Estimation of Densities and Derivatives of Densities at a Point with Spherical Data

Jussi Klemelä

Rolf Nevanlinna Institute
University of Helsinki

*This work was supported in part by the Academy of Finland under Grant no. 1015998 and by the University of Helsinki under a research grant.

Proposed running head: Estimation of Derivatives with Spherical Data.

Correspondence author:

Jussi Klemelä
Rolf Nevanlinna Institute
P.O. Box 4
SF-00014 UNIVERSITY OF HELSINKI
FINLAND

Abstract

Estimating the density function and the Laplacian of the density function is considered, when the observations are vectors on the d-dimensional unit sphere. Also higher order derivatives and iterated Laplacians of the density are estimated. The performance of estimators is measured by the mean squared error at a point. Density function is estimated by the kernel estimator. The Laplacian of the density is estimated with two estimators. The other is a kernel estimator and the other is the Laplacian of a kernel estimator.

AMS 1991 subject classifications

Primary: 62G07
Secondary: 62H11

Key words and phrases

estimation of derivatives, kernel estimator, Laplace operator, nonparametric density estimation, spherical data.

1 Introduction

Statistical data often consists of measurements which are (unrestricted) vectors of real numbers. Sometimes, however, the observations are concentrating on some Euclidean manifold. In this article the case when the observations are concentrating on the d-dimensional unit sphere $S_d = \{x \in \mathbf{R}^{d+1} : \|x\| = 1\}$, $d \geq 2$, is considered. These kind of observations arise when the data consists of directions and directions are represented as unit vectors.

It is assumed to have n independent identically distributed observations from a distribution which is absolutely continuous with respect to the Lebesgue measure of the sphere, which will be denoted by $\mu = \mu_d$. Kernel density estimator with the kernel function $L : [0, \infty[\to \mathbf{R}$ and the smoothing parameter $\kappa > 0$ is

$$\hat{f}_n(x) = \hat{f}_n(x, \kappa, L) = \frac{c(\kappa)}{n} \sum_{i=1}^{n} L(\kappa \arccos(x' X_i)), \tag{1}$$

where $x \in S_d$ and

$$c(\kappa)^{-1} = \int_{S_d} L(\kappa \arccos(x' y)) d\mu(y). \tag{2}$$

This estimator has been proposed in Watson (1970), Beran (1979). This definition of the kernel estimator with data in S_d is similar to the definition of the kernel estimator with data in $\mathbf{R}^{d+1}$, except that restriction to symmetric kernels is made.

This article is mainly concerned with the estimation of the Laplacian of a density. The estimation of the Laplacian of a density is motivated because it will be seen that the formula for the asymptotically optimal smoothing parameter of the kernel estimator, when the mean squared error criterion is used, involves such functional of a density. For functions $S_d \to \mathbf{R}$ there is an important concept of higher order derivative, different from the iterated Laplacian. This article discusses also the estimation of higher order derivatives and iterated Laplacians.

Estimation of derivatives with Euclidean data has been studied by Bhattacharya (1967), Singh (1976), (1979), (1987), Sheather (1983). In the Euclidean case, derivatives of the kernel estimator are also kernel estimators. In the spherical case, however, the Laplacian of the kernel estimator is not a kernel estimator. Thus we have two natural estimators for the Laplacian of the density. The other is the kernel estimator and the other is the Laplacian of the kernel estimator.

In Section 2 the definition of the Laplace operator and of the derivatives are given. Then two expansions of a convolution are formulated, which are basic for the study of the asymptotics of the bias of the estimators. In Section 3 an estimator will be constructed for an iterated Laplacian of a density. The estimator is an iterated Laplacian of the kernel estimator. The definition of a class s kernel is given to develop a fast convergence rate theory for this estimator. The asymptotic mean squared error and the asymptotically optimal smoothing parameter are given. In Section 4 an estimator will be constructed for higher order derivatives of a density. The estimator is a linear combination of kernel estimators. The definition of a class (r,s) kernel is given to develop a fast convergence rate theory for this estimator. The asymptotic mean squared error and the asymptotically optimal smoothing parameter are given. As a special case, the estimators of Section 3 and 4 define two different estimators for the Laplacian of a density. In section 5 some comparisons between the estimators are made. The proofs are given in the Appendix.

2 Preliminaries

Let us start with definitions of two parameterizations of S_d. Let $\eta \in S_d$ and $T_\eta = \{\xi \in S_d \,|\, \xi \perp \eta\}$. Let $\phi_\eta : S_d \backslash \{\eta, -\eta\} \to T_\eta \times\,]0,\pi[$ be a parameterization of S_d defined by

$$\phi_\eta^{-1}(\xi,\theta) = \eta\cos\theta + \xi\sin\theta. \tag{3}$$

Let us also use ϕ_η to denote a mapping $S_d \setminus \{\eta,-\eta\} \to T_\eta \times\,]-1,1[$ defined by

$$\phi_\eta^{-1}(\xi,t) = \eta t + \xi(1-t^2)^{1/2}.$$

To following integration formula corresponds to the first parameterization. For $f : S_d \to \mathbf{R}$ it is true that

$$\int_{S_d} f(x)d\mu(x) = \int_0^\pi d\theta \sin^{d-1}\theta \int_{T_\eta} f\left(\phi_\eta^{-1}(\xi,\theta)\right) d\mu_{d-1}(\xi) \tag{4}$$

Let us define the Laplace operator. Let $f : S_d \to \mathbf{R}$ and let $\eta^1,\ldots,\eta^{d+1} \in S_d$ be orthogonal. The Laplace operator is defined recursively by

$$\begin{aligned}\Delta f(x) &= \Delta_d f(x) \\ &= \left[(1-t^2)\frac{\partial^2}{\partial t^2} - td\frac{\partial}{\partial t} + \frac{1}{1-t^2}\Delta_{d-1}\right] f\left(\phi_{\eta^{d+1}}^{-1}(\xi,t)\right)\Big|_{(\xi,t)=\phi_{\eta^{d+1}}(x)}\end{aligned} \tag{5}$$

for $d \geq 2$ and $\Delta_1 f(x) = \frac{\partial^2}{\partial \theta^2} f\left(\phi_{\eta^1,\eta^2}^{-1}(\theta)\right)\Big|_{\theta=\phi_{\eta^1,\eta^2}(x)}$ where $\phi_{\eta^1,\eta^2} : S_1 \setminus \{\eta^2\} \to]0, 2\pi[$, $\eta^1, \eta^2 \in S_1$, $\eta^1 \perp \eta^2$, is defined by $\phi_{\eta^1,\eta^2}^{-1}(\theta) = \eta^2 \cos\theta + \eta^1 \sin\theta$. For $s \geq 4$ even, define $\Delta^{s/2} f = \Delta \Delta^{(s-2)/2} f$.

