African Diaspora Journal of Mathematics
Volume 21, Number 1, pp. 87–97

ISSN 1539-854X
www.african-j-math.org

Asymptotic Results For Continuous Associated Kernel Estimators of Density Functions

Célestin C. Kokonendji*
Laboratoire de Mathématiques de Besançon - UMR 6623 CNRS-UFC
Université Bourgogne Franche-Comté - UFR Sciences et Techniques
16, route de Gray - 25030 Besançon cedex, France.

Francial Giscard B. Libengué Dobélé-Kpoka†
Département de Mathématiques et Informatique.
Université de Bangui, Faculté des Sciences.
BP: 908, Bangui République Centrafricaine.

Abstract

Symmetric kernel estimators of an unknown density function on a partial or totally bounded support suffer from edge effects and several authors considered specific asymmetric kernels, belonging in the large class of continuous associated kernels. Asymptotic properties of the corresponding estimators have been examined on a case-by-case basis. In this paper, it is proposed general asymptotic results for continuous associated kernel estimators; in particular, weak and strong global convergences are shown with respect to both uniform and L^1 norms. Three lognormal kernel estimators have used for illustrations and discussions. Finally, some concluding remarks are made.

AMS Subject Classification: 62G07; 62G20; 62G99.

Keywords: Asymmetric kernel, Boundary bias, Convergence, Lognormal distribution, Nonparametric kernel estimator.

1 Introduction

Let $X_1,\ldots,X_n$ be independent and identically distributed (i.i.d.) random variables with an unknown density function f with respect to the Lebesgue measure on its support $\mathbb{T} \subset \mathbb{R}$. The support $\mathbb{T}$ will be supposed to be a convex set of $\mathbb{R}$ and it might be partially or totally bounded. Because of symmetry, the classical or symmetric kernels, not depending on any parameter, are not appropriate smoothers for this

*E-mail address: celestin.kokonendji@univ-fcomte.fr
†E-mail address: libengue@gmail.com

density. Our scope in this work is to investigate the consistency of the following continuous associated kernel estimator,

$$\widehat{f_n}(x) = \frac{1}{n}\sum_{i=1}^{n} K_{x,h}(X_i), \;\; x \in \mathbb{T} := \text{support}(f), \tag{1.1}$$

where $h = h_n > 0$ is an arbitrary sequence of smoothing parameters that fulfills $\lim_{n\to\infty} h_n = 0$, while $K_{x,h}(\cdot)$ is a suitably chosen continuous kernel function which intrinsically depends on the bandwidth h and on the target point x where the density f is estimated. This estimator has been proposed in [11] for discrete functions (see also [1], [10], [21], [22]) and extended to the continuous functions in [9] which generalizes the particular cases of Cheng ([4], [5]) for beta kernels with $\mathbb{T} = [0,1]$ and gamma kernels with $\mathbb{T} = (0,\infty)$. The reader can refer to [8], [13] and [18] for other special continuous cases and also to [7] for a family of asymmetric kernels. It is known in the previous papers that the estimators (1.1) are without of edge effect; however, it is necessary to reduce the pointwise bias inside of $\mathbb{T}$. Several authors have investigated and compared their performances; e.g., [2], [19] and [20].

More precisely, for any $x \in \mathbb{T}$ and $h > 0$, a continuous associated kernel $K_{x,h}$ is also a probability density function (uniformly bounded) with respect to the Lebesgue measure on its support $\mathbb{S}_{x,h} \subseteq \mathbb{T} \subset \mathbb{R}$ satisfying the following three conditions:

$$x \in \mathbb{S}_{x,h}, \;\; \mathbb{E}\left(\mathcal{Z}_{x,h}\right) = x + A(x,h) \;\text{ and } Var\left(\mathcal{Z}_{x,h}\right) = B(x,h), \tag{1.2}$$

where $\mathcal{Z}_{x,h}$ is a real random variable with density $K_{x,h}$, both $A(x,h)$ and $B(x,h)$ tend to 0 when $h = h_n$ goes to 0 (as $n \to \infty$ afterward). Generally built in a "Do It Yourself" way (e.g., [4], [5], [8], [13] and [18]), a continuous associated kernel can be now constructed through the mode-dispersion method introduced in [14]. Indeed, starting from a suitable parametric probability density function (so-called type of continuous kernel) K_θ on $\mathbb{S}_\theta$ with $\theta \in \Theta \subseteq \mathbb{R}^k$ for $k \geq 2$, the use of the mode-dispersion method for constructing an appropriate associeted kernel $K_{x,h} := K_{\theta(x,h)}$ on $\mathbb{S}_{x,h} := \mathbb{S}_{\theta(x,h)}$ that matches with the support $\mathbb{T}$ of the target density consists to solve the following system of two equations: $x = M(a,b)$ and $h = D(a,b)$ where $M(a,b)$ and $D(a,b)$ are, respectively, the only mode and a dispersion parameter of $K_{\theta(a,b)}$. Allowing that $\mathbb{T}$ is partially or totally bounded (e.g., $[t_1,t_2]$, (t_1,∞), $(-\infty,t_2)$ with $t_1 < t_2$), one of choices of the continuous associated kernel $K_{x,h}$ is such that $\mathbb{S}_{x,h}$ matches with $\mathbb{T}$ for all x and h or $\cup_{x,h}\mathbb{S}_{x,h} \subseteq \mathbb{T}$. This is the case of kernels from beta and its extension, gamma and its inverse, inverse Gaussian and its reciprocal version, Weibull, lognormal and so on. See also Figure 1 and [12] for a multivariate version of the mode-dispersion approach. Note that the symmetric kernel estimator of f introduced by Rosenblatt [17] and Parzen [16] can be obtained from (1.1) with

$$K_{x,h}(\cdot) = (1/h)K\{(\cdot - x)/h\}, \;\; x \in \mathbb{T} := \mathbb{R}, \; h > 0,$$

where K is the so-called classical kernel function of bounded variation on its support $\mathbb{S}$ with zero mean and unit variance: $\mathbb{S}_{x,h} = x - h\mathbb{S}$, $A(x,h) = 0$ and $B(x,h) = h^2$ in (1.2).

