Ann. Inst. Statist. Math.
32 (1980), Part A, 351-362

ESTIMATING A DENSITY ON THE POSITIVE HALF LINE BY THE METHOD OF ORTHOGONAL SERIES

PETER HALL

(Received Jan. 30, 1980; revised May 23, 1980)

Abstract

The kernel method of density estimation is not so attractive when the density has its support confined to $(0, \infty)$, particularly when the density is unsmooth at the origin. In this situation the method of orthogonal series is competitive. We consider three essentially different orthogonal series—those based on the even and odd Hermite functions, respectively, and that based on Laguerre functions—and compare them from the point of view of mean integrated square error.

1. Introduction and summary

In the past a considerable amount of attention has been devoted to the problem of estimating a density which is smooth on the whole real line. There are several competitive classes of nonparametric estimators, the two most popular being kernel estimators (introduced by Rosenblatt [6]; see also Rosenblatt [7], Parzen [5] and Watson and Leadbetter [13]), and orthogonal series estimators (see Cencov [2], Kronmal and Tarter [3] and Watson [12]). A kernel estimator usually has greater efficiency than an estimator based on orthogonal series, although it may be more tedious to calculate and update.

If the density has its support confined to the positive half line and is not smooth at the origin then the kernel method will not be so attractive. Define the *partial* mean integrated square error of a kernel estimator $\hat{f}_n$ by

$$J_n(a)=\int_a^\infty \mathrm{E}\,[\hat{f}_n(x)-f(x)]^2 dx\ , \qquad a>0\ .$$

Under suitable conditions on the density f, $J_n(a)$ will converge at a

AMS 1970 subject classification: Primary 62G05.
Key words: Density estimate, Hermite series, Laguerre series, mean square error, orthogonal polynomials.

rate of $O(n^{-4/5})$ for all $a>0$. However, the estimator may perform poorly in the neighbourhood of the origin, and its mean integrated square error, $J_n(0)$, may converge at a slower rate.

In this situation an estimator based on an orthogonal series may be very competitive. We shall consider three essentially different series —those based on the even and odd Hermite functions, respectively, and that based on Laguerre functions. Of course, Hermite functions have a simple expression in terms of Laguerre functions, but they give rise to estimators with significantly different properties. In Section 2 we consider the Hermite series expansion, and show that for many purposes the even series gives a more efficient estimator. Indeed, the partial mean integrated square error, defined by

$$J_n(a, b)=\int_a^b \mathrm{E}\,[\hat{f}_n(x)-f(x)]^2dx\ , \qquad 0<a<b<\infty\ ,$$

converges at a rate of $O(n^{-4/5})$. To obtain this rate it is necessary to impose more stringent conditions on the smoothness of the density than in the case of a kernel estimator, but we feel that this is compensated for by the comparative ease with which an orthogonal series estimator may be calculated. In Section 3 we examine estimators based on Laguerre series, and show that these estimators are generally inferior to those based on the even Hermite series.

Schwartz [9] and Walter [11] studied the properties of the Hermite series used to estimate densities which are smooth on the whole real line. In this case it was also necessary to impose more stringent conditions to obtain rates of convergence comparable with those of a kernel estimator. As Schwartz and Walter pointed out, the properties of orthogonal series estimators in one dimension extend easily to any number of dimensions, and the rates of convergence are preserved. Therefore our estimators will be more attractive than kernel estimators in a multidimensional situation.

2. Estimation by Hermite polynomial expansion

Let $c_m=(\pi^{1/2}2^m m!)^{-1/2}$. The orthonormal Hermite system on $(-\infty, \infty)$ is defined by

$$h_m(x)=c_m H_m(x)e^{-x^2/2}\ , \qquad m\geqq 0\ ,$$

where the H_m are the Hermite polynomials. Suppose that the bounded density f has its support confined to $(0, \infty)$, and is continuous and of bounded variation on each interval $(0, \lambda)$, $\lambda>0$. Then the even and odd Hermite expansions of f, defined respectively by

$$f(x)=2\sum_0^\infty a_{2m}h_{2m}(x)=2\sum_0^\infty a_{2m+1}h_{2m+1}(x) ,$$

with $a_m=\int_0^\infty f(x)h_m(x)dx$, converge to f at each point $x>0$. (See Sansone [8], pp. 381–382). If $X_1, X_2, \cdots, X_m$ is an independent sample with common density f then

$$\hat{a}_m=n^{-1}\sum_{i=1}^n h_m(X_i)$$

is an unbiased estimate of a_m, and two estimates of f are given by

$$\hat{f}_{n1}(x;m)=\sum_{j=0}^m \hat{a}_{2j}h_{2j}(x) \quad \text{and} \quad \hat{f}_{n2}(x;m)=\sum_{j=0}^m \hat{a}_{2j+1}h_{2j+1}(x) , \qquad x>0 .$$