There exists a concept of higher order derivatives for functions $S_d \to \mathbf{R}$, different from a iterated Laplacian. When $g : \mathbf{R}^{d+1} \to \mathbf{R}$ and $x, \xi \in \mathbf{R}^{d+1}$, define the derivative of g at x in the direction of ξ to be $D_\xi g(x) = \lim_{h\to 0} h^{-1}[g(x + h\xi) - g(x)]$ and $D_\xi^s g = D_\xi D_\xi^{s-1} g$, for $s \geq 2$ integer. When $f : S_d \to \mathbf{R}$, define $\bar{f} : \mathbf{R}^{d+1} \to \mathbf{R}$ by $\bar{f}(x) = f(x/\|x\|)$. Define $D^s f : S_d \to \mathbf{R}$ by

$$D^s f(x) = \omega_{d-1}^{-1} \int_{T_x} D_\xi^s \bar{f}(x) d\mu_{d-1}(\xi), \tag{6}$$

where $d \geq 2$, $T_x = \{\xi \in S_d : \xi \perp x\}$, and $\omega_{d-1} = \mu_{d-1}(S_{d-1})$. This concept was first defined by Hall, Watson, and Cabrera (1987). It holds that $D^2 f = d^{-1}\Delta f$, when $\bar{f}$ and its partial derivatives are differentiable (Klemelä, 1997, Theorem 3.1). It also holds that $D^s f(x) = 0$ for all $x \in S_d$, when $s \geq 1$ is odd (Klemelä, 1997, equation 26). By considering a function of form $f : S_d \to \mathbf{R}$, $f(x) = g(x'\eta)$, where $\eta \in S_d$, $g : [-1, 1] \to \mathbf{R}$, it can be shown that there exist no constant C for which $D^4 = C\Delta^2$.

Let us define the smoothness class used in this study. Let $\tilde{D}^s f : S_d \times \mathbf{R} \to \mathbf{R}$ be defined by

$$\tilde{D}^s f(x, \theta) = \omega_{d-1}^{-1} \int_{T_x} D_{\phi_x^{-1}(\xi,\theta+\pi/2)}^s f\left(\phi_x^{-1}(\xi, \theta)\right) d\mu_{d-1}(\xi), \tag{7}$$

where ϕ_x was defined in (3). Note that $\phi_x^{-1}(\xi, \theta + \pi/2) \perp \phi_x^{-1}(\xi, \theta)$. Let $s \geq 2$ be even and $x_0 \in S_d$. Let $\mathbf{F}(s, x_0)$ be the set of such functions $f : S_d \to \mathbf{R}$ that $D^i f(x_0)$ is defined for $i = 1, \ldots, s$, for all $\xi \in T_{x_0} = \{\xi \in S_d \,|\, \xi \perp x_0\}$, $\partial^s/\partial\theta^s f(\phi_{x_0}^{-1}(\xi, \theta))$ is continuous as a function of $\theta \in \mathbf{R}$, $\left|\tilde{D}^s f(x_0, \theta)\right|$ is bounded for $\theta \in \mathbf{R}$, and $\lim_{\theta\to 0} \tilde{D}^s f(x_0, \theta) = D^s f(x_0)$.

Let us define a certain linear combination of derivatives. First, define for $i \geq 0$ integer,

$$\gamma_i = \sum_{\alpha_1+\cdots+\alpha_{d-1}=i} \frac{(-1)^{\alpha_1}}{(2\alpha_1 + 1)!} \cdots \frac{(-1)^{\alpha_{d-1}}}{(2\alpha_{d-1} + 1)!}, \ \gamma_0 = 1. \tag{8}$$

The linear combination of derivatives is defined by

$$\mathcal{D}_i^s f = \sum_{j=i}^{s/2} \frac{1}{(2j)!} \gamma_{s/2-j} D^{2j} f. \tag{9}$$

where $s \geq 2i$ even and $0 \leq i \leq s$. Let us define $\mathcal{D}_i^0 f = f$.

The expectation of the kernel estimator is

$$\mathrm{E}(\hat{f}_n(x)) = c(\kappa) \int_{S_d} f(y) L(\kappa \arccos(x'y)) d\mu(y).$$

This integral is called a convolution. In studying the bias of the estimators, we will need two expansions of a convolution. Define

$$\alpha_s(L) = \int_0^\infty t^{s+d-1} L(t) dt, \tag{10}$$

for $s \geq 0$.

Lemma 2.1 *Let $s \geq 2$ be even and $f \in \mathbf{F}(s, x_0)$.*

(i) *Suppose that*

$$\int_0^\kappa t^{2i+d-1} L(t) dt = \alpha_{2i}(L) + o(\kappa^{2i-s}), \tag{11}$$

for $i = 1, \ldots, s/2$. Then

$$\begin{aligned} &\kappa^d \int_{S_d} L(\kappa \arccos(x_0'y)) f(y) d\mu(y) \\ = \quad & d_0(\kappa, L) f(x_0) + \omega_{d-1} \sum_{j=1}^{s/2} \kappa^{-2j} \alpha_{2j}(L) \mathcal{D}_1^{2j} f(x_0) + o(\kappa^{-s}) \end{aligned}$$

where

$$d_0(\kappa, L) = \kappa^d c(\kappa)^{-1} = \kappa^d \int_{S_d} L(\kappa \arccos(x'y)) d\mu(y). \tag{12}$$

(ii) *Suppose that (11) holds for $i = 0, \ldots, s/2$. Then*

$$\kappa^d \int_{S_d} L(\kappa \arccos(x_0'y)) f(y) d\mu(y) = \omega_{d-1} \sum_{j=0}^{s/2} \kappa^{-2j} \alpha_{2j}(L) \mathcal{D}_0^{2j} f(x_0) + o(\kappa^{-s}).$$

The proof of Lemma 2.1 is given in the Appendix.

3 The First Estimator

In this section an estimator for $\Delta^{r/2} f(x_0)$, $r \geq 0$ even, will be constructed, where Δ is the Laplace operator defined in (5). The estimator to be defined is the Laplacian of the kernel estimator.

