The Annals of Mathematical Statistics
1971, Vol. 42, No. 6, 1870–1886

A POLYNOMIAL ALGORITHM FOR DENSITY ESTIMATION[1]

BY GRACE WAHBA

University of Wisconsin, Madison

An algorithm for density estimation based on ordinary polynomial (Lagrange) interpolation is studied. Let $F_n(x)$ be $n/(n+1)$ times the sample c.d.f. based on n order statistics, $t_1, t_2, \cdots t_n$, from a population with density $f(x)$. It is assumed that $f^{(\nu)}$ is continuous, $\nu = 0, 1, 2, \cdots, r$, $r = m-1$, and $f^{(m)} \in L_2(-\infty, \infty)$. $F_n(x)$ is first locally interpolated by the mth degree polynomial passing through $F_n(t_{ik_n})$, $F_n(t_{(i+1)k_n})$, $\cdots$ $F_n(t_{(i+m)k_n})$, where k_n is a suitably chosen number, depending on n. The density estimate is then, locally, the derivative of this interpolating polynomial. If $k_n = O(n^{(2m-1)/(2m)})$, then it is shown that the mean square convergence rate of the estimate to the true density is $O(n^{-(2m-1)/(2m)})$. Thus these convergence rates are slightly better than those obtained by the Parzen kernel-type estimates for densities with r continuous derivatives.

If it is assumed that $f^{(m)}$ is bounded, and $k_n = O(n^{2m/(2m+1)})$, then it is shown that the mean square convergence rates are $O(n^{-2m/(2m+1)})$, which are the same as those of the Parzen estimates for m continuous derivatives. An interesting theorem about Lagrange interpolation, concerning how well a function can be interpolated knowing only its integral at nearby points, is also demonstrated.

1. Introduction and summary. Let $t_1, t_2, \cdots t_n$ be the order statistics from a random sample of size n from a population with unknown density $f(x)$. We are interested in estimating the density $f(x)$. Suppose that f has r bounded derivatives in the neighborhood of x. Then the Parzen or kernel-type estimate $f_n(x)$, for $f(x)$, (see Parzen (1962)) has the property that

$$E(f_n(x)-f(x))^2 = O(n^{-2r/(2r+1)}), \qquad r = 1, 2, \cdots \tag{1.1}$$

In this note we consider a very simple type of density estimate as follows. Let f possess r continuous derivatives and suppose $f^{(m)} \in L_2(-\infty, \infty)$, with $m = r+1$. Let $F_n(x)$ be $n/(n+1)$ times the sample cumulative distribution function. Let k_n be an appropriately chosen sequence depending on $n(k_n = O(n^{(2m-1)/2m}))$. Let l be the greatest integer in $(n-1)/k_n$. Let

$$\begin{aligned} \hat{f}_{n,m}(x) &= 0, && x < t_{2k_n} \\ &= \frac{d}{dx}\hat{F}_{n,m}(x), && t_{2k_n} \leqq x < t_{(l-m+1)k_n} \\ &= 0, && t_{(l-m+1)k_n} \leqq x \end{aligned} \tag{1.2}$$

Received May 26, 1970; revised April 1, 1971.

[1] Research supported in part by the Wisconsin Alumni Research Foundation.

where $\hat{F}_{n,m}(x)$ is defined as follows:

For $m = 1$,

$$\hat{F}_{n,1}(x) = F_n(t_{ik_n}) + x\frac{F_n(t_{(i+1)k_n}) - F_n(t_{ik_n})}{t_{(i+1)k_n} - t_{ik_n}}, \qquad t_{ik_n} \leqq x < t_{(i+1)k_n}; i = 2, 3, \cdots, l-1.$$

For $m \geqq 2$, let $\hat{F}_{n,m,i}(x)$, $i = 1, 2, \cdots l-m-1$, be the mth degree polynomial which interpolates to $F_n(x)$ at the $m+1$ points $x = t_{ik_n}, t_{(i+1)k_n}, \cdots t_{(i+m)k_n}$. For $x \in [t_{(i+1)k_n}, t_{(i+2)k_n})$, define $\hat{F}_{n,m}(x)$ to coincide with $\hat{F}_{n,m,i}(x)$, $i = 1, 2, \cdots l-m-1$. A more symmetric positioning of the local interpolating polynomial may be made, the present choice is primarily for notational convenience. Similarly, the definition of $\hat{f}_{n,m}(x)$ for $x \notin [t_{2k_n}, t_{(l-m+1)k_n})$ is arbitrarily chosen for notational convenience, and to simplify the proofs.