Since $h_m(0)=0$ for odd m then use of the estimator $\hat{f}_{n2}$ forces the estimated density to pass through the origin. If $f(0+)\neq 0$, the estimator $\hat{f}_{n2}$ will have a larger bias (and consequently a larger mean square error) than the estimator $\hat{f}_{n1}$. However, in real situations there may be theoretical or experimental evidence to suggest that $f(0+)=0$, and in this case the estimator $\hat{f}_{n2}$ will often perform better than $\hat{f}_{n1}$. These heuristic conclusions are made precise by

THEOREM 1. *Assume that* $\mathrm{E}(X_1^5)<\infty$. *If* f *has a bounded derivative on* $(0, \infty)$, *if* $f(0+)=f(0)$ *is well defined and*

$$\int_0^\infty x^{5/2}|f'(x)|dx<\infty ,$$

then

$$\int_0^\infty \mathrm{E}\,[\hat{f}_{n2}(x;m)-f(x)]^2dx=(2/n\pi)m^{1/2}+\pi^{-1}m^{-1/2}f(0)^2+o(n^{-1}m^{1/2}+m^{-1/2})$$

as m *and* $n\rightarrow\infty$. *If* f *has three bounded derivatives on* $(0, \infty)$, *if* $f(0+)=0$ *and* $f''(0)\equiv f''(0+)$ *is well defined and*

$$\text{(2.1)} \qquad \int_0^\infty x^{5/2}|f'''(x)-3xf''(x)+3(x^2-1)f'(x)+x(3-x^2)f(x)|dx<\infty ,$$

then

$$\int_0^\infty \mathrm{E}\,[\hat{f}_{n2}(x;m)-f(x)]^2dx$$
$$=(2/n\pi)m^{1/2}+(80\pi)^{-1}f''(0)^2m^{-5/2}+o(n^{-1}m^{1/2}+m^{-5/2})$$

as m *and* $n\rightarrow\infty$. *If* f *has two bounded derivatives on* $(0, \infty)$, *if* $f'(0)\equiv f'(0+)$ *is well defined and*

$$\int_0^\infty x^{5/2}|f''(x)-2xf'(x)|dx<\infty ,$$

then

$$\int_0^\infty \mathrm{E}\,[\hat{f}_{n1}(x;m)-f(x)]^2dx$$
$$=(2/n\pi)m^{1/2}+(12\pi)^{-1}m^{-3/2}f'(0)^2+o(n^{-1}m^{1/2}+m^{-3/2}) .$$

The optimal value of the mean integrated square error may be determined in the usual way. For example, if $f'(0)\neq 0$ then

$$\inf_m \int_0^\infty \mathrm{E}\,[\hat{f}_{n1}(x;m)-f(x)]^2dx\sim(4/3\pi)[2^{1/2}|f'(0)|]^{1/2}n^{-3/4} ,$$

and the infemum is attained with

$$m\sim(1/2)|f'(0)|(n/2)^{1/2} .$$

The results described above have obvious analogues if we assume that $f^{(i)}(0+)=0$ for $0\leqq i\leqq r$, say. However, in applications there will usually be little evidence to permit even the assumption $f'(0+)=0$.

Hermite polynomial expansions of functions often converge more quickly away from the origin than in the neighbourhood of the origin. In this respect the mean integrated square error, which is influenced by behaviour in the region of the origin almost as much as by behaviour in the region of any other point, may give an unduly pessimistic view of the performance of the estimator. If we are prepared to sacrifice some accuracy of our estimator near the origin and towards infinity we would construct it to minimise the *partial* mean integrated square error,

$$\int_a^b \mathrm{E}\,[\hat{f}_{n1}(x;m)-f(x)]^2dx ,$$

where $0<a<b<\infty$. This would lead us to choose a smaller value of m. We shall consider only the estimator $\hat{f}_{n1}$; similar results to the following may be obtained for $\hat{f}_{n2}$.