Let $L : [0, \infty[\to \mathbf{R}$ be such that $\alpha_0(|L|) < \infty$ and L has r derivatives. Define

$$\hat{g}_n(x) = \Delta^{r/2}\left(\hat{f}_n(x, \kappa, L)\right) = \frac{c(\kappa)}{n}\sum_{i=1}^{n} L^{(r,\kappa)}(x'X_i),$$

where $L^{(r,\kappa)} : [-1, 1] \to \mathbf{R}$ is defined by $L^{(0,\kappa)}(t) = L(\kappa \arccos t)$ and

$$L^{(r,\kappa)}(t) = \left[(1-t^2)\frac{\partial^2}{\partial t^2} - dt\frac{\partial}{\partial t}\right] L^{(r-2,\kappa)}(t)$$

if $r \geq 2$. When $r = 0$, $\hat{g}_n$ is the usual kernel estimator of a density.

For example, $L^{(2,\kappa)}(t) = \kappa^2 L^{(2)}(\kappa \arccos t) + \kappa(d-1)L^{(1)}(\kappa \arccos t)(1-t^2)^{-1/2}t$ and $L^{(4,\kappa)}(t) = \kappa^4 L^{(4)}(\kappa \arccos t) + \kappa^3(d-1)L^{(3)}(\kappa \arccos t)(1-t^2)^{-1/2}2t + \kappa^2(d-1)L^{(2)}(\kappa \arccos t)(1-t^2)^{-1}[(d-1)t^2-2] + \kappa^1(d-1)L^{(1)}(\kappa \arccos t)(1-t^2)^{-3/2}t(3-d)$. We will see that when $d = 1$, then $\hat{g}_n$ is a kernel estimator. This follows from a general formula for $L^{(r,\kappa)}(t)$, given in the next lemma.

Lemma 3.1 *Let $r \geq 2$ be even. Then*

$$L^{(r,\kappa)}(t) = \sum_{i=1}^{r} \kappa^i L^{(i)}(\kappa \arccos t)(1-t^2)^{(r-i)/2} P_{r,i}(t),$$

where $P_{r,i}$ are polynomials, $P_{r,r} \equiv 1$. When $d = 1$, $P_{r,i} \equiv 0$ for $i = 1, \ldots, r-1$.

The proof of Lemma 3.1 is given in the Appendix. The definition of a class s kernel is needed to make the bias of the estimators to converge faster.

Definition 3.2 *Let $s \geq 0$ be even. A class s kernel is a measurable function $L : [0, \infty[\to \mathbf{R}$ which satisfies*

(i) $\alpha_i(|L|) < \infty$ *for $i = 0, s$,*

(ii) $\int_0^\kappa t^{2i+d-1} L(t)dt = \alpha_{2i}(L) + o(\kappa^{2i-s})$ *for $i = 1, \ldots, s/2$,*

(iii) $\alpha_0(L) \neq 0$ *and $\alpha_{2i}(L) = 0$ for $i = 1, \ldots, s/2 - 1$.*

The next lemma gives a convergence property for the expectation of the estimator.

Lemma 3.3 *Let $0 \leq r < s$ be even. Assume that $\Delta^{r/2} f \in \mathbf{F}(s-r, x_0)$. (The definition of $\mathbf{F}$ was given in Section 2.) Let L be a class $s-r$ kernel. Then*

$$\mathrm{E}\left(\hat{g}_n(x_0)\right) = \Delta^{r/2} f(x_0) + \kappa^{r-s}\alpha_0(L)^{-1}\alpha_{s-r}(L)\mathcal{D}_1^{s-r}\Delta^{r/2} f(x_0) + o(\kappa^{r-s})$$

when $\kappa \to \infty$.

Proof. Because the Laplace operator is symmetric,

$$\begin{aligned}
\mathrm{E}\left(\hat{g}_n(x_0)\right) &= c(\kappa)\mathrm{E}(L^{(r,\kappa)}(x_0'X_1)) = c(\kappa)\int_{S_d} L^{(r,\kappa)}(x_0'y)f(y)d\mu(y) \\
&= c(\kappa)\int_{S_d} L(\kappa\arccos(x_0'y))\Delta^{r/2}f(y)d\mu(y).
\end{aligned}$$

The assertion follows from Lemma 2.1 (i), because $c(\kappa) = \kappa^d d_0(\kappa)^{-1}$ and $d_0(\kappa) \sim \omega_{d-1}\alpha_0(L)$. □

Let us next give the variance of the estimator.

Lemma 3.4 *Let $f : S_d \to \mathbf{R}$ be a bounded density which is continuous at $x_0 \in S_d$. Let $r \geq 0$ be even. Let L be such that for $m = 1, 2$,*

$$\begin{aligned}
\textstyle\int_0^\infty t^{d-1}|L(t)|dt &< \infty, \\
\textstyle\int_0^\infty t^{m(i-r)+d-1}|L^{(i)}(t)|^m dt &< \infty, \ \textit{for } i = 1, \ldots, r.
\end{aligned} \tag{13}$$

Then

$$\mathrm{Var}\left(\hat{g}_n(x_0)\right) = \frac{\kappa^{2r+d}}{n}\frac{\beta_{r,2}(L)}{\omega_{d-1}\alpha_0(L)^2}f(x_0) + \frac{o(\kappa^{2r+d})}{n}$$

when $\kappa \to \infty$, where

$$\beta_{r,2}(L) = \int_0^\infty t^{d-1}\left\{\sum_{i=1}^r P_{r,i}(1)t^{i-r}L^{(i)}(t)\right\}^2 dt.$$

The proof of Lemma 3.4 is given in the Appendix. From Lemma 3.3 and Lemma 3.4 it follows that

$$\begin{aligned}
&\mathrm{E}\left(\hat{g}_n(x_0,\kappa) - \Delta^{r/2}f(x_0)\right)^2 = \left(\mathrm{E}(\hat{g}_n(x_0,\kappa)) - \Delta^{r/2}f(x_0)\right)^2 + \mathrm{Var}\left(\hat{g}_n(x_0,\kappa)\right) \\
&= \kappa^{2(r-s)}\frac{\alpha_{s-r}(L)^2}{\alpha_0(L)^2}\left(\mathcal{D}_1^{s-r}\Delta^{r/2}f(x_0)\right)^2 + \frac{\kappa^{2r+d}}{n}\frac{\beta_{r,2}(L)}{\omega_{d-1}\alpha_0(L)^2}f(x_0) \\
&\quad + o(\kappa^{2(r-s)}) + \frac{o(\kappa^{2r+d})}{n}.
\end{aligned}$$

