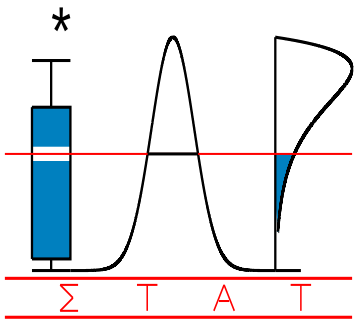
T E C H N I C A L

R E P O R T

0339

L_1 rate of convergence of Asymmetric Kernel Density Estimators and Smoothed Histograms

T.BOUEZMARNI, M. Mesfioui and J.M.ROLIN



I A P S T A T I S T I C S

N E T W O R K

INTERUNIVERSITY ATTRACTION POLE

http://www.stat.ucl.ac.be/IAP

L_1 rate of convergence of Asymmetric Kernel Density Estimators and Smoothed Histograms

T. Bouezmarni *, M. Mesfioui † and J.M. Rolin *

December 17, 2003

Abstract

The authors consider the Smoothed histograms (Gawronski and Stadtmüller (1980)) and the Gamma kernel density estimator (Chen 2000) for an iid sample of a density defined on $[0, +\infty)$. They give the asymptotic behavior, the lower and upper bound of the expected average absolute error of this estimator on each compact. The general formula of the asymptotic behavior, the lower and upper bound of the expected average absolute error for the generalized kernel estimates will be established.

Key words and phrases: Asymptotic properties, Smoothed histograms, Gamma kernel, Nonparametric density estimation, L_1 error.
MSC 2000: Primary 62G07; secondary 62G20.

We would like to thank I. Gijbels for her stimulating remarks. Bouezmarni and Rolin gratefully ackowledge financial support from the "Projet d'Actions de Recherche Concertées" n^0 98/03 – 217, and from the IAP research network n^0 P5/24 of the Belgian state. The second author acknowledges the financial support of the Natural Sciences and Engineering Research Council of Canada. Part of this research was done when Bouezmarni was visiting Université du Québec a Trois Rivière.

*Institute of Statistics, Université Catholique de Louvain.
†Departement de mathémathique et d'informatique, Université du Québec a Trois Rivière.

1 Introduction

Let f be an unknown density function defined on the positive real line. Without correction, the standard kernel estimates suffer from boundary effects. In fact, using a fixed symmetric kernel is not appropriate for fitting densities with bounded supports as a weight is given outside the support. Many methods are suggested to correct this boundary problem: Boundary modification Rice(1985), Smooth optimum kernels Muller (1991), local linear estimates Lejeune and Sarda (1992) and boundary correction using generalized jackknifing Jones (1993), but these methods allow the corrected estimator to become negative. Recently, to correct this negativity, Marron and Ruppert (1994) propose transformations to reduce boundary bias, Jones and Foster (1996) introduce a nonnegative boundary correction for kernel density estimator and Chen (2000) the Gamma kernel estimator.

In this paper, we study the performance of two estimators: the smoothed histogram using the Poisson distribution proposed by Gawronski and Stadtmüller (1980) and the Gamma kernel estimator introduced by Chen (2000). The two estimators are proposed to remove the boundary bias of the standard kernel density estimator. They are very simple in implementation, free of boundary bias, non-negative and their rate of convergence for the mean integrated squared error is $O(n^{-4/5})$. One goal of the paper is to find a correspondant result in L_1 for the gamma kernel estimate, for the smoothed histogram and for the generalized kernel under some conditions. The monograph of Devroye and Gyorfi (1985) contains lots of results for the standard kernel density estimator in the L_1 case. The choice of L_1 distance is motivated by many reasons: it is a natural metric on the space of density functions, proportional to the total variation metric, invariant under monotone transformation of the coordinate axes. To avoid a slow convergence result, a mixture of continuity and tail conditions seems necessary, and even the non differentiability of the density function at one point suffices to obtain slow convergence result. For more detail see Devroye and Gyorfi (1985).

The paper is structured as follows. In section 2, we present the gamma kernel density estimator and the smoothed histogram, and some quantities closely linked to the rate of convergence of the expected average absolute error of these estimators. For the class of density function defined on a compact support,

the rate of convergence, the lower and upper bound of the expected average absolute error of these estimators are proved in Section 3. In Section 4, we extend the results obtained in Section 3 to the generalized kernel estimators.