THEOREM 2. *Suppose* $0<a<b<\infty$. *If* f *has three bounded derivatives on* $(0,\infty)$, *if* $\mathrm{E}\,(X_1^5)<\infty$, $f'(0)\equiv f'(0+)$ *is well defined and if* (2.1) *holds then*

$$\int_a^b \mathrm{E}\,[\hat{f}_{n1}(x;m)-f(x)]^2dx=O(n^{-1}m^{1/2}+m^{-2}) .$$

Setting $m=O(n^{2/5})$ we obtain an estimator whose partial mean integrated square error converges at a rate of $O(n^{-4/5})$ on any compact subset of $(0,\infty)$. This is of course the rate of convergence of the mean integrated square error for a kernel estimator.

PROOF OF THEOREM 1. The following lemma is an easy corollary to the Riemann-Lebesgue lemma and the result (10) of [8], p. 324.

LEMMA 1. *Let g be a measurable function on the real line. If*

$$\int_{-\infty}^{\infty}(1+|x|^{5/2})|g(x)|dx<\infty$$

then $\int_{-\infty}^{\infty} g(x)h_m(x)dx=o(m^{-1/4})$, *and if*

$$\int_{-\infty}^{\infty}(1+|x|^{5})|g(x)|dx<\infty \qquad \text{and} \qquad \int_{-\infty}^{\infty} g(x)dx\neq 0$$

then $\int_{-\infty}^{\infty} g(x)h_m^2(x)dx\sim(2\pi^2 m)^{-1/2}\int_{-\infty}^{\infty} g(x)dx.$

Integrating by parts and using the formula $H'_{m+1}=2(m+1)H_m$ we see that

$$\begin{aligned}(2.2)\qquad c_m^{-1}a_m=&-\frac{1}{2}(m+1)^{-1}f(0)H_{m+1}+[4(m+1)(m+2)]^{-1}f'(0)H_{m+2}(0)\\ &+[8(m+1)(m+2)(m+3)]^{-1}[f''(0)-f(0)]H_{m+3}(0)\\ &-[8(m+1)(m+2)(m+3)]^{-1}\int_0^{\infty}[f'''(x)-3xf''(x)\\ &+3(x^2-1)f'(x)+x(3-x^2)f(x)]e^{-x^2/2}H_{m+3}(x)dx\ .\end{aligned}$$

Since $H_{2m}(0)=(-1)^m(2m)!/m!$ then under the appropriate conditions of Theorem 1,

$$a_{2m-1}=\frac{1}{2}(-1)^{m+1}\pi^{-1/2}m^{-3/4}f(0)+r_m\ ,$$

$$a_{2m}=\frac{1}{4}(-1)^{m+1}\pi^{-1/2}m^{-5/4}f'(0)+s_m\ ,$$

and if $f(0)=0$, $a_{2m-1}=(1/8)(-1)^{m+1}\pi^{-1/2}m^{-7/4}f''(0)+t_m$, where

$$|r_m|\leqq m^{-1/2}\left|\int_0^{\infty}[xf(x)-f'(x)]e^{-x^2/2}h_{2m}(x)dx\right|+o(m^{-3/4})\ ,$$

$$|s_m|\leqq m^{-1}\left|\int_0^{\infty}[f''(x)-2xf'(x)+(x^2-1)f(x)]e^{-x^2/2}h_{2m+2}(x)dx\right|+o(m^{-5/4})$$

and

$$\begin{aligned}|t_m|\leqq m^{-3/2}\Bigg|\int_0^{\infty}&[f'''(x)-3xf''(x)+3(x^2-1)f'(x)+x(3-x^2)f(x)]\\ &\times e^{-x^2/2}h_{2m+2}(x)dx\Bigg|+o(m^{-7/4})\ .\end{aligned}$$

It follows from Lemma 1 that under appropriate conditions, $r_m=o(m^{-3/4})$, $s_m=o(m^{-5/4})$ and $t_m=o(m^{-7/4})$, so that

$$\int_0^\infty [\mathrm{E}\,\hat{f}_{n1}(x\,;\,m)-f(x)]^2dx=2\sum_{m+1}^{\infty} a_{2j}^2=(12\pi)^{-1}f'(0)^2m^{-3/2}+o(m^{-3/2})$$

and

$$\int_0^\infty [\mathrm{E}\,\hat{f}_{n2}(x\,;\,m)-f(x)]^2dx=(80\pi)^{-1}f''(0)^2m^{-5/2}+o(m^{-5/2}) \text{ if } f(0)=0\ .$$

