The Annals of Statistics
1975, Vol. 3, No. 1, 30–48

INTERPOLATING SPLINE METHODS FOR DENSITY ESTIMATION I. EQUI-SPACED KNOTS[1]

BY GRACE WAHBA

University of Wisconsin

Statistical properties of a variant of the histospline density estimate introduced by Boneva–Kendall–Stefanov are obtained. The estimate we study is formed for x in a finite interval, $x \in [a, b] = [0, 1]$ say, by letting $\hat{F}_n(x)$, $x \in [0, 1]$ be the unique cubic spline of interpolation to the sample cumulative distribution function $F_n(x)$ at equi-spaced points $x = jh$, $j = 0, 1, \cdots, l+1$, $(l+1)h = 1$, which satisfies specified boundary conditions $\hat{F}_n'(0) = a$, $\hat{F}_n'(1) = b$. The density estimate $\hat{f}_n(x)$ is then $\hat{f}_n(x) = d/dx\,\hat{F}_n(x)$. It is shown how to estimate a and b. A formula for the optimum h is given. Suppose f has its support on [0, 1] and $f^{(m)} \in \mathscr{L}_p[0, 1]$. Then, for $m = 1, 2, 3$ and certain values of p, it is shown that

$$E(f_n(x) - f(x))^2 = O(n^{-(2m-2/p)/(2m+1-2/p)}).$$

Bounds for the constant covered by the "O" are given. An extension to the $\mathscr{L}_p$ case of known convergence properties of the derivative of an interpolating spline is found, as part of the proofs.

1. Introduction. In this paper we study the convergence properties of a histospline density estimate of the type introduced by Boneva, Kendall and Stefanov (BKS) [3] and discussed by Schoenberg [16], [17]. Although BKS considered the estimation of densities supported on the entire real line as well as on a finite interval, we consider here only densities supported on a finite interval, say, [0, 1].

Let $W_p^{(m)}$ be the Sobolev space of functions

$$\{f : f^{(\nu)} \text{ abs. cont.}, \nu = 0, 1, \cdots, m-1, f^{(m)} \in \mathscr{L}_p[0, 1]\}.$$

Let $h > 0$ satisfy $1/h = l + 1$, where l is a positive integer. Let h_j be the fraction of independent observations from some density f, falling between jh and $(j+1)h$, $j = 0, 1, \cdots, l$. As a density estimate $\hat{f}$, BKS seek the unique function in the space $W_2^{(1)}$ which minimizes

$$\int_0^1 (g'(x))^2\,dx \tag{1.1}$$

subject only to

$$\int_{jh}^{(j+1)h} g(x)\,dx = h_j, \qquad j = 0, 1, \cdots, l. \tag{1.2}$$

Let $\hat{F}$ be the unique function in the space $W_2^{(2)}$ which minimizes

$$\int_0^1 (G''(x))^2\,dx \tag{1.3}$$

Received April 1973; revised June 1974.

[1] This report was supported by the Air Force Office of Scientific Research under Grant No. AFOSR 72-2363.

Key words and phrases. Spline, histospline, density estimate, optimal convergence rates.

subject to

$$(1.4) \qquad \begin{aligned} G(0) &= 0 \\ G(ih) &= \textstyle\sum_{j=0}^{i-1} h_j\,, \qquad i = 1, 2, \cdots, l+1\,. \end{aligned}$$

Clearly, $\hat{F}$ satisfies $\hat{F}' = \hat{f}$, where $\hat{f}$ is the solution associated with the problem of (1.1) and (1.2). Therefore this BKS histospline density estimate is the derivative of the so-called natural cubic spline of interpolation $\hat{F}$ to the sample cdf F_n at the points ih, $i = 0, 1, \cdots, l + 1$. See Schoenberg [15], Section 3 for a discussion of the natural cubic splines of interpolation. The spline $\hat{F}$ is a cubic polynomial in each interval $[jh, (j + 1)h]$, $j = 0, 1, \cdots, l$, uniquely characterized on [0, 1] by (1.4) and the conditions $\hat{F}$, $\hat{F}'$, $\hat{F}''$ continuous and $F''(0) = F''(1) = 0$. The drawback to using the natural cubic spline is that maximum possible convergence rates in the cases $m = 2$ and $m = 3$ defined below will not obtain in a neighborhood of the boundaries unless F also satisfies $F''(0) = F''(1) = 0$. (See [9].) Another histospline was considered by BKS and subsequently Schoenberg [16], [17]. It is the derivative of the solution $\hat{F}$ to the problem: Find the unique function in the space $W_2^{(2)}$ which minimizes (1.3) subject to (1.4) and the additional conditions

$$(1.5) \qquad \begin{aligned} G'(0) &= 0 \\ G'(1) &= 0\,. \end{aligned}$$

The histospline we study is a variation of this. We replace (1.5) by

$$(1.6) \qquad \begin{aligned} G'(0) &= \hat{a}_1 \\ G'(1) &= \hat{b}_1 \end{aligned}$$

where $\hat{a}_1$ and $\hat{b}_1$ are estimates of $f(0)$ and $f(1)$ formed from the sample cdf in a manner to be described. Thus, if f has its support on [0, 1] we let $\hat{F}_n$ be the solution to the problem: Find the unique function in the space $W_2^{(2)}$ which minimizes (1.3) subject to

$$\begin{aligned} G(jh) &= F_n(jh)\,, \qquad j = 0, 1, \cdots, l+1\,, \\ G'(0) &= \hat{a}_1 \\ G'(1) &= \hat{b}_1 \end{aligned}$$

and our density estimate $\hat{f}_n(x)$ is

$$(1.7) \qquad \hat{f}_n(x) = \frac{d}{dx}\hat{F}_n(x)\,.$$

This estimate can be less smooth than the first (unconstrained) BKS estimate. Of course, we may replace $\hat{a}_1$ and $\hat{b}_1$ by 0 if it is known that $f(0) = 0$, $f(1) = 0$.

As do BKS, ([3], pages 34–35), we consider that h is a parameter to be chosen. Our criterion is minimum mean square error at a point. Our theorems provide results on the optimum choice of h.

Let $W_p^{(m)}(M)$ be given by

$$W_p^{(m)}(M) = \{f : f \in W_p^{(m)}, ||f^{(m)}||_p \leqq M\}$$

where

$$||f^{(m)}||_p = [\int |f^{(m)}(\xi)|^p \, d\xi]^{1/p} , \qquad p \geqq 1$$
$$||f^{(m)}||_\infty = \sup_\xi |f^{(m)}(\xi)| .$$

It is shown in Wahba [21], based on a result of Farrell [5], that if $\hat{f}_n(x)$, $n = 1, 2, \cdots$ is any sequence of estimates of the true density f at the point x, and ε is any positive number, that if

$$\sup_{f \in W_p^{(m)}(M)} E(\hat{f}_n(x) - f(x))^2 = b_n n^{-(2m-2/(p+\varepsilon))/(2m+1-2/(p+\varepsilon))}$$

then there exists $D_0 > 0$ such that $b_n \geqq D_0$ for infinitely many n. Thus, the best possible mean square convergence rate uniform over $W_p^{(m)}(M)$ is not better than $n^{-(2m-2/(p+\varepsilon))/(2m+1-2/(p+\varepsilon))}$ for arbitrarily small ε. It is known for various density estimates that the rate $n^{-(2m-2/p)/(2m+1-2/p)}$ is achieved, that is,

$$\sup_{f \in W_p^{(m)}(M)} E(\hat{f}_n(x) - f(x))^2 \leqq Dn^{-(2m-2/p)/(2m+1-2/p)} , \tag{1.8}$$

where D depends on m, p, M, the method (and, possibly, bounds on f). See, for example [21], where the Parzen kernel type estimate ([12]), the Kronmal-Tarter orthogonal series estimate ([11]) and the polynomial algorithm for density estimation ([20]) are studied. The ordinary histogram with optimally chosen "bin" size also satisfies (1.8) with $m = 1$.

It is the purpose (and main theorem) of this note to prove that the histospline density estimate of (1.6) with optimally chosen h shares the rate of convergence property (1.8) of these other estimates. The result (1.8) is proved for $m = 1, 2$ and 3, and several sets of values for p. An upper bound for D is given. In particular, for $m = 1$, $p = 2$, $h_{\text{opt}} \sim cn^{-\frac{1}{4}}$. Recent work indicates that the $m = 1$, $p = 2$ result is true for the BKS histospline of (1.1) and (1.2). (R. Kuhn personal communication.)