The optimal rate of convergence is achieved by choosing $\kappa = Cn^{1/(2s+d)}$ and then it is true that

$$\begin{aligned}
&\lim_{n\to\infty} n^{2(s-r)/(2s+d)}E\left(\hat{g}_n(x_0,\kappa) - \Delta^{r/2}f(x_0)\right)^2 \\
&= C^{2(r-s)}\frac{\alpha_{s-r}(L)^2}{\alpha_0(L)^2}\left(\mathcal{D}_1^{s-r}\Delta^{r/2}f(x_0)\right)^2 + C^{2r+d}\frac{\beta_{r,2}(L)}{\omega_{d-1}\alpha_0(L)^2}f(x_0).
\end{aligned}$$

This expression is minimized with respect to C, when $C = C^*$, where

$$C^* = \left[A(L) \frac{\left(\mathcal{D}_1^{s-r} \Delta^{r/2} f(x_0)\right)^2}{f(x_0)} \right]^{1/(2s+d)}$$

and $A(L) = 2(s-r)\omega_{d-1}\alpha_{s-r}(L)^2/((2r+d)\beta_{r,2}(L))$. The expression evaluated at C^* has the value

$$\left(\mathcal{D}_1^{s-r}\Delta^{r/2} f(x_0)\right)^{2(2r+d)/(2s+d)} f(x_0)^{2(s-r)/(2s+d)} \alpha_0(L)^{-2} \alpha_{s-r}(L)^{2(2r+d)/(2s+d)} \tag{14}$$

$$\times \beta_{r,2}(L)^{2(s-r)/(2s+d)} \omega_{d-1}^{2(r-s)/(2s+d)} \left(\frac{2(s-r)}{2r+d}\right)^{(4r-2s+d)/(2s+d)} .$$

4 The Second Estimator

In this section an estimator for $D^r f(x)$, $r \geq 0$ even, will be constructed, when $D^r f(x)$ is defined in (6). It is true that $D^2 f = d^{-1}\Delta f$ for functions f whose radial extensions and their partial derivatives are differentiable (Klemelä, 1997, Theorem 3.1). Thus Δf will be estimated again, now with a different estimator than in the previous section. The estimator to be defined is a linear combination of kernel estimators.

Let $L_i : [0,\infty[\to \mathbf{R}$, $i = 0, \ldots, r/2$, be such that $\alpha_{2i}(|L_i|) < \infty$ and $\alpha_{2i}(L_i) \neq 0$. Let $a = (a_0, \ldots, a_{r/2}) \in \mathbf{R}^{r/2+1}$ and define

$$\hat{h}_{n,a}(x) = \frac{\kappa^d}{n} \sum_{i=1}^{n} M_{\kappa,a}(x' X_i)$$

where

$$M_{\kappa,a}(t) = \sum_{i=0}^{r/2} a_i C_i(L_i) \kappa^{2i} L_i(\kappa \arccos t), \;\; t \in [-1,1]$$

$C_i(L_i) = [\omega_{d-1}\alpha_{2i}(L_i)]^{-1}$.

It will be seen in the proof of Lemma 4.2 that $\hat{h}_{n,a}$ is an estimator for $\sum_{i=0}^{r/2} a_i \mathcal{D}_0^{2i} f$, where $\mathcal{D}_0^{2i} f$ was defined in equation (9). Thus, a will be chosen in such a way that $\sum_{i=0}^{r/2} a_i \mathcal{D}_0^{2i} f = D^r f$. Thus, let a be such that

$$\frac{1}{(2i)!} \sum_{j=i}^{r/2} \gamma_{j-i} a_j = \delta_{r/2,i}, \;\; i = 0, \ldots, r/2, \tag{15}$$

where $\delta_{r/2,r/2} = 1$, $\delta_{r/2,i} = 0$ when $i \neq r/2$, and γ_{j-i} were defined in (8). For example, when $r = 0$, then $a_0 = 1$, and when $r = 2$, then $(a_0, a_1) = (-2\gamma_1, 2) = (d/3, 2)$. When $r = 0$, estimator $\hat{h}_{n,a}$ is otherwise similar to the kernel estimator of a density but the normalization constant is different. The bias of the estimator can be reduced by using kernels of a special form, given in the following definition.

Definition 4.1 *Let $0 \leq r \leq s$ be even. A class (r, s) kernel is a measurable function $L : [0, \infty[\to \mathbf{R}$ which satisfies*

(i) $\alpha_i(|L|) < \infty$ *for* $i = 0, s$,

(ii) $\int_0^\kappa t^{2i+d-1} L(t)dt = \alpha_{2i}(L) + o(\kappa^{2i-s})$ *for* $i = 0, \ldots, s/2$,

(iii) $\alpha_r(L) \neq 0$ *and* $\alpha_{2i}(L) = 0$ *for* $i = 0, \ldots, r/2 - 1, r/2 + 1, \ldots, s/2 - 1$.

A class s kernel in the sense of Definition 3.2, for which $\int_0^\kappa t^{d-1} L(t)dt = \alpha_0(L) + o(\kappa^{-s})$, is a class $(0, s)$ kernel in the sense of Definition 4.1. When L has a compact support, condition (ii) is always true. Also, when L has a compact support, a class $(r, r+2)$ kernel is the same as a class (r, r) kernel. To construct a class (r, s) kernel, L can be fit into a polynomial model on $[0, 1]$. The next lemma gives a convergence property for the expectation of the estimator.

Lemma 4.2 *Let $0 \leq r < s$ be even. Assume that $f \in \mathbf{F}(s, x_0)$. Let L_i be a class $(2i, s)$ kernel for $i = 0, \ldots, r/2$ and let a be chosen as in (15). Then*

$$\mathrm{E}\left(\hat{h}_{n,a}(x_0)\right) = D^r f(x_0) + \kappa^{r-s} \alpha_r(L_{r/2})^{-1} \alpha_s(L_{r/2}) \mathcal{D}_0^s f(x_0) + o(\kappa^{r-s})$$

when $\kappa \to \infty$.