Following [1] and [3], this paper focuses on some convergences in the family of continuous associated kernel estimators (1.2) for the unknown density function f

on $\mathbb{T} \subset \mathbb{R}$. The assumption of uniform continuity of f on $\mathbb{T} = \mathbb{R}$ is relaxed to a simple continuity of f on any compact set of $\mathbb{T}$. The following Section 2 is devoted to the main results for which we show two kinds of weak and strong consistencies (pointwise and global) of these estimators in the sense of both uniform and $\mathbb{L}^1$ norms. We also present some remarks related to some well-known situations. Finally, Section 3 provides illustrations on three different lognormal kernel estimators for $\mathbb{T} = (0, \infty)$ which is partially bounded to the left.

2 Results and remarks

Let us first show some pointwise properties of the estimator (1.1).

Proposition 2.1. *Under assumptions (1.1) and (1.2), for any fixed x in $\mathbb{T}$ and $h = h_n > 0$, one has*

$$\mathbb{E}\{\widehat{f_n}(x)\} = \mathbb{E}\left\{f(\mathcal{Z}_{x,h})\right\}. \tag{2.1}$$

Furthermore, for f in the class $\mathscr{C}^2(\mathbb{T})$, we have

$$Bias\left\{\widehat{f_n}(x)\right\} = A(x,h)f'(x) + \frac{1}{2}\left\{A^2(x,h) + B(x,h)\right\}f''(x) + o(h^2) \tag{2.2}$$

and, for f bounded on $\mathbb{T}$,

$$Var\left\{\widehat{f_n}(x)\right\} = \frac{1}{n}f(x)\left\|K_{x,h}\right\|_2^2 + o\left(\frac{1}{nh^{r_2}}\right), \tag{2.3}$$

where $r_2 = r_2\left(K_{x,h}\right) > 0$ is a real largest number such that $\left\|K_{x,h}\right\|_2^2 = \int_{\mathbb{S}_{x,h}\cap\mathbb{T}} K_{x,h}^2(u)du \leq c_2(x)h_n^{-r_2}$ and $0 < c_2(x) < \infty$.

Proof. The first result (2.1) is straightforward obtained as follows:

$$\mathbb{E}\left\{\widehat{f_n}(x)\right\} = \int_{\mathbb{S}_{x,h}\cap\mathbb{T}} K_{x,h}(t)f(t)dt = \int_{\mathbb{S}_{x,h}\cap\mathbb{T}} f(t)K_{x,h}(t)dt = \mathbb{E}\left\{f(\mathcal{Z}_{x,h})\right\}.$$

From (2.1) and by using the Taylor-Lagrange formula successively around $\mathbb{E}\left(\mathcal{Z}_{x,h}\right)$ and x, the second result (2.2) can be shown as:

$$\begin{aligned}
Bias\left\{\widehat{f_n}(x)\right\} &= \mathbb{E}\{\widehat{f_n}(x)\} - f(x) = \mathbb{E}\left\{f(\mathcal{Z}_{x,h})\right\} - f(x) \\
&= f\left\{\mathbb{E}(\mathcal{Z}_{x,h})\right\} + \frac{1}{2}Var(\mathcal{Z}_{x,h})f''\left\{\mathbb{E}(\mathcal{Z}_{x,h})\right\} - f(x) + o\left(\mathbb{E}\left\{\mathcal{Z}_{x,h} - \mathbb{E}(\mathcal{Z}_{x,h})\right\}^2\right) \\
&= f\left\{x + A(x,h)\right\} + \frac{1}{2}B(x,h)f''\left\{x + A(x,h)\right\} - f(x) + o\left\{B(x,h)\right\} \\
&= A(x,h)f'(x) + \frac{1}{2}\left\{A^2(x,h) + B(x,h)\right\}f''(x) + o(h^2).
\end{aligned}$$

In fact, the rest $o(h^2)$ comes from (1.2) and $o\left(\mathbb{E}\left\{\mathcal{Z}_{x,h}-\mathbb{E}(\mathcal{Z}_{x,h})\right\}^2\right)=\mathbb{E}\left(o_P\left\{\mathcal{Z}_{x,h}-\mathbb{E}(\mathcal{Z}_{x,h})\right\}^2\right)$ where $o_P(\cdot)$ is the probability rate of convergence. Concerning the variance (2.3) we have

$$\begin{aligned} Var\left\{\widehat{f_n}(x)\right\} &= \frac{1}{n}\mathbb{E}\left\{K^2_{x,h}(X_1)\right\}-\frac{1}{n}\left[\mathbb{E}\left\{K_{x,h}(X_1)\right\}\right]^2 \\ &\simeq \frac{1}{n}\int_{\mathbb{S}_{x,h}\cap\mathbb{T}} K^2_{x,h}(u)f(u)du =: I_1. \end{aligned}$$