In order to achieve the rate (1.8) for higher m, it is necessary (and doubtless, sufficient) to use higher degree splines. A proof of (1.8) for $m > 3$ is not forthcoming at this time, however, due to the complexity of the formula for higher degree splines.

2. Explicit expressions for the histospline estimate, and outline of proof of the main theorem. Our development of an explicit formula for an interpolating spline will be slightly unorthodox, for the purpose of easing the proofs of the main theorem. The reader may consult [2], [6], [7], [15] and the bibliography [14] for additional background on splines.

We endow $W_2^{(2)}$ with the inner product

$$\langle F, G\rangle = F(0)G(0) + F'(0)G'(0) + \int_0^1 F''(x)G''(x) \, dx . \tag{2.1}$$

$W_2^{(2)}$ is then a reproducing kernel Hilbert space (RKHS) with the reproducing kernel

$$
\begin{aligned}
Q(s,t) &= 1 + st + \int_0^{\min(s,t)} (s-u)(t-u)\,du \\
&= 1 + st + \left(\frac{ts^2}{2} - \frac{s^3}{6}\right), && s < t \qquad (2.2)\\
&= 1 + st + \left(\frac{st^2}{2} - \frac{t^3}{6}\right), && s > t\,.
\end{aligned}
$$

Denote this Hilbert space $\mathscr{H}_Q$, with norm $\|\cdot\|_Q$. The true cdf F is always assumed to be in $\mathscr{H}_Q$, that is, $f = F' \in W_2^{(1)}$.

Let Q_t be the function defined on [0, 1] by

$$Q_t(s) = Q(s,t)\,, \qquad s, t \in [0,1]\,, \tag{2.3}$$

and let Q_t' be the function defined on [0, 1] by

$$
\begin{aligned}
Q_t'(s) &= \frac{d}{du} Q(s,u)\Big|_{u=t}\,, && s, t \in [0,1] \\
&= s + \frac{s^2}{2}\,, && s \leqq t \qquad (2.4)\\
&= s + st - \frac{t^2}{2}\,, && s \geqq t\,.
\end{aligned}
$$

By the properties of RKHS (see, e.g. [10]), $Q_t, Q_{t'} \in \mathscr{H}_Q$ for each t, and

$$
\begin{aligned}
\langle G, Q_t\rangle &= G(t)\,, && G \in \mathscr{H}_Q,\ t \in [0,1] \qquad (2.5)\\
\langle G, Q_t'\rangle &= G'(t)\,.
\end{aligned}
$$

Consider the solution to the problem:
Find $G \in \mathscr{H}_Q$ to min $\|G\|_Q$ subject to

$$
\begin{aligned}
G'(0) &= \langle G, Q_0'\rangle = a_1 \\
G(0) &= \langle G, Q_0\rangle = a_0 \\
G(s_i) &= \langle G, Q_{s_i}\rangle = y_i\,, && i = 1, 2, \cdots, l \qquad (2.6)\\
G(1) &= \langle G, Q_1\rangle = b_0 \\
G'(1) &= \langle G, Q_1'\rangle = b_1
\end{aligned}
$$

where $0 < s_1 < \cdots < s_l < 1$.
Denoting $\bar{s} = (s_1, s_2, \cdots, s_l)$, $\bar{y} = (y_1, y_2, \cdots, y_l)$, $\bar{a} = (a_1, a_0)$, $\bar{b} = (b_0, b_1)$, let $S(x) = S(x; \bar{s}; \bar{a}, \bar{y}, \bar{b})$ be the solution to this problem. Then, by observing that $S \in \mathscr{S}_l(\bar{s})$ defined by

$$\mathscr{S}_l(\bar{s}) = \text{span}\,\{Q_0', Q_{s_i}, i = 0, 1, \cdots, l+1, Q_1'\} \tag{2.7}$$

where $s_0 = 0$, $s_{l+1} = 1$, it may be established that

$$
\begin{aligned}
&S(x; \bar{s}; \bar{a}, \bar{y}, \bar{b}) \qquad (2.8)\\
&\quad = (Q_0'(x), Q_0(x), Q_{s_1}(x), \cdots, Q_{s_l}(x), Q_1(x), Q_1'(x))Q_{l+4}^{-1}(\bar{a}; \bar{y}; \bar{b})'
\end{aligned}
$$

where Q_{l+4} is the $(l+4) \times (l+4)$ Grammian matrix of the basis for $\mathscr{S}_l(\bar{s})$.

Q_{l+4} is of full rank (see for example [19]) and the entries may be found from (2.5). By observing the nature of the inner product in $\mathscr{H}_Q$, it is easily seen that $S(x; \bar{s}; \bar{a}, \bar{y}, \bar{b})$ is also the solution to: Find $G \in \mathscr{H}_Q$ to

$$\min \int_0^1 (G''(x))^2\, dx$$

subject to (2.6). The solution to this problem is well known ([15]) to be the unique cubic spline satisfying (2.6). It may easily be checked from (2.3), (2.4) and (2.8) that S has the characteristic properties of a cubic spline, viz. S is a polynomial of degree less than or equal to three in each interval $[s_i, s_{i+1}]$, $i = 0, 1, \cdots, l$, and S, S' and S'' are continuous.

The density estimate $\hat{f}_n(x)$ that we study is thus given by

$$\hat{f}_n(x) = \frac{d}{dx}\hat{F}_n(x)\,, \tag{2.9}$$

$$\hat{F}_n(x) = S(x; \bar{s}_h; \hat{a}, \bar{F}_n, \hat{b})$$

with

$$\begin{aligned} \bar{s}_h &= (h, 2h, \cdots, lh)\,, \qquad (l+1)h = 1 \\ \hat{a} &= (\hat{a}_1, 0) \\ \bar{F}_n &= (F_n(h), F_n(2h), \cdots, F_n(lh)) \\ \hat{b} &= (1, \hat{b}_1)\,. \end{aligned}$$

Equation (2.8) is not the computationally best method for computing $\hat{F}_n$, because Q_{l+4} is ill-conditioned for large l; however, computing routines for $S(x)$ and $S'(x)$ are commonly available. See, for example, [1]. The estimates $\hat{a}_1$ and $\hat{b}_1$ depend on m, ($= 1, 2$, or 3) and are defined as follows: Let $l_{0,\nu}(x)$ be the polynomial of degree m satisfying

$$\begin{aligned} l_{0,\nu}(x) &= 1\,, \qquad x = \nu h\,, \\ &= 0\,, \qquad x = jh\,, \quad j \neq \nu\,, \quad j = 0, 1, \cdots, m\,. \end{aligned} \tag{2.10}$$

and let $l_{1,\nu}(x)$ be the polynomial of degree m satisfying

$$\begin{aligned} l_{1,\nu}(x) &= 1\,, \qquad x = 1 - \nu h \\ &= 0\,, \qquad x = 1 - jh\,, \quad j \neq \nu\,, \quad j = 0, 1, \cdots, m\,. \end{aligned} \tag{2.11}$$

Let

$$\hat{a}_1 = \hat{F}_n'(0) = \hat{f}_n(0) = \frac{d}{dx}\sum_{\nu=0}^{m} l_{0,\nu}(x)\Big|_{x=0} F_n(\nu h) \tag{2.12}$$

$$\hat{b}_1 = \hat{F}_n'(1) = \hat{f}_n(1) = \frac{d}{dx}\sum_{\nu=0}^{m} l_{1,\nu}(x)\Big|_{x=1} F_n(1 - \nu h)\,. \tag{2.13}$$

$\hat{a}_1$ is the derivative at 0 of the mth degree polynomial interpolating the sample cdf at $0, h, \cdots, mh$, and similarly for $\hat{b}_1$. It follows from (2.8) that $S(x; \bar{s}; \bar{a}, \bar{y}, \bar{b})$ is linear in the entries of $\bar{a}$, $\bar{y}$, and $\bar{b}$, that is

$$S(x; \bar{s}; \bar{a} + \bar{\varepsilon}_a, \bar{y} + \bar{\varepsilon}, \bar{b} + \bar{\varepsilon}_b) = S(x; \bar{s}; \bar{a}, \bar{y}, \bar{b}) + S(x; \bar{s}; \bar{\varepsilon}_a, \bar{\varepsilon}, \bar{\varepsilon}_b)$$

where $\bar{\varepsilon}_a$, $\bar{\varepsilon}$ and $\bar{\varepsilon}_b$ are 2-, l- and 2-vectors, respectively. $(d/dx)S(x; \bar{s}, \bar{a}, \bar{y}, \bar{b})$ also has this linearity property.