Suppose, further, that f is bounded, $f(u) > 0$ for u in a neighborhood of x and $|u(1-F(u))| \leqq M$, $u \geqq x$, $|uF(u)| \leqq M$, $u \leqq x$. We prove

THEOREM 1.

$$E|f(x) - \hat{f}_{n,m}(x)|^2 = O(n^{-(2m-1)/2m}) \qquad m = 1, 2, \cdots \tag{1.3}$$

Thus with the main added assumption of the square-integrability of the $m = (r+1)$st derivative, this simple algorithm improves upon the rate of the Parzen estimates.

If, in addition, we assume $f^{(m)}$ bounded, and let $k_n = O(n^{2m/(2m+1)})$, we prove

THEOREM 2.

$$E|f(x) - \hat{f}_{n,m}(x)|^2 = O(n^{-2m/(2m+1)}). \tag{1.4}$$

Thus, this algorithm achieves the same convergence rate as the Parzen estimates.

The proofs proceed by breaking the mean square error into two major terms. These terms might be viewed as the squared bias and the variance. The bias term may be viewed as the error made in approximating a smooth density at a point by differentiating a polynomial which interpolates to actual values of the c.d.f. in the neighborhood of x. The variance term then results from the fact that the c.d.f. is not known but estimated. We use the following theorem about polynomial (Lagrange) interpolation which tells us about the bias error.

We suppose $x_0 < x_1 < \cdots x_m$ are $m+1$ real numbers, and $f^{(\nu)}$, $\nu = 0, 1, 2, \cdots r$ absolutely continuous on $[x_0, x_m]$, $f^{(m)} \in L_2[x_0, x_m]$. Let $l_\nu(x; x_0, x_1, \cdots x_m) = l_\nu(x)$ be the mth degree polynomials satisfying $l_\nu(x_\mu) = \delta_{\mu,\nu}$, μ, $\nu = 0, 1, 2 \cdots m$. Then we have

THEOREM 3.

$$\left|f(x) - \sum_{\nu=0}^{m} \frac{d}{dx} l_\nu(x) \int_{x_0}^{x_\nu} f(\xi)\, d\xi\right|^2 \leqq a(m) \int_{x_0}^{x_m} [f^{(m)}(\xi)]^2\, d\xi\, |x_m - x_0|^{2m-1} \tag{1.5}$$

$$x \in [x_0, x_m], m = 1, 2; x \in [x_1, x_{m-1}], m \geqq 3$$

where $a(m)$ is a constant depending on m.

To minimize the mean square error, k_n is chosen so that the bounds for the squared bias and variance terms are of the same order of magnitude.

The polynomial algorithm for $m = 1$ $(r = 0)$ coincides with an algorithm recently studied by Van Ryzin (1970) ("unsymmetric case"). He obtained the interesting result that if $k_n = o(n^{\frac{4}{5}})$, and x is a point at which f' exists and is continuous, then

$$(k_n^{\frac{1}{2}}(f(x)-\hat{f}_{m,1}(x)) \to \mathcal{N}(0,f^2(x)). \tag{1.6}$$

Van Ryzin's theorem tells us what happens if we proceed here as though f' were only square integrable (e.g. $k_n' = O(n^{\frac{1}{2}})$) but in fact f' exists and is continuous at x.

We remind the reader that an extensive literature exists on density estimation. For a bibliography, see Wegman (1970).

2. Description of the algorithm and the main theorems. It is convenient to have some general formulae for interpolating polynomials. Let $x_0, x_1, \cdots x_m$ be $m+1$ distinct real numbers. Let $l_\nu(x)$ be defined by

$$l_\nu(x) = l_\nu(x; x_0, x_1, \cdots x_m) = \frac{\prod_{\mu=0,\mu\neq\nu}^{m}(x-x_\mu)}{\prod_{\mu=0,\mu\neq\nu}^{m}(x_\nu-x_\mu)}, \quad \nu = 0,1,2,\cdots m. \tag{2.1}$$

It is easily seen that $l_\nu(x)$ is the mth degree polynomial satisfying

$$\begin{aligned} l_\nu(x_\mu) &= 1, & \mu &= \nu \\ &= 0, & \mu &\neq \nu. \end{aligned} \tag{2.2}$$

