n(u) \, |f(x-u) - f(x)| \, du \\ &\leq \sup_{|u|\leq\delta_n} |f(x-u) - f(x)| + 2\|f\|_\infty \int_{|u|\geq\delta_n} K_n(u)du. \end{aligned}$$

For n large enough, $\delta_n < \min(a, 1-b)$ and then $|u| \leq \delta_n$ entails $(x-u) \in [0,1]$. Now, since f is α-Lipschitzian on $[0,1]$, it yields

$$|g_n(x) - f(x)| \leq L_f \delta_n^\alpha + 2\|f\|_\infty \int_{|u|\geq\delta_n} K_n(u)du. \tag{9.9}$$

Two cases arise:

(i) If $u \to u^2 K(u)$ is integrable then

$$|g_n(x) - f(x)| \leq L_f \delta_n^\alpha + 2\|f\|_\infty \int_{\mathbb{R}} u^2 K(u)du \left(\frac{h_n}{\delta_n}\right)^2.$$

Considering $\delta_n = h_n^{\frac{2}{\alpha+2}}$ in this inequality (which can also be found page 61 in [3] under different hypotheses) gives the result.

(ii) If K has a compact support, let $A > 0$ such that supp $(K) \subset [-A, A]$. Then, considering $\delta_n = Ah_n$, the second term in (9.9) vanishes and the result is proved. ■

Proof of Lemma 5. Consider $x \in C$. As a consequence of the definitions

$$\begin{aligned} \left|\mathrm{E}(\hat{f}_n(x)) - f_n(x)\right| &\leq \frac{1}{k_n}\sum_{r=1}^{k_n} K_n(x - x_r) \max_r \left|\mathrm{E}(X_{n,r}^\star) - f(x_r)\right| \\ &\leq \left(1 + \left\|\frac{1}{k_n}\sum_{r=1}^{k_n} K_n(. - x_r) - 1\right\|_\infty^C\right) \max_r \left|\mathrm{E}(X_{n,r}^\star) - f(x_r)\right| \\ &\leq (1 + o(1)) \max_r \left|\mathrm{E}(X_{n,r}^\star) - f(x_r)\right|, \end{aligned}$$

with Corollary 2. Besides, we have

$$\left|\mathrm{E}(X_{n,r}^{\star}) - f(x_r)\right| \leq \frac{k_n}{nc} + \left|\mathrm{E}(X_{n,r}^{\star}) - k_n\lambda_{n,r} + \frac{k_n}{nc}\right| + |k_n\lambda_{n,r} - f(x_r)|,$$

and Lemma 1 yields

$$\left\|\mathrm{E}(\hat{f}_n) - f_n\right\|_{\infty}^{C} = O\left(\frac{k_n}{n}\right) + O\left(\frac{n}{k_n^{1+2\alpha}}\right) + O\left(\frac{1}{k_n^{\alpha}}\right) = O\left(\frac{k_n}{n}\right),$$

under **(C)**. ■

Proof of Lemma 7. Let $x \in C$. In view of the independence of the $X_{n,r}^{\star}$, $r = 1, \ldots, k_n$,

$$\mathrm{Var}(\hat{f}_n(x)) = \frac{1}{k_n^2}\sum_{r=1}^{k_n} K_n^2(x - x_r)\,\mathrm{Var}(X_{n,r}^{\star}).$$

Introducing

$$\Delta V_n = \frac{n^2}{k_n^2}\max_r\left|\mathrm{Var}(X_{n,r}^{\star}) - \frac{k_n^2}{n^2c^2}\right| \text{ and } \Delta K_n = \left\|\frac{1}{h_nk_n}\sum_{r=1}^{k_n} K^2\left(\frac{.-x_r}{h_n}\right) - \|\,K\,\|_2^2\right\|_{\infty}^{C},$$

we have

$$\left\|\frac{\mathrm{Var}(\hat{f}_n)}{\sigma_n^2} - \sigma^2\right\|_{\infty}^{C} \leq \Delta K_n(\Delta V_n + 1/c^2) + \Delta V_n\|\,K\,\|_2^2.$$

Lemma 1 shows that $\Delta V_n \to 0$, Corollary 2 applied to the kernel $K^2/\|\,K\,\|_2^2$ shows that $\Delta K_n \to 0$ as $n \to \infty$, and the conclusion follows. ■