Let F be the true cdf, and let $\tilde{F}$ be the cubic spline of interpolation to F, with knots jh, $j = 1, 2, \cdots, l$, and matching F and F' at the boundaries, that is

$$\tilde{F}(x) = S(x; \bar{s}_h; \bar{F}_a, \bar{F}_h, \bar{F}_b)$$

where

$$\begin{aligned} \bar{F}_a &= (F'(0), 0) \\ \bar{F}_h &= (F(h), F(2h), \cdots, F(lh)) \\ \bar{F}_b &= (1, F'(1)) . \end{aligned}$$

(Note, by the nature of the minimization problem (2.6), that $\tilde{F}$ is the projection of F onto $\mathscr{S}_l(\bar{s})$.)

We may write

$$f(x) - \hat{f}_n(x) = \frac{d}{dx}(F(x) - \hat{F}_n(x))$$

$$= \frac{d}{dx}(F(x) - \tilde{F}(x)) + \frac{d}{dx}(\tilde{F}(x) - \hat{F}_n(x)) \tag{2.14}$$

$$= \frac{d}{dx}(F(x) - \tilde{F}(x)) + \frac{d}{dx}H_n(x) , \tag{2.15}$$

where

$$H_n(x) = S(x; \bar{s}_h; \bar{\varepsilon}_a, \bar{\varepsilon}, \bar{\varepsilon}_b) \tag{2.16}$$

and

$$\begin{aligned} &\bar{\varepsilon}_a = (\varepsilon_0', 0) , \qquad \varepsilon_0' = F'(0) - \hat{a}_1 , \\ &\bar{\varepsilon} = (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_l) , \qquad \varepsilon_j = F(jh) - F_n(jh) , \qquad j = 1, 2, \cdots, l , \\ &\bar{\varepsilon}_b = (0, \varepsilon'_{l+1}) , \qquad \varepsilon'_{l+1} = F'(1) - \hat{b}_1 . \end{aligned} \tag{2.17}$$

The first term on the right of (2.15), which we shall call the bias term, is nonrandom and depends only on how well F can be approximated by an interpolating cubic spline. The second, or variance term is a (linear) function of the random variables ε_0', ε'_{l+1} and ε_i, $i = 1, 2, \cdots, l$.

Then, as usual,

$$E(f(x) - \hat{f}_n(x))^2 \leqq 2\left(\frac{d}{dx}(F(x) - \tilde{F}(x))\right)^2 + 2E\left(\frac{d}{dx}H_n(x)\right)^2 . \tag{2.18}$$

Bounds on the absolute bias, $|d/dx(F(x) - \tilde{F}(x))|$ appear in the approximation theory literature in various forms, for equally spaced, as well as arbitrarily spaced knots. If $F^{(m+1)} \in \mathscr{L}_2[0, 1]$, then it is known ([18], Theorems 5.1, 5.2 and 5.3) that, for $m = 1, 2, 3$,

$$\sup_x \left|\frac{d}{dx}(F(x) - \tilde{F}(x))\right| \leqq K_2(m)\|F^{(m+1)}\|_2 h^{m-\frac{1}{2}} \tag{2.19}$$

where $\|\cdot\|_p$ is the $\mathscr{L}_p$ norm, and $K_2(m)$ is a constant depending on m. Generalizations of (2.19) to arbitrary m are given when $\tilde{F}$ is replaced by an interpolating spline of higher degree. For $F^{(m+1)} \in \mathscr{L}_\infty[0, 1]$, $m = 3$, ([9]) gives

$$\sup_x \left|\frac{d}{dx}(F(x) - \tilde{F}(x))\right| \leqq K_\infty(m)\|F^{(m+1)}\|_\infty h^m , \tag{2.20}$$

and [9] is easily extendable to $m = 1, 2$. (For earlier results, see [13].) Some information about generalizations of (2.20) up to, but not beyond $m = 5$ are known ([4]). We would like to have the result

$$F^{(m+1)} \in \mathscr{L}_p \Rightarrow \sup_x \left|\frac{d}{dx}(F(x) - \tilde{F}(x))\right| \leqq K_p(m)\|F^{(m+1)}\|_p h^{m-1/p} , \quad p \geqq 1 , \tag{2.21}$$

or, equivalently, $f \in W_p^{(m)} \Rightarrow$

$$\sup_x \left(\frac{d}{dx}(F(x) - \tilde{F}(x))\right)^2 \leqq \frac{A'}{2}\|f^{(m)}\|_p^2 h^{2m-2/p} , \qquad p \geqq 1 \tag{2.22}$$

where $A' = A'(m, p)$. We are not aware of such results for $p \neq 2$ or ∞. We provide a proof of (2.22) good for $m = 2, 3$, $1 \leqq p \leqq 2$. In the proof, the dependency on the knots $\{s_i\}_{i=1}^{l}$ is retained so that the results may be used in a sequel paper where the knots are determined by the order statistics. Combining these results will give us a bound on the bias for

$$\begin{aligned} m &= 1 , & p &= 2 , \ \infty \\ m &= 2 , & 1 &\leqq p \leqq 2 , \ \infty \\ m &= 3 , & 1 &\leqq p \leqq 2 , \ \infty . \end{aligned}$$

The establishment of bounds on the variance term is tedious for cubic splines, and we are unable to do it for higher degree splines. It will be shown that

$$E\left(\frac{d}{dx}H_n(x)\right)^2 \leqq \frac{B}{2}\frac{1}{nh} + \frac{A''}{2}\|f^{(m)}\|_p^2 h^{2m-2/p} \tag{2.23}$$

where A'', B are constants to be given. Then we will have

$$\sup_{f \in W_p^{(m)}(M)} E(f(x) - \hat{f}_n(x))^2 \leqq AM^2 h^{2m-2/p} + B\frac{1}{nh} \tag{2.24}$$

where $A = A' + A''$.
The right-hand side of (2.24) is minimized by taking $h = k_n/n$ with

$$k_n = \left[\frac{1}{(2m - 2/p)}\frac{B}{M^2 A}\right]^{1/(2m+1-2/p)} \cdot n^{(2m-2/p)/(2m+1-2/p)} . \tag{2.25}$$

Then, we will have the main result, which is:

$$\sup_{f \in W_p^{(m)}(M)} E[f(x) - \hat{f}_n(x)]^2 \leqq Dn^{-(2m-2/p)/(2m+1-2/p)} \tag{2.26}$$

where

$$D = \frac{(2m + 1 - 2/p)}{(2m - 2/p)^{(2m-2/p)}}(M^2 A B^{2m-2/p})^{1/(2m+1-2/p)} . \tag{2.27}$$

Details of these assertions are in the next section.

3. Proof of the main theorem.

3.1. *Bounds on the bias term.* The case $m = 3$, $p = \infty$ is covered by

PROPOSITION 1. *Let* $F^{(iv)} \in \mathscr{L}_\infty[0, 1]$. *Then*

$$\left(\frac{d}{dx}(F(x) - \tilde{F}(x))\right)^2 \leqq (\tfrac{3}{16})^2 \|F^{(iv)}\|_\infty^2 h^6 . \tag{3.1}$$

PROOF. This is Theorem 2 of [9]; the proof there may be extended from $F^{(iv)}$ continuous to $F^{(iv)} \in \mathscr{L}_\infty$.