Let $t_{ik_n}, t_{(i+1)k_n}, \cdots t_{(i+m)k_n}$ be the order statistics indicated by the subscripts, and, for convenience, define ${}_{li,\nu}(x)$ by

$${}_{li,\nu}(x) = l_\nu(x; t_{ik_n}, t_{(i+1)k_n}, \cdots t_{(i+m)k_n}). \tag{2.3}$$

The estimate $\hat{f}_{n,m}$ defined in (1.2) is given by

$$\begin{aligned} \hat{f}_{n,m}(x) &= \frac{d}{dx}\sum_{\nu=0}^{m} {}_{li,\nu}(x)\frac{(i+\nu)k_n+1}{(n+1)}, \quad i = i(x), x \in [t_{2k_n}, t_{(l-m+1)k_n}) \\ &= 0 \quad \text{otherwise} \end{aligned} \tag{2.4a}$$

where $i(x)$ is defined for $x \in [t_{2k_n}, t_{(l-m+1)k_n})$ as that value i which satisfies

$$t_{(i+1)k_n} \leqq x < t_{(i+2)k_n} \tag{2.4b}$$

for $m \geqq 2$, and by that value i which satisfies

$$t_{ik_n} \leqq x < t_{(i+1)k_n} \tag{2.4c}$$

when $m = 1$.

That is to say, $\sum_{v=0}^{m} l_{i,v}(x)[(i+v)k_m+1]/(n+1)$ is the mth degree polynomial which interpolates to $F_n(t_{(i+v)k_n})$, $v = 0, 1, 2 \cdots m$. In view of the fact that

$$\sum_{v=0}^{m} l_{i,v}(x) \equiv 1 \tag{2.5}$$

we may rewrite (2.4) as

$$\hat{f}_{n,m}(x) = \frac{d}{dx}\sum_{v=0}^{m} l_{i,v}(x)\frac{vk_n}{(n+1)}, \quad x \in [t_{2k_n}, t_{(l-m+1)k_n}), i = i(x) \tag{2.6}$$
$$= 0 \quad \text{otherwise.}$$

We may now write

$$\begin{aligned} f(x) - \hat{f}_{n,m}(x) &= \left\{ f(x) - \sum_{v=1}^{m} \frac{d}{dx} l_{i,v}(x) \int_{t_{ik_n}}^{t_{(i+v)k_n}} f(\xi)\,d\xi \right\} \\ &\quad + \left\{ \sum_{v=1}^{m} \frac{d}{dx} l_{i,v}(x)\psi_{i,v} \right\} \qquad x \in [t_{2k_n}, t_{(l-m+1)k_n}) \\ &= f(x) \qquad x \notin [t_{2k_n}, t_{(l-m+1)k_n}) \end{aligned} \tag{2.7}$$

where

$$i = i(x) \tag{2.8}$$
$$\psi_{i,v} = F(t_{(i+v)k_n}) - F(t_{ik_n}) - \frac{vk_n}{n+1}$$

and $F(t) = \int_{-\infty}^{t} f(\xi)d\xi$. It is appropriate to view the two terms in brackets in (2.7) as the bias and the variance terms, respectively.

From (2.7) we may write

$$\begin{aligned} |f(x) - \hat{f}_{n,m}(x)|^2 &\leqq 2\left| f(x) - \sum_{v=1}^{m} \frac{d}{dx} l_{i,v}(x) \int_{t_{ik_n}}^{t_{(i+v)k_n}} f(\xi)\,d\xi \right|^2 \\ &\quad + 2m \sum_{v=1}^{m} \left(\frac{d}{dx} l_{i,v}(x)\right)^2 \psi_{i,v}^2, \qquad x \in [t_{2k_n}, t_{(l-m+1)k_n}) \\ &= f^2(x) \qquad x \notin [t_{2k_n}, t_{(l-m+1)k_n}). \end{aligned}$$

The bias term may be studied via Theorem 3, which we state below and prove in Section 3.

