ON THE EFFECT OF DENSITY SHAPE ON THE PERFORMANCE OF ITS KERNEL ESTIMATE

Luc Devroye
School of Computer Science
McGill University

Matthew P. Wand
Australian Graduate School of Management
University of New South Wales

ABSTRACT. We introduce a universal functional $Q(f)$ which measures the difficulty a given density f poses to the standard kernel density estimate if one uses the optimal smoothing factor. The functional is well-defined (but possibly infinite) for all densities, regardless of their smoothness or tail properties. It is proportional to the limit of $n^{2/5}\mathbf{E}\int|f_n - f|$ where f_n is the optimal kernel estimate. This paper settles some questions left unanswered in Devroye and Györfi (1985) and Hall and Wand (1988).

1991 MATHEMATICS SUBJECT CLASSIFICATIONS: Primary: 62G07. Secondary: 62G05, 62F12, 60F25.

Authors' addresses: Luc Devroye, School of Computer Science, McGill University, 3480 University Street, Montreal, Canada H3A 2A7. Matthew P. Wand, Australian Graduate School of Management, University of New South Wales, P.O.Box 1, Kensington, NSW 2033, Australia. Luc Devroye's research was sponsored by NSERC Grant A3456 and FCAR Grant 90-ER-0291.

1. Introduction.

Let $X_1, \ldots, X_n$ be i.i.d. random variables with common density f on the real line. We consider the kernel estimate

$$f_n(x) = \frac{1}{n} \sum_{i=1}^{n} K_h(x - X_i)$$

where $K_h(x) = (1/h)K(x/h)$, $h > 0$ is the smoothing factor depending upon n only, K, the *kernel*, is a given density symmetric about zero and satisfying $\int K^2 < \infty$, $\int x^2 K(x)\,dx < \infty$ (Akaike, 1954; Rosenblatt, 1956; Parzen, 1962). Sometimes we will write f_{nh} to make the dependence upon h explicit. The L_1 error given by

$$J_{nh} = \int |f_{nh} - f|$$

measures in many situations the quality of the estimate f_n. The expected L_1 error $\mathbf{E}J_{nh}$ is a function of n, f, h and K. Of these factors, the user can only choose K and h. The choices of h and K have led to extensive discussions, especially data-dependent choices for h for fixed K. The question addressed here is: which density asymptotically minimizes

$$\inf_h \mathbf{E}J_{nh} \ ?$$

In addition, we would like to see how f influences the asymptotic behavior of this performance measure. Note that $\inf_h \mathbf{E}J_{nh}$ can be considered as a measure of the difficulty of estimating f by the kernel estimate. If we have more information about this, then we could apply nonlinear transformations to the data as in Devroye and Györfi (1985) and try to knead the data into a cloud that has a density close to the optimum. This question has been addressed to some extent in the book of Devroye and Györfi (1985) and in Wand, Marron and Ruppert (1991) and Ruppert and Cline (1993). Devroye and Györfi showed that the isosceles triangular density minimizes both an upper bound and a lower bound for $\inf_h \mathbf{E}J_{nh}$. It turns out however that the isosceles triangular density does not minimize the $\inf_h \mathbf{E}J_{nh}$ asymptotically. The optimal density is the solution of a difficult differential equation. We derive precise asymptotic expressions for $\inf_h \mathbf{E}J_{nh}$ that are valid for all densities. This forces us to replace derivatives (which may not exist) by generalized derivatives, thus rendering the derivation more involved.

2. Main result.

Devroye and Györfi (1985) considered the following subclass of densities: $\mathcal{F}$ is the class of all densities f with compact support, such that f is absolutely continuous, f' is absolutely continuous and there exists a version of f'' that is bounded and continuous on the real line. Define

$$\alpha = \sqrt{\int K^2}\ ,\ \beta = \int x^2 K(x)\,dx$$

and

$$A(K) = \alpha^{4/5}\beta^{1/5}\ .$$

We also introduce the function $\psi(u) \stackrel{\text{def}}{=} \mathbf{E}|N-u|$, where N is a normal $(0,1)$ random variable.

LEMMA 1 (DEVROYE AND GYÖRFI, 1985). *If $f \in \mathcal{F}$ and*

$$\lim_{n\to\infty} h = 0\ ,\ \lim_{n\to\infty} nh = \infty\ ,$$

then

$$\left|\mathbf{E}J_{nh} - \alpha \int \sqrt{\frac{f}{nh}}\psi\left(\sqrt{nh^5}\frac{\beta|f''|}{2\alpha\sqrt{f}}\right)\right| \le o(h^2) + o(1/\sqrt{nh})\ .$$

As noted by Hall and Wand (1988), this implies the following.

LEMMA 2 (HALL AND WAND, 1988). *For $f \in \mathcal{F}$,*

$$n^{2/5}\inf_h \mathbf{E}J_{nh} \to 2^{-1/5}A(K)Q(f)\ ,$$

where

$$Q(f) \stackrel{\text{def}}{=} \inf_{u>0}\int \frac{\sqrt{f}}{u^{1/5}}\psi\left(\frac{u|f''|}{\sqrt{f}}\right)\ .$$

One of the goals of this paper is to derive a generalization of this result that is valid even if $f \notin \mathcal{F}$, e.g., when f is the isosceles triangular density or the Laplace density.

As we know from Devroye and Györfi (1985), some densities are such that an $O(n^{-2/5})$ rate of convergence is unattainable. Basically, these include all unsmooth densities — even one simple discontinuity suffices to put f in this class; it even contains continuous densities with an infinite first derivative at some point, such as the beta (a,a) density with $1 < a < 2$ — and all densities with a heavy tail, such as the Cauchy density and, in fact, all densities for which $\int\sqrt{f} = \infty$. We first consider densities that can insure a convergence rate of $O(n^{-2/5})$. To describe the main result, it is necessary to generalize the notion of a second derivative, without which Lemmas 1 and 2 cannot be extended. We let φ be a mollifier, i.e., a density with support on $[-1,1]$, having four continuous derivatives. If $*$ denotes the convolution operator, $f * \varphi_a$ is a density with all the smoothness properties required to apply Lemmas 1 and 2. Furthermore, as $a \to 0$, $f * \varphi_a \to f$ at almost all x. These two observations are at the basis of our generalization for $Q(f)$:

$$Q(f) \stackrel{\text{def}}{=} \inf_{u>0} \limsup_{a \to 0} \int \frac{\sqrt{f}}{u^{1/5}} \psi \left(\frac{|(f * \varphi_a)''| u}{\sqrt{f}} \right) .$$

For $f \in \mathcal{F}$, this definition of $Q(f)$ coincides with the definition of Lemma 2. In particular, the definition does not depend upon the choice of the mollifier φ. This property is also true for most densities outside $\mathcal{F}$. We will define a class $\mathcal{H}$ containing $\mathcal{F}$, for which, among other things, $Q(f)$ does not depend upon the choice of the mollifier. The idea of a mollifier goes back to Devroye and Györfi (1985) and Devroye (1987). Other equivalent generalizations of derivatives were pointed out by Mokkadem (1989a,b) and Karunamuni and Mehra (1988). For multidimensional studies of the L_1 error of the kernel density estimate, we refer to Holmström and Klemelä (1992).