Proof of Lemma 8. The bias expands as

$$\left\|\mathrm{E}(\tilde{f}_n) - f\right\|_{\infty}^{C} \leq \|f_n - f\|_{\infty}^{C} + \left\|\mathrm{E}(\tilde{f}_n) - f_n\right\|_{\infty}^{C},$$

which first term is controlled by Proposition 1. Consider the second term:

$$\begin{aligned}\left|\mathrm{E}(\tilde{f}_n(x)) - f_n(x)\right| &\leq \left(\frac{1}{k_n}\sum_{r=1}^{k_n} K_n(x - x_r)\right)\max_r\left|\mathrm{E}(X_{n,r}^{\star}) + \mathrm{E}(Z_n) - f(x_r)\right| \qquad (9.10)\\ &\leq \left(1 + \left\|\frac{1}{k_n}\sum_{r=1}^{k_n} K_n(.-x_r) - 1\right\|_{\infty}^{C}\right)\max_r\left|\mathrm{E}(X_{n,r}^{\star}) + \mathrm{E}(Z_n) - f(x_r)\right|.\end{aligned}$$

Corollary 2 shows that it is sufficient to consider

$$\left|\mathrm{E}(X^\star_{n,r}) + \mathrm{E}(Z_n) - f(x_r)\right| \le \left|\mathrm{E}(X^\star_{n,r}) - k_n\lambda_{n,r} + \frac{k_n}{nc}\right| + \left|\mathrm{E}(Z_n) - \frac{k_n}{nc}\right| + \left|k_n\lambda_{n,r} - f(x_r)\right|.$$

Lemma 1 and (6.1) yield

$$\left\|\mathrm{E}(\tilde{f}_n) - f_n\right\|_\infty^C = O\left(\frac{n}{k_n^{1+2\alpha}}\right) + O\left(\frac{1}{k_n^{2\alpha}}\right) + O\left(\frac{1}{k_n^{\alpha}}\right) = O\left(\frac{1}{k_n^{\alpha}}\right)$$

under **(C)**, and the conclusion follows. ∎

Acknowledgement

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

References

[1] Abbar, H. and Suquet, Ch. (1993) Estimation L^2 du contour d'un processus de Poisson homogène sur le plan. *Pub. IRMA Lille*, **31**, **II**.

[2] Bosq, D. (1977) Contribution à la théorie de l'estimation fonctionnelle. *Publications de l'Institut de Statistique de l'Université de Paris*, **XIX**, 2, 1–96.

[3] Bosq, D. and Lecoutre, J. P. (1987) *Théorie de l'estimation fonctionnelle.* Economica, Paris.

[4] Cowling, A. and Hall, P. (1996) On pseudodata methods for removing boundary effects in kernel density estimation. *Journal of the Royal Statistical Society B*, **58**, 551–563.

[5] 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.

[6] 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.

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

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

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

[10] Hall, P., Park, B. U. and Stern, S. E. (1998) On polynomial estimators of frontiers and boundaries. *J. Multiv. Analysis*, **66**, 71–98.

[11] Knight, K. (2001) Limiting distributions of linear programming estimators. *E*xtremes, **4**(2), 87–103.

[12] Girard, S. (2004) On the asymptotic normality of the L_1 error for Haar series estimates of Poisson point processes boundaries, *Statistics and Probability Letters*, **66**, 81–90.

[13] Härdle, W., Park, B. U. and Tsybakov, A. B. (1995) Estimation of a non sharp support boundaries. *J. Multiv. Analysis*, **43**, 205–218.

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

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

[16] Jacob, P. (1984) Estimation du contour discontinu d'un processus ponctuel sur le plan. *Publications de l'Institut de Statistique de l'Université de Paris*, **XXIX**, 3–4, 1–26.

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

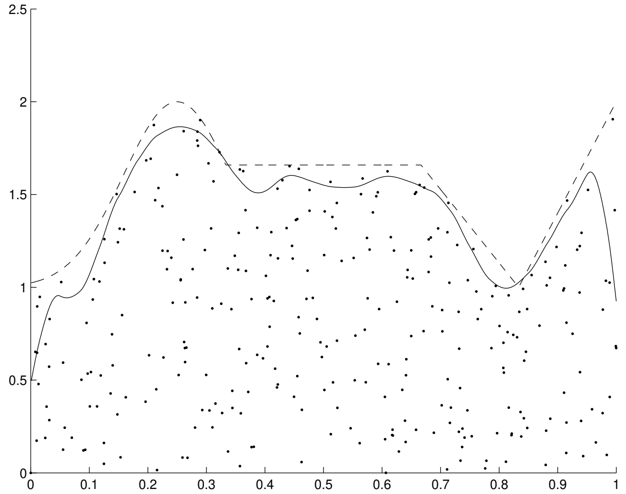
[18] Korostelev, A., Simar, L. and Tsybakov, A. B. (1995) Efficient estimation of monotone boundaries. *The Annals of Statistics*, **23**, 476–489.