Since $\mathrm{E}\,(X_1^5)<\infty$ then by Lemma 1,

$$n\int_0^\infty \mathrm{var}\,[\hat{f}_{n1}(x\,;\,m)]dx\sim 2\sum_1^m (4\pi^2 j)^{-1/2}\sim 2\pi^{-1}m^{1/2}\ .$$

Similarly $n\int_0^\infty \mathrm{var}\,[\hat{f}_{n2}(x\,;\,m)]dx\sim 2\pi^{-1}m^{1/2}$, and Theorem 1 follows.

PROOF OF THEOREM 2. We first prove

LEMMA 2. *Let* $A(z)$ *denote either* $\cos z$ *or* $\sin z$, *and suppose* a, b, c *and* d *are real numbers with* $0<a<b<\infty$ *and* $0\leqq|d|<c<\infty$. *Then* $\sum_{n=1}^{m} A[(cn+d)^{1/2}x]=O(m^{1/2})$ *uniformly in* $x\in[a,b]$ *as* $m\to\infty$.

PROOF. We treat only the case $A\equiv\cos$. By making a scale change we may assume without loss of generality that $0<a<b<2\pi$. Let $p=p(n)$ be the smallest integer such that $p^2\geqq cn+d$. Then $(cn+d)^{1/2}=p-(p^2-cn-d)/2p+r_n$ where for a constant C, $|r_n|\leqq Cp[p^2-(p-1)^2]^2/p^4=O(n^{-1/2})$ as $n\to\infty$. Therefore

$$\begin{aligned}\cos[(cn+d)^{1/2}x]&=\cos(px)\cos[(p^2-cn-d)x/2p]\\&\quad+\sin(px)\sin[(p^2-cn-d)x/2p]+O(n^{-1/2})\end{aligned}$$

uniformly in $x\in[a,b]$.

Let B_k be the set of indices n such that $(k-1)^2<cn+d\leqq k^2$, let a_k+1 be the number of elements of B_k, n_k be the largest element and $b_k=cn_k+d$. Then $|a_k-2c^{-1}k|<3+c^{-1}$ and $|b_k-k^2|<c$. With $q=p(m)-1$ and

$$S(m)=\sum_{n=1}^{q}\sum_{n\in B_k}\cos[(cn+d)^{1/2}x]$$

we have, after some algebra,

$$\begin{aligned}S(m)&=\sum_{k=1}^{q}\cos(kx)\sum_{j=0}^{a_k}\cos(cjx/2k)+\sum_{k=1}^{q}\sin(kx)\sum_{j=1}^{a_k}\sin(cjx/2k)+O(m^{1/2})\\&=(2/cx)\Big[\sin x\sum_{k=1}^{q}k\cos(kx)+(1-\cos x)\sum_{k=1}^{q}k\sin(kx)\Big]+O(m^{1/2})\end{aligned}$$

$$=O(m^{1/2})$$

uniformly in $x \in [a, b]$, since, for example,

$$\left| \sum_{k=1}^{q} \cos(kx) \sum_{j=0}^{a_k} \sin(cjx/2k) \sin[(k^2-b_k)x/2k] \right| = O(m^{1/2}) ,$$

$$2 \sum_{0}^{n} \cos(jz) = 1 + \cos(nz) + \sin(nz) \sin z/(1-\cos z)$$

and

$$\left| \sum_{1}^{q} k \cos(kx) \right| \leqq q \max_{1 \leqq j \leqq q} \left| \sum_{k=j}^{q} \cos(kx) \right| .$$

Now, $\Delta(m, x) = \left| \sum_{1}^{m} A[(cn+d)^{1/2}x] - S(m) \right|$ is dominated by the number of indices n such that $[p(m)-1]^2 < cn+d \leqq cm+d$. Therefore $\Delta(m, x) < 2p(m)/c+1$, and Lemma 2 follows immediately.