The case $m = 1$ or 2 and $p = \infty$ is covered by

PROPOSITION 2. *Let* $F^{(iii)} \in \mathscr{L}_\infty[0, 1]$. *Then*

$$\left(\frac{d}{dx}(F(x) - \tilde{F}(x))\right)^2 \leqq (\tfrac{9}{4})^2 \|F^{(iii)}\|_\infty^2 h^4 . \tag{3.2}$$

Suppose only that $F^{(ii)} \in \mathscr{L}_\infty[0, 1]$. Then

$$\left(\frac{d}{dx}(F(x) - \tilde{F}(x))\right)^2 \leqq (\tfrac{9}{2})^2 \|F^{(ii)}\|_\infty^2 h^2 . \tag{3.3}$$

PROOF. This may be proved from the argument in [9] by following the proof of Theorem 2 in [9], and noting that, if $F^{(iii)} \in \mathscr{L}_\infty$, then r_i of [9], equation (8) is bounded by $3h|\sup_\xi F^{(iii)}(\xi)|$, if only $F^{(ii)} \in \mathscr{L}_\infty$, then r_i of [9], equation (8) is bounded by $6|\sup_\xi F^{(ii)}(\xi)|$.

The next series of lemmas result in a theorem which provides bounds on the bias for $m = 1$, $p = 2$, and $m = 2, 3$, $1 \leqq p \leqq 2$.

LEMMA 1.

$$\left(\frac{d}{dx}(F(x) - \tilde{F}(x))\right)^2 \leqq \|Q_x' - \tilde{Q}_x'\|_Q^2 \|F - \tilde{F}\|_Q^2 \tag{3.4}$$

where $\tilde{Q}_x'$ *is the projection of* Q_x' *in* $\mathscr{H}_Q$ *onto* $\mathscr{S}_l(\bar{s})$.

PROOF. Since $\tilde{F}$ is the projection of F onto $\mathscr{S}_l(\bar{s})$,

$$\left|\frac{d}{dx}(F(x) - \tilde{F}(x))\right| = |\langle Q_x', F - \tilde{F}\rangle| = |\langle Q_x' - \tilde{Q}_x', F - \tilde{F}\rangle| . \tag{3.5}$$

LEMMA 2.

$$\|Q_x' - \tilde{Q}_x'\|_Q^2 \leqq \tfrac{1}{3} h . \tag{3.6}$$

PROOF. See Appendix.

LEMMA 3. *Let* $F^{(iv)} \in \mathscr{L}_p[0, 1]$, $1 \leqq p \leqq 2$. *Then*

$$\|F - \tilde{F}\|_Q^2 \leqq \tfrac{1}{48} \|F^{(iv)}\|_p^2 h^{5-2/p} . \tag{3.7}$$

PROOF. See Appendix.

LEMMA 4. *Let* $F^{(iii)} \in \mathscr{L}_p[0, 1]$, $1 \leqq p \leqq 2$. *Then*

$$\|F - \tilde{F}\|_Q^2 \leqq \tfrac{1}{3} \|F^{(iii)}\|_p^2 h^{3-2/p} . \tag{3.8}$$

PROOF. See Appendix.

THEOREM 1. *Let $f^{(m)} \in \mathscr{L}_p$ for $m = 1$, $p = 2$, or $m = 2, 3$, $1 \leqq p \leqq 2$. Then*

$$\left(\frac{d}{dx}(F(x) - \tilde{F}(x))\right)^2 \leqq A\|f^{(m)}\|_p^2 h^{2m-2/p} \tag{3.9}$$

where

$$\begin{aligned} A &= \tfrac{1}{3}, & m &= 1, & p &= 2 \\ A &= \tfrac{1}{9}, & m &= 2, & 1 &\leqq p \leqq 2 \\ A &= \tfrac{1}{144}, & m &= 3, & 1 &\leqq p \leqq 2 . \end{aligned}$$

PROOF. The result follows upon combining Lemmas 1, 2, 3 and 4, and noting, in the case $m = 1$, $p = 2$, that $\|F - \tilde{F}\|_Q^2 \leqq \|f^{(1)}\|_2^2$.

3.2. *Bounds on the variance term.* We seek a bound on

$$E\left[\frac{d}{dx} H_n(x)\right]^2$$

where

$$H_n(x) = S(x; \bar{s}_h, \bar{\varepsilon}_a, \bar{\varepsilon}, \bar{\varepsilon}_b) \tag{3.10}$$

and

$$\begin{aligned} \bar{s}_h &= (h, 2h, \cdots, lh), \\ \bar{\varepsilon}_a &= (\varepsilon_0', \varepsilon_0), & \varepsilon_0 &= 0, \quad \varepsilon_0' = F'(0) - \hat{a}_1, \\ \bar{\varepsilon} &= (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_l), & \varepsilon_j &= \tilde{F}(jh) - \hat{F}_n(jh) = F(jh) - \hat{F}_n(jh), \quad j = 1, 2, \cdots, \\ \bar{\varepsilon}_b &= (\varepsilon_{l+1}, \varepsilon_{l+1}'), & \varepsilon_{l+1} &= 0, \quad \varepsilon_{l+1}' = F'(1) - \hat{b}_1 . \end{aligned}$$

Lemma 5 bounds the derivative of a cubic spline in terms of h and the data $\bar{\varepsilon}_a$, $\bar{\varepsilon}$, $\bar{\varepsilon}_b$.

LEMMA 5. *For $jh \leqq x < (j+1)h$, $j = 0, 1, \cdots, l$,*

$$\left|\frac{d}{dx} H_n(x)\right| \leqq 8 \left\{\sum_{i=0}^{l} c_i \frac{|\psi_i|}{h} + \frac{1}{2^{j+1}} |\varepsilon_0'| + \frac{1}{2^{l+2-j}} |\varepsilon_{l+1}'|\right\} \tag{3.11}$$

where

$$\begin{aligned} \psi_i &= \varepsilon_{i+1} - \varepsilon_i = [F((j+1)h) - F(jh)] - [F_n((j+1)h) - F_n(jh)] \\ c_i &= \frac{1}{2^{|i-j|+1}} + \frac{1}{2^{|i+1-j|+1}}, \qquad i = 0, 1, \cdots, l, \ i \neq j \\ c_j &= \frac{1}{2} + \frac{1}{2^2} + \frac{1}{8} . \end{aligned}$$

PROOF. See Appendix.

Bounds on $E[d/dx\, H_n(x)]^2$ may now be found by bounding the random variables on the right of (3.11).

Now

$$\begin{aligned}\psi_j = \varepsilon_{j+1} - \varepsilon_j &= [F((j+1)h) - F(jh)] - [F_n((j+1)h) - F_n(jh)] \\ &= p_j - \frac{\#\text{ of observations between } jh \text{ and } (j+1)h}{n}\end{aligned}$$

where

$$p_j = \textstyle\int_{jh}^{(j+1)h} f(\xi)\,d\xi < h\Lambda$$

and

$$\Lambda = \sup_\xi f(\xi)\,.$$

Thus $n(p_j - \psi_j)$ is binomial $B(n, p_j)$ and so

$$E(\psi_j)^2 = \frac{1}{n} p_j(1 - p_j) \leqq \frac{\Lambda h}{n}\,. \tag{3.12}$$

To complete the bound on the variance term, we need to know $E(\varepsilon_0')^2 \equiv E(F'(0) - \hat{a}_1)^2$ and $E(\varepsilon_{l+1}')^2 \equiv E(F'(1) - \hat{b}_1)^2$. The answer is given by

LEMMA 6. *Let $\hat{a}_1$ and $\hat{b}_1$ be given by* (2.12) *and* (2.13) *for $m = 1, 2, 3$. Then*

$$\left.\begin{matrix} E(F'(0) - \hat{a}_1)^2 \\ E(F'(1) - \hat{b}_1)^2 \end{matrix}\right\} \leqq \alpha \|f^{(m)}\|_p^2 h^{2m-2/p} + \beta \frac{\Lambda}{nh} \tag{3.13}$$

where

$$\begin{aligned}\alpha &= 8m^{2m-2/p}/[(m-1)!]^2 \\ \beta &= 2m^3(m+1)^2\Lambda\,.\end{aligned} \tag{3.14}$$

PROOF. See Appendix.