2 The gamma kernel and the smoothed histogram

Let $X_1, \cdots, X_n$ be a random sample from an unknown density function f defined on the positive real line. Near the origin, the standard kernel estimates have very often a considerable bias. In fact, the standard kernel density function has no knowledge of the boundary, and in general assigns probability mass outside the support. This is due to the use of fixed and symmetric kernel. To correct the boundary problem a simple idea is to use a flexible kernel, with varying shape and the amount of smoothing, which never assigns weight outside the support of the density function. Chen (2000) proposed the gamma kernel estimator, namely

$$\hat{f}_b(x) = \frac{1}{n}\sum_{i=1}^{n} K_1(x,b)(X_i) \tag{1}$$

where b is the bandwidth and $K_1(x,b)$ denotes the asymmetric kernel which is the gamma density with parameters $(x/b+1,b)$ given by

$$K_1(x,b)(t) = \Gamma_{(x/b+1,b)}(t) = \frac{t^{x/b}e^{-t/b}}{b^{x/b+1}\Gamma(x/b+1)}. \tag{2}$$

The second estimator considered here is a smoothed histogram developed by Gawronski and Stadtmüller (1980, 1981). This estimator is based on Poisson distribution and it is defined as follows

$$\hat{f}_k(x) = k\sum_{i=0}^{+\infty} \omega_{i,k}\, p_{ki}(x). \tag{3}$$

The weights $\omega_{i,k}$ are random and defined by

$$\omega_{i,k} = F_n\left(\frac{i+1}{k}\right) - F_n\left(\frac{i}{k}\right) \tag{4}$$

where F_n denotes the empirical distribution function. The integer k denotes the smoothing parameter and $p_{ki}(.)$ is a Poisson distribution function with

parameter kx, namely

$$p_{ki}(x) = e^{-kx}\frac{(kx)^i}{i!} \qquad i = 0, 1, \ldots.$$

For this choice, the smoothed histogram $\hat{f}_k$ can be expressed as a mixture of gamma densities:

$$\hat{f}_k(x) = \sum_{i=0}^{+\infty} \omega_{i,k}\, \Gamma_{(i+1,1/k)}(x)$$

where $\Gamma_{(i+1,1/k)}$ denotes gamma distribution defined as in (2). Note that this estimator can be also reformulated as a standard expression:

$$\hat{f}_k(x) = \frac{1}{n}\sum_{i=0}^{n} K_2(x,k)(X_i) \tag{5}$$

where

$$K_2(x,k)(X_i) = \sum_{j=0}^{+\infty} \mathbb{1}_{\left\{X_i \in A_{k,j}\right\}}\, \Gamma_{(j+1,1/k)}(x) \tag{6}$$

and

$$A_{k,j} = \left]\frac{j}{k}, \frac{j+1}{k}\right].$$

The gamma kernel density estimator and the smoothed histogram are a particular case of the generalized kernel density estimator (Foldes and Revesz (1974) and, Walter and Blum (1979)), but do not take the form of convolution product as the standard kernel. Note also that both are usually free of boundary bias, nonnegative and attain the optimal rate of convergence error for the mean integrated squared error. Bouezmarni and Scaillet (2003) proved their L_1 and their uniform convergence for continuous densities, in this paper we give the rate of convergence in L_1 of the two estimators given by (1) and (5). Similar results are also given for the generalized kernel estimate under some conditions.

2.1 Preliminaries

In this section, we are going to show some elementary results used throughout this paper. We start by the useful function R defined by

$$R(x) = \sqrt{2\pi}x^{x+1/2}e^{-x}/\Gamma(x+1), \quad x > 0.$$

Now, in order to obtain upper bounds of the kernels $\Gamma_{(x/b+1,b)}(t)$ and $\Gamma_{(j+1,1/k)}(x)$, let us recall the following important properties of R, see Chen(1999):

$$R(x) < 1, \text{ for all } x > 0$$

and

$$R(x) \longrightarrow 1 \text{ as } x \longrightarrow \infty.$$

Combining these properties with Stirling's formula and using the fact that x is the mode of the kernel $K(x, b)$ in (2), we get that

$$\Gamma_{(x/b+1,b)}(t) \leq \sqrt{\frac{1}{2\pi x b}} \tag{7}$$

and

$$\Gamma_{(j+1,1/k)}(x) \leq \sqrt{\frac{k}{2\pi x}} \quad j \geq 0. \tag{8}$$

In the remainder of this paper, we need the next quantities

$$\alpha(x) = \sqrt{\frac{f(x)}{2\sqrt{\pi x}}} \tag{9}$$

and

$$\beta(x) = af'(x) + \frac{1}{2}xf''(x). \tag{10}$$

To examine the gamma kernel, we will choose $a = 1$ in (10), but for the smoothed histogram, the appropriate value to choose is then $a = 1/2$. Now, the formula

$$B_M(f) = \left(\int_0^M \alpha(x)dx\right)^{4/5} \left(\int_0^M \mid \beta(x) \mid dx\right)^{1/5} \tag{11}$$

is closely linked to the rate of convergence ofhe expected average absolute error for twise continuously differentiable densities f with bounded support $[0, M]$, i.e $f \in C2([0, 1])$.