From the result (2.2) and the expansion (16_1) of [8], p. 326 we see that

$$a_{2m}h_{2m}(x) = b_m \left\{ \cos[(4m+1)^{1/2}x] + (4m+1)^{-1/2} \frac{1}{6} x^3 \sin[(4m+1)^{1/2}x] \right\} + O(m^{-2})$$

uniformly in $x \in [a, b]$, where $b_m = -(4\pi)^{-1} m^{-3/2} f'(0)[1+O(m^{-1})]$. Using Abel's method of summation and Lemma 2 we find that

$$\sum_{m+1}^{\infty} a_{2j}h_{2j}(x) = \sum_{m+1}^{\infty} (b_j - b_{j+1}) U_j(x) + \frac{1}{6} x^3 \sum_{m+1}^{\infty} [b_j(4j+1)^{-1/2} - b_{j+1}(4j+5)^{-1/2}] V_j(x) + O(m^{-1}) , \tag{2.3}$$

where $U_m(x) = \sum_{1}^{m} \cos[(4j+1)^{1/2}x]$ and $V_m(x) = \sum_{1}^{m} \sin[(4j+1)^{1/2}x]$. The constant coefficients in the series on the right in (2.3) are $O(j^{-5/2})$, and so in view of Lemma 2,

$$f(x) - \mathrm{E}[\hat{f}_{n1}(x; m)] = \sum_{m+1}^{\infty} a_{2j}h_{2j}(x) = O(m^{-1})$$

uniformly in $a \leqq x \leqq b$. Therefore $\int_a^b [\mathrm{E}\,\hat{f}_{n1}(x; m) - f(x)]^2 dx = O(m^{-2})$, and from the proof of Theorem 1 we find that $\int_a^b \operatorname{var}[\hat{f}_{n1}(x; m)] dx = O(m^{1/2}/n)$, completing the proof of Theorem 2.

3. Estimation by Laguerre polynomial expansion

Suppose $\alpha > -1$ and let $c_m^{(\alpha)} = [m!/\Gamma(m+\alpha+1)]^{1/2}$. The (generalized) orthonormal Laguerre system on $(0, \infty)$ is given by

$$l_m^{(\alpha)}(x)=c_m^{(\alpha)}L_m^{(\alpha)}(x)x^{\alpha/2}e^{-x/2}\,, \qquad m\geqq 0\,,$$

where the $L_m^{(\alpha)}$ are the (generalized) Laguerre polynomials. The even and odd numbered Hermite polynomials have a simple expression in terms of $L_m^{(-1/2)}$ and $L_m^{(1/2)}$, respectively, but the different polynomial classes lead to estimators with quite different properties.

Suppose that the bounded density f has its support confined to $(0, \infty)$, and is continuous and of bounded variation on each interval $(0, \lambda)$, $\lambda>0$. Then the Laguerre expansion of f, defined by

$$f(x)=\sum_0^\infty a_m^{(\alpha)}l_m^{(\alpha)}(x)$$

with $a_m^{(\alpha)}=\int_0^\infty f(x)l_m^{(\alpha)}(x)dx$, converges to f at each point $x>0$. (See [8], p. 384.) If $X_1, X_2, \cdots, X_m$ is an independent sample with common density f then an unbiased estimate of $a_m^{(\alpha)}$ is

$$\hat{a}_m^{(\alpha)}=n^{-1}\sum_{i=1}^n l_m^{(\alpha)}(X_i)\,,$$

and an estimate of f is

$$\hat{f}_{n\alpha}(x\,;\,m)=\sum_{j=0}^m \hat{a}_j^{(\alpha)}l_j^{(\alpha)}(x)\,.$$

Let $l=l(\alpha)$ be the smallest integer strictly greater than $(13+28|\alpha-1/2|)/6$, suppose f is differentiable on $(0, \infty)$ and define

$$g_\alpha(x)=xf(x)+2f'(x)-\alpha f(x)\,.$$

THEOREM 3. *Suppose f is differentiable and of bounded variation on $(0, \infty)$, g_α is of bounded variation on $(0, 1)$, $x^{-1/4}g_\alpha(x)$ is of bounded variation on $(1, \infty)$, and*

$$\mathrm{E}\,(X_1^{3l-1/2})<\infty \qquad \text{and} \qquad \int_0^\infty x^{3l/2}|g_\alpha(x)|dx<\infty\,.$$

Then for $\alpha>-1$,

$$\int_0^\infty \mathrm{E}\,[\hat{f}_{n\alpha}(x\,;\,m)-f(x)]^2dx$$
$$=(n\pi)^{-1}m^{1/2}\int_0^\infty x^{-1/2}f(x)dx+O(m^{-1})+o(n^{-1}m^{1/2})$$

as m and $n\to\infty$.