Note that, if $x = 0$, or $x = 1$,

$$\begin{aligned}E(f(0) - \hat{f}_n(0))^2 &= E(F'(0) - \hat{a}_1)^2 \\ E(f(1) - \hat{f}_n(1))^2 &= E(F'(1) - \hat{b}_1)^2\end{aligned}$$

and the mean square error is given by the right-hand side of (3.13). For $x \neq 0, 1$, we combine Lemmas 5, 6 and (3.12) to obtain, for $jh \leqq x < (j+1)h$,

$$\begin{aligned}E\left(\frac{d}{dx} H_n(x)\right)^2 &\leqq 8^2 \left\{2 \textstyle\sum_{r,s=0}^{l} c_r c_s E \dfrac{|\psi_r \psi_s|}{h^2} + 4\left(\dfrac{1}{2^{2j+2}} + \dfrac{1}{2^{2l+4-2j}}\right)\right. \\ &\qquad \left.\times \left(\alpha \|f^{(m)}\|_p^2 h^{2m-2/p} + \beta \frac{\Lambda}{nh}\right)\right\} \\ &\leqq 8^2 \left\{2(3\tfrac{1}{8})^2 \frac{\Lambda}{nh} + 4\left(\frac{1}{2^{2(x/h)}} + \frac{1}{2^{2(1-x)/h}}\right)\right. \\ &\qquad \left.\times \left(\alpha \|f^{(m)}\|_p^2 h^{2m-2/p} + \beta \frac{\Lambda}{nh}\right)\right\}.\end{aligned} \tag{3.15}$$

Note that if x is bounded away from 0 and 1, then $[1/2^{2(x/h)} + 1/2^{2(1-x)/h}] \to 0$ rapidly as $h \to 0$.

3.3. *Final result.* Summarizing the results from (3.1), (3.2), (3.3), (3.9) and

(3.15) gives

$$E(f(x) - \hat{f}_n(x))^2 \leqq A\|f^{(m)}\|_p^2 h^{2m-2/p} + B\frac{1}{nh}$$

where

$$A = 2\left[A' + 64\alpha\left(\frac{4}{2^{2x/h}} + \frac{4}{2^{2(1-x)/h}}\right)\right] \tag{3.16}$$

$$B = 2\left[B' + 64\beta\left(\frac{4\Lambda}{2^{2x/h}} + \frac{4\Lambda}{2^{2(1-x)/h}}\right)\right] \tag{3.17}$$

and

$$\begin{aligned} A' &= (\tfrac{9}{2})^2 & m &= 1\,, & p &= \infty \\ & (\tfrac{9}{4})^2 & m &= 2\,, & p &= \infty \\ & (\tfrac{3}{16})^2 & m &= 3\,, & p &= \infty \\ & \tfrac{1}{3} & m &= 1\,, & p &= 2 \\ & \tfrac{1}{9} & m &= 2\,, & 1 &\leqq p \leqq 2 \\ & \tfrac{1}{144} & m &= 3\,, & 1 &\leqq p \leqq 2 \\ B' &= 2\cdot(8\cdot 3\tfrac{1}{8})^2 \end{aligned}$$

and α and β are given by (3.14). It can be shown easily that there exists $\Lambda = \Lambda(m, p, M) < \infty$ such that $\sup_{f \text{ a density} \in W_p^{(m)}(M)} \sup_\xi f(\xi) \leqq \Lambda$, so that our results are uniform over $W_p^{(m)}(M)$. This demonstration is omitted. We have proved

THEOREM 2. *Suppose f has its support on $[0, 1]$ and $f \in W_p^{(m)}(M)$ for one of the following cases:*

$$\begin{aligned} m &= 1\,, & p &= 2\,, & p &= \infty \\ m &= 2\,, & 1 &\leqq p \leqq 2\,, & p &= \infty \\ m &= 3\,, & 1 &\leqq p \leqq 2\,, & p &= \infty\,. \end{aligned}$$

Let F_n be the sample cdf *based on n independent observations from F, and let $\hat{F}_n(x)$ be the cubic spline of interpolation to F_n at the points jh, $j = 0, 1, \cdots, l+1$; $(l+1)h = 1$, which satisfies the boundary conditions $\hat{F}_n'(0) = \hat{a}_1$, $\hat{F}_n'(1) = \hat{b}_1$, where $\hat{a}_1$ and $\hat{b}_1$ are given by* (2.12) *and* (2.13). *Let $\hat{f}_n(x) = d/dx\, \hat{F}_n(x)$, and suppose h is chosen as $h = k_n/n$,*

$$k_n = \left[\frac{1}{(2m - 2/p)}\frac{B}{M^2A}\right]^{1/(2m+1-2/p)} n^{(2m-2/p)/(2m+1-2/p)}$$

where A and B are given by (3.16) *and* (3.17). *Then*

$$\sup_{f\in W_p^{(m)}(M)} E[f(x) - \hat{f}_n(x)]^2 \leqq Dn^{-(2m-2/p)/(2m+1-2/p)} \tag{3.18}$$

with

$$D = \frac{(2m + 1 - 2/p)}{(2m - 2/p)^{(2m-2/p)}}(M^2AB^{2m-2/p})^{1/(2m+1-2/p)}\,. \tag{3.19}$$

APPENDIX

This appendix contains the proofs of Lemmas 2–6. The proofs are carried out where the knots $\{s_i\}_{i=1}^{l}$ do not necessarily satisfy $s_{i+1} - s_i = h$, but only $0 = s_0 < s_1 < \cdots < s_l < s_{l+1} = 1$. The purpose of this generality is to allow the lemmas to be referenced for a later report which deals with the situation where the knots are determined by the order statistics. Let I_j be the interval $[s_j, s_{j+1}]$, for $j = 0, 1, \cdots, l$.

LEMMA 2. *Let $\tilde{Q}_x'$ be the projection of Q_x' onto $\mathscr{S}_l(\bar{s})$, and let $x \in I_j$. Then*

$$||Q_x' - \tilde{Q}_x'||_Q^2 \leqq \tfrac{1}{3}(s_{j+1} - s_j)\,, \qquad j = 0, 1, \cdots, l\,.$$

PROOF. For $x \in I_j$, define R_x' in $\mathscr{H}_Q$ by

$$R_x' = \frac{1}{(s_{j+1} - s_j)}(Q_{s_{j+1}} - Q_{s_j})\,.$$

Since $R_x' \in \mathscr{S}_l(\bar{s})$ and $\tilde{Q}_x'$ is the projection of Q_x' onto $\mathscr{S}_l(\bar{s})$

$$||Q_x' - \tilde{Q}_x'||_Q \leqq ||Q_x' - R_x'||_Q\,. \tag{A1.1}$$

To compute the square of the right side of (A1.1), note from (2.4) that

$$Q_x'(0) = 0 \tag{A1.2}$$

$$\frac{d}{ds}Q_x'(s)\Big|_{s=0} = 1 \tag{A1.3}$$

$$\frac{d^2}{ds^2}Q_x'(s) = 1 \quad s < x; \qquad = 0 \quad s > x\,. \tag{A1.4}$$

After some calculations,

$$R_x'(0) = 0 \tag{A1.5}$$

$$\frac{d}{ds}R_x'(s)\Big|_{s=0} = 1 \tag{A1.6}$$

$$\frac{d^2}{ds^2}R_x'(s) = \begin{cases} 1\,, & 0 \leqq s \leqq s_j \\ \dfrac{(s_{j+1} - s)}{(s_{j+1} - s_j)}\,, & s_j \leqq s \leqq s_{j+1} \\ 0\,, & s_{j+1} \leqq s \leqq 1\,. \end{cases} \tag{A1.7}$$

$$\frac{d^2}{ds^2}Q_x'(s) - \frac{d^2}{ds^2}R_x'(s) = 0\,, \qquad \text{for} \quad s \notin I_j \tag{A1.8}$$

and

$$||Q_x' - R_x'||_Q^2 = \frac{1}{(s_{j+1} - s_j)^2}\left[\int_{s_j}^{x}(u - s_j)^2\,du + \int_{x}^{s_{j+1}}(s_{j+1} - u)^2\,du\right] \leqq \tfrac{1}{3}(s_{j+1} - s_j)\,.$$

LEMMA 3. *Let $F \in \mathscr{H}_Q$ satisfy $F^{(\mathrm{iv})} = \rho \in \mathscr{L}_p[0, 1]$. Let $\tilde{F}$ be the projection of F onto $\mathscr{S}_l(\bar{s})$. Then*

$$||F - \tilde{F}||_Q^2 \leqq \tfrac{1}{48}\textstyle\sum_{j=0}^{l}(s_{j+1} - s_j)^{5-2/p}\left[\int_{s_j}^{s_{j+1}}|\rho(\xi)|^p\,d\xi\right]^{2/p}\,. \tag{A2.1}$$