Finally, to describe the exact asymptotic behavior of $I\!E(J_{bn})$, let us introduce as in Devroye & Penrod (1984), the following function

$$\psi(u) = \sqrt{\frac{2}{\pi}}\left(u\int_0^u e^{-x^2/2}dx + e^{-u^2/2}\right), \qquad u \geq 0\,.$$

Note that ψ is a monotone increasing convex function satisfying the following properties

$$\psi(u) = I\!E(\mid Z - u \mid), \quad \text{for all} \quad u \geq 0$$

where $Z \sim N(0,1)$ and

$$\max\left\{u, \sqrt{\frac{2}{\pi}}\right\} \leq \psi(u) \leq u + \sqrt{\frac{2}{\pi}}. \tag{12}$$

3 L_1- rates of convergence for the gamma kernel and for the smoothed histogram

The optimal rate of convergence for the mean integrated squared error for gamma kernel estimate and the smoothed histogram is $O(n^{-4/5})$ provided that the underlying density is twice derivable, for more details see Chen(2000). Our aim in this section is to find a similar result in L_1. For that, we will examine the asymptotic expression for the mean integrated absolute deviation

$$I\!E(J_{hn}) = I\!E\left[\int \mid \hat{f}(x) - f(x) \mid dx\right] \tag{13}$$

where $\hat{f}$ is either the smoothed histogram or the gamma kernel estimate, h is either the bandwidth parameter of gamma kernel b or the inverse of the bandwidth parameter of the smoothed histogram. More precisely, we are going to find the asymptotic behavior, the lower and upper bounds of $I\!E\{J_{hn})\}$ for $\hat{f}$. We prove also that there exists a quantity $B_M(f)$, in equation(11), depending only on f such that

$$\inf_{h>0} n^{2/5} I\!E(J_{hn}) \leq C^* B_M(f) + o(1)$$

and

$$\inf_{h>0} n^{2/5} I\!E\left(J_{hn}\right) \geq C B_M(f) + o(1).$$

where $C^* = 5(8\pi)^{-2/5}$ and $C = \inf_{u>0} \dfrac{\psi(u)}{u^{1/5}}$ are universal constants. Note that the slow convergence theorem of Devroye (1983) is a useful tool in the context presented here. Throughout this section, we make the following hypotheses

$$\lim_{n\longrightarrow\infty} h = 0 \quad \text{and} \quad \lim_{n\longrightarrow\infty} n\, h^{1/2} = +\infty. \tag{14}$$

The next result states that for any density f in $C^2([0,M])$ with finite $B_M(f)$, the mean integrated absolute deviation $I\!E(J_{hn})$ decreases at rate $n^{-2/5}$ when h is chosen in proportion to $n^{-2/5}$.

Theorem 3.1 *Let f be a density function in $C^2([0,M])$, $0 < M < \infty$, and let $\hat{f}$ be either the smoothed histogram or the gamma kernel estimates. Then, under* (14)*, the asymptotic behavior of $I\!E\{J_{hn}\}$ is given by*

$$I\!E(J_{hn}) = J(h,n) + o\left(h + n^{-1/2}h^{-1/4}\right) \tag{15}$$

where

$$J(h,n) = \int_0^M (n^{1/2}h^{1/4})^{-1}\,\alpha(x)\,\psi\left((n^{1/2}h^{5/4})\frac{|\,\beta(x)\,|}{\alpha(x)}\right)\,dx. \tag{16}$$

An upper bound of $J(h,n)$ is

$$\sqrt{\frac{2}{\pi}}\int_0^M (n^{1/2}h^{1/4})^{-1}\,\alpha(x)dx + h\int_0^M (\,|\,\beta(x)\,|\,)\,dx. \tag{17}$$

In particular,

$$\limsup_{n\to\infty}\inf_{h>0} n^{2/5} I\!E(J_{hn}) \le C^*\,B_M(f) \tag{18}$$

where $C^* = 5(8\pi)^{-2/5} = 1.3768102\cdots$ is a universal constant. Noting that for any density function f defined on $[0,M]$ such that $B_M(f) < \infty$, the upper bound given by (17) *is not exceeded when h is chosen to be equal to*

$$h_n = \left[\frac{\frac{1}{4}\sqrt{\frac{2}{\pi}}\int_0^M \alpha(x)\,dx}{\int_0^M (|\,\beta(x)\,|)\,dx}\right]^{4/5} n^{-2/5}.$$

A similar result was obtained for the standard kernel estimates by Rosenblatt (1956) and Devroye & Penrod (1984), Tenbusch (1998) for Bernstein estimates and Bouezmarni and Rolin (2002) for Beta kernel estimates. Note that the last two estimators are designed for density functions defined on $[0,1]$.