By using the Taylor-Lagrange expansion around x, the term I_1 gives

$$I_1 := \frac{1}{n}\int_{\mathbb{S}_{x,h}\cap\mathbb{T}} K^2_{x,h}(u)f(u)du = \frac{1}{n}f(x)\int_{\mathbb{S}_{x,h}\cap\mathbb{T}} K^2_{x,h}(u)du + R(x,h),$$

with

$$R(x,h) = \frac{1}{n}\int_{\mathbb{S}_{x,h}\cap\mathbb{T}} K^2_{x,h}(u)\left[(u-x)f'(x)+\frac{(u-x)^2}{2}f''(x)+o\{(u-x)^2\}\right]du.$$

Under the assumption of $\left\|K_{x,h}\right\|_2^2 \leq c_2(x)h_n^{-r_2}$ we deduce successively

$$0 \leq R(x,h) \leq \frac{1}{nh^{r_2}}\int_{\mathbb{S}_{x,h}\cap\mathbb{T}} c_2(x)\left\{(u-x)f'(x)+\frac{(u-x)^2}{2}f''(x)\right\}du \simeq o(n^{-1}h^{-r_2}).\blacksquare$$

The first consistency results concern two pointwise convergences of $\widehat{f_n}$ almost surely ($a.s.$) and by mean square error (i.e., $\mathbb{L}^2$) implying in probability ($\mathbb{P}$).

Theorem 2.2. *Let $f \in \mathscr{C}^2(\mathbb{T})$ and $\widehat{f_n}$ its estimator defined by (1.1). For any fixed x in $\mathbb{T}$ then one has $\widehat{f_n}(x) \xrightarrow{a.s.} f(x)$ as $n \to \infty$; furthermore, if there is a real largest number $r_2 = r_2(K_{x,h_n}) > 0$ such that*

$$h_n^{r_2}\left\|K_{x,h_n}\right\|_2^2 \leq c_2(x) < \infty \text{ and } \lim_{n\to\infty} nh_n^{r_2} = \infty$$

then $\widehat{f_n}(x) \xrightarrow{\mathbb{L}^2} f(x)$ as $n \to \infty$.

Proof. Since typically the continuous associated kernels $K_{x,h_n}(\cdot)$ are uniformly bounded in $h_n > 0$ and $x \in \mathbb{T}$, and by the Hoeffding inequality, one has

$$\mathbb{P}\left(\left|\sum_{i=1}^n\left[K_{x,h_n}(X_i)-\mathbb{E}\{K_{x,h_n}(X_i)\}\right]\right| \geqslant \varepsilon\right) \leqslant 2\exp\left(-2\varepsilon^2/n\right)$$

and, hence, $\mathbb{P}(|\widehat{f_n}(x)-\mathbb{E}\{\widehat{f_n}(x)\}| \geqslant \varepsilon) \leqslant 2\exp\left(-2n\varepsilon^2\right)$. Then, $|\widehat{f_n}(x)-\mathbb{E}\{\widehat{f_n}(x)\}| \to 0$ as $n \to \infty$ by applying the Borel-Cantelli lemma. Since $|\mathbb{E}\{\widehat{f_n}(x)\}-f(x)| = O(h_n)$ from (1.2) and (2.2), the result $\widehat{f_n}(x) \xrightarrow{a.s.} f(x)$ as $n \to \infty$ follows by the triangular inequality. The second convergence is immediate from (2.2), (2.3) and assumptions.$\blacksquare$

In what follows, our idea of *soft global* convergence is related to the continuity of f on any compact set of $\mathbb{T}$; thus, the continuity on $\mathbb{T}$ denoted by $\mathscr{C}^0(\mathbb{T})$ implies uniform continuity on all compact subsets of $\mathbb{T}$. Before showing the first soft global convergences of $\widehat{f_n}$ by uniform norm, we need the following result on a global bias.

Proposition 2.3. *Let f be in $\mathscr{C}^0(\mathbb{T})$ and bounded, $\widehat{f_n}$ its estimator given in (1.1). For any compact set I included in $\mathbb{T}$, one has: $\sup_{x\in I}|\mathbb{E}\{\widehat{f_n}(x)\}-f(x)|\to 0$ as $n\to\infty$.*

Proof. Let us fix an $\varepsilon>0$. From (1.2) and for $\delta>0$, there exist $n_0,n_1\in\mathbb{N}$ such that: $n>n_0\Rightarrow|\mathbb{E}(\mathcal{Z}_{x,h})-x|<\delta/2$ and $n>n_1\Rightarrow Var(\mathcal{Z}_{x,h})<\delta^2/4$. Writing

$$\left|\mathbb{E}\{\widehat{f_n}(x)\}-f(x)\right|\leqslant\int_{\mathbb{S}_{x,h}\cap\mathbb{T}}\left|f(u)-f(x)\right|K_{x,h_n}(u)du\leqslant J_1+J_2$$

and since $\left\{x\in\mathbb{T};|\mathbb{E}\left(\mathcal{Z}_{x,h}\right)-x|\leqslant\delta/2\right\}\subseteq\mathbb{S}_{x,h}$ with f uniformly continu on I, one has for $n>n_0$

$$J_1=\int_{|u-\mathbb{E}\left(\mathcal{Z}_{x,h}\right)|\leqslant\delta/2}\left|f(u)-f(x)\right|K_{x,h_n}(u)du\leqslant\frac{\varepsilon}{2}\int_{|u-\mathbb{E}\left(\mathcal{Z}_{x,h}\right)|\leqslant\delta/2}K_{x,h_n}(u)du\leqslant\frac{\varepsilon}{2}.$$