THEOREM 3. *Let $x_0 < x_1 < \cdots x_m$ be $m+1$ real numbers and suppose $f(x)$ satisfies $f^{(r)}(x)$ absolutely continuous on $[x_0, x_m]$, $f^{(m)}(x) \in L_2[x_0, x_m]$, $m = r+1$. Then*

$$\left| f(x) - \sum_{v=1}^{m} \frac{d}{dx} l_v(x; x_0, x_1, \cdots x_m) \int_{x_0}^{x_v} f(\xi)\,d\xi \right|^2 \tag{2.10}$$
$$\leqq a(m) \int_{x_0}^{x_m} [f^{(m)}(\xi)]^2 \,d\xi \,|x_m - x_0|^{2m-1}$$

with

$$a(1) = 1 \qquad x \in [x_0, x_m], m = 1, 2$$
$$a(2) = (\tfrac{5}{2})^2 \qquad x \in [x_1, x_{m-1}], m = 3, 4, \cdots^2$$
$$a(m) = \left[\frac{2(m+3)}{(m-1)}\right]^2, \qquad m \geqq 3.$$

Then, applying (2.10) to (2.9) we may write

$$(2.11) \quad |f(x)-f_{n,m}(x)|^2 < 2a(m)\int_{t_{ik_n}}^{t_{(i+m)k_n}} [f^{(m)}(\xi)]^2\, d\xi \,|t_{(i+m)k_n} - t_{ik_n}|^{2m-1} + 2m \sum_{\nu=1}^{m} \left[\frac{d}{dx} l_{i,\nu}(x)\right]^2 \psi_{i,\nu}^2, \quad i = i(x), x \in [t_{2k_n}, t_{(l-m+1)k_n})$$
$$\leqq f^2(x) \qquad x \notin [t_{2k_n}, t_{(l-m+1)k_n}).$$

In the case $|f^{(m)}(\xi)| \leqq c$, $-\infty < \xi < \infty$ we may write

$$(2.12) \quad |f(x)-\hat{f}_{n,m}(x)|^2 \leqq 2a(m)c^2|t_{(i+m)k_n} - t_{ik_n}|^{2m} + 2m \sum_{\nu=1}^{m} \left(\frac{d}{dx} l_{i,\nu}(x)\right)^2 \psi_{i,\nu}^2, \quad i = i(x), x \in [t_{2k_n}, t_{(l-m+1)k_n})$$
$$= f^2(x), \qquad x \notin [t_{2k_n}, t_{(l-m+1)k_n}).$$

Therefore

$$(2.13) \quad E|f(x)-\hat{f}_{n,m}(x)|^2 \leqq \max_i 2a(m)\int_{-\infty}^{\infty} [f^{(m)}(\xi)]^2\, d\xi\, E|t_{(i+m)k_n} - t_{ik_n}|^{2m-1} + 2m \sum_{\nu=1}^{m} E^{\frac{1}{2}}\left[\frac{d}{dx} l_{i,\nu}(x)\right]^4 E^{\frac{1}{2}}\psi_{i,\nu}^4 + f^2(x) \cdot P_r\{x \notin [t_{2k_n}, t_{(l-m+1)k_n})\}$$

$$(2.14) \quad \leqq \max_i 2a(m) \sup_{-\infty<\xi<\infty} |f^{(m)}(\xi)|^2 E|t_{(i+m)k_n} - t_{ik_n}|^{2m} + 2m \sum_{\nu=1}^{m} E^{\frac{1}{2}}\left[\frac{d}{dx} l_{i,\nu}(x)\right]^4 E^{\frac{1}{2}}\psi_{i,\nu}^4 + f^2(x) \cdot P_r\{x \notin [t_{2k_n}, t_{(l-m+1)k_n})\}.$$

We now proceed to bound the expressions on the right of (2.13) and (2.14).

Since

$$(2.15) \quad \frac{d}{dx} l_{i,\nu}(x) = \sum_{\mu=0,\mu\neq\nu}^{m} \frac{\prod_{\xi=0,\xi\neq\mu,\xi\neq\nu}^{m}(x - t_{(i+\xi)k_n})}{\prod_{\xi=0,\xi\neq\nu}^{m}(t_{(i+\nu)k_n} - t_{(i+\xi)k_n})}.$$

We have, as a loose upper bound, good for $t_{ik_n} \leqq x \leqq t_{(i+m)k_n}$,

$$(2.16) \quad \left|\frac{d}{dx} l_{i,\nu}(x)\right| \leqq m(t_{(i+m)k_n} - t_{ik_n})^{m-1} \frac{1}{\min_{\nu=0,1,\cdots,m-1}(t_{(i+\nu+1)k_n} - t_{(i+\nu)k_n})^m}$$