We define the biais $B(x)$ and the variance $\sigma_n^2(x)$ of $\hat{f}(x)$ at x by

$$B(x) = I\!E\{\hat{f}(x)\} - f(x) \quad \text{and} \quad \sigma_n^2(x) = I\!E\left(\left[\hat{f}(x) - I\!E\{\hat{f}(x)\}\right]^2\right).$$

To prove theorem 3.1, we need the next lemmas concerning the bias and the variance of the gamma kernel and the smoothed histogram estimators. We start by the first lemma which describes the asymptotic behavior of $B(x)$. The proof of this result can be found in Chen (2000) and Gawronski and Stadtmüller (1980).

Lemma 3.1 *For all $f \in C^2([0,M])$, we have*

$$\int_0^M \mid |B(x)| - h|\beta(x)| \mid dx = o(h) \quad as \quad h \longrightarrow 0\,.$$

Hereafter, we give the asymptotic behavior of $\sigma_n(x)$. We examine only the gamma kernel estimator. The proof is almost the same for smoothed histogram.

Lemma 3.2 *For all $f \in C^2([0,M])$, we have*

$$\int_0^M \left|\sigma_n(x) - \alpha(x)\sqrt{n^{-1}h^{-1/2}}\right| dx = o\left(n^{-1/2}h^{-1/4}\right),\ as\ n\sqrt{h} \longrightarrow \infty\,.$$

Proof. Let $\hat{f}$ be the gamma kernel estimator associated to $f \in C^2([0,M])$. From Chen (2000), one can get

$$Var(\hat{f}) = \frac{1}{\sqrt{2\pi}}\frac{n^{-1}b^{-1/2}}{x^{1/2}}f(x) + o(n^{-1}b^{-1/2}) \qquad \text{if} \qquad x \geq b$$

and

$$Var(\hat{f}) = \frac{\Gamma(2x/b+1)}{2^{1+x/b}\Gamma^2(x/b+1)}(nb)^{-1}f(x) + o((nb)^{-1}) \qquad \text{if} \qquad x < b.$$

Choosing $\delta = b^{1-\epsilon}$ with $0 < \epsilon < \frac{3}{4}$, we get then

$$\begin{aligned}\int_0^M \left|\sigma_n(x) - \alpha(x)\sqrt{n^{-1}b^{-1/2}}\right| dx &= \int_0^\delta + \int_\delta^M \left|\sigma_n(x) - \alpha(x)\sqrt{n^{-1}b^{-1/2}}\right| dx \\ &= o\left(n^{-1/2}b^{-1/4}\right).\end{aligned}$$

Finally, we need the useful lemma due to Devroye & Penrod (1984).

Lemma 3.3 *For nonnegative numbers u, v, w, z we have*

$$\left|u\psi\left(\frac{v}{u}\right) - w\psi\left(\frac{z}{w}\right)\right| \leq \mid v - z \mid + \sqrt{\frac{2}{\pi}} \mid u - w \mid .$$

3.1 Proof of Theorem 3.1

We are now in position to prove the theorem 3.1. We start by proving (15). For that, combining (7) and (8) with lemma (8) in Devroye and Gyorfi (1985) and using the result in lemma 3.1 , we get that

$$\left| I\!E \left\{ \mid \hat{f}(x) - f(x) \mid \right\} - \sigma_n(x)\psi \left\{ \frac{\mid B(x) \mid}{\sigma_n(x)} \right\} \right| \leq \frac{c\, h^{-1/2}}{n\sqrt{x}} \tag{19}$$

for some positive constant c. It follows from (13) that

$$\begin{aligned} \mid I\!E(J_{hn}) - J(h,n) \mid \leq & \left| \int_0^M I\!E \mid \hat{f}(x) - f(x) \mid dx - \int_0^M \sigma_n(x)\, \psi \left(\frac{\mid B(x) \mid}{\sigma_n(x)} \right) dx \right| \\ & + \left| \int_0^M \sigma_n(x)\, \psi \left(\frac{\mid B(x) \mid}{\sigma_n(x)} \right) dx - J(h,n) \right|. \end{aligned}$$