On the other hand, the Bienaymé-Tchebychev inequality allows to obtain, for $n>n_1$,

$$\begin{aligned}J_2&=\int_{|u-\mathbb{E}\left(\mathcal{Z}_{x,h}\right)|>\delta/2}\left|f(u)-f(x)\right|K_{x,h_n}(u)du\\&\leqslant\sup_{u\in I}\left|f(u)-f(x)\right|\int_{|u-\mathbb{E}\left(\mathcal{Z}_{x,h}\right)|>\delta/2}K_{x,h_n}(u)du\leqslant\frac{2\varepsilon}{\delta^2}Var\left(\mathcal{Z}_{x,h}\right)\leqslant\frac{\varepsilon}{2}.\end{aligned}$$

Consequently, for $n>\max(n_0,n_1)$ one gets $\sup_{x\in I}|\mathbb{E}\{\widehat{f_n}(x)\}-f(x)|\leqslant\varepsilon.$■

Here we show the results of weak ($\mathbb{P}$) and strong (*a.s.*) soft global convergences with uniform norm.

Theorem 2.4. *Let $f\in\mathscr{C}^0(\mathbb{T})$ and bounded, $\widehat{f_n}$ its estimator given in (1.1). Assume that there is a real largest number $r_0=r_0(K)>0$ such that, for all $x\in\mathbb{T}$,*

$$h_n^{r_0}\int_{\mathbb{S}_{x,h}}|dK_{x,h_n}(s)|\leq c_0$$

with c_0 bounded on any compact containing x. For any compact set I included in $\mathbb{T}$, one has:

$$(i)\quad \text{if }\lim_{n\to\infty}nh_n^{2r_0}=\infty\ \text{ then }\sup_{x\in I}\left|\widehat{f_n}(x)-f(x)\right|\xrightarrow{\mathbb{P}}0\text{ as }n\to\infty;$$

$$(ii)\quad \text{if }\lim_{n\to\infty}nh_n^{2r_0}/\log n=\infty\ \text{ then }\sup_{x\in I}\left|\widehat{f_n}(x)-f(x)\right|\xrightarrow{a.s.}0\text{ as }n\to\infty.$$

Proof. The triangular inequality and Proposition 2.3 lead to: for all $\varepsilon>0$, there exists $n_0\in\mathbb{N}$ such that for $n>n_0$ we have

$$\sup_{x\in I}\left|\widehat{f_n}(x)-f(x)\right|\leqslant\sup_{x\in I}\left|\widehat{f_n}(x)-\mathbb{E}\{\widehat{f_n}(x)\}\right|+\frac{\varepsilon}{2}.\tag{2.4}$$

From F and $F_n=(1/n)\sum_{i=1}^n\delta_{X_I}$ the theoretical and empirical distribution functions of X_i, one can express:

$$\widehat{f_n}(x)=\int_{\mathbb{S}_{x,h}\cap\mathbb{T}}K_{x,h_n}(u)dF_n(u)=K_{x,h_n}(u)F_n(u)\Big|_{\mathbb{S}_{x,h}\cap\mathbb{T}}-\int_{\mathbb{S}_{x,h}\cap\mathbb{T}}F_n(u)dK_{x,h_n}(u)$$

and then

$$\mathbb{E}\{\widehat{f_n}(x)\} = K_{x,h_n}(u)F(u)\Big|_{\mathbb{S}_{x,h}\cap\mathbb{T}} - \int_{\mathbb{S}_{x,h}\cap\mathbb{T}} F(u)dK_{x,h_n}(u).$$

Since $F_n(u) \xrightarrow{a.s.} F(u)$ as $n \to \infty$ by Glivenko-Cantelli, one has

$$\left|\widehat{f_n}(x) - \mathbb{E}\{\widehat{f_n}(x)\}\right| \leqslant \int_{\mathbb{S}_{x,h}\cap\mathbb{T}} |F_n(u) - F(u)|\, dK_{x,h_n}(u).$$

By (2.4) and assumption $h_n^{r_0}\int_{\mathbb{S}_{x,h}} |dK_{x,h_n}(s)| \leqslant c_0(x)$, one gets

$$\begin{aligned}\sup_{x\in I\subseteq\mathbb{T}}\left|\widehat{f_n}(x) - f(x)\right| &\leqslant \left(\sup_{x\in I\subseteq\mathbb{T}} |F_n(x) - F(x)|\right)\left(\sup_{x\in I\subseteq\mathbb{T}} \int_{\mathbb{S}_{x,h}\cap\mathbb{T}} \left|dK_{x,h_n}(u)\right|\right)\\ &\leqslant C_0 h_n^{-r_0} \sup_{x\in I\subseteq\mathbb{T}} |F_n(x) - F(x)|\end{aligned}$$

with $C_0 = \sup_{x\in I\subseteq\mathbb{T}} c_0 < \infty$. For given $h_n > 0$ and adapting now the Massart [15] inequality (which is an extension of the Dvoretzky-Kiefer-Wolfowitz inequality) one has, for all $\varepsilon > 0$, $\mathbb{P}(\sup_{x\in I\subseteq\mathbb{T}} |F_n(x) - F(x)| \geqslant \varepsilon) \leqslant 2\exp\left(-2\varepsilon^2 n h_n^{2r_0}/C_0^2\right) \to 0$; and Part (*i*) ensues from it. According to the previous proof of Part (i), we have

$$\sum_{n\geqslant 1} \mathbb{P}\left(\sup_{x\in I\subseteq\mathbb{T}} |\widehat{f_n}(x) - f(x)| \geqslant \varepsilon\right) \leqslant 2\sum_{n\geqslant 1} \exp\left(-2\varepsilon^2 n h_n^{2r_0}/C_0^2\right).$$