[2] We believe that the Theorem is also true for $x \in [x_0, x_m]$, $m \geqq 3$, but have been unable to obtain a general proof.

and

$$(2.17)\quad E^{\frac{1}{2}}\left|\frac{d}{dx}l_{i,v}(x)\right|^4 \leqq m^2 E^{\frac{1}{4}}(t_{(i+m)k_n}-t_{ik_n})^{8(m-1)} \cdot E^{\frac{1}{4}}\left(\frac{1}{\min_{v=0,1,\cdots,m-1}(t_{(i+v+1)k_n}-t_{(i+v)k_n})^{8m}}\right).$$

We will use the following Lemma 1, proved in the Appendix:

LEMMA 1. (a) *Suppose* $f(u) \geqq \lambda$ *for* $u \in [x-\varepsilon, x+\varepsilon]$, *and* $\max_{u \geqq x+\varepsilon} |u(1-F(u))| \leqq M$, $\max_{u \leqq x-\varepsilon}|uF(u)| \leqq M$. *Let* $i = i(x)$ *be the random integer defined by* (2.4b) *and* (2.4c) *for* $x \in [t_{2k_n}, t_{(l-m+1)k_n})$, *and* $i(x) = -(m+1)$ *otherwise. Let* $t_\xi = 0$ *for* $\xi < 0$. *Then, for fixed* p $(1 \leqq p \leqq k_n/2)$,

$$(2.18a)\qquad E(t_{(i+m)k_n}-t_{ik_n})^p \leqq \frac{1}{\lambda^p}\left(\frac{mk_n}{n+1}\right)^p (m)\left(1+O\left(\frac{1}{k_n}\right)\right).$$

(b) *Suppose* $f(u) \leqq \Lambda$. *Then, for any* j, *and fixed* $q(1 \leqq q \leqq k_n)$, *and* $m < k_n$,

$$(2.18b)\qquad E(t_{(j+m)k_n}-t_{jk_n})^{-q} \leqq \Lambda^q\left(\frac{n+1}{(mk_n)}\right)^q\left(1+O\left(\frac{1}{k_n}\right)\right).$$

Thus, assuming the hypotheses of the Lemma,

$$(2.19)\qquad E^{\frac{1}{2}}\left|\frac{d}{dx}l_{i,v}(x)\right|^4 \leqq m^{2m+\frac{1}{4}}\frac{\Lambda^{2m}}{\lambda^{2(m-1)}}\left[\left(1+O\left(\frac{1}{k_n}\right)\right)\right]\cdot\left(\frac{n+1}{k_n}\right)^2.$$

The $\{\psi_{i,v}\}_{v=1}^m$ are centered coverages, that is

$$(2.20)\qquad \psi_{i,v} \sim \rho_v - \frac{vk_n}{n+1}$$

where

$$(2.21)\qquad \rho_v \sim Be(vk_n, n-vk_n+1)$$

$$E\rho_v = \frac{vk_n}{(n+1)}.$$

In the Appendix we show the following

LEMMA 2.

$$(2.22)\qquad E^{\frac{1}{2}}\psi_{i,v}^4 \leqq \frac{3^{\frac{1}{2}}vk_n}{(n+1)^2}\left(1+O\left(\frac{vk_n}{n+2}\right)\right)^{\frac{1}{2}}.$$

We next invoke Lemma 3, proved in the Appendix:

LEMMA 3. *Let* $n \to \infty$, $k_n/n \to 0$, x *such that* $F(x) > 0$, l *the greatest integer in* $(n-1)/k_n$, *and* m *fixed. Then*

$$(2.23)\qquad P_r\{x \notin [t_{2k_n}, t_{(l-m+1)k_n})\} = O\left(\frac{k_n}{n^2}\right).$$