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

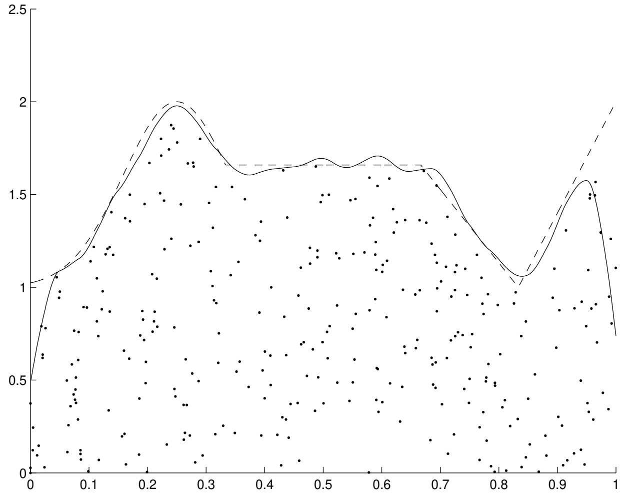
[20] Mammen, E. and Tsybakov, A. B. (1995) Asymptotical minimax recovery of set with smooth boundaries. *The Annals of Statistics*, **23**(2), 502–524.

[21] Reiss, R.D. (1993) A course on point processes. *Springer series in statistics.*

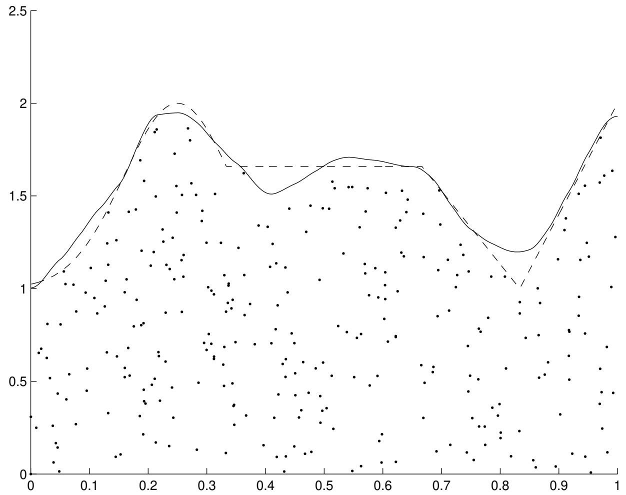
[22] Renyi, A. and Sulanke, R. (1963) Uber die konvexe Hülle von n zufälligen gewählten Punkten. *Z. Wahrscheinlichkeitstheorie verw. Geb.* **2**, 75–84.