If $s_{j+1} - s_j = h$, $j = 0, 1, \cdots, l$, *and* $1 \leqq p \leqq 2$,

$$\|F - \tilde{F}\|_Q^2 \leqq \tfrac{1}{48} h^{5-2p} \|F^{(\text{iv})}\|_p^2 .$$

PROOF. First we show that $F^{(\text{iv})} = \rho$ implies that

$$F(t) = \textstyle\int_0^1 Q(t, s)\rho(s)\, ds + c_1 Q_0(t) + c_2 Q_1(t) + c_3 Q_0'(t) + c_4 Q_1'(t) \tag{A2.2}$$

for some $\{c_i\}$. But $Q_0(t, s) = \int_0^1 (s - u)_+(t - u)_+\, du$ is the Green's function for the operator D^4, with boundary conditions

$$G^{(\nu)}(0) = 0 , \qquad \nu = 0, 1$$
$$G^{(\nu)}(1) = 0 , \qquad \nu = 2, 3 .$$

Thus, F always has a representation

$$F(t) = \textstyle\int_0^1 Q_0(t, s)\rho(s)\, ds + \sum_{i=0}^3 d_i t^i . \tag{A2.3}$$

But

$$\textstyle\int_0^1 Q_0(t, s)\rho(s)\, ds = \int_0^1 Q(t, s)\rho(s)\, ds - \int_0^1 (1 + st)\rho(s)\, ds .$$

Since $Q_0(t)$, $Q_0'(t)$, $Q_1(t)$ and $Q_1'(t)$ span the same space as $\{1, t, t^2, t^3\}$, $\{c_i\}$ can always be found so that (A2.2) equals (A2.3).

Next, if v is any element in $\mathscr{H}_Q$ of the form

$$v = \textstyle\sum_{i=0}^{l+1} c_i Q_{s_i} + aQ_0' + bQ_1' .$$

Then, since $v \in \mathscr{S}_l(\bar{s})$,

$$\|F - \tilde{F}\|_Q \leqq \|F - v\|_Q . \tag{A2.4}$$

The proof now proceeds by finding an element $v \in \mathscr{S}_l(\bar{s})$ so that the right-hand side of (A2.4) is bounded by the right-hand side of (A2.1). For $x \in I_j$, define $R_x \in \mathscr{S}_l(\bar{s})$ by

$$R_x = \frac{(s_{j+1} - x)}{(s_{j+1} - s_j)} Q_{s_j} + \frac{(x - s_j)}{(s_{j+1} - s_j)} Q_{s_{j+1}} , \qquad j = 0, 1, \cdots, l .$$

Define $v \in \mathscr{S}_l(\bar{s})$ by

$$\begin{aligned} v &= \textstyle\int_0^1 R_x \rho(x)\, dx + c_1 Q_0 + c_2 Q_0' + c_3 Q_1 + c_4 Q_1' \\ &\equiv \sum_{j=0}^l \left\{ Q_{s_j} \int_{s_j}^{s_{j+1}} \frac{(s_{j+1} - x)}{(s_{j+1} - s_j)} \rho(x)\, dx + Q_{s_{j+1}} \int_{s_j}^{s_{j+1}} \frac{(x - s_j)}{(s_{j+1} - s_j)} \rho(x)\, dx \right\} \\ &\quad + c_1 Q_0 + c_2 Q_0' + c_3 Q_1 + c_4 Q_1' . \end{aligned}$$

Now

$$F - v = \textstyle\int_0^1 (Q_x - R_x)\rho(x)\, dx$$

and, by the properties of the reproducing kernel, it can be shown that

$$\|F - v\|_Q^2 = \textstyle\int_0^1 \int_0^1 \rho(x)\rho(x') \langle Q_x - R_x, Q_{x'} - R_{x'} \rangle\, dx\, dx' .$$

Since

$$Q_x(0) - R_x(0) = 0$$

$$\frac{d}{ds}(Q_x(s) - R_x(s))\Big|_{s=0} = 0$$

$$\frac{d^2}{ds^2}(Q_x(s) - R_x(s)) = (x-s)_+ - \frac{(x-s_j)(s_{j+1}-s)}{(s_{j+1}-s_j)}, \qquad x \in I_j;\ s \in I_j$$

$$= 0, \qquad x \in I_j,\ s \in I_k,\ k \neq j,$$

it follows that

$$\langle Q_x - R_x, Q_{x'} - R_{x'} \rangle = 0$$

if $x \in I_j$, $x' \in I_k$ with $j \neq k$. Thus

$$||F - v||_Q^2 \leqq \sum_{j=0}^{l} \{\int_{s_j}^{s_{j+1}} |\rho(x)| ||Q_x - R_x||_Q \, dx\}^2 .$$

Furthermore

$$||Q_x - R_x||_Q^2 = \int_{s_j}^{s_{j+1}} \left[(x-s)_+ - \frac{(x-s_j)_+}{(s_{j+1}-s_j)}(s_{j+1}-s)_+\right]^2 ds$$

$$\leqq \tfrac{1}{48}(s_{j+1}-s_j)^3, \qquad x \in I_j,$$

so that

$$||F - v||_Q^2 \leqq \tfrac{1}{48} \sum_{j=0}^{l} (s_{j+1}-s_j)^3 [\int_{s_j}^{s_{j+1}} |\rho(t)| \, dt]^2 .$$

For $1/p + 1/p' = 1$, a Hölder inequality gives

$$\int_{I_j} |\rho(t)| \, dt \leqq [\int_{I_j} dt]^{1/p'} [\int_{I_j} |\rho(t)|^p \, dt]^{1/p}$$

$$= (s_{j+1}-s_j)^{1-1/p} [\int_{I_j} |\rho(t)|^p \, dt]^{1/p} .$$

Thus,

$$||F - v||_Q^2 \leqq \tfrac{1}{48} \sum_{j=0}^{l} (s_{j+1}-s_j)^{5-2/p} [\int_{I_j} |\rho(t)|^p \, dt]^{2/p} .$$

If $(s_{j+1} - s_j) = h$ and $1 \leqq p \leqq 2$, then

$$||F - v||_Q^2 \leqq \tfrac{1}{48} h^{5-2/p} [\int_0^1 |\rho(t)|^p \, dt]^{2/p} .$$

LEMMA 4. *Let $F \in \mathscr{H}_Q$ satisfy $F^{(\text{iii})} = \eta \in \mathscr{L}_p[0, 1]$. Let $\tilde{F}$ be the projection of F onto $\mathscr{S}_l(\bar{s})$. Then*

$$||F - \tilde{F}||_Q^2 \leqq \tfrac{1}{3} \sum_{j=0}^{l} (s_{j+1}-s_j)^{3-2/p} [\int_{s_j}^{s_{j+1}} |\eta(\xi)|^p \, d\xi]^{2/p} .$$

If $(s_{j+1} - s_j) = h$, and $1 \leqq p \leqq 2$, then

$$||F - \tilde{F}||_Q^2 \leqq \tfrac{1}{3} h^{3-2/p} ||F^{(\text{iii})}||_p^2 .$$

PROOF. As in the proof of Lemma 3, by the Green's function properties of $Q_x'(s)$ there exist c_1, c_2, c_3, c_4 such that

$$F(t) = \int_0^1 Q_x'(t)\eta(x) \, ds + c_1 Q_0(t) + c_2 Q_0'(t) + c_3 Q_1(t) + c_4 Q_1'(t) .$$

Let

$$v = \int_0^1 R_x' \eta(x) \, dx + c_1 Q_0 + c_2 Q_0' + c_3 Q_1 + c_4 Q_1'$$

where R_x' is defined as in the proof of Lemma 2,

$$R_x' = \frac{1}{(s_{j+1}-s_j)}(Q_{s_{j+1}} - Q_{s_j}) \qquad \text{for} \quad x \in I_j .$$

Then

$$||F - \tilde{F}||_Q^2 \leqq ||F - v||_Q^2 = \int_0^1 \int_0^1 \eta(x)\eta(x')\langle Q_x' - R_x', Q_{x'}' - R_{x'}'\rangle \, dx \, dx' .$$

Also, it can be shown that

$$\langle Q_x' - R_x', Q_{x'}' - R_{x'}'\rangle = 0 \qquad \text{if} \quad x \in I_j , \quad x' \in I_k , \quad j \neq k ;$$

so that

$$||F - v||_Q^2 \leqq \sum_{j=0}^{l} \{\int_{s_j}^{s_{j+1}} |\eta(x)| ||Q_x' - R_x'|| \, dx\}^2 .$$

By Lemma 2,

$$||Q_x' - R_x'||^2 \leqq \tfrac{1}{3}(s_{j+1} - s_j) \qquad \text{for } x \in I_j$$

from which the result follows as in Lemma 3.