The class $\mathcal{H}$ contains all densities f for which

$$\sup_{a>0} \int |(f * \varphi_a)''| < \infty$$

and for which the following condition (1) holds: there exists a cup-shaped unimodal symmetric positive function H satisfying

$$\int \frac{1}{H(x)} \, dx < \infty \, , \; \mathbf{E} H(X-a) < \infty \, , \mathbf{E} H(X+a) < \infty \tag{1}$$

for some $a > 0$, where X is a random variable with density f. (A function H is cup-shaped if there exists a point θ such that H is nonincreasing on $(-\infty, \theta]$ and nondecreasing on (θ, ∞).)

Remark 1. Sufficient condition for $\mathcal{H}$. (1) is satisfied when

$$\mathbf{E}|X| \log^{1+\epsilon}(1+|X|) < \infty$$

for some $\epsilon > 0$, or when the $1+\epsilon$-th moment of X exists. (Hint: take $H(x) = |x| \log^{1+\epsilon}(1+|x|)$.) $\square$

REMARK 2. A NECESSARY AND SUFFICIENT CONDITION FOR MOST DENSITIES. When f can be written as a finite mixture of unimodal densities (and most densities appearing in applications can), then condition (1) holds if and only if $\int \sqrt{f} < \infty$. We prove this for symmetric unimodal f, leaving the general case to the reader: first, (1) implies $\int \sqrt{f} < \infty$ in all cases, by the Cauchy-Schwarz inequality:

$$\int \sqrt{f} = \int \frac{\sqrt{fH}}{\sqrt{H}} \leq \left(\int fH \int (1/H) \right)^{1/2} < \infty \ .$$

Next, when $\int \sqrt{f} < \infty$, choose

$$H(x) = \begin{cases} 1/\sqrt{f(a)} & \text{if } |x| < 2a \\ 1/\sqrt{f(|x|-a)} & \text{if } |x| \geq 2a \end{cases}$$

when verifying (1). □

THEOREM 1. *Let $f \in \mathcal{H}$, and assume that K is a kernel with compact support. Then*

$$n^{2/5} \inf_{h>0} \mathbf{E} \int |f_{nh} - f| \to 2^{-1/5} A(K) Q(f) \ .$$

Furthermore, if $nh^5 \to \gamma > 0$,

$$n^{2/5} \mathbf{E} \int |f_{nh} - f| \to \limsup_{a \to 0} \int \frac{\alpha \sqrt{f}}{\gamma^{1/10}} \psi \left(\frac{|(f * \varphi_a)''| \beta \sqrt{\gamma}}{2\alpha \sqrt{f}} \right) \ .$$

THEOREM 2. *Consider a density f for which $\int \sqrt{f} = \infty$ or*

$$\limsup_{a \to 0} \int |(f * \varphi_a)''| = \infty \ .$$

Then $Q(f) = \infty$, so that the statement of Theorem 1 remains formally valid:

$$\lim_{n \to \infty} n^{2/5} \inf_{h>0} \mathbf{E} J_{nh} = \infty \ .$$

PROOF. Direct from Lemma 6 below. This result was already noted by Devroye and Györfi (1985), and we refer to the discussion given there. □

There are a few densities for which $\int \sqrt{f} < \infty$, yet condition (1) is not satisfied. Such densities are not covered by the results of this paper. We should point out that this class only contains pathological cases. To see this, we apply remark 2 above, and obtain without work:

THEOREM 3. *Let f be a density that can be written as a finite mixture of unimodal densities. Then*

$$n^{2/5} \inf_{h>0} \mathbf{E} \int |f_{nh} - f| \to 2^{-1/5} A(K) Q(f) \ .$$

Furthermore, $Q(f) < \infty$ if and only if $\int \sqrt{f} < \infty$ and

$$\limsup_{a \to 0} \int |(f * \varphi_a)''| < \infty \ .$$

For the vast class of densities covered by Theorem 3, we have a complete answer to the problem we set out to solve. In particular, Theorem 3 gives necessary and sufficient conditions for the rate of convergence to be $O(n^{-2/5})$. Theorem 3 reaffirms that a poor rate of convergence is due to one of two factors: either the distribution is too spread out ($\int \sqrt{f} = \infty$) or not smooth enough ($\limsup_{a \to 0} \int |(f * \varphi_a)''| = \infty$).

REMARK 3. UNSMOOTH DENSITIES. When $\limsup_{a \to 0} \int |(f * \varphi_a)''| = \infty$, one may still be able to obtain rates of convergence for the kernel estimate. For example, if $\limsup_{a \to 0} \int |(f * \varphi_a)'| < \infty$, we get at least a rate of the order of $n^{-1/3}$. Unfortunately, a simple general theory is impossible to get because the rate of convergence depends upon the (lack of) smoothness of f. A similar phenomenon occurs for densities with large tails. To put it more succinctly, every beta (a, a) density with $a < 2$ induces a different rate of convergence, as does every Student t density with parameter $\nu \leq 1$. However, by the results of this paper, for $a \geq 2$ and $\nu > 1$, the rates are all $O(n^{-2/5})$.

3. Another measure of difficulty.

The difficulty related to a density f on $\mathbb{R}$ for the kernel estimate is measured by

$$Q(f) = \inf_{u>0} \int \frac{\sqrt{f}}{u^{1/5}} \psi \left(\frac{|f''| u}{\sqrt{f}} \right)$$

whenever f is absolutely continuous, f' is absolutely continuous, f'' is absolutely integrable, and $\int \sqrt{f} < \infty$. The question that imposes itself is: which density minimizes $Q(f)$? In what follows, we concentrate on the family $\mathcal{F}$ of all densities for which $\int \sqrt{f} < \infty$, f is absolutely continuous, f' is absolutely continuous, and f'' is absolutely integrable.

The following properties of ψ help us in the ensuing analysis. We refer to Devroye and Györfi (1985, p. 77) for details.

LEMMA 3 (THE FUNCTION ψ). *Let N be a standard normal random variable.*

- $\psi(u) \geq \mathbf{E}|N| = \sqrt{2/\pi}$.
- $\psi(u) \geq |\mathbf{E}N - u| = |u|$.
- $\psi(u) \leq \mathbf{E}|N| + |u| = \sqrt{2/\pi} + |u|$.
- $\lim_{u \downarrow 0} \psi(u) = \sqrt{2/\pi}$.
- $|\psi(u) - \psi(v)| \leq |u - v|$.
- *For $u > 0$, $\psi'(u) = \mathbf{P}\{|N| \leq u\} > 0$.*
- *ψ is convex and symmetric about zero.*
- $|u\psi(v/u) - u^*\psi(v^*/u^*)| \leq |v - v^*| + \sqrt{2/\pi}|u - u^*|$.
- *If $\xi(u) = \psi(u) - |u|$, then ξ is monotonically decreasing on $[0, \infty)$ from $\sqrt{2/\pi}$ to 0.*
- $\inf_{u>0} \psi(u)/u^{1/5} \stackrel{\text{def}}{=} \gamma = 1.028493\ldots$.