Proof. By Lemma 2.1 (ii),

$$\begin{aligned}&\kappa^d \int_{S_d} L_i(\kappa \arccos(x_0' y)) f(y) d\mu(y) \\ &\quad = \omega_{d-1} \left[\kappa^{-2i} \alpha_{2i}(L_i) \mathcal{D}_0^{2i} f(x_0) + \kappa^{-s} \alpha_s(L_i) \mathcal{D}_0^s f(x_0)\right] + o(\kappa^{-s}),\end{aligned}$$

for $i = 0, \ldots, r/2$. Thus,

$$\mathrm{E}\left(\hat{h}_{n,a}(x_0)\right) = \kappa^d \mathrm{E}\left(M_{\kappa,a}(x_0' X_1)\right) = \kappa^d \int_{S_d} M_{\kappa,a}(x_0' y) f(y) d\mu(y)$$

$$
\begin{aligned}
&= \sum_{i=0}^{r/2} a_i C_i(L_i)\kappa^{2i+d} \int_{S_d} L_i(\kappa \arccos(x_0' y)) f(y) d\mu(y) \\
&= \sum_{i=0}^{r/2} a_i \mathcal{D}_0^{2i} f + \kappa^{r-s} a_{r/2} \alpha_r(L_{r/2})^{-1} \alpha_s(L_{r/2}) \mathcal{D}_0^s f(x_0) + o(\kappa^{r-s}).
\end{aligned}
$$

□

Next, let us give the asymptotic variance of $\hat{h}_{n,a}$.

Lemma 4.3 *Let $f : S_d \to \mathbf{R}$ be a bounded density which is continuous at $x_0 \in S_d$. Let $r \geq 0$ be even. Let $L_i : [0, \infty[\to \mathbf{R}$, $i = 0, \ldots, r/2$ be such that*

$$
\alpha_{2i}(|L_i|) < \infty, \ \ \alpha_{2i}(L_i) \neq 0, \ \ \alpha_0\left(|L_i|^m\right) < \infty, \ \ m = 1, 2 \tag{16}
$$

Let $a = (a_0, \ldots, a_{r/2}) \in \mathbf{R}^{r/2+1}$. Then

$$
\mathrm{Var}\left(\hat{h}_{n,a}(x_0)\right) = \frac{\kappa^{2r+d}}{n} \frac{a_{r/2}^2 \alpha_0(L_{r/2}^2)}{\omega_{d-1}\alpha_r(L_{r/2})^2} f(x_0) + \frac{o(\kappa^{2r+d})}{n}
$$

when $\kappa \to \infty$.

The proof of Lemma 4.3 is given in the Appendix. From Lemma 4.2 and Lemma 4.3 it follows, when a is chosen as in (15), that

$$
\begin{aligned}
&\mathrm{E}\left(\hat{h}_{n,a}(x_0,\kappa) - D^r f(x_0)\right)^2 = \left(\mathrm{E}(\hat{h}_{n,a}(x_0,\kappa)) - D^r f(x_0)\right)^2 + \mathrm{Var}\left(\hat{h}_{n,a}(x_0,\kappa)\right) \\
&= \kappa^{2(r-s)} \frac{\alpha_s(L_{r/2})^2}{\alpha_r(L_{r/2})^2} \left(\mathcal{D}_0^s f(x_0)\right)^2 + \frac{\kappa^{2r+d}}{n} \frac{\alpha_0(L_{r/2}^2)}{\omega_{d-1}\alpha_r(L_{r/2})^2} f(x_0) \\
&\quad + o(\kappa^{2(r-s)}) + \frac{o(\kappa^{2r+d})}{n}.
\end{aligned}
$$

The optimal rate of convergence is achieved by choosing $\kappa = Cn^{1/(2s+d)}$ and then it is true that

$$
\begin{aligned}
&\lim_{n\to\infty} n^{2(s-r)/(2s+d)} E\left(\hat{h}_{n,a}(x_0,\kappa) - D^r f(x_0)\right)^2 \\
&\quad = C^{2(r-s)} \frac{\alpha_s(L_{r/2})^2}{\alpha_r(L_{r/2})^2} \left(\mathcal{D}_0^s f(x_0)\right)^2 + C^{2r+d} \frac{\alpha_0(L_{r/2}^2)}{\omega_{d-1}\alpha_r(L_{r/2})^2} f(x_0).
\end{aligned}
$$

This expression is minimized with respect to C, when $C = C^*$, where

$$
C^* = \left[A(L) \frac{(\mathcal{D}_0^s f(x_0))^2}{f(x_0)} \right]^{1/(2s+d)}
$$

and $A(L) = 2(s-r)\omega_{d-1}\alpha_s(L_{r/2})^2/((2r+d)\alpha_0(L^2_{r/2}))$. The expression evaluated at C^* has the value

$$(\mathcal{D}^s_0 f(x_0))^{2(2r+d)/(2s+d)} f(x_0)^{2(s-r)/(2s+d)} \alpha_r(L_{r/2})^{-2}\alpha_s(L_{r/2})^{2(2r+d)/(2s+d)} \tag{17}$$
$$\times \alpha_0(L^2_{r/2})^{2(s-r)/(2s+d)}\omega_{d-1}^{2(r-s)/(2s+d)}\left(\frac{2(s-r)}{2r+d}\right)^{(4r-2s+d)/(2s+d)}.$$

5 Comparison of the Estimators

Let us compare estimators $\hat{g}_n(x_0)$ and $\hat{h}_{n,a}(x_0)$ when $r=0$ and $a=a_0=1$. In this case $\hat{g}_n(x_0)$ and $\hat{h}_{n,a}(x_0)$ are estimators for $f(x_0)$. Assume further, that f satisfies the smoothness conditions of Lemma 3.3 and Lemma 4.2 with the smoothness index $s=2$.

Now expression (17) gives the asymptotic mean squared error for $\hat{g}_n(x_0)$ as

$$\left(D^2 f(x_0)\right)^{2d/(d+4)} f(x_0)^{4/(d+4)} M(L)$$

where

$$M(L) = \left[\left(\frac{\alpha_2(L)}{\alpha_0(L)}\right)^{2d}\left(\frac{\alpha_0(L^2)}{\alpha_0(L)^2}\right)^4\right]^{1/(d+4)} \times \omega_{d-1}^{-4/(d+4)}\left(\frac{4}{d}\right)^{(d-4)/(d+4)}.$$

and L is a class 2 kernel. The asymptotic mean squared error for $\hat{h}_{n,a}(x_0)$ is

$$\left(\frac{1}{2}D^2 f(x_0) + \frac{-d}{6}f(x_0)\right)^{(2d)/(d+4)} f(x_0)^{4/(d+4)} M(L)$$

where L is a class $(0,2)$ kernel.