For an appropriate choice of $\varepsilon > 0$ and n large (i.e., $n > n_0$), one gets $\sum_{n\geqslant 1}\exp(-2\varepsilon^2 n h_n^{2r_0}/C_0^2) \simeq o(n^{-(1+\delta)}), \forall\delta > 0$. Thus, one has the almost complete convergence and Part (*ii*) is deduced.■

Concerning the second soft global convergences through $\mathbb{L}^1$ norm with respect to the Lebesgue measure, we first prove that the bias of $\widehat{f_n}$ converges.

Proposition 2.5. *Let $f \in \mathscr{C}^0(\mathbb{T})$ and $\widehat{f_n}$ its estimator given in (1.1). Then:*

$$\int_{\mathbb{T}} \left|\mathbb{E}\{\widehat{f_n}(x)\} - f(x)\right| dx \to 0 \ \text{ as } n \to \infty.$$

Proof. In a similar way to Devroye [6], we use the following decomposition

$$\left\|\mathbb{E}\left(\widehat{f_n}\right) - f\right\|_1 \leq \left\|\mathbb{E}\left(\widehat{f_n}\right) - \mathbb{E}\left(\widehat{f_n^\star}\right)\right\|_1 + \left\|\mathbb{E}\left(\widehat{f_n^\star}\right) - f^\star\right\|_1 + \left\|f^\star - f\right\|_1,$$

where $f^\star$ is continue and such that $\left\|(f - f^\star)\mathbf{1}_I\right\|_1 < \varepsilon, \forall\varepsilon > 0$ and $f^\star = 0$ on $I^c := \mathbb{T}\setminus I$. For any given event A, $\mathbf{1}_A$ denotes the indicator function of A that takes the value 1 if the event A occurs and 0 otherwise. By considering $\widehat{f_n^\star} = \mathbf{1}_I\widehat{f_n}$ one has, on one hand

$$\begin{aligned}\left\|\mathbb{E}\left(\widehat{f_n}\right) - \mathbb{E}\left(\widehat{f_n^\star}\right)\right\|_1 &= \int_{\mathbb{T}}\left|\int_{\mathbb{S}_{x,h}\cap\mathbb{T}} K_{x,h_n}(u)\{f(u) - \mathbf{1}_I(x)f(u)\}du\right|dx\\ &\leq \Lambda_n \int_{\mathbb{S}_{x,h}\cap\mathbb{T}} \left|f(u) - f^\star(u)\right| du \leq \Lambda_n\varepsilon\end{aligned}$$

with the total mass $\Lambda_n = \int_{\mathbb{T}} K_{x,h_n}(u)dx \simeq 1, \forall u$, and on the other hand

$$\begin{aligned}\left\|\mathbb{E}\left(\widehat{f_n^\star}\right) - f^\star\right\|_1 &\leq \int_I \left| \int_{\mathbb{S}_{x,h}\cap\mathbb{T}} K_{x,h_n}(u)\left\{f^\star(u) - f^\star(x)\right\}du\right| dx \\ &\leq \varepsilon \int_I \int_{\mathbb{S}_{x,h}\cap\mathbb{T}} K_{x,h_n}(u)dxdu \leq \Lambda_n \ell(I)\varepsilon,\end{aligned}$$

where $\ell(I)$ denotes the length of I. Since $\left\|f^\star - f\right\|_1 = \left\|(f - f^\star)\mathbf{1}_I\right\|_1 < \varepsilon$, one deduces $\left\|\mathbb{E}\left(\widehat{f_n}\right) - f\right\|_1 \leq [\Lambda_n\{1 + \ell(I)\} + 1]\,\varepsilon.$ ■

Finally, we only state the last soft global convergences of $\widehat{f_n}$ because the proofs are similar to those of Theorem 2.4, and we omit them.

Theorem 2.6. *Let $f \in \mathscr{C}^0(\mathbb{T})$ and bounded, $\widehat{f_n}$ its estimator (1.1). Suppose that there is a real largest number $r_1 = r_1(K) > 0$ such that, for all $x \in \mathbb{T}$,*

$$h_n^{r_1} \int_{\mathbb{S}_{x,h}} |dK_{x,h_n}(s)| \leq c_1$$

with c_1 in $\mathbb{L}^1$ on any compact set containing x. Then one has:

$$(i) \quad \textit{if } \lim_{n\to\infty} nh_n^{2r_1} = \infty \quad \textit{then} \quad \int_{\mathbb{T}} \left|\widehat{f_n}(x) - f(x)\right| dx \xrightarrow{\mathbb{P}} 0 \textit{ as } n \to \infty;$$

$$(ii) \quad \textit{if } \lim_{n\to\infty} nh_n^{2r_1}/\log n = \infty \quad \textit{then} \quad \int_{\mathbb{T}} \left|\widehat{f_n}(x) - f(x)\right| dx \xrightarrow{a.s.} 0 \textit{ as } n \to \infty.$$

The following remark completes our idea to make a study for pointing out the sensibility of the continuous associated kernel, which depends on the target of estimation.