Lemma 3 immediately yields the following estimates:

LEMMA 4 (BOUNDS). *For $f \in \mathcal{F}$, if*

$$B(f) = \int^{4/5} \sqrt{f} \int^{1/5} |f''| ,$$

then

$$\gamma \leq \frac{Q(f)}{B(f)} \leq 5(8\pi)^{-2/5} = 1.3768102\ldots .$$

PROOF. Clearly, for fixed $u > 0$,

$$\begin{aligned}
&\int \frac{\sqrt{f}}{u^{1/5}} \psi\left(\frac{|f''|u}{\sqrt{f}}\right) \\
&\qquad \leq \sqrt{\frac{2}{\pi}} u^{-1/5} \int \sqrt{f} + u^{4/5} \int |f''| \\
&\qquad \leq 5(8\pi)^{-2/5} \int^{4/5} \sqrt{f} \int^{1/5} |f''| .
\end{aligned}$$

Also,

$$\begin{aligned}
&\int \frac{\sqrt{f}}{u^{1/5}} \psi\left(\frac{|f''|u}{\sqrt{f}}\right) \\
&\qquad \geq \max\left(\sqrt{\frac{2}{\pi}} u^{-1/5} \int \sqrt{f}, u^{4/5} \int |f''|\right) \\
&\qquad \geq (2/\pi)^{2/5} \int^{4/5} \sqrt{f} \int^{1/5} |f''| .
\end{aligned}$$

This bound can easily be eclipsed by using Jensen's inequality. Again for fixed u, note that by the convexity of ψ,

$$
\begin{aligned}
&\inf_{u>0} \int \frac{\sqrt{f}}{u^{1/5}} \psi\left(\frac{|f''|u}{\sqrt{f}}\right) \\
&\qquad \geq \inf_{u>0} \int \frac{\sqrt{f}}{u^{1/5}} \psi\left(\frac{\int |f''|u}{\int \sqrt{f}}\right) \\
&\qquad = \int^{4/5} \sqrt{f} \int^{1/5} |f''| \ \times \inf_{u>0} u^{-1/5} \psi(u) \\
&\qquad = \gamma \int^{4/5} \sqrt{f} \int^{1/5} |f''| \ .
\end{aligned}
$$

This concludes the proof of Lemma 4. □

Lemma 4 sets the stage for our problem. Indeed, it is relatively easy to minimize $B(f)$: it is minimized in $\mathcal{F}$ by a sequence of densities best described by $f_t * \varphi_a$ as $a \downarrow 0$, where f_t is the isosceles triangular density (Devroye and Györfi, 1985). In fact, with our generalized definition of a second derivative, f_t is the minimizer; note that $f_t \notin \mathcal{F}$. It is interesting that f_t minimizes both the upper and lower bound of Lemma 4, yet it does not minimize $Q(f)$. The reason is that at f_t, we attain the upper bound. More formally, for any piecewise linear density with a finite number of pieces, we have

$$
Q(f) = 5(8\pi)^{-2/5} B(f) \ ,
$$

where $Q(f)$ is defined above and

$$
B(f) = \left(\int \sqrt{f}\right)^{4/5} \left(\limsup_{a \to 0} \int |(f * \varphi_a)''|\right)^{1/5}
$$

generalizes the definition of B given above. The upper bound of Lemma 4 is reached on this family. At the other side, we will exhibit a density for which the lower bound of Lemma 4 is attained as well.

The interest in $B(f)$ is not only historical—we are also driven by practical considerations: $B(f)$ is much easier to compute and estimate than $Q(f)$. This may be important when one designs a data-based method for choosing the smoothing factor. Estimating the two integrals in the definition of $B(f)$ should be straightforward. The minimization in $Q(f)$ is problematic. Furthermore, the close relationship between $B(f)$ and $Q(f)$ tells us that minimizing one would be almost equivalent to minimizing the other. In the next section, we explore this relationship in more detail.

4. Hunting for the best density.

It helps to visualize each density in the plane with coordinates $(B(f), Q(f))$. We note that the bounds given above, which remain valid with generalized definitions of second derivatives as well, cut out an infinite linear wedge in the plane. Furthermore, the absolute lower bound

$$B(f) \geq \frac{4}{3^{4/5}} = 1.660\ldots$$

shown in Devroye and Györfi (1985) limits this wedge from the left. To illustrate things, we show many densities in this plane. Densities with finite coordinates necessarily have small tails and bounded (but possibly discontinuous) first derivatives. For more stable computations, we employed the relationship

$$Q(f) = \inf_{u>0} \left(u^{4/5} \int |f''| + \int \frac{\sqrt{f}}{u^{1/5}} \xi \left(\frac{|f''|u}{\sqrt{f}} \right) \right) .$$

For densities with f' monotonically increasing to a peak, then decreasing to a low, and then increasing again as $x \to \infty$, with a finite number of discontinuities allowed, we may replace $\int |f''|$ in the above formula by $4 \sup |f'|$. This follows from the results of this paper, and is one of the reasons the results are so useful. Also, if f' is absolutely continuous on a finite number of open intervals (which is the case for all densities considered here), then the formula for $Q(f)$ remains valid as well, with the understanding that the first term is replaced by $u^{4/5}(\int |f''| + J(f))$, where $J(f)$ is the sum of the absolute values of the jumps at the points of discontinuity of f'. This too is a simple consequence of Theorem 1.

Put Figure 1 about here.

The figure shown above is geometrically exact—it was drawn in POSTSCRIPT from precise computations. It requires a considerable amount of computations, as each point represents the result of an optimization with respect to u of an integral, whose integrand ψ in turn is an integral of a non-standard function. This was repeated for thousands of densities. The idea is to identify densities with small values for $Q(f)$ that we may then take as target densities for possible transformations in the transformed kernel estimate. For more work on transformed kernel estimates, we refer to Devroye and Györfi (1985), Ruppert and Wand (1992), Ruppert and Cline (1993), and Wand, Marron and Ruppert (1991). At the same time, we would like this target density to be simple in form, for otherwise it would not be attractive to users. A list of densities and families of densities is listed below.

- THE NORMAL DENSITY. We computed $B(f)$ and $Q(f)$ for the exponential power densities given by

$$f(x) = \frac{e^{-|x|^a}}{2\Gamma(1+1/a)} \ ,$$

where $a \geq 1$ to have a finite value for $B(f)$. The family is sometimes attributed to Subbotin. Prominent members include the Laplace density ($a = 1$) and the normal density ($a = 2$). We note that for the normal density,

$$B(f) = \left(\frac{512\pi}{e}\right)^{1/10} = 1.893\ldots \ , \ Q(f) = 1.986\ldots \ .$$

Even within this small family, the normal density is neither optimal with respect to B nor Q. The optimal $B(f)$ is obtained for $a \approx 3.01$, while the optimal $Q(f) \approx 1.95$ is obtained for $a \approx 2.45$. The family forms a lasso in the (B, Q) plane, with the Laplace density at one end. Normal transformations as advocated in Ruppert and Cline (1993) should nevertheless provide us with expedient kernel estimates.