Comparison of the estimators $\hat{g}_n(x_0)$ and $d\cdot\hat{h}_{n,a}(x_0)$ when $r=2$ and $a=(a_0,a_1)=(d/3,2)$ is more difficult. In this case $\hat{g}_n(x_0)$ and $d\cdot\hat{h}_{n,a}(x_0)$ are estimators for $\Delta f(x_0)$. Assume, that f satisfies the smoothness conditions of Lemma 3.3 and Lemma 4.2 with the smoothness index $s=4$. Then expression (14) gives the asymptotic mean squared error for $\hat{g}_n(x_0)$ as

$$\left(\mathcal{D}^2_1\Delta f(x_0)\right)^{2(d+4)/(d+8)} f(x_0)^{4/(d+8)} M_1(L)$$

where

$$M_1(L) = \left[\left(\frac{\alpha_2(L)}{\alpha_0(L)} \right)^{2(d+4)} \left(\frac{\beta_{2,2}(L)}{\alpha_0(L)^2} \right)^4 \right]^{1/(d+8)} \omega_{d-1}^{-4/(d+8)} \left(\frac{4}{d+4} \right)^{d/(d+8)}$$

and L is a class 2 kernel which is twice differentiable. Expression (17) gives the asymptotic mean squared error for $d \cdot \hat{h}_{n,a}(x_0)$ as

$$d^2 \left(\mathcal{D}_0^4 f(x_0) \right)^{2(d+4)/(d+8)} f(x_0)^{4/(d+8)} M_2(L_1)$$

where

$$M_1(L_1) = \left[\left(\frac{\alpha_4(L_1)}{\alpha_2(L_1)} \right)^{2(d+4)} \left(\frac{\alpha_0(L_1^2)}{\alpha_2(L_1)^2} \right)^4 \right]^{1/(d+8)} \omega_{d-1}^{-4/(d+8)} \left(\frac{4}{d+4} \right)^{d/(d+8)}$$

and L_1 is a class $(2,4)$ kernel.

6 Appendix

Proof of Lemma 2.1. The proof of (i) goes similarily as the proof of Theorem 2.5 (ii) in Klemelä (1997) by using the fact that

$$\int_0^\pi \tilde{L}_\kappa^{(s)}(\theta) \tilde{D}^s f(x,\theta) d\theta = \frac{1}{s!} d_s(\kappa, L) D^s f(x) + o(\kappa^{-s}).$$

The expansion (ii) follows from expansion (i) by the fact that

$$d_0(\kappa, L) = \omega_{d-1} \sum_{i=0}^{s/2} \kappa^{-2i} \gamma_i \alpha_{2i}(L) + o(\kappa^{-s}),$$

which is proved in Klemelä (1997, Lemma 2.3). $\square$

Proof of Lemma 3.1. The proof is by induction. Assume that the assertion holds for s. Put $g(t) = L(\kappa \arccos t)$. Now

$$\begin{aligned}
&\left[(1-t^2)\frac{\partial^2}{\partial t^2} - td\frac{\partial}{\partial t} \right] \left[\sum_{i=1}^s \kappa^i P_{s,i}(t)(1-t^2)^{(r-i)/2} g^{(i)}(t) \right] \\
&= \sum_{i=1}^s (1-t^2)^{(r-i+2)/2} \{ \kappa^i g^{(i)}(t) \\
&\quad \Big[(s-i+2)t \left(P_{s,i}^{(1)}(t)(1-t^2) + P_{s,i}(t)(s-i)t \right)
\end{aligned}$$

$$
\begin{aligned}
&+(1-t^2)\left(P_{s,i}^{(2)}(t)(1-t^2)-2tP_{s,i}^{(1)}(t)+P_{s,i}(t)(s-i)+P_{s,i}^{(1)}(t)(s-i)t\right)\\
&-td\left(P_{s,i}^{(1)}(t)(1-t^2)+P_{s,i}(t)(s-i)t\right)\Big]\\
&-\kappa^{i+1}g^{(i+1)}(t)\Big[(s-i+2)tP_{s,i}(t)(1-t^2)^{1/2}\\
&+(1-t^2)^{1/2}\left(P_{s,i}^{(1)}(t)(1-t^2)+P_{s,i}(t)(s-i)t\right)\\
&+(1-t^2)\left(P_{s,i}^{(1)}(t)(1-t^2)^{1/2}-P_{s,i}(t)t(1-t^2)^{-1/2}\right)-tdP_{s,i}(t)(1-t^2)^{1/2}\Big]\\
&+\kappa^{i+2}g^{(i+2)}(t)(1-t^2)P_{s,i}(t)\}\\
= \;&\kappa^{s+2}g^{(s+2)}(t)+\kappa^{s+1}g^{(s+1)}(t)(1-t^2)^{-1/2}\left[(d-1)t+P_{s,s-1}(t)\right]\\
&+\sum_{i=1}^{s}\kappa^i g^{(i)}(t)\Big[(s-i+2)t\left(P_{s,i}^{(1)}(t)(1-t^2)+P_{s,i}(t)(s-i)t\right)\\
&+(1-t^2)\left(P_{s,i}^{(2)}(t)(1-t^2)-2tP_{s,i}^{(1)}(t)+P_{s,i}(t)(s-i)+P_{s,i}^{(1)}(t)(s-i)t\right)\\
&-td\left(P_{s,i}^{(1)}(t)(1-t^2)+P_{s,i}(t)(s-i)t\right)+(s-i+3)tP_{s,i-1}(t)(1-t^2)^{1/2}\\
&+(1-t^2)^{1/2}\left(P_{s,i-1}^{(1)}(t)(1-t^2)+P_{s,i-1}(t)(s-i+1)t\right)\\
&+(1-t^2)\left(P_{s,i-1}^{(1)}(t)(1-t^2)^{1/2}-P_{s,i-1}(t)t(1-t^2)^{-1/2}\right)\\
&-tdP_{s,i-1}(t)(1-t^2)^{1/2}+(1-t^2)P_{s,i-2}(t)\Big]
\end{aligned}
$$

where $P_{s,0}=0=P_{s,-1}$. □

The proof of Lemma 3.4 is based on the following two lemmas. The following lemma states that $\kappa^d\left(d_0(\kappa)^{-1}L^{(r,\kappa)}(x'y)\right)^m$ is an approximate identity up to the constant of normalization, where $d_0(\kappa)$ was defined in (12) and by an approximate identity we mean such family of functions $K_\kappa : S_d \to \mathbf{R}$, $\kappa > 0$, that $\lim_{\kappa\to\infty}\int_{S_d} K_\kappa d\mu = 1$ and for every $\delta > 0$, $\lim_{\kappa\to\infty}\int_{||x-\eta||>\delta} K_\kappa(x)d\mu(x) = 0$ for some $\eta \in S_d$.