Putting together (2.13) and (2.14) with (2.18), (2.19), (2.22) and (2.23) gives

$$E|f(x)-\hat{f}_{n,m}(x)|^2 \leqq \left\{A\left(\frac{k_n}{(n+1)}\right)^{2m-1} + B\frac{1}{k_n}\right\} + O\left(\frac{k_n}{n^2}\right) \tag{2.24}$$

$$\leqq \left\{C\left(\frac{k_n}{(n+1)}\right)^{2m} + B\frac{1}{k_n}\right\} + O\left(\frac{k_n}{n^2}\right) \tag{2.25}$$

where

$$A = 2a(m)\int_{-\infty}^{\infty} [f^{(m)}(\xi)]^2\, d\xi \cdot \left(\frac{m}{\lambda}\right)^{2m-1} m\left(1+O\left(\frac{1}{k_n}\right)\right) \tag{2.26a}$$

$$B = m^{2m+3\frac{1}{4}} \frac{\Lambda^{2m}}{\lambda^{2(m-1)}} 3^{\frac{1}{2}}\left(1+O\left(\frac{1}{k_n}\right)+O\left(\frac{k_n}{n}\right)\right) \tag{2.26b}$$

$$C = 2a(m)\sup_{-\infty<\xi<\infty} |f^{(m)}(\xi)|^2 \left(\frac{m}{\lambda}\right)^{2m} m\left(1+O\left(\frac{1}{k_n}\right)\right). \tag{2.26c}$$

A lemma given by Parzen (1962, Lemma 4a) tells us how to choose k_n to minimize the terms in brackets on the right-hand side of (2.24) and (2.25), namely, take[3]

$$k_n = \left(\frac{B}{(2m-1)A}\right)^{1/2m} (n+1)^{(2m-1)/2m}, \tag{2.27}$$

for (2.24), and

$$k_n = \left(\frac{B}{2mC}\right)^{1/(2m+1)} (n+1)^{2m/(2m+1)}, \tag{2.28}$$

for (2.25).

We then have

$$E|f(x)-\hat{f}_{n,m}(x)|^2 \leqq Dn^{-(2m-1)/2m} + o(n^{-(2m-1)/2m}) \tag{2.29}$$

$$\leqq Gn^{-(2m)/(2m+1)} + o(n^{-(2m)/(2m+1)}) \tag{2.30}$$

where

$$D = \frac{2m}{(2m-1)^{2m-1}}(AB^{2m-1})^{1/2m} \tag{2.31}$$

$$G = \frac{2m+1}{2m^{2m}}(CB^{2m})^{1/(2m+1)}. \tag{2.32}$$

We have thus proved:

---

[3] We assume $A,\ C \neq 0$. The dominant term of A and C equals 0 if f is a polynomial of degree $\leqq m-1$ on its support set. In this case we would like k_n as large as possible, which happens if exactly m order statistics are used to estimate the density.

THEOREM 1. *Let* $f(u) \leqq \Lambda$, *all* u, *let* $f(u) \geqq \lambda$ *for* u *in a neighborhood of* x, *let* $|u(1-F(u))|$ *and* $|uF(u)|$ *be bounded, respectively, for* $u \geqq x$ *and* $u \leqq x$. *Let* $f^{(\nu)}$, $\nu = 0, 1, 2, \cdots r$ *be continuous, let* $f^{(m)} \in L_2[\infty, \infty]$, $m = r+1$, *and let the estimates* $\hat{f}_{n,m}(x)$ *be given by* (2.4), *with* k_n *chosen as in* (2.27). *Then*

$$E|f(x)-f_{n,m}(x)|^2 \leqq Dn^{-(2m-1)/2m}+o(n^{-(2m-1)/2m}) \tag{2.33}$$

where D *is given by* (2.31).

THEOREM 2. *Let* $f(x)$ *satisfy the assumptions of Theorem* 1, *and in addition suppose* $\sup_{\xi \in [\infty,\infty]} |f(\xi)|^2 < \infty$. *Then, if* k_n *is chosen as in* (2.28),

$$E|f(x)-\hat{f}_{n,m}(x)|^2 \leqq Gn^{-2m/(2m+1)}+o(n^{-2m/(2m+1)}) \tag{2.34}$$

where G *is given by* (2.32).

3. The interpolation theorem. This section is given over to the proof of the following:

THEOREM 3. *Let* $x_0 < x_1 < \cdots < x_m$ *be* $m+1$ *real numbers and suppose* $f(x)$ *satisfies* $f^{(\nu)}(x)$ *absolutely continuous on* $[x_0, x_m]$, $\nu = 0, 1, 2, \cdots r$, $f^{(m)}(x) \in L_2[x_0, x_m]$, $m = r+1$. *Let* $l_\nu(x) = l_\nu(x; x_0, x_1, \cdots x_m)$ *be the* mth *degree polynomial with* $l_\nu(x_\mu) = \delta_{\mu,\nu}$, $\mu, \nu = 0, 1, \cdots m$. *Then