Now from (19), we see that

$$\left| \int_0^M I\!E \mid \hat{f}(x) - f(x) \mid dx - \int_0^M \sigma_n(x)\, \psi \left(\frac{\mid B(x) \mid}{\sigma_n(x)} \right) dx \right| \leq c \frac{2\sqrt{M}}{nh^{1/2}}$$

which implies that

$$\mid I\!E(J_{hn}) - J(h,n) \mid \leq c \frac{2\sqrt{M}}{nh^{1/2}} + \left| \int_0^M \sigma_n(x)\, \psi \left(\frac{\mid B(x) \mid}{\sigma_n(x)} \right) dx - J(h,n) \right|$$

Putting $u = \sigma_n(x)$, $v = \mid B(x) \mid$, $w = n^{-1/2} h^{-1/4} \alpha(x)$ and $z = h \mid \beta(x) \mid$, and using Lemma (3.1) and Lemma (3.2), we obtain

$$\int_0^M \mid v - z \mid dx = o(h) \quad \text{and} \quad \int_0^M \mid u - w \mid = o(n^{-1/2} h^{-1/4}).$$

Now, combining the above informations with lemma 3.3, we see that the upper bound on $\mid I\!E(J_{hn}) - J(h,n) \mid$ is then equal to

$$c \frac{\pi}{nh^{1/2}} + \int_0^M \mid\mid B(x) \mid - h \mid \beta(x) \mid\mid dx + \sqrt{\frac{2}{\pi}} \int_0^M \left| \sigma(x) - \alpha(x) \sqrt{n^{-1}\, h^{-1/2}} \right| dx$$

which can be reduced to

$$c \frac{\pi}{nh^{1/2}} + o(h) + o(h^{-1/4} n^{-1/2}) = o(h + h^{-1/4} n^{-1/2}).$$

Therefore,

$$I\!E(J_{hn}) = J(h,n) + o(h + h^{-1/4}n^{-1/2}) \quad \text{for all} \quad f \in C^2[0,M].$$

Note that the upper bound (17) is an immediate consequence of (12) and (16). Finally, the upper bound (18) is obtained by taking the value of h which minimizes the upper bound (17). We obtain then

$$J(h,n) \leq C^* \frac{B(f)}{n^{2/5}}$$

where $C^* = (4^{1/5} + 4^{-4/5}) \left(\sqrt{\frac{2}{\pi}}\right)^{4/5} = 5(8\pi)^{-2/5} = 1.3768102\cdots$

■

Hereafter, we derive a lower bound of the mean integrated absolute error $I\!E(J_{hn})$.

Theorem 3.2 *Let f be a density function in $C^2([0,M])$, and $\hat{f}_h$ either the gamma histogram or the gamma kernel estimate. Suppose that*

$$\lim_{n\longrightarrow\infty} h = 0 \quad \textit{and} \quad \lim_{n\longrightarrow\infty} n\, h^{1/2} = +\infty$$

then,

$$\liminf_{n\to\infty} \inf_{h>0} n^{2/5} I\!E\,(J_{hn}) \geq C\, B_M(f)$$

where $C = \inf_{u>0} \dfrac{\psi(u)}{u^{1/5}} = 1.028493\cdots$ is a universal constant.

Proof. We adapt the proof given in Devroye and Gyorfi (1985) for the standard kernel at our context, we can then write from (19)

$$\begin{aligned} \int_0^M I\!E|\hat{f}_h(x) - f(x)|dx &\geq \int_0^M \sigma_n(x)\, \psi\left(\frac{|\, B(x)\,|}{\sigma_n(x)}\right) dx - \int_0^M \frac{c\, h^{-1/2}}{n\sqrt{x}} dx \\ &= \int_0^M \sigma_n(x)\, \psi\left(\frac{|\, B(x)\,|}{\sigma_n(x)}\right) dx - o(n^{-1/2}b^{-1/4}) \quad (20) \end{aligned}$$

Now, let $g(x)$ be the density function defined by

$$g(x) = \frac{\sigma_n(x)}{\int_0^M \sigma_n(x)dx}$$

and put

$$a = \int_0^M \sigma_n(x)dx.$$

We observe then that

$$\int_0^M \sigma_n(x)\,\psi\left(\frac{\mid B(x)\mid}{\sigma_n(x)}\right)dx \;=\; \int_0^M a\psi\left(\frac{\mid B(x)\mid}{\sigma_n(x)}\right)g(x)dx. \qquad (21)$$

Using the fact that the function $a\psi$ is convex, it follows then from Jensen's inequality that

$$\int_0^M a\psi\left(\frac{\mid B(x)\mid}{\sigma_n(x)}\right)g(x)dx \;\geq\; \int_0^M \sigma_n(x)\,\psi\left(\frac{\int_0^M \mid B(x)\mid dx}{\int_0^M \sigma_n(x)dx}\right)dx. \quad (22)$$