- THE BETA FAMILY. Terrell (1990) showed that the beta (4,4) density minimizes the properly standardized expected L_2 error for the kernel estimate. The led us to look at the symmetric beta (a, a) densities. To have finite $B(f)$, we need $a \geq 2$. The limit as $a \to \infty$ is the normal density, for which we already had a good value for $Q(f)$. All the beta densities in our family have smaller values for B, and most have smaller values for Q, with the minimal $Q(f) \approx 1.9235$ obtained at $a \approx 5.4$. Note that Terrell's optimum density for L_2 is no longer optimal here. At the other extreme of the family is beta $(2, 2)$ density $f(x) = 6x(1-x)$, which is but an affine transformation of Epanechnikov's kernel. This yields the value $B(f) = 1.830419599\ldots$. The value $Q(f) \approx 2.2294$ is the worst in the symmetric beta family with finite $Q(f)$ values, despite the fact that f is similar in form to the best positive kernel K in the kernel estimate. The quartic density (a linearly transformed beta $(3, 3)$ density) is given by $f(x) = (15/16)(1-x^2)_+^2$ on $[-1, 1]$. We have $\int \sqrt{f} = \sqrt{15}/3$,

$\int |f''| = 10/\sqrt{3}$, and thus $B(f) = 5^{3/5}2^{1/5}3^{-1/2} = 1.741916666\ldots$. We have seen that Q is not minimal here.

• Powers of a triangle. Consider the family of densities

$$f(x) = \frac{\alpha+1}{2}(1-|x|)_+^\alpha$$

for $\alpha \geq 1$. For $\alpha < 1$, we have $B(f) = \infty$. At $\alpha = 1$, we obtain the isosceles triangular density, for which $B(f)$ is minimal. As $\alpha \to \infty$, the densities become more and more peaked, and B increases monotonically. The optimal value of Q is obtained for $\alpha \approx 1.14$. By and large, this family is uninteresting. Compare however the isosceles triangle with the Laplace density. The discontinuous derivative of f at the origin is the main contributor to unsmoothness in both cases. But the large tails of the Laplace density have a disastrous effect on the value of $B(f)$ and $Q(f)$.

• The trapezoidal family. Let f be trapezoidal in form on $[-1,1]$, with flat part on $[-a,a]$, where $a \in [0,1)$ is a parameter. As we noted, within this family, we must attain the upper bound for B/Q at all points. The best density, minimizing both Q and B, is the isosceles triangle. As $a \to 1$, we move closer to the uniform $[0,1]$ density, for which both B and Q are infinite.

• The cosine family. In our search for the best density, we came across the cosine family with densities given by

$$f(x) = \frac{\Gamma^2(1+a/2)2^a(a+1)\cos^a(x)}{\Gamma(a+2)\pi},\ |x| \leq \frac{\pi}{2}\ .$$

Interestingly, the curve carved out in the (B,Q) plane lies underneath and to the left of that of the beta family. In the limited number of examples described here, this family contains the best density thus far, at $a \approx 1.7962$. The minimal recorded $Q(f)$ is ≈ 1.9231.

• The Sakai density. The inequality $Q(f) \geq \gamma B(f)$ was obtained by appealing to Jensen's inequality. As is well-known, equality is reached if the integrand takes a constant value. In our case, this would occur when $|f''|/\sqrt{f}$ is held fixed. Among the symmetric unimodal densities, this leads to a unique density modulo a linear transformation. This density is symmetric about 0 and vanishes off $[-1,1]$. We introduce the constants

$$B = \int_{2^{-2/3}}^{1} \frac{1}{\sqrt{1-v^{3/2}}}\, dv = 1.010059626\ldots$$

and

$$\alpha = \frac{B}{2^{11/6}+B} = 0.2208430100\ldots \ .$$

On $[-1,-\alpha]$, we define

$$f(x) = (c/12)^2 (x+1)^4 \ ,$$

where c is a normalization constant to be determined further on. On $[-\alpha, 0]$, we define $f(x)$ as the unique solution of the equation

$$x+\alpha = D\left(F(f(x)2^{-2/3}/f(-\alpha)) - F(2^{-2/3})\right) \ ,$$

where F is defined by

$$F(z) = \int_0^z \frac{1}{\sqrt{1-u^{3/2}}}\, du \ ,$$

and $D = (1-\alpha)/2^{11/6} = 0.2186435377\ldots$. On $[0,1]$, f is defined by symmetry. Note that f is bell-shaped, with convex pieces for $|x| \geq \alpha$ and a concave piece on $[-\alpha,\alpha]$. The name Sakai refers to the Japanese word for boundary, as f falls on the border of the admissible region in the (B,Q) plane. It is quickly verified that both f and f' are continuous on $\mathbb{R}$. A bit of work shows that

$$f''(x) = \begin{cases} -c\sqrt{f(x)} & , |x| \leq \alpha \\ c\sqrt{f(x)} & , |x| > \alpha \end{cases} \ .$$

Note that $f(x)/f(-\alpha)$ does not depend upon c. Standard computations show that $\int f = 1$ if $c = 12(1-\alpha)^{-2}/\sqrt{C}$, where

$$\begin{aligned} C &= \int \frac{f(x)}{f(-\alpha)}\, dx \\ &= 2^{5/3} D \int_{2^{-2/3}}^{1} \frac{v}{\sqrt{1-v^{3/2}}}\, dv + \frac{2(1-\alpha)}{5} = 0.9251260780\ldots \ . \end{aligned}$$

Therefore, we determine that $c = 20.55089981\ldots$. More computations lead us to

$$\int \sqrt{f} = c(1-\alpha)^3 \left(\frac{1}{18} + \frac{G}{\sqrt{8}} \right) ,$$

where

$$G = \int_{2^{-2/3}}^{1} \frac{\sqrt{v}}{\sqrt{1-v^{3/2}}}\, dv = 0.9428090416\ldots .$$

This yields $\int \sqrt{f} = 1.080098210\ldots$.

Also, $B(f) = c^{1/5} \int \sqrt{f} = 1.977103546\ldots$. Furthermore,

$$Q(f) = \gamma B(f) = \gamma c^{1/5} \int \sqrt{f} = \gamma c^{-4/5} \int |f''| = 2.033437\ldots .$$

This is not as good as we anticipated. There are no densities with $Q = \gamma B < 2.033437\ldots$.