Lemma 6.1 *Let $r \geq 0$ be even and $m \geq 1$ an integer. Let $L : [0,\infty[\to \mathbf{R}$ be such that condition (13) is satisfied.*

(i) *Then,*

$$\kappa^d\int_{S_d}\left(d_0(\kappa)^{-1}L^{(r,\kappa)}(x'y)\right)^m d\mu(y) = \kappa^{mr}\omega_{d-1}^{1-m}\alpha_0(L)^{-m}\beta_{r,m}(L)+o(\kappa^{mr})$$

when $\kappa\to\infty$, where

$$\beta_{r,m}(L)=\int_0^\infty t^{d-1}\left\{\sum_{i=1}^{r}P_{r,i}(1)t^{i-r}L^{(i)}(t)\right\}^m dt$$

and $P_{r,i}$ were defined in Lemma 3.1.

(ii) *For every $\delta > 0$,*

$$\kappa^d \int_{\|y-x\|>\delta} \left| d_0(\kappa)^{-1} L^{(r,\kappa)}(x'y) \right|^m d\mu(y) = o(\kappa^{mr})$$

when $\kappa \to \infty$.

Proof. By equation (4) and the substitution $\theta = \kappa^{-1} u$,

$$\begin{aligned}
\kappa^d \int_{S_d} \left(L^{(r,\kappa)}(x'y) \right)^m d\mu(y) &= \kappa^d \omega_{d-1} \int_0^\pi \left(L^{(r,\kappa)}(\cos\theta) \right)^m \sin^{d-1}\theta d\theta \\
&= \kappa^d \omega_{d-1} \int_0^{\pi\kappa} \left(L^{(r,\kappa)}(\cos(\kappa^{-1}u)) \right)^m \sin^{d-1}(\kappa^{-1}u)\kappa^{-1} du.
\end{aligned}$$

By Lemma 3.1,

$$L^{(r,\kappa)}(\cos(\kappa^{-1}u)) = \sum_{i=1}^r \kappa^i P_{r,i}(\cos(\kappa^{-1}u)) \sin^{i-r}(\kappa^{-1}u) L^{(i)}(u).$$

The first assertion follows from the dominated convergence theorem and the fact that $d_0(\kappa) \sim \omega_{d-1}\alpha_0(L)$. The second assertion follows similarly, because

$$\begin{aligned}
&\kappa^d \int_{\|y-x\|>\delta} \left| L^{(r,\kappa)}(x'y) \right|^m d\mu(y) \\
&= \kappa^d \omega_{d-1} \int_{\kappa \arccos(1-\delta/2)}^{\pi\kappa} \left| L^{(r,\kappa)}(\cos(\kappa^{-1}u)) \right|^m \sin^{d-1}(\kappa^{-1}u)\kappa^{-1} du.
\end{aligned}$$

□

The following lemma describes how the convolution of $\left(L^{(r,\kappa)}\right)^m$ and f approximates f.

Lemma 6.2 *Let $f : S_d \to \mathbf{R}$ be continuous at $x_0 \in S_d$ and assume f is bounded. Let $r \geq 0$ be even. Let $L : [0, \infty[\to \mathbf{R}$ be such that the condition (13) is satisfied. Then,*

$$\kappa^d \int_{S_d} \left(d_0(\kappa)^{-1} L^{(r,\kappa)}(x_0'y) \right)^m f(y) d\mu(y) = \kappa^{mr} \omega_{d-1}^{1-m} \alpha_0(L)^{-m} \beta_{r,m}(L) f(x_0) + o(\kappa^{mr}),$$

when $\kappa \to \infty$, where $\beta_{r,m}(L)$ was defined in Lemma 6.1.

Proof. Let $\epsilon > 0$. By the continuity of f at x_0 and Lemma 6.1 there exists such $\delta > 0$ and $\kappa_0 > 0$ that for $\kappa \geq \kappa_0$,

$$\sup_{\{y \,|\, \|y-x_0\| \leq \delta\}} |f(y) - f(x_0)| \leq \frac{\epsilon}{2} \left[\kappa^{d-mr} \int_{S_d} \left| d_0(\kappa)^{-1} L^{(r,\kappa)}(x_0'y) \right|^m d\mu(y) \right]^{-1}$$

and

$$\kappa^{d-mr}\int_{\|y-x_0\|>\delta}\left|d_0(\kappa)^{-1}L^{(r,\kappa)}(x_0'y)\right|^m d\mu(y)\le\frac{\epsilon}{2}\left[2\sup_{y\in S_d}|f(y)|\right]^{-1}.$$

Now for $\kappa\ge\kappa_0$,

$$\begin{aligned}
&\left|\kappa^{d-mr}\int_{S_d}\left(d_0(\kappa)^{-1}L^{(r,\kappa)}(x_0'y)\right)^m(f(y)-f(x_0))d\mu(y)\right| \\
\le\ &\sup_{\{y\,|\,\|y-x_0\|\le\delta\}}|f(y)-f(x_0)|\kappa^{d-mr}\int_{S_d}\left|d_0(\kappa)^{-1}L^{(r,\kappa)}(x'y)\right|^m d\mu(y) \\
&+2\sup_{y\in S_d}|f(y)|\kappa^{d-mr}\int_{\|y-x_0\|>\delta}\left|d_0(\kappa)^{-1}L^{(r,\kappa)}(x_0'y)\right|^m d\mu(y) \\
\le\ &\epsilon.
\end{aligned}$$

Thus

$$\begin{aligned}
&\kappa^d\int_{S_d}\left(d_0(\kappa)^{-1}L^{(r,\kappa)}(x_0'y)\right)^m f(y)d\mu(y) \\
=\ &\kappa^d\int_{S_d}\left(d_0(\kappa)^{-1}L^{(r,\kappa)}(x'y)\right)^m d\mu(y)f(x_0)+o(\kappa^{mr})
\end{aligned}$$

and the assertion follows from Lemma 6.1. □

Proof of Lemma 3.4. Firstly,

$$\begin{aligned}
\mathrm{Var}\left(\hat{g}_n(x_0)\right) &= n^{-1}c(\kappa)^2\mathrm{Var}\left(L^{(r,\kappa)}(x_0'X_1)\right) \\
&= n^{-1}c(\kappa)^2\left\{\mathrm{E}\left(L^{(r,\kappa)}(x_0'X_1)\right)^2-\left(\mathrm{E}\,L^{(r,\kappa)}(x_0'X_1)\right)^2\right\}.
\end{aligned}$$