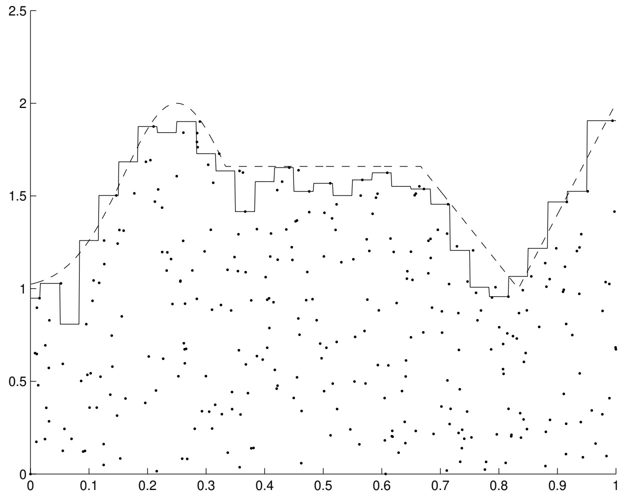
(a) $\hat{f}_n$ estimate



(b) $\tilde{f}_n$ estimate

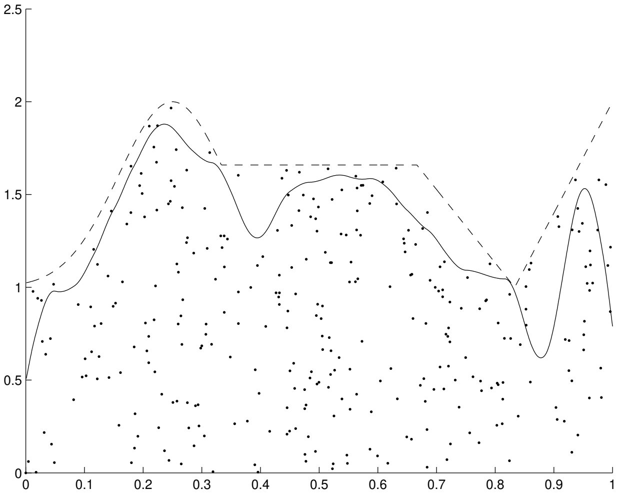


(c) $\check{f}_n$ estimate

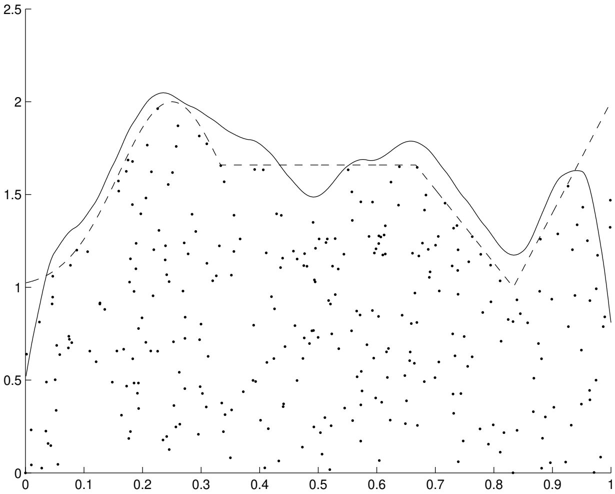


(d) $\hat{f}_n^G$ estimate

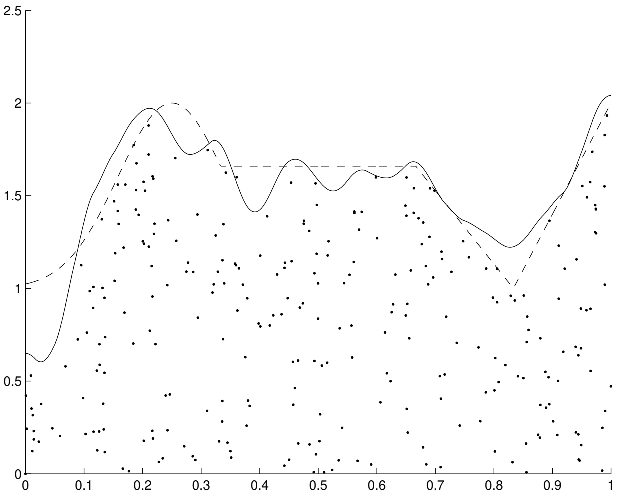
Figure 1: Comparison in the best cases. Continuous line: estimate, dashed line: true frontier.



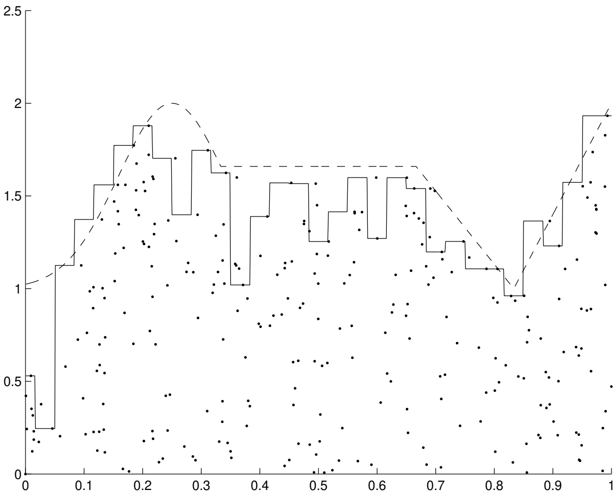
(a) $\hat{f}_n$ estimate



(b) $\tilde{f}_n$ estimate



(c) $\check{f}_n$ estimate



(d) $\hat{f}_n^G$ estimate

Figure 2: Comparison in the worst cases. Continuous line: estimate, dashed line: true frontier.