• A QUADRATIC SPLINE. The family considered here consists of three quadratic splines pieced together:

$$f(x) = \begin{cases} b - cx^2/2 & |x| \le \theta \ ; \\ (|x| - \epsilon)^2/2 & \theta < |x| \le \epsilon \ ; \\ 0 & \epsilon < |x| \ . \end{cases}$$

Here b, c, θ and ϵ are chosen in such a way that f and f' are continuous, while $\int f = 1$. This yields $\theta^3 = 1/(c(c + 2/3 + c^2/3))$, $\epsilon = \theta(1 + c)$, $b = (1 + c)c\theta^2/2$. We are free to pick c. At one end of the spectrum $(c \to 0)$, we find the beta $(2, 2)$ density. As c increases, the value of $Q(f)$ initially decreases to reach a minimum at $c \approx 1.6$ of about 1.946. Then it increases again to the density at the other extremum, consisting of two parabolic pieces leaning against each other with a discontinuity in f' at the origin.

• A QUARTIC SPLINE. The family considered here consists of three quartic splines pieced together:

$$f(x) = \begin{cases} ax^4 & 0 \le x \le \theta \ ; \\ b - c(x-1/2)^4/2 & \theta \le x \le 1-\theta \ ; \\ a(1-x)^4 & \theta < 1-\theta \le x \le 1 \ ; \\ 0 & |x-1/2| > 1/2 \ . \end{cases}$$

Here a, b, c and θ are chosen in such a way that f and f' are continuous, while $\int f = 1$. This yields for $\theta \in (0, 2/3)$: $c = 5/(2-3\theta)$, $b = (c/2)(1/2-\theta)^3$, $a = c(1/2-\theta)^3/\theta^3$. The minimal value $Q(f) \approx 2.133$ occurs at $\theta \approx 0.33$.

• OTHER DENSITIES. The following densities are far from optimal. We provide some values for $B(f)$ and $Q(f)$ for future reference. The logistic density has $B(f) = 2.064\ldots$ and $Q(f) = 2.176\ldots$. The extreme value density is close: $B(f) = 2.060\ldots$ and $Q(f) = 2.272\ldots$. Similarly, the gamma (4) density is close: $B(f) = 2.016\ldots$ and $Q(f) = 2.290\ldots$. Because of its heavy tail, we note the dismal performance for Student's t3 density: $B(f) = (2000\pi/(27\sqrt{5}))^{1/5} = 2.531\ldots$, $Q(f) = 2.746\ldots$. At $Q(f) = 4.580\ldots$, the situation is even worse for the lognormal distribution. Finally, we recall that for the Cauchy and uniform densities, $B(f) = Q(f) = \infty$.

5. Proof of theorem 1.

We recall some properties from Devroye and Györfi (1985, p. 79 and p. 93):

LEMMA 5 (DEVROYE AND GYÖRFI, 1985, P. 91). *If K is bounded and has compact support, then there exists a constant c such that*

$$\left| \mathbf{E}|f_n(x) - f(x)| - \sigma_n(x)\psi\left(\frac{|B_n(x)|}{\sigma_n(x)}\right) \right| \le \frac{c}{nh} \ ,$$

where

$$\sigma_n^2(x) = \mathbf{E}(f_n(x) - \mathbf{E}f_n(x))^2$$

and

$$B_n(x) = \mathbf{E}f_n(x) - f(x) = f * K_h(x) - f(x) \ .$$

LEMMA 6. *For any density f,*

$$\int \sqrt{f * K_h} \geq \int \sqrt{f} \ .$$

When f satisfies condition (1) and K is a kernel with compact support, then

$$\lim_{h \to 0} \int |\sqrt{f * K_h} - \sqrt{f}| = 0$$

and $\int \sqrt{f * K_h} < \infty$ (all $h > 0$).

PROOF. The first statement is immediate from Jensen's inequality applied to the convolution integral. For the second result, assume without loss of generality that $a = s$ in condition (1), and that K vanishes off $[-s, s]$. From the Cauchy-Schwarz inequality, we have for large finite T:

$$\begin{aligned}
&\int_T^\infty |\sqrt{f * K_h} - \sqrt{f}| \\
&\quad = \int_T^\infty \sqrt{H(x)} |\sqrt{f * K_h(x)} - \sqrt{f(x)}| \times H(x)^{-1/2} \, dx \\
&\quad \leq \int_T^\infty H(x)(f * K_h(x) + f(x)) \, dx \times \int H(x)^{-1} \, dx \\
&\quad = \mathbf{E}\{H(X + hY) I_{X+hY>T} + H(X) I_{X>T}\} \times \int H(x)^{-1} \, dx \\
&\quad \leq \mathbf{E}\{(H(X+s) + H(X-s)) I_{X+s>T} + H(X) I_{X>T}\} \times \int H(x)^{-1} \, dx \ ,
\end{aligned}$$

where Y has density K and X has density f, and $h \leq 1$. The last expression on the right-hand-side of the chain is as small as desired by our choice of T and the finiteness of $\mathbf{E}H(X+s)$ and $\mathbf{E}H(X-s)$. Again by the Cauchy-Schwarz inequality, if H_2 denotes the Hellinger distance,

$$\begin{aligned}
&\int_{-T}^T |\sqrt{f * K_h} - \sqrt{f}| \\
&\quad \leq \sqrt{2T \int_{-T}^T |\sqrt{f * K_h} - \sqrt{f}|^2} \\
&\quad = \sqrt{2T} H_2(f * K_h, f) \\
&\quad \leq \sqrt{2T \int |f * K_h - f|} \\
&\quad = o(1)
\end{aligned}$$

(see Devroye, 1987, p. 7). This concludes the proof of Lemma 6. □

Lemma 7 (Devroye and Györfi, 1985, p. 79). *When* $\int \sqrt{f} = \infty$ *or*

$$\limsup_{a\to 0} \int |(f * \varphi_a)''| = \infty \ ,$$

where φ *is a mollifier, then*

$$\lim_{n\to\infty} n^{2/5} \inf_{h>0} \mathbf{E} J_{nh} = \infty \ .$$

The class $\mathcal{H}$ captures nearly all densities in the complement of the set described in Lemma 7. It was also shown by Devroye and Györfi (1985) that

$$\sup_{a>0} \int |(f * \varphi_a)''| = \liminf_{a\to 0} \int |(f * \varphi_a)''| < \infty \ .$$

Lemma 8. *Let* K *have compact support. For* $f \in \mathcal{H}$,

$$\inf_h \mathbf{E} J_{nh} = O(n^{-2/5}) \ .$$

Proof. Choose h such that $nh^5 = 1$. Assume that K has support in $[-s, s]$. Let $S_n = [-s_n, s_n]$ be such that $s_n = o(n^{2/5})$ and at the same time $\mathbf{P}\{|X| \geq s_n - 1\} = o(n^{-2/5})$ where X has density f. If f has compact support, the existence of such a sequence is obvious. If f does not, but condition (1) holds, then for the function H in that condition, $H(n)/n \to \infty$ as $n \to \infty$ along the integers. This implies that we can find a sequence of integers v_n for which

$$\frac{v_n}{n^{2/5}} \to 0 \ , \ \frac{n^{2/5}}{H(v_n)} \to 0 \ .$$

Take $s_n = v_n + 1$. Then $s_n = o(n^{2/5})$ and

$$\mathbf{P}\{|X| \geq s_n - 1\} \leq \frac{\mathbf{E}H(|X|)}{H(s_n - 1)} = \frac{\mathbf{E}H(|X|)}{H(v_n)} = o(n^{-2/5}) \ .$$