By Lemma 6.4,

$$\begin{aligned}
\kappa^d\mathrm{E}\left(d_0(\kappa)^{-1}L^{(r,\kappa)}(x_0'X_1)\right)^2 &= \kappa^d\int_{S_d}\left(d_0(\kappa)^{-1}L^{(r,\kappa)}(x_0'y)\right)^2 f(y)d\mu(y) \\
&= \kappa^{2r}\omega_{d-1}^{-1}\alpha_0(L)^{-2}\beta_{r,2}(L)f(x_0)+o(\kappa^{2r}),
\end{aligned}$$

and

$$\begin{aligned}
c(\kappa)^2\left[\mathrm{E}\,L^{(r,\kappa)}(x_0'X_1)\right]^2 &= \left[\kappa^d\int_{S_d}d_0(\kappa)^{-1}L^{(r,\kappa)}(x_0'y)f(y)d\mu(y)\right]^2 \\
&= [O(\kappa^r)]^2=o(\kappa^{2r+d}).
\end{aligned}$$

□

The proof of Lemma 4.3 is based on the following two lemmas. The following lemma states that $\kappa^{d-mr}M^m_{\kappa,a}$ is an approximate identity up to the constant of normalization.

Lemma 6.3 *Let $r \geq 0$ be even, $a = (a_0, \ldots, a_{r/2}) \in \mathbf{R}^{r/2+1}$ and $m \geq 1$ integer. Let $L_i : [0, \infty[\to \mathbf{R}$, $i = 0, \ldots, r/2$, be such that the condition (16) is satisfied.*

(i) *Then,*

$$\kappa^d \int_{S_d} M_{\kappa,a}^m(x'y) d\mu(y) = a_{r/2}^m \kappa^{mr} \omega_{d-1}^{1-m} \alpha_r(L_{r/2})^{-m} \alpha_0(L_{r/2}^m) + o(\kappa^{mr})$$

when $\kappa \to \infty$.

(ii) *For every $\delta > 0$,*

$$\kappa^d \int_{\|y-x\|>\delta} M_{\kappa,a}^m(x'y) d\mu(y) = o(\kappa^{mr})$$

when $\kappa \to \infty$.

Proof. By equation (4) and the substitution $t = \kappa\theta$,

$$\begin{aligned}
\kappa^d \int_{S_d} M_{\kappa,a}^m(x'y) d\mu(y) &= \kappa^d \omega_{d-1} \int_0^\pi M_{\kappa,a}^m(\cos\theta) \sin^{d-1}\theta d\theta \\
&= \kappa^d \omega_{d-1} \int_0^{\pi\kappa} M_{\kappa,a}^m(\cos(\kappa^{-1}t)) \sin^{d-1}(\kappa^{-1}t)\kappa^{-1} dt \\
&= \kappa^d \omega_{d-1} \int_0^{\pi\kappa} \left[\sum_{i=0}^{r/2} a_i C_i(L_i) \kappa^{2i} L_i(t) \right]^m \sin^{d-1}(\kappa^{-1}t)\kappa^{-1} dt.
\end{aligned}$$

The assertion follows from the dominated convergence theorem. The second assertion follows similarly, because

$$\begin{aligned}
&\kappa^d \int_{\|y-x\|>\delta} M_{\kappa,a}^m(x'y) d\mu(y) \\
&= \kappa^d \omega_{d-1} \int_{\kappa \arccos(1-\delta/2)}^{\pi\kappa} M_{\kappa,a}^m(\cos(\kappa^{-1}u)) \sin^{d-1}(\kappa^{-1}u)\kappa^{-1} du.
\end{aligned}$$

□

The following lemma describes how the convolution of $M_{\kappa,a}^m$ and f approximates f.

Lemma 6.4 *Let $f : S_d \to \mathbf{R}$ be continuous at $x_0 \in S_d$ and bounded. Let $r \geq 0$ be even and $a = (a_0, \ldots, a_{r/2}) \in \mathbf{R}^{r/2+1}$. Let $L_i : [0, \infty[\to \mathbf{R}$, $i = 0, \ldots, r/2$, be such that the condition (16) is satisfied. Then,*

$$\kappa^d \int_{S_d} M_{\kappa,a}^m(x_0'y) f(y) d\mu(y) = a_{r/2}^m \kappa^{mr} \omega_{d-1}^{1-m} \alpha_r(L_{r/2})^{-m} \alpha_0(L_{r/2}^m) f(x_0) + o(\kappa^{mr}),$$

when $\kappa \to \infty$.

Proof. In view of Lemma 6.3, the proof is similar to the proof of Lemma 6.2. □

Proof of Lemma 4.3. The proof follows from Lemma 6.4 in the same way as Lemma 3.4 was proved. □

Acknowledgments

This paper is based on part of the author's doctoral dissertation, which was written at the University of Helsinki under the supervision of Dr. Lasse Holmström. His support is greatly acknowledged.

References

[1] Beran, R. (1979). Exponential models for directional data. *Ann. of Statist.* **7** 1162–1178.

[2] Bhattacharya, P. K. (1967). Estimation of a probability density function and its derivatives. *Sankhyā Ser. A* **29** 373–382.

[3] Hall, P., Watson, G. S., and Cabrera, J. (1987). Kernel density estimation with spherical data. *Biometrika* **74** 751–62.

[4] Klemelä, J. (1997). Integrated risk of the kernel estimator with spherical data. Submitted to *J. Multivariate Anal.*

[5] Sheather, S. J. (1983). A data-based algorithm for choosing the window width when estimating the density at a point. *Comp. Statist. Anal.* **1** 229–238.

[6] Singh, R. S. (1976). Nonparametric estimation of mixed partial derivatives of a multivariate density. *J. Multivariate Anal.* **6** 111–122.

[7] Singh, R. S. (1979). Mean squared errors of estimates of a density and its derivatives. *Biometrika* **66** 177–180.

[8] Singh, R. S. (1987). MISE of kernel estimates of a density and its derivatives. *Statist. Probab. Lett.* **5** 153–159.