The moment restrictions may be relaxed by a more judicious choice of ε in the expansion (32) of [8], p. 345; see the techniques in the proof below. However, this leads to a very complex expression for the optimal l.

The rate of convergence obtained is the best possible for distributions satisfying the conditions of the theorem. The term $O(m^{-1})$ stands for $\sum_{m+1}^{\infty} [a_j^{(\alpha)}]^2$, and if $a_m^{(\alpha)}$ behaves "regularly" as $m\to\infty$, the rate can only be reduced if $a_m^{(\alpha)}$ converges to zero at a faster rate than $O(m^{-1})$. However, consider the density $f(x)=(1/2)e^{-x/2}$, $x>0$, for which

$$\begin{aligned}2[c_m^{(\alpha)}]^{-1}a_m^{(\alpha)}&=\int_0^\infty L_m^{(\alpha)}(x)x^{\alpha/2}e^{-x}dx\\&=\Gamma\left(\frac{1}{2}\alpha+1\right)\Gamma(m+\alpha+1)_2F_1\left(-m,\frac{1}{2}\alpha+1;\alpha+1,1\right)\Big/m!\Gamma(\alpha+1)\\&=\frac{1}{2}\alpha\Gamma\left(m+\frac{1}{2}\alpha\right)\Big/m!\sim\frac{1}{2}\alpha m^{\alpha/2-1}\end{aligned}$$

if $\alpha\neq 0$. (We refer the reader to Magnus, Oberhettinger and Soni [4], p. 245, Abramowitz and Stegun [1], p. 556 for standard results on Laplace transforms and hypergeometric functions.) But $c_m^{(\alpha)}\sim m^{-\alpha/2}$, and so $a_m^{(\alpha)}\sim\alpha/4m$. Therefore for densities satisfying the conditions of Theorem 3, the rate of convergence of the mean integrated square error may be no better than $O(n^{-2/3})$.

PROOF OF THEOREM 3. The proof is largely contained in two lemmas. First we establish

LEMMA 3. *Under the conditions of Theorem 3, $a_m^{(\alpha)}=O(m^{-1})$ for any $\alpha>-1$.*

PROOF. (We shall often omit the notation α.) Integrating by parts and using the recursive properties of Laguerre polynomials we obtain

$$c_m^{-1}a_m=-(2m)^{-1}\int_0^\infty g(x)x^{\alpha/2}e^{-x/2}L_{m-1}^{(\alpha+1)}(x)dx\ . \tag{3.1}$$

Let $N=4m-3$, $\beta=(2\alpha+3)\pi/4$ and choose $0<\tau<\varepsilon=1/6$ such that

$$l>\left[\left(\frac{13}{2}\right)\left(\frac{1}{3}-\tau\right)+2\left|\alpha-\frac{1}{2}\right|\left(\frac{4}{3}-\tau\right)\right]\Big/3\tau\ .$$

From the expansion (32) of [8], p. 345 we deduce that for constants $\delta_m\to 0$,

$$\begin{aligned}c_m x^{\alpha/2}e^{-x/2}L_{m-1}^{(\alpha+1)}(x)\\=(1+\delta_m)\pi^{-1/2}m^{1/4}x^{-3/4}\{\cos[(Nx)^{1/2}-\beta]+p_1(x)(Nx)^{-1/2}\\\times\sin[(Nx)^{1/2}-\beta]+R(m,x,\alpha)(Nx)^{-1}p_2(x^{1/2})\}\ ,\end{aligned} \tag{3.2}$$

where p_1 and p_2 are polynomials of degrees 2 and $3l+8$, respectively, and $|R(m, x, \alpha)|\leqq 1$ for $N^{-1}\leqq x\leqq N^{1/3}$. Let $A(z)$ denote either $\cos z$ or $\sin z$, and $B(z)=A(\pi/2-z)$.