By Lemmas 3 and 5,

$$\begin{aligned}
&\mathbf{E} \int |f_n - f| \\
&\leq \int_{S_n} \frac{c}{nh} + \mathbf{E} \int_{S_n^c} |f_n - f| + \int_{S_n} \sigma_n \psi(\frac{B_n}{\sigma_n}) \\
&\leq \frac{2cs_n}{nh} + 2\mathbf{P}\{|X| \geq s_n - sh\} + \int |B_n| + \sqrt{\frac{2}{\pi}} \int \sigma_n \\
&= I + II + III + IV \ .
\end{aligned}$$

By our choice of s_n and h, we see that $I = o(n^{-2/5})$. Since $sh < 1$ for all n large enough, we also have $II = o(n^{-2/5})$. Thirdly,

$$\limsup_{n\to\infty} n^{2/5} \int |B_n| = \limsup_{n\to\infty} n^{2/5} \int |f * K_h - f|$$

$$
\begin{aligned}
&= \limsup_{n\to\infty} \int \frac{|f*K_h - f|}{h^2} \\
&\le \limsup_{h\downarrow 0} \int \frac{|f*K_h - f|}{h^2} \\
&= \sup_{a>0} \int |(f*\varphi_a)''| < \infty
\end{aligned}
$$

(Devroye and Györfi, 1985, p. 86). Hence $III = O(n^{-2/5})$. Finally, recalling that

$$
\begin{aligned}
\sigma_n^2 &= \frac{1}{n}((K_h)^2 * f - (K_h * f)^2) \\
&= \frac{\alpha^2}{nh}(L_h * f - h\alpha^{-2}(K_h * f)^2)
\end{aligned}
$$

where $L = K^2/\alpha^2$, we have by Lemma 6,

$$
\begin{aligned}
n^{2/5} \int \sigma_n &\le \alpha \int \sqrt{L_h * f} \\
&\le \alpha \int |\sqrt{L_h * f} - \sqrt{f}| + \alpha \int \sqrt{f} \\
&= o(1) + \alpha \int \sqrt{f}\ .
\end{aligned}
$$

This concludes the proof of Lemma 8. □

LEMMA 9. *Let h be a sequence of positive real numbers with the property that*

$$
\mathbf{E}J_{nh} \sim \inf_{u>0} \mathbf{E}J_{nu}\ .
$$

Assume that K has compact support. When $f \in \mathcal{H}$, then there exist constants a, b such that

$$
0 < a \le \liminf_{n\to\infty} nh^5 \le \limsup_{n\to\infty} nh^5 \le b < \infty\ .
$$

PROOF. If $h \to 0$ or $nh \to \infty$ do not hold along a subsequence, then $\limsup \mathbf{E} J_{nh} > 0$ by the equivalence result of Devroye and Györfi (1985, p. 12). Assume next that $nh^5 \to \infty$, $h \to 0$, along a subsequence. We will show that along this subsequence, $n^{2/5}\mathbf{E} J_{nh} \to \infty$, which contradicts Lemma 8.

By Lemma 5,

$$\begin{aligned}
\mathbf{E}\int |f_{nh} - f| &\geq \mathbf{E}\int_{-n^{1/5}}^{n^{1/5}} |f_{nh} - f| \\
&\geq \int_{-n^{1/5}}^{n^{1/5}} \sigma_n \psi\left(\frac{B_n}{\sigma_n}\right) - \frac{2cn^{1/5}}{nh} \\
&\geq \int_{-n^{1/5}}^{n^{1/5}} |B_n| - \frac{2cn^{1/5}}{nh} \\
&= h^2 \left(\int_{-n^{1/5}}^{n^{1/5}} \frac{|f * K_h - f|}{h^2} - \frac{2cn^{1/5}}{nh^3} \right) .
\end{aligned}$$

Thus, if every $\limsup$ and $\liminf$ is defined for n increasing along our subsequence,

$$\begin{aligned}
&\limsup_{n\to\infty} n^{2/5}\mathbf{E}\int |f_{nh} - f| \\
&\qquad \geq \limsup_{n\to\infty} (nh^5)^{2/5} \liminf_{n\to\infty} \int_{-n^{1/5}}^{n^{1/5}} \frac{|f * K_h - f|}{h^2} - \limsup_{n\to\infty} \frac{2c}{n^{2/5}h} .
\end{aligned}$$

But

$$\liminf_{n\to\infty} \int_{-n^{1/5}}^{n^{1/5}} \frac{|f * K_h - f|}{h^2} = \sup_{a>0} \int |(f * \varphi_a)''| > 0 .$$

Also,

$$\limsup_{n\to\infty} \frac{2c}{n^{2/5}h} < \infty$$

because $h > n^{-1/5}$ for n large enough, along our subsequence. We conclude that along the entire sequence,

$$\limsup_{n\to\infty} n^{2/5}\mathbf{E}\int |f_{nh} - f| = \infty ,$$

as required.

Assume finally that $nh^5 \to 0$, $nh \to \infty$, along a subsequence. We will show that along this subsequence, $n^{2/5}\mathbf{E} J_{nh} \to \infty$, which once again contradicts Lemma 8. That would conclude the proof of Lemma 9. Find a constant $t > 1$ with the property that

$$\int_{|x|\geq t-1} \sqrt{f} \leq \frac{1}{2}\int \sqrt{f} .$$

By Lemma 5,

$$\mathbf{E}\int |f_{nh} - f| \geq \mathbf{E}\int_{-t}^{t} |f_{nh} - f|$$

$$
\begin{aligned}
&\geq \int_{-t}^{t} \sigma_n \psi\left(\frac{B_n}{\sigma_n}\right) - \frac{2ct}{nh} \\
&\geq \sqrt{\frac{2}{\pi}} \int_{-t}^{t} \sigma_n - \frac{2ct}{nh} \ .
\end{aligned}
$$

We see that

$$
n^{2/5} \times (nh)^{-1} = (nh^5)^{-1/10} \times (nh)^{-1/2} = o((nh^5)^{-1/10})
$$

along the subsequence in question. Also, if K (and thus L) vanishes off $[-s, s]$,

$$
\begin{aligned}
n^{2/5} \int_{-t}^{t} \sigma_n &\geq \frac{n^{2/5}\alpha}{\sqrt{nh}} \left(\int_{-t}^{t} \sqrt{L_h * f} - \int \frac{\sqrt{h} K_h * f}{\alpha} \right) \\
&\geq \frac{\alpha}{(nh^5)^{1/10}} \left(\int_{-t}^{t} L_h * \sqrt{f} - \frac{\sqrt{h}}{\alpha} \right) \\
&\geq \frac{\alpha}{(nh^5)^{1/10}} \left(\int_{-t+sh}^{t-sh} \sqrt{f} - \frac{\sqrt{h}}{\alpha} \right) \\
&\geq \frac{\alpha + o(1)}{2(nh^5)^{1/10}} \ .
\end{aligned}
$$