Now from (12), lemma 3.1 and lemma 3.2 , and using the fact that $C = \inf_u \frac{\psi(u)}{u^{1/5}}$, we observe that

$$\begin{aligned}\int_0^M \sigma_n(x)\,\psi\left(\frac{\int_0^M \mid B(x)\mid dx}{\int_0^M \sigma_n(x)dx}\right)dx &\geq C\left(\int_0^M \sigma_n(x)dx\right)^{4/5}\left(\int_0^M \mid B(x)\mid dx\right)^{1/5}\\ &\geq n^{-2/5}CB_M(f). \qquad (23)\end{aligned}$$

Finally the result is deduced from (20), (21), (22) and (23).

4 L_1- rates of convergence of the generalized kernel estimate

Most of the density estimators are a special case of the generalized kernel density estimator discussed by Foldes and Revesz (1974), and Walter and Blum (1979). This estimator is expressed by

$$\hat{f}_h(x) = \frac{1}{n}\sum_{i=1}^{n} K\left(x, h, X_i\right).$$

The aim of this section is to obtain the asymptotic of the expected average absolute error of the generalized kernel density estimator. The results found by Devroye and Gyorfi (1985) for the standard kernel estimate, by Tenbusch (1998) for the Bernstein estimator and by Bouezmarni and Rolin(2003) for

beta kernel are special cases of the results found here. We discuss also as a special case, the reflection method (Schuster(1985)), the local linear estimate Lejeune and Sarda(1992) and the Nonnegative boundary correction Jones and Foster (1996).

The following conditions will be necessary.

Conditions A:

(A.1) $\sup_t K(x,h,t) \leq h^{-\psi_1} g_1(x), \psi_1 > 0.$

(A.2) $\int_a^b \left| B_h(x) - h^{\psi_2} g_2(x) \right| dx = o(n^{-1}h^{-\psi_1}),\ \psi_2 > 0.$

(A.3) $\int_a^b \left| \sigma_h(x) - h^{-\psi_3/2} g_3(x) \right| = o(h^{-\psi_3/2}),\ 0 < \psi_3 < 2\psi_1.$

(A.4) $\lim h^{\psi_2} = 0$ and $\lim n^{-1}h^{-\psi_3} = 0.$

(A.5) $g_1,\ g_2$ and $g_3 \in L_1([a,b])$.

where

$$B_h(x) = \int_0^\infty \{f(t) - f(x)\} K(x,h,t)\, dt$$

and

$$\sigma_h(x) = \left(\int_0^\infty K^2(x,h,t) f(t)\, dt - \left(\int_0^\infty K(x,h,t) f(t)\, dt \right)^2 \right)^{1/2}$$

4.1 Main Result of the Section

Hereafter, we derive the asymptotic behavior of the expected average absolute error of the generalized kernel density estimator.

Theorem 4.1 *Let f be a density function in $C^2([a,b])$, and let $\hat{f}_h$ be its generalized kernel density estimator. Under conditions A, we have*

$$I\!E \left[\int_a^b \mid \hat{f}_h(x) - f(x) \mid dx \right] = J(h,n) + o(h^{\psi_2} + n^{-1/2}h^{-\psi_3/2})$$

where

$$J(h,n) = \int_a^b n^{-1/2} h^{-\psi_3/2}\, g_3(x)\, \psi \left(n^{1/2} h^{(\psi_3+2\psi_2)/2} \frac{\mid g_2(x) \mid}{g_3(x)} \right) dx.$$

In particular, we have

$$\limsup_{n\to\infty} \inf_{h>0} n^{1/(2\psi_2+\psi_3)} I\!E(J_{hn}) \leq C^* \, B(f)$$

where

$$C^* = \left\{ \left(\frac{\psi_3}{2\psi_2} \right)^{\frac{2\psi_2}{2\psi_2+\psi_3}} + \left(\frac{\psi_3}{2\psi_2} \right)^{\frac{-\psi_3}{2\psi_2+\psi_3}} \right\} \left(\frac{2}{\pi} \right)^{\frac{\psi_2}{2\psi_2+\psi_3}}$$

and

$$B(f) = \left(\int_a^b g_3(x)dx \right)^{\frac{2\psi_2}{2\psi_2+\psi_3}} \left(\int_a^b \mid g_2(x) \mid dx \right)^{\frac{\psi_3}{2\psi_2+\psi_3}}.$$

The upper bound is not exceeded for the following choice of h, when f is a density on $[a,b]$ and $B(f) < \infty$

$$h = \left(\frac{\psi_3 \int_a^b g_3(x)dx}{2\psi_2 \int_a^b |g_2(x)|dx} \right)^{2/(\psi_3+2\psi_2)} n^{-1/(\psi_3+2\psi_2)}.$$

The next theorem gives a lower bound of the asymptotic mean absolute error of the generalized kernel density.