Remark 2.7. Let $x \in \mathbb{T}$. (i) If c_0 of Theorem 2.4 garanteeing a kind of the pointwise bounded variation of the associated kernel is integrable on all compact containing x then $r_1 = r_0$, where r_1 is given in Theorem 2.6. (ii) The function c_2 of Theorem 2.2 means the pointwise square integrability of the continuous associated kernel.

We conclude this section by the four following observations.

[A] For classical (continuous associated) kernels K, one easily checks that c_j are constants for $j = 0,1,2$ in Theorems 2.2, 2.4 and 2.6; they provide the same sense to the respective hypothesis of square integrable of K for $j = 2$, bounded variation of K for $j = 0$, and integrable variation of K for $j = 1$. Moreover, one has $r_0 = r_1 = r_2 = 1$.

[B] As regards the (non-classical and well-known) beta and gamma kernels, one has $r_0 = r_1 = 2r_2 = 1$ for their first versions and $\widetilde{r_0} = \widetilde{r_1} = \widetilde{r}_{2,-1} = 2\widetilde{r}_{2,0} = 1$ for their modified versions in the sense of Chen ([4], [5]), where $\widetilde{r}_{2,-1}$ and $\widetilde{r}_{2,0}$ represent r_2 for the left boundary (−1) and inside (0) points respectively; see also [3].

[C] Figure 1 illustrates two behaviours of some continuous associated kernels on $\mathbb{T} = (0,\infty)$: (a) at the left boundary with $x = 0.5 = h$ and (b) inside area with $x = 1.5$ and $h = 0.3$. Lognormal, gamma and reciprocal inverse Gaussian kernels appear here to be interesting smoothers on this positive real line, while inverse gamma and inverse Gaussian can be used only for a left boundary area of $\mathbb{T} = (0,\infty)$. In general, the choice of any continuous associated kernel for the boundary part (left or right) must be crucial. However, we can require a classical one such as Gaussian or Epanechnikov for inside points.

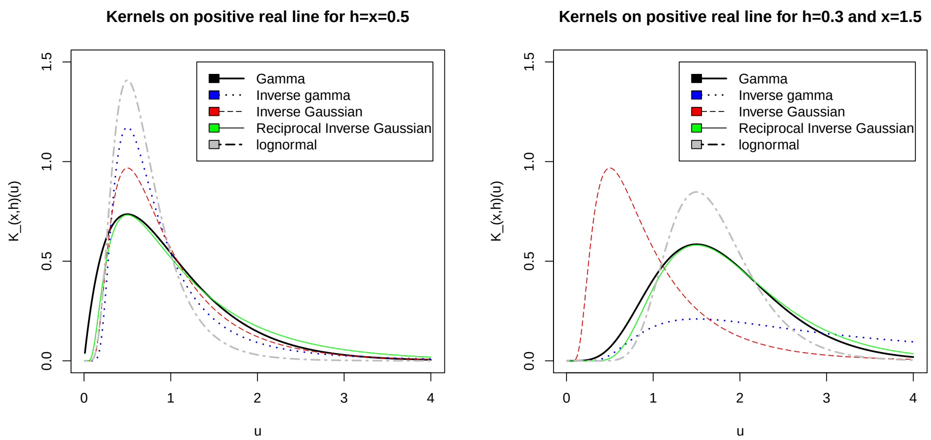

Figure 1. Some continuous associated kernels on positive real line for a given h and x

[D] The practical selection of bandwidth in this family of continuous associated kernel estimators must be done through an adaptative (like Lepski) method (e.g., [2]) or a Bayesian procedure (e.g., [22]). Because the cross-validation method is so slow at boundary regions. Work in this direction by simulations and an application is in progress.

3 Applications to lognormal kernel estimators

We now examine three families of (associated) lorgnormal kernel estimators.

First, we recall the basical properties of lognormal distribution (or type of kernel) that we need. The density function of lognormal with two parameters $a \in \mathbb{R}$ and $b > 0$ is defined as follows:

$$LN_{\theta(a,b)}(u) = \frac{1}{ub\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{\log u - a}{b}\right)^2\right\} \mathbf{1}_{(0,\infty)}(u).$$

Since its mode and a dispersion parameter are given by $\exp(a-b^2)$ and b respectively, the mode-dispersion method provides

$$\theta(x,h) = \left(\log x + h^2, h\right) \text{ for all } x \in (0,\infty) \text{ and } h > 0.$$

Note also that its mean $\exp(a+b^2/2)$ and its variance $\left(\exp b^2-1\right)\exp(2a+b^2)$ allow to calculate

$$A_\theta(x,h)=x\left(\exp(3h^2/2)-1\right) \text{ and } B_\theta(x,h)=x^2\exp(3h^2)\left(\exp h^2-1\right)$$

as introduced in (1.2) with $\mathbb{S}_{LN_{\theta(x,h)}}=(0,\infty)=\mathbb{S}_{LN_{\theta(a,b)}}$.