Thus, along the given subsequence, $n^{2/5} \mathbf{E} J_{nh} \to \infty$. This concludes the proof of Lemma 9. $\square$

Lemma 10. *Let φ be a mollifier, let K be a compact support kernel, and let f be any density with $\int \sqrt{f} < \infty$. Then, for any constant γ,*

$$
\lim_{h \downarrow 0} \int \sqrt{f} \psi \left(\frac{\gamma |f * K_h - f|}{h^2 \sqrt{f}} \right) = \lim_{a \downarrow 0} \int \sqrt{f} \psi \left(\frac{\gamma |(f * \varphi_a)''|}{\sqrt{f}} \right) \ .
$$

PROOF. We follow the proof of Lemma 4 on page 84 of Devroye and Györfi (1985).
Part 1. First assume that f has two continuous derivatives and that $\int |f''| < \infty$. Let L be the kernel associated to a symmetric kernel K (page 104 of Devroye, 1987):

$$L(x) = \int_{|x|}^{\infty} (y - |x|) K(y)\, dy \ .$$

It is a nonnegative function integrating to $\beta/2$. We also have (p. 108 of Devroye, 1987):

$$f * K_h - f = h^2 f'' * L_h \ .$$

As $h \to 0$, we see that $(f * K_h - f)/h^2 \to f''\beta/2$ at almost all x by the Lebesgue density theorem and the integrability of f''. Thus, as $h \to 0$,

$$\begin{aligned} &\int \sqrt{f} \left| \frac{\gamma |f * K_h - f|}{h^2 \sqrt{f}} \right| \\ &\quad = \int \frac{\gamma |f * K_h - f|}{h^2} \\ &\quad = \int \gamma |f'' * L_h| \\ &\quad \le \int \gamma |f''| \beta/2 \ . \end{aligned}$$

Furthermore, since $\int \sqrt{f} < \infty$,

$$\int \sqrt{f} \xi \left(\frac{\gamma |f * K_h - f|}{h^2 \sqrt{f}} \right) \to \int \sqrt{f} \xi \left(\frac{\gamma |f''| \beta}{2 \sqrt{f}} \right) \ .$$

Putting this together, we see that

$$\limsup_{h \downarrow 0} \int \sqrt{f} \psi \left(\frac{\gamma |f * K_h - f|}{h^2 \sqrt{f}} \right) \le \int \sqrt{f} \psi \left(\frac{\gamma |f''| \beta}{2 \sqrt{f}} \right) \ .$$

Part 2. For $f \in \mathcal{H}$, we next show that

$$\limsup_{h \downarrow 0} \int \sqrt{f} \psi \left(\frac{\gamma |f * K_h - f|}{h^2 \sqrt{f}} \right) \le \liminf_{a \downarrow 0} \int \sqrt{f} \psi \left(\frac{\gamma |(f * \varphi_a)''| \beta}{2 \sqrt{f}} \right) \ .$$

Using $\psi(a+b) \le \psi(a) + \psi(b)$, we have

$$\begin{aligned} &\int \sqrt{f} \psi \left(\frac{\gamma |f * K_h - f|}{h^2 \sqrt{f}} \right) \\ &\quad \le \int \sqrt{f} \psi \left(\frac{\gamma |f - f * \varphi_a|}{h^2 \sqrt{f}} \right) \\ &\qquad + \int \sqrt{f} \psi \left(\frac{\gamma |f * K_h * \varphi_a - f * \varphi_a|}{h^2 \sqrt{f}} \right) \\ &\qquad + \int \sqrt{f} \psi \left(\frac{\gamma |f * K_h * \varphi_a - f * K_h|}{h^2 \sqrt{f}} \right) \\ &\quad = I + II + III \ . \end{aligned}$$

When $a \to 0$,

$$I = \int \sqrt{f}\psi\left(\frac{\gamma|f - f * \varphi_a|}{h^2\sqrt{f}}\right) \le \int \frac{\gamma|f - f * \varphi_a|}{h^2} \to 0 \ .$$

In a similar fashion, $II \to 0$. By part 1 of this proof, we have

$$\liminf_{a\to 0} III \le \liminf_{a\to 0} \int \sqrt{f}\psi\left(\frac{\gamma|(f * \varphi_a)''|\beta}{2\sqrt{f}}\right) \ .$$

<u>Part 3.</u> We show that

$$\liminf_{h\downarrow 0} \int \sqrt{f}\psi\left(\frac{\gamma|f * K_h - f|}{h^2\sqrt{f}}\right) \ge \limsup_{a\downarrow 0} \int \sqrt{f}\psi\left(\frac{\gamma|(f * \varphi_a)''|\beta}{2\sqrt{f}}\right) \ .$$

We need an auxiliary result. Let $f \ge 0$ and g be arbitrary functions, and let φ be a density. Then

$$\begin{aligned}
\int f\psi(g/f) &= \int \mathbf{E}|Nf(x) - g(x)| \\
&= \mathbf{E}\int |Nf(x) - g(x)| \\
&\ge \mathbf{E}\int |Nf * \varphi(x) - g * \varphi(x)| \\
&= \int f * \varphi\psi(g * \varphi / f * \varphi) \ .
\end{aligned}$$

Armed with this fact, we see that

$$\int \sqrt{f}\psi\left(\frac{\gamma|f * K_h - f|}{h^2\sqrt{f}}\right) \ge \int \sqrt{f * \varphi_a}\psi\left(\frac{\gamma|f * \varphi_a * K_h - f * \varphi_a|}{h^2\sqrt{f * \varphi_a}}\right) \ .$$

Take the limit infimum as $h \to 0$ on both sides of this inequality. Note that for fixed a,

$$\frac{f * \varphi_a * K_h - f * \varphi_a}{h^2} = (f * \varphi_a)'' * L_h \to \frac{(f * \varphi_a)''\beta}{2}$$

at almost all x. Applying Fatou's lemma on the right-hand-side, and the continuity of ψ, we have

$$\liminf_{h\to 0} \int \sqrt{f * \varphi_a}\psi\left(\frac{\gamma|f * \varphi_a * K_h - f * \varphi_a|}{h^2\sqrt{f * \varphi_a}}\right) \ge \int \sqrt{f * \varphi_a}\psi\left(\frac{\gamma|(f * \varphi_a)''|\beta}{2\sqrt{f * \varphi_a}}\right) \ .$$

Next take the limit supremum as $a \to 0$ on the right-hand side, and note that

$$\left|\int \sqrt{f * \varphi_a}\psi\left(\frac{\gamma|(f * \varphi_a)''|\beta}{2\sqrt{f * \varphi_a}}\right) - \int \sqrt{f}\psi\left(\frac{\gamma|(f * \varphi_a)''|\beta}{2\sqrt{f}}\right)\right| \le \sqrt{\frac{2}{\pi}} \int |\sqrt{f} - \sqrt{f * \varphi_a}| \to 0 \ .$$