Theorem 4.2 *Let f be a density function in $C^2([a,b])$, and $\hat{f}_h$ an estimate of f. Under conditions A, we have*

$$\liminf_{n\to\infty} \inf_{h>0} n^{\frac{\psi_2}{2\psi_2+\psi_3}} I\!E \left(\int \mid \hat{f}_h(x) - f(x) \mid dx \right) \geq C \, B_M(f)$$

where

$$C = \inf_u \frac{\psi(u)}{u^{\psi_3/(2\psi_2+\psi_3)}}.$$

The proofs of theorems 4.1 and 4.2 are similar to those of the theorems 3.1 and 3.2. The start point is the inequality

$$\left| I\!E \left\{ \mid \hat{f}_h(x) - f(x) \mid \right\} - n^{-1/2}\sigma_h(x)\psi \left\{ \frac{\mid B_h(x) \mid}{n^{-1/2}\sigma_h(x)} \right\} \right| \leq c_1 \, h^{-\psi_1} n^{-1} g_1(x) \quad (24)$$

which is a consequence of lemma (8) in Devroye and Gyorfi (1985) and condition (A.1).

Example 1 *For $f \in C^2([a,b])$, conditions A are satisfied for the standard kernel with a symmetric and bounded kernel. In fact, it suffices the take:*

$$\psi_1 = \psi_3 = 1, \quad \psi_2 = 2, \quad g_2(x) = \frac{1}{2}f''(x) \quad \textit{and} \quad g_3(x) = f(x).$$

Consequently, we obtain

$$B(f) = \left(\int_a^b \sqrt{f(x)}dx\right)^{\frac{4}{5}} \left(\frac{1}{2}\int_a^b \mid f''(x) \mid dx\right)^{\frac{1}{5}}$$

and,

$$C^* = 5(8\pi)^{-2/5} = 1.3768102\cdots \quad \textit{and} \quad C = 1.028493\cdots$$

Example 2 *For $f \in C^1([a,b])$, conditions A are satisfied for the histogram estimate. For that, we choose*

$$\psi_1 = \psi_3 = \psi_2 = 1, \quad \int g_2(x) = \frac{1}{4}\int f'(x) \quad \textit{and} \quad g_3(x) = f(x).$$

It follows that

$$B(f) = \left(\int_a^b \sqrt{f(x)}dx\right)^{\frac{2}{3}} \left(\frac{1}{4}\int_a^b \mid f'(x) \mid dx\right)^{\frac{1}{3}}$$

and,

$$C^* = \left(\frac{27}{2\pi}\right)^{1/3} = 1.625778 \quad \textit{and} \quad C = 1.290381\cdots\cdots$$

4.2 Formulas of $B(f)$ for some density function estimates

The table bellow gives the formulas of $B(f)$ for the local linear estimates obtained by Lejeune and Sarda (1992), the nonnegative boundary correction examined by Jones and Foster (1996), the gamma histogram and the gamma kernel estimate studied in details in this paper. We need the following quantities $p = xh$,

$$B(p) = \frac{\mu_2(p)\nu_1(p) - \mu_1(p)\nu_2(p)}{\mu_0(p)\nu_1(p) - \mu_1(p)\nu_0(p)}$$

and

$$V(p) = \frac{\nu_1(p)^2 b(p) - \nu_1(p)\mu_1(p)e(p) + \mu_1^2(p)g(p)}{\{\mu_0(p)\nu_1(p) - \mu_1(p)\nu_0(p)\}^2}$$

where for $k = 0, 1, 2$

$$\mu_k = \int t^k K(t)dt, \quad \nu_k = \int t^{k+1} K(t)dt, \quad b(p) = \int_{-1}^{\min(p,1)} K^2(t)dt,$$

$$e(p) = \int_{-1}^{\min(p,1)} t\, K^2(t)dt \quad \text{and} \quad g(p) = \int_{-1}^{\min(p,1)} t^2\, K^2(t)dt.$$