Then, to estimate an unknown density function f with support $\mathbb{T}=(0,\infty)$ we can first consider the lognormal kernel $LN_{\theta(x,h)}$ obtained by the above mode-dispersion method and used in Figure 1:

$$LN_{\theta(x,h)}(u)=\frac{1}{uh\sqrt{2\pi}}\exp\left(\frac{-1}{2h^2}\left[\log u-\log\left\{x\exp\left(h^2\right)\right\}\right]\right)\mathbf{1}_{u>0},\quad x>0,h>0.$$

For $\alpha_1(h)>0$ bounding both left edge and inside regions of $\mathbb{T}=(0,\infty)=(0,\alpha_1(h)]\cup(\alpha_1(h),\infty)$, a modified version $LN_{\widetilde{\theta}(x,h)}:=\widetilde{LN}_{\theta(x,h)}$ of $LN_{\theta(x,h)}$ is given by

$$LN_{\widetilde{\theta}(x,h)}=LN_{\widetilde{\theta}_{-1}(x,h)}+LN_{\widetilde{\theta}_0(x,h)}$$

with the left boundary part

$$\widetilde{\theta}_{-1}(x,h)=\theta\left(x^{-2}\alpha_1^3(h)\exp\left(-3h^2/2\right),h\right) \text{ for } x\in(0,\alpha_1(h)]=:\mathbb{T}_{-1}$$

and the inside part

$$\widetilde{\theta}_0(x,h)=\theta\left(x\exp\left(-3h^2/2\right),h\right) \text{ for } x\in(\alpha_1(h),\infty)=:\mathbb{T}_0.$$

Let us mention here that, for both modified beta and gamma kernels, $\alpha_1(h)=2h$ has been chosen by Chen ([4], [5]) and $\alpha_1(h)=h$ by Zhang [19] and Zhang and Karunamuni [20]. The third version of lognormal kernel defined as

$$LN_{\theta^*(x,h)}=LN_{\theta\left(x\exp h^2,2\sqrt{\log(1+h)}\right)}$$

has been considered by Jin and Kawczak [8]. From (1.1) we have three estimators $\widehat{f_n}$, $\widetilde{f_n}$ and f_n^* of f corresponding to the above three lognormal kernels LN_θ, $LN_{\widetilde{\theta}}$ and LN_{θ^*} respectively. Thus, denoting by $\widehat{r_j}=r_j(\widehat{f_n})$, $\widetilde{r_j}=r_j(\widetilde{f_n})$ and $r_j^*=r_j(f_n^*)$ the r_j of these lognormal kernel estimators as in Theorems 2.2, 2.4 and 2.6 for $j=0,1,2$, we have the following results:

$$\widehat{r_0}=\widetilde{r_0}=r_0^*=\widehat{r_1}=\widetilde{r_1}=r_1^*=2, \tag{3.1}$$

$$\widehat{r_2}=\widetilde{r}_{2,0}=2r_0^*=1 \text{ and } \widetilde{r}_{2,-1}=\widetilde{r}_{2,-1}(\alpha_1); \tag{3.2}$$

in particular, if $\alpha_1(h)=a_1h^\beta$ then $\widetilde{r}_{2,-1}(\alpha_1)=(3\beta\sqrt{2}+2)/2$ for $\beta>-\sqrt{2}/3$.

Finally, we show the two situations $\widetilde{r_0}=2$ of (3.1) and $\widetilde{r}_{2,-1}=\widetilde{r}_{2,-1}(\alpha_1)$ depending on the bounding $\alpha_1:=\alpha_1(h)$ of (3.2). Indeed, one has

$$\begin{aligned}\int|dLN_{\widetilde{\theta}(x,h_n)}(u)| &= \frac{1}{{h_n}^2}\int\left|\frac{1}{u}\log\left\{\frac{ux^2}{\alpha_1(h_n)}\right\}LN_{\widetilde{\theta}_{-1}(x,h_n)}(u)+\frac{1}{u}\log\left(\frac{u}{x}\right)LN_{\widetilde{\theta}_0(x,h_n)}(u)\right|du\\
&\leqslant \frac{\kappa_{-1}}{{h_n}^2}\int\left|\frac{2}{u}\log(x)+\frac{1}{u}\log(u)\right|\mathbf{1}_{u\leqslant\alpha_1(h_n)}du+\frac{\kappa_0}{{h_n}^2}\int\left|\frac{2}{u}\log(x)\right|\mathbf{1}_{u>\alpha_1(h_n)}du\\
&\leqslant {h_n}^{-2}\log^2(x)\left\{(5\kappa_{-1}/2)\mathbf{1}_{x\leqslant\alpha_1(h_n)}+\kappa_0\mathbf{1}_{x>\alpha_1(h_n)}\right\},\end{aligned}$$

where κ_{-1} and κ_0 are positive constants; this leads to $\widetilde{r}_0 = 2$. As for $\widetilde{r}_{2,-1} = \widetilde{r}_{2,-1}(\alpha_1)$ and since

$$\mathbb{E}\left(Y^m\right) = \exp\left[m\left\{h^2 + \log(x)\right\} + (mh)^2/2\right],\ \forall m \in \mathbb{R},$$

for any lognormal random variable $Y \sim LN_{\theta(x,h)}$, one gets

$$\begin{aligned}
\left\|LN_{\widetilde{\theta}_{-1}(x,h_n)}\right\|_2^2 &= \int_0^{\{\alpha_1(h_n)\}^{\sqrt{2}}} \frac{v^{-\sqrt{2}/2}dv}{vh_n^2\pi\sqrt{8}} \exp\left(\frac{-1}{2h_n^2}\left[\log v - \log\left\{\frac{\{\alpha_1(h_n)\}^{3\sqrt{2}}}{x^{2\sqrt{2}}e^{\sqrt{2}h_n^2/2}}\right\}\right]^2\right) \\
&\leqslant \frac{1}{2h_n\sqrt{2\pi}}\mathbb{E}\left(Y_\star^{-\sqrt{2}/2}\right) = \frac{1}{2h_n\sqrt{2\pi}}\left\{\frac{x^2}{\alpha_1^3(h_n)}\right\}^{\sqrt{2}/2}\exp\left(\frac{\sqrt{2}+1}{4}h_n^2\right) \\
&\leqslant \frac{x^{\sqrt{2}/2}\left[1+\left\{\left(\sqrt{2}+1\right)/4\right\}h_n^2\right]}{2\sqrt{\pi}h_n\{\alpha_1(h_n)\}^{3\sqrt{2}/2}},
\end{aligned}$$