Let $p_0(x)\equiv 1$. For $j=0$ or 1 we estimate

$$I(N)=\int_{N^{-1}}^{1} x^{-3/4}(Nx)^{-j/2}p_j(x)g(x)A[(Nx)^{1/2}]dx .$$

Noting that p_jg is of bounded variation on (0, 1), and using the second mean value theorem, we deduce that

$$I(N)=O(m^{-1/4}) . \tag{3.3}$$

A similar argument can be used to show that with $M=N^{1/3}$,

$$J(N)=\int_{1}^{M} x^{-3/4}g(x)A[(Nx)^{1/2}]dx=O(m^{-1/2}) . \tag{3.4}$$

It is easily seen that

$$\left|\int_{1}^{M} x^{-3/4}(Nx)^{-1/2}p_1(x)g(x)A[(Nx)^{1/2}]dx\right|=O(m^{-1/2}) , \tag{3.5}$$

$$\left|\int_{N^{-1}}^{1} x^{-3/4}R(m, x, \alpha)(Nx)^{-1}p_2(x^{1/2})g(x)dx\right|=O(m^{-1/4}) , \tag{3.6}$$

$$\left|\int_{1}^{M} x^{-3/4}R(m, x, \alpha)(Nx)^{-1}p_2(x^{1/2})g(x)dx\right|=O(m^{-1/4}) . \tag{3.7}$$

Combining (3.2)–(3.7) we find that

$$c_m\int_{N^{-1}}^{M} x^{\alpha/2}e^{-x/2}L_{m-1}^{(\alpha+1)}(x)g(x)dx=O(1) . \tag{3.8}$$

Since $|L_{m-1}^{(\alpha+1)}(x)|\leqq Cm^{\alpha+1}e^{x/2}$ (Abramowitz and Stegun [1], p. 786) then

$$\left|c_m\int_{0}^{N^{-1}} x^{\alpha/2}e^{-x/2}L_{m-1}^{(\alpha+1)}(x)g(x)dx\right|=O(1) , \tag{3.9}$$

and since $\sup_{x\geqq 1} e^{-x/2}x^{\alpha/2+7/12}|L_{m-1}^{(\alpha+1)}(x)|\leqq Cm^{\alpha/2+1/4}$ (Szegö [10], p. 235) then

$$\left|c_m\int_{M}^{\infty} x^{\alpha/2}e^{-x/2}L_{m-1}^{(\alpha+1)}(x)g(x)dx\right|=o(1) . \tag{3.10}$$

From (3.8)–(3.10) it follows that $c_m\int_0^\infty x^{\alpha/2}e^{-x/2}L_{m-1}^{(\alpha+1)}(x)g(x)dx=O(1)$ as $m\rightarrow\infty$, and Lemma 3 follows immediately from (3.1).

LEMMA 4. *Under the conditions of Theorem 3,*

$$\int_0^\infty f(x)l_m^{(\alpha)}(x)^2dx\sim(2\pi)^{-1}m^{-1/2}\int_0^\infty x^{-1/2}f(x)dx$$

for any $\alpha>-1$.

PROOF. From the expansion (32) of [8], p. 345 we see that for constants $\delta_m \to 0$,

$$(3.11) \qquad l_m^{(\alpha)}(x)^2 = \frac{1}{2}(1+\delta_m)\pi^{-1}m^{-1/2}x^{-1/2}\{1+\cos[2(Nx)^{1/2}-2\beta] + (Nx)^{-1/2}R_1q_1(x) + (Nx)^{-1}R_2q_2(x^{1/2}) + (Nx)^{-3/2}R_3q_3(x^{1/2}) + (Nx)^{-2}R_4q_4(x^{1/2})\}\ ,$$

where $N=4m+1$, $\beta=(2\alpha+1)\pi/4$, q_1 through q_4 are polynomials of degrees 2, $3l+8$, $3l+12$ and $6l+16$, respectively, and $|R_i|=|R_i(m, x, \alpha)|\leqq 1$ for $N^{-1}\leqq x\leqq N^{1/3}$. (Here ε, τ and l are chosen as in the proof of Lemma 3. In the case $\alpha\leqq -1/2$, note the comments of [8], pp. 346–348.) Let $M=N^{1/3}$ and note that $(Nx)^{-j/2}=(x/M)^{3j/2}x^{-2j}$. It is readily seen that

$$\int_1^M x^{-1/2}f(x)\{(Nx)^{-1/2}|R_1q_1(x)|+(Nx)^{-j/2}|R_jq_j(x^{1/2})|\}dx=o(1)$$

as $M\to\infty$ for $j=2$, 3 and 4, and

$$\int_{N^{-1}}^1 x^{-1/2}(Nx)^{-j/2}|R_jq_j|f(x)dx\leqq CN^{-j/2}\int_{N^{-1}}^1 x^{-(j+1)/2}dx=O(m^{-1/2})\ .$$