Estimates	$B(f)$
local linear estimates	$\left(\int_0^M \sqrt{V(p)f(x)}dx\right)^{4/5} \left(\int_0^M \left\lvert\frac{1}{2}(B(p)f''(x))\right\rvert dx\right)^{1/5}$
Nonnegative correction	$\left(\int_0^M \sqrt{V(p)f(x)}dx\right)^{4/5} \left(\int_0^M \left\lvert\frac{1}{2}(B(p)f''(x)) + \frac{\mu_1^2(p)}{\mu_0^2(p)}\frac{f'(x)^2}{f(x)}\right\rvert dx\right)^{1/5}$
Gamma histogram	$\left(\int_0^M \sqrt{\frac{f(x)}{2\sqrt{\pi x}}}dx\right)^{4/5} \left(\int_0^M \left\lvert\frac{1}{2}(f'(x) + xf''(x))\right\rvert dx\right)^{1/5}$
Gamma kernel	$\left(\int_0^M \sqrt{\frac{f(x)}{2\sqrt{\pi x}}}dx\right)^{4/5} \left(\int_0^M \left\lvert f'(x) + \frac{1}{2}xf''(x)\right\rvert dx\right)^{1/5}$

Table 1: formula of $B(f)$ for some density function estimates.

5 Conclusion

The optimal rate of convergence for the expected absolute error for the gamma kernel and the smoothed histogram have been obtained. The lower and upper bound of this error were also discussed. The similar results have been obtained, under some conditions, for the generalized kernel density estimator. The results given by Devroye and Gyorfi (1985) for the standard kernel estimate, by Tenbusch (1998) for the Bernstein estimator and by Bouezmarni and Rolin (2003) for beta kernel can be viewed as a special cases of the results concerning the generalized kernel density estimator considered her.

References

Bouezmarni, T. and Rolin, J. M. (2003). *Consistency of Beta kernel Density Function Estimator.* The Canadian Journal of Statistics, Vol. 31, No 1, pp. 89-98.

Bouezmarni, T. and Scaillet , O. (2003). *Consistency of Asymmetric Kernel Density Estimators and Smoothed Histograms with Application to Income Data.* Institute de Statistique, Université Catholique de Louvain, mimeo#0306.

Bouezmarni, T. and Rolin, J. M. (2003). *Bernstein Estimator For Unbounded Density Function.* Institute de Statistique, Université Catholique de Louvain, mimeo#0308.

Chen, S. X. (1999). *A Beta Kernel Estimator for Density Functions.* Comp. Stat. and Data anal., **31**, 131-145.

Chen, S. X. (2000). *Probability Density Functions Estimation Using Gamma kernels.* Ann. Inst. Statis. Math., **52**, 471-480.

Devroye, L. (1997). *Universal Smoothing Factor Selection in Density Estimation: Theory and Practice.* Test, **6**, 223-320.

Devroye, L. and Gyorfi, L. (1985). *Nonparametric Density Estimation: The L_1-View.* Wiley, New York.

Devroye, L. and Lugosi, G. (1996). *A Universally Acceptable Smoothing Factor for Kernel Density Estimates.* Ann. of Statist., **24**, 2499-2512.

Foldes, A. and Revesz, P. (1974). *A General Method for Density Estimation.* Sc. Math. Hung., **7**, 90-94.

Gawronski, N. and Stadtmüller, U. (1980). *On Density Estimation by Means of Poisson's Distribution.* Scand. J. Statist., **7**, 90-94.

Gawronski, N. and Stadtmüller, U. (1981). *Smoothing histograms by means of lattice and continuous distributions.* Metrika, **28**, 155-164.

Hall, P. and Wand, M. (1988). *Minimizing L1 Distance in Nonparametric Density Estimation.* J. of Mult. Anal., **26**, 59-88.

Jones, M. (1993). *Simple Boundary Correction For Kernel Density Estimation.* Statist. Comp., **3**, 135-146.

Jones, M. and Foster, P. (1996). *A Simple Nonnegative Boundary Correction Method For Kernel Density Estimation.* Statist. Sin., **6**, 1005-1013.

Marron, J. S. and Ruppert, D. (1994). *Transformations to reduce boundary bias in kernel density estimation.* J. Roy. Statist. Soc. Ser. B **56** , 653-671.

Massart, P. (1990). *The Tight Constant in Dvoretzky-Kiefer-Woolfowitz Inequality.* Ann. of Prob., **18**, 1269-1283.

Müller, H. (1991). *Smooth Optimum Kernel Estimators Near Endpoints.* Biometrika, **78**, 521-530.

Rice, J. (1984). *Boundary Modification for Kernel Regression.* Comm. Statist. Theor. Meth., **13**, 893-900.

Stadtmüller, U. (1983). *Asymptotic Distribution of Smoothed Histograms.* Metrika, **30**, 145-158.

A. Tenbusch (1998). *Nonparametric Curve Estimation With Bernstein Polynomial.* Universitätsverlag Rasch, Osnabrück.