LEMMA 5. *Let $S(x) = S(x, \bar{s}; \bar{\varepsilon}_a, \bar{\varepsilon}, \bar{\varepsilon}_b)$ be the cubic spline of interpolation defined by* (2.8) *with*

$$\begin{aligned} \bar{s} &= (s_1, s_2, \cdots, s_l) \\ \bar{\varepsilon}_a &= (\varepsilon_0', \varepsilon_0) , && \varepsilon_0 = 0 \\ \bar{\varepsilon} &= (\varepsilon_1, \varepsilon_2, \cdots, \varepsilon_l) \\ \bar{\varepsilon}_b &= (\varepsilon_{l+1}, \varepsilon_{l+1}') , && \varepsilon_{l+1} = 0 . \end{aligned}$$

Let

$$\begin{aligned} \Delta_i &= (s_{i+1} - s_i) , && i = 0, 1, \cdots, l , \\ \phi_i &= (\varepsilon_{i+1} - \varepsilon_i) , && i = 0, 1, \cdots, l . \end{aligned}$$

Then, for $x \in I_j$,

$$\left|\frac{d}{dx} S(x)\right| \leqq 8 \left\{\sum_{i=0}^{l} c_i \frac{\Delta_j}{\Delta_i} \frac{|\phi_i|}{\Delta_i} + \frac{1}{2^{j+1}} \frac{\Delta_j}{\Delta_0} |\varepsilon_0'| + \frac{1}{2^{l+2-j}} \frac{\Delta_j}{\Delta_l} |\varepsilon_{l+1}'|\right\} \tag{A4.1}$$

where

$$\begin{aligned} c_i &= \frac{1}{2^{|i-j|+1}} + \frac{1}{2^{|i+1-j|+1}} && i = 0, 1, \cdots, l, \ i \neq j \\ &= \frac{1}{2} + \frac{1}{2^2} + \frac{1}{8} , && i = j . \end{aligned}$$

PROOF. Define

$$R_x' = \frac{1}{\Delta_j} (Q_{s_{j+1}} - Q_{s_j}) \qquad \text{for} \quad x \in I_j .$$

Then

$$\frac{d}{dx} S(x) = \langle S, Q_x'\rangle_Q = \langle S, R_x'\rangle_Q + \langle S, Q_x' - R_x'\rangle .$$

Since $S(s_j) = \varepsilon_j, j = 0, 1, \cdots, l + 1$,

$$\langle S, R_x'\rangle = \frac{1}{\Delta_j} (\varepsilon_{j+1} - \varepsilon_j) = \frac{\phi_j}{\Delta_j} . \tag{A4.2}$$

Since the spline S is a cubic between the knots $0, s_1, \cdots, s_l, 1$, $d^2/dx^2 \, S(x)$ is linear between the knots, and by the properties of cubic splines, continuous. Define κ_i by

$$\left.\frac{d^2}{ds^2} S(s)\right|_{s=s_i} = \kappa_i .$$

Thus

$$\frac{d^2}{ds^2} S(s) = \frac{1}{\Delta_j} [\kappa_j(s_{j+1} - s) + \kappa_{j+1}(s - s_j)] \qquad \text{for} \quad s \in I_j \, .$$

Combining this with (A1.2)—(A1.8) gives

$$\begin{aligned} &\langle S, Q_x' - R_x' \rangle_Q \\ &\quad = \frac{1}{\Delta_j^2} \int_{s_j}^{x} (s - s_j)[\kappa_j(s_{j+1} - s) + \kappa_{j+1}(s - s_j)]\, ds \\ &\qquad + \frac{1}{\Delta_j^2} \int_{x}^{s_{j+1}} (s_{j+1} - s)[\kappa_j(s_{j+1} - s) + \kappa_{j+1}(s - s_j)]\, ds \, , \qquad \text{for} \quad x \in I_j \end{aligned}$$

and so

$$\text{(A4.3)} \qquad |\langle S, Q_x' - R_x' \rangle_Q| \leqq \Delta_j \max(|\kappa_j|, |\kappa_{j+1}|) \, , \qquad x \in I_j \, .$$

To proceed, we need to know the relationship between the κ_i and the data $\bar{s}$, $\bar{\varepsilon}_a$, $\bar{\varepsilon}$, $\bar{\varepsilon}_b$. By using a formula found in Kershaw, [9], equation (5), we may express this relationship for cubic splines. It is

$$(\kappa_0, \kappa_1, \cdots, \kappa_l, \kappa_{l+1}) = 6A^{-1}(\xi_0, \xi_1, \cdots, \xi_l, \xi_{l+1})$$

where A is the $(l+2) \times (l+2)$ matrix given by

$$A = \begin{bmatrix} 2 & 1 & & & & & \\ \alpha_1 & 2 & 1 - \alpha_1 & & & 0 & \\ & \alpha_2 & 2 & 1 - \alpha_2 & & & \\ & & \cdot & & & & \\ & & & \cdot & & & \\ & & & & \cdot & & \\ 0 & & & \alpha_l & 2 & 1 - \alpha_l \\ & & & & 1 & 2 \end{bmatrix}$$

where

$$\alpha_i = \frac{\Delta_{i-1}}{\Delta_i - \Delta_{i-1}} \qquad i = 1, 2, \cdots, l$$

and

$$\begin{aligned} \xi_0 &= (\psi_0 - \Delta_0 \varepsilon_0')/\Delta_0^2 \\ \xi_i &= \left(\frac{\psi_i}{\Delta_i} - \frac{\psi_{i-1}}{\Delta_{i-1}}\right)\Big/(\Delta_i + \Delta_{i-1}) \, , \qquad i = 1, 2, \cdots, l \\ \xi_{l+1} &= -(\psi_l - \Delta_l \varepsilon_{l+1}')/\Delta_l^2 \, . \end{aligned}$$

For $i = 1, 2, \cdots, l$, ξ_i is the second divided difference of $S(x)$ at (s_{i-1}, s_i, s_{i+1}). We are now going to appeal to another result of Kershaw's, which gives bounds on the entries of A^{-1}. Let a^{rs}, $r, s = 0, 1, \cdots, l+1$, be the r, sth entry of A^{-1}. According to [8],

$$|a^{rs}| \leqq \frac{4}{3} \frac{1}{2^{|r-s|+1}} \, , \qquad r, s = 0, 1, \cdots, l+1 \, .$$

Therefore, since $\kappa_j = 6 \sum_{i=0}^{l+1} a^{ji} \xi_i$,

$$\text{(A4.4)} \qquad |\kappa_j| \leqq 6 \cdot \tfrac{4}{3} \cdot \sum_{i=0}^{l+1} \frac{1}{2^{|i-j|+1}} |\xi_i|$$

and, combining (A4.2), (A4.3), and (A4.4) gives

$$|\langle S, Q_x' - R_x' \rangle| \leqq \Delta_j \cdot 6 \cdot \tfrac{4}{3} \cdot \textstyle\sum_{i=0}^{l+1} \frac{1}{2^{|i-j|+1}} |\xi_i| , \qquad x \in I_j$$

and

$$\begin{aligned} |\langle S, Q_x' \rangle| &\leqq \frac{|\phi_j|}{\Delta_j} + 8 \cdot \textstyle\sum_{i=1}^{l} \frac{1}{2^{|i-j|+1}} \left\{ \frac{|\phi_i|}{\Delta_i} + \frac{|\phi_{i-1}|}{\Delta_{i-1}} \right\} \frac{\Delta_j}{\Delta_i + \Delta_{i-1}} \\ &\quad + 8 \frac{1}{2^{j+1}} \frac{\Delta_j}{\Delta_0} \left\{ \frac{|\phi_0|}{\Delta_0} + |\varepsilon_0'| \right\} + 8 \frac{1}{2^{l+2-j}} \frac{\Delta_j}{\Delta_l} \left\{ \frac{|\phi_l|}{\Delta_l} + |\varepsilon_{l+1}'| \right\} \\ &\leqq 8 \cdot \textstyle\sum_{i=0}^{l} c_i \frac{|\phi_i|}{\Delta_i} \frac{\Delta_j}{\Delta_i} + \frac{8}{2^{j+1}} \frac{\Delta_j}{\Delta_0} |\varepsilon_0'| + \frac{8}{2^{l+2-j}} \frac{\Delta_j}{\Delta_l} |\varepsilon_{l+1}'| \end{aligned}$$

where

$$c_i = \frac{1}{2^{|i-j|+1}} + \frac{1}{2^{|i+1-j|+1}} , \qquad i = 0, 1, 2, \cdots, l, \; i \neq j$$

$$c_j = \frac{1}{2} + \frac{1}{2^2} + \frac{1}{8} .$$

We remark that the lack of a generalization of this lemma for higher degree splines is the stumbling block in generalizing the main theorem to higher m.