Writing $A(z)$ for $\cos z$ or $\sin z$ we have

$$\int_{N^{-1}}^M x^{-1/2}A[2(Nx)^{1/2}]f(x)dx=2N^{-1/2}\int_1^{(NM)^{1/2}} A(2y)f(y^2/N)dy=O(m^{-1/2})$$

by the second mean value theorem. Using the expansion (3.11) and the estimates above we see that

$$(3.12) \qquad \int_{N^{-1}}^M f(x)l_m^{(\alpha)}(x)^2dx=(2\pi)^{-1}m^{-1/2}\int_{N^{-1}}^M x^{-1/2}f(x)dx+o(m^{-1/2})\ .$$

Since $|L_m^{(\alpha)}(x)|\leqq Cm^\alpha e^{x/2}$ then

$$(3.13) \qquad \int_0^{N^{-1}} f(x)l_m^{(\alpha)}(x)^2dx\leqq Cm^\alpha\int_0^{N^{-1}} x^\alpha dx=O(m^{-1})\ ,$$

and since $\sup_{x\geqq 1} x^{\alpha/2+1/12}e^{-x/2}|L_m^{(\alpha)}(x)|\leqq Cm^{\alpha/2-1/4}$ ([10], p. 235) then

$$(3.14) \qquad \int_M^\infty f(x)l_m^{(\alpha)}(x)^2dx\leqq Cm^{-1/2}\int_M^\infty x^{-1/6}f(x)dx=o(m^{-1/2})\ .$$

Combining (3.12)–(3.14) we deduce Lemma 4.

It follows from Parseval's equality and Lemma 3 that

$$\int_0^\infty [\mathrm{E}\,\hat{f}_{n\alpha}(x;m)-f(x)]^2dx=\sum_{m+1}^\infty a_j^2=O(m^{-1})\ ,$$

and from Lemma 4 that $n\int_0^\infty \operatorname{var}[\hat{f}_{n\alpha}(x; m)]dx \sim \pi^{-1}m^{1/2}$, proving Theorem 3.

AUSTRALIAN NATIONAL UNIVERSITY

REFERENCES

[1] Abramowitz, M. and Stegun, I. A. (1965). *Handbook of Mathematical Functions*, Dover, New York.

[2] Cencov, N. N. (1962). Evaluation of an unknown distribution density from observations, *Soviet Math. Dokl.*, **3**, 1559-1562.

[3] Kronmal, R. and Tarter, M. (1968). The estimation of probability densities and cumulatives by Fourier series methods, *J. Amer. Statist. Ass.*, **63**, 925-952.

[4] Magnus, W., Oberhettinger, F. and Soni, R. P. (1966). *Formulas and Theorems for the Special Functions of Mathematical Physics*, Springer, Berlin.

[5] Parzen, E. (1962). On the estimation of a probability density function and mode, *Ann. Math. Statist.*, **33**, 1065-1076.

[6] Rosenblatt, M. (1956). Remarks on some non-parametric estimates of a density function, *Ann. Math. Statist.*, **27**, 832-837.

[7] Rosenblatt, M. (1971). Curve estimates, *Ann. Math. Statist.*, **42**, 1815-1841.

[8] Sansone, G. (1959). *Orthogonal Functions*, Interscience, New York.

[9] Schwartz, S. C. (1967). Estimation of a probability density by an orthogonal series, *Ann. Math. Statist.*, **38**, 1262-1265.

[10] Szegö, G. (1939). *Orthogonal Polynomials*, Amer. Math. Soc. Coll. Publ., **32**, Providence.

[11] Walter, G. G. (1977). Properties of Hermite series estimation of probability density, *Ann. Statist.*, **5**. 1258-1264.

[12] Watson, G. S. (1969). Density estimation by orthogonal series, *Ann. Math. Statist.*, **38**, 1262-1265.

[13] Watson, G. S. and Leadbetter, M. R. (1963). On the estimation of the probability density, I, *Ann. Math. Statist.*, **34**, 480-491.