with $Y_\star \sim LN_{\theta(x^\star,h_n)}$ such that $x^\star = x^{-\sqrt{2}/2}\{\alpha_1(h_n)\}^{3\sqrt{2}}\exp(-\sqrt{2}h_n^2/2 - h_n^2)$; and, therefore, $\widetilde{r}_{2,-1}(\alpha_1)$ is the largest power in h_n of $h_n\{\alpha_1(h_n)\}^{3\sqrt{2}/2}$.

Acknowledgements. The authors sincerely thank the Associate Editor and the anonymous referee for their valuable comments and suggestions. We also thank the Agence Universitaire de la Francophonie (AUF) for having financed this work. Part of this work was done while the second author was visiting the Laboratoire de Mathématique de Besançon, France.

References

[1] B. Abdous and C.C. Kokonendji, Consistency and asymptotic normality for discrete associated-kernel estimator. *African Diaspora Journal of Mathematics*, **8** (2009), 63-70.

[2] K. Bertin and N. Klutchnikoff, Minimax properties of beta kernel estimators. *Journal of Statistical Planning and Inference*, **141** (2011), 2287-2297.

[3] T. Bouezmarni and J.M. Rolin, Consistency of the beta kernel density function estimator. *The Canadian Journal of Statistics*, **31** (2003), 89-98.

[4] S.X. Chen, Beta kernel estimators for density functions. *Computational Statistics and Data Analysis*, **31** (1999), 131-145.

[5] S.X. Chen, Gamma kernel estimators for density functions. *Annals of the Institute of Statistical Mathematics*, **52** (2000), 471-480.

[6] L. Devroye, The equivalence of weak, strong and complete convergence in L_1 for kernel density estimate. *The Annals of Statistics*, **11** (1983), 896-904.

[7] M. Hirukawa, *Asymmetric Kernel Smoothing - Theory and Applications in Economics and Finance*. Springer, Singapore 2018.

[8] X. Jin and J. Kawczak, Birnbaum-Saunder and lognormal kernel estimators for modelling durations in high frequency financial data. *Annals of Economics and Finance*, **4** (2003), 103-1024.

[9] C.C. Kokonendji and T. Senga Kiessé, Discrete associated kernels method and extensions. *Statistical Methodology*, **8** (2011), 497-516.

[10] C.C. Kokonendji, T. Senga Kiessé and N. Balakrishnan, Semiparametric estimation for count data through weighted distributions. *Journal of Statistical Planning and Inference*, **139** (2009), 3625-3638.

[11] C.C. Kokonendji, T. Senga Kiessé and S.S. Zocchi. Discrete triangular distributions and non-parametric estimation for probability mass function. *Journal of Nonparametric Statistics*, **19** (2007), 241-254.

[12] C.C. Kokonendji and M.S. Somé. On multivariate associated kernels to estimate general density functions. *Journal of the Korean Stat. Soc.* **47** (2018), 112-126.

[13] C.N. Kuruwita, K.B. Kulasekera and W.J. Padgett, Density estimation using asymmetric kernels and Bayes bandwidths with censored data. *Journal of Statistical Planning and Inference*, **140** (2010), 1765-1774.

[14] F. G. B. Libengué Dobélé-Kpoka and C. C. Kokonendji, The mode-dispersion approach for constructing continuous associated kernels. *Afrika Statistika*, **12** (2017), 1417-1446.

[15] P. Massart, The tight constant in Dvoretzky-Kiefer-Wolfowitz inequality. *The Annals of Probability*, **18** (1990), 1269-1283.

[16] E. Parzen, On estimation of a probability density function and mode. *Annals of Mathematical Statistics*, **33** (1962), 1065-1076.

[17] M. Rosenblatt, Remarks on some nonparametric estimates of a density function, *Annals of Mathematical Statistics*, **27** (1956), 832-837.

[18] O. Scaillet, Density estimation using inverse and reciprocal inverse Gaussian kernels. *Journal of Nonparametric Statistics*, **16** (2004), 217-226.

[19] S. Zhang, A note on the performance of the gamma kernel estimators at the boundary. *Statistics and Probability Letters*, **80** (2010), 548-557.

[20] S. Zhang and R.J. Karunamuni, Boundary performance of the beta-kernel estimator. *Journal of Nonparametric Statistics*, **22** (2009), 81-104.

[21] N. Zougab, S. Adjabi and C.C. Kokonendji, Binomial kernel and Bayes local bandwidth in discrete functions estimation. *Journal of Nonparametric Statistics*, **24** (2012), 783-795.

[22] N. Zougab, S. Adjabi and C.C. Kokonendji, Adaptive smoothing in associated kernel discrete functions estimation using Bayesian approach. *Journal of Statistical Computation and Simulation*, **83** (2013), 2219-2231.