LEMMA 6. *Suppose* $f^{(m)} \in \mathscr{L}_p$ *on* $[0, 1]$. *Let* F_n *be the sample* cdf *for* n *independent observations from* f. *Let*

$$\hat{f}_n(0) = \frac{d}{dx} \textstyle\sum_{\nu=0}^{m} l_{0,\nu}(x) \Big|_{x=0} F_n(\nu h)$$

$$\hat{f}_n(1) = \frac{d}{dx} \textstyle\sum_{\nu=0}^{m} l_{1,\nu}(x) \Big|_{x=1} F_n(1 - \nu h)$$

where $l_{0,\nu}$ *and* $l_{1,\nu}$ *are the Lagrange polynomials defined in* (2.10) *and* (2.11). (*For* $m = 1$, $\hat{f}_n(0) = (1/n)F_n(h)$.) *Then*

$$\text{(A5.1)} \qquad \left.\begin{matrix} E(f(0) - \hat{f}_n(0))^2 \\ E(f(1) - \hat{f}_n(1))^2 \end{matrix}\right\} \leqq \frac{8m^{2m-2/p}}{[(m-1)!]^2} \, ||f^{(m)}||_p^2 h^{2m-2/p} + 2m^3(m+1)^2 \frac{\Lambda}{nh} .$$

PROOF.

$$\text{(A5.2)} \qquad \begin{aligned} |f(0) - \hat{f}_n(0)| &\leqq \left| f(0) - \frac{d}{dx} \textstyle\sum_{\nu=0}^{m} l_{0,\nu}(x) \Big|_{x=0} F(\nu h) \right| \\ &\quad + \left| \frac{d}{dx} \textstyle\sum_{\nu=0}^{m} l_{0,\nu}(x)[F(\nu h) - F_n(\nu h)] \right| . \end{aligned}$$

By combining Lemma 3.1 of [21], and Theorem 3 of [20], and noting that $\prod_{j=0}^{m} (0 - jh) = 0$ in equation (3.28) of [20], it can be shown that

$$\begin{aligned} &\left| f(0) - \frac{d}{dx} \textstyle\sum_{\nu=0}^{m} l_{0,\nu}(x) \Big|_{x=0} F(\nu h) \right|^2 \\ \text{(A5.3)} \qquad &\leqq \left[\frac{2}{(m-1)!} \right]^2 [\textstyle\int_0^{mh} |f^{(m)}(\xi)|^p \, d\xi]^{2/p} (mh)^{m-1/p} , \\ &\qquad\qquad p \geqq 1, \; m = 1, 2, \cdots . \end{aligned}$$

It can be verified, for $m = 1, 2, 3$, that

$$\left|\frac{d}{dx} l_{0,\nu}(x)\Big|_{x=0}\right| \leqq \frac{m}{h}. \tag{A5.4}$$

Now

$$F_n(\nu h) = \frac{\#\ \text{observations in } [0, \nu h]}{n},$$

and hence $nF_n(\nu h)$ is binomial $B(n, \sum_{j=1}^{\nu} p_j)$, where $p_j = \int_{(j-1)h}^{jh} f(\xi)\, d\xi$, and hence

$$E[F(\nu h) - F_n(\nu h)]^2 = (\textstyle\sum_{j=1}^{\nu} p_j)(1 - \sum_{j=1}^{\nu} p_j)/n \leqq \frac{\Lambda m h}{n}. \tag{A5.5}$$

Putting together (A5.2) (A5.3) (A5.4) and (A5.5) gives

$$E(f(0) - \hat{f}_n(0))^2 \leqq \frac{8m^{2m-2/p}}{((m-1)!)^2} \|f^{(m)}\|_p^2 h^{2m-2/p} + 2m^3(m+1)^2 \frac{\Lambda}{nh}.$$

The proof is carried out similarly for $x = 1$.

REFERENCES

[1] Academic Computing Center. *Approximation and Interpolation, Reference Manual for the* 1108. Chapter 6. MACC, Univ. of Wisconsin.

[2] Ahlberg, J. H., Nilson, E. N. and Walsh, J. L. (1967). *The Theory of Splines and their Applications.* Academic Press, New York.

[3] Boneva, L., Kendall, D. and Stefanov, I. (1971). Spline transformations: Three new diagnostic aids for the statistical data analyst. *J. Roy. Statist. Soc.* **33** 1-70.

[4] DeBoor, Carl (1968). On the convergence of odd degree spline interpolation. *J. Approx. Theor.* **1** 452-463.

[5] Farrell, R. H. (1972). On best obtainable asymptotic rates of convergence in estimation of a density function at a point. *Ann. Math. Statist.* **43** 170-180.

[6] Greville, T. N. E. (1969). Introduction to spline functions, in *Theory and Applications of Spline Functions* (T. N. E. Greville, ed.). Academic Press, New York. 1-35.

[7] Greville, T. N. E. (1971). Another look at cubic spline interpolation of equidistant data. Univ. of Wisconsin MRC Technical Summary Report #1148, Madison.

[8] Kershaw, D. (1970). Inequalities on the elements of the inverse of a certain tridiagonal matrix. *Math. Comp.* **24** 155-158.

[9] Kershaw, D. (1971). A note on the convergence of interpolatory cubic splines. *SIAM J. Numer. Anal.* **8** 67-74.

[10] Kimeldorf, George, and Wahba, Grace (1971). Some results on Tchebycheffian spline functions, *J. Math. Anal. Appl.* **33** 82-95.

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

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

[13] Sharma, A. and Meir, A. (1966). Degree of approximation of spline interpolation. *J. Math. Mech.* **15** 759-767.

[14] Van Rooy, P. L. J. and Schurer, F. (1971). A bibliography on spline functions. Dept. of Mathematics, Technological Univ. Eindhoven, Netherlands, T. H.—Report 71-WSK-02.

[15] Schoenberg, I. J. (1967). On spline functions, in *Inequalities,* (O. Shisha, ed.). 255-291. Academic Press, New York.

[16] SCHOENBERG, I. J. (1972). Notes on spline functions II. On the smoothing of histograms. Univ. of Wisconsin MRC Technical Summary Report #1222, Madison.

[17] SCHOENBERG, I. J. (1972). Splines and histograms. Univ. of Wisconsin MRC Technical Summary Report #1273, Madison.

[18] SCHULTZ, MARTIN (1970). Error bounds for polynomial spline interpolation. *Math. Comp.* **24** 507–515.

[19] WAHBA, GRACE (1971). A note on the regression design problem of Sacks and Ylvisaker. *Ann. Math. Statist.* **42** 1035–1053.

[20] WAHBA, GRACE (1971). A polynomial algorithm for density estimation. *Ann. Math. Statist.* **42** 1870–1886.

[21] WAHBA, GRACE (1975). Optimal convergence properties of variable knot kernel, and orthogonal series methods for density estimation. *Ann. Statist.* **3** 15–29.

DEPARTMENT OF STATISTICS
UNIVERSITY OF WISCONSIN
1210 WEST DAYTON STREET
MADISON, WISCONSIN 53706