This concludes the proof of part 3. Now combine with part 2, and obtain Lemma 10. □

PROOF OF THEOREM 1. Take a subsequence h with the property that $nh^5 \to \gamma$. Then, by Lemma 9, if K has compact support, and $f \in \mathcal{H}$,

$$\mathbf{E}\int |f_{nh} - f| = \int \sigma_n \psi\left(\frac{B_n}{\sigma_n}\right) + o(n^{-2/5}) = \frac{\alpha}{\sqrt{nh}}\int \sqrt{f}\psi\left(\frac{B_n\sqrt{nh}}{\alpha\sqrt{f}}\right) + o(n^{-2/5}) \ .$$

To see the first equality, use some estimates from the proof of Lemma 9. Take $s_n = o(n^{2/5})$ such that $\mathbf{P}\{|X| \geq s_n - sh\} = o(n^{-2/5})$ where $[-s, s]$ is the support of K, and X has density f. Such sequence s_n can be found for $f \in \mathcal{H}$ as was demonstrated in the proof of Lemma 9. Next, with $S_n = [-s_n, s_n]$, $\int_{S_n} c/(nh) = 2cs_n/(nh) = o(n^{-2/5})$. Also, if $L = K^2/\alpha^2$,

$$\begin{aligned}
\int_{S_n^c} \sigma_n \psi\left(\frac{B_n}{\sigma_n}\right) &\leq \sqrt{\frac{2}{\pi}} + \int_{S_n^c} |B_n| \\
&\leq \sqrt{\frac{2}{\pi nh}}\int_{S_n^c} \sqrt{L_h * f} + o(n^{-2/5}) \\
&\leq O(n^{-2/5})\int |\sqrt{L_h * f} - \sqrt{f}| + O(n^{-2/5})\int_{S_n^c} \sqrt{f} + o(n^{-2/5}) \\
&= o(n^{-2/5}) \ .
\end{aligned}$$

The second part of the equality can be obtained as follows:

$$\begin{aligned}
&\left|\int \sigma_n \psi\left(\frac{B_n}{\sigma_n}\right) - \int \frac{\alpha\sqrt{f}}{\sqrt{nh}}\psi\left(\frac{B_n\sqrt{nh}}{\alpha\sqrt{f}}\right)\right| \\
&\quad \leq \sqrt{\frac{2}{\pi}}\int \left|\sigma_n - \frac{\alpha\sqrt{f}}{\sqrt{nh}}\right| \\
&\quad \leq \sqrt{\frac{2}{\pi}}\int \left|\frac{\alpha\sqrt{f}}{\sqrt{nh}} - \frac{\alpha\sqrt{f * L_h}}{\sqrt{nh}}\right| + \sqrt{\frac{2}{\pi}}\frac{\alpha}{\sqrt{n}} \\
&\quad = o\left(\frac{1}{\sqrt{nh}}\right) = o(n^{-2/5}) \ .
\end{aligned}$$

Similarly, along our subsequence,

$$n^{2/5}\mathbf{E}\int |f_{nh} - f| = \int \frac{\alpha\sqrt{f}}{\gamma^{1/10}}\psi\left(\frac{B_n\sqrt{\gamma}}{h^2\alpha\sqrt{f}}\right) + o(1) \ .$$

By Lemma 10, we thus have along the subsequence,

$$n^{2/5}\mathbf{E}\int |f_{nh} - f| \to \limsup_{a\to 0}\int \frac{\alpha\sqrt{f}}{\gamma^{1/10}}\psi\left(\frac{|(f * \varphi_a)''|\beta\sqrt{\gamma}}{2\alpha\sqrt{f}}\right) \ .$$

From Lemma 9, we recall that for any optimal sequence h stays asymptotically in $[dn^{-1/5}, bn^{-1/5}]$ for some positive constants d, b. Thus, nh^5 has at least one subsequence tending to a constant $\gamma \in [d, b]$. The "best" subsequence is one that minimizes the limit with respect to γ. A simple reparametrization, setting $\gamma = 4\alpha^2 u^2/\beta^2$, yields the following:

$$n^{2/5} \inf_{h>0} \mathbf{E} \int |f_{nh} - f| \to 2^{-1/5} A(K) Q(f) \ . \ \square$$

6. Acknowledgment.

We thank both referees.

7. References.

H. Akaike, "An approximation to the density function," *Annals of the Institute of Statistical Mathematics*, vol. 6, pp. 127–132, 1954.

L. Devroye, "A note on the L1 consistency of variable kernel estimates," *Annals of Statistics*, vol. 13, pp. 1041–1049, 1985.

L. Devroye, *A Course in Density Estimation*, Birkhäuser, Boston, 1987.

L. Devroye, "Coupled samples in simulation," *Operations Research*, vol. 38, pp. 115–126, 1990.

L. Devroye and L. Györfi, *Nonparametric Density Estimation: The L1 View*, John Wiley, New York, 1985.

P. Hall and M. P. Wand, "Minimizing L_1 distance in nonparametric density estimation," *Journal of Multivariate Analysis*, vol. 26, pp. 59–88, 1988.

L. Holmström and J. Klemelä, "Asymptotic bounds for the expected L_1 error of a multivariate kernel density estimator," *Journal of Multivariate Analysis*, vol. 0, pp. 0–0, 1992.

R. J. Karunamuni and K. L. Mehra, "Optimal convergence properties of kernel density estimators without differentiability conditions," Technical Report 88-16, The University of Alberta, 1988.

A. Mokkadem, "Study of risks of kernel estimators," *Theory of Probability and its Applications*, vol. 34, pp. 341–350, 1989a.

A. Mokkadem, "Estimation of the entropy and information of absolutely continuous random variables," *IEEE Transactions on Information Theory*, vol. IT-35, pp. 193–196, 1989b.

E. Parzen, “On the estimation of a probability density function and the mode,” *Annals of Mathematical Statistics*, vol. 33, pp. 1065–1076, 1962.

M. Rosenblatt, “Remarks on some nonparametric estimates of a density function,” *Annals of Mathematical Statistics*, vol. 27, pp. 832–837, 1956.

D. Ruppert and D. B. H. Cline, “Transformation-kernel density estimation — bias reduction by empirical transformations,” *Annals of Statistics*, vol. 0, pp. 0–0, 1993.

D. Ruppert and M. Wand, “Correcting for kurtosis in density estimation,” *Australian Journal of Statistics*, vol. 34, pp. 19–29, 1992.

G. R. Terrell, “The maximal smoothing principle in density estimation,” *Journal of the American Statistical Association*, vol. 85, pp. 470–477, 1990.

M. P. Wand and L. Devroye, “How easy is a given density to estimate?” *Computational Statistics and Data Analysis*, vol. 0, pp. 0–0, 1992.

M. P. Wand, J. S. Marron, and D. Ruppert, “Transformations in density estimation,” *Journal of the American Statistical Association*, vol. 86, pp. 343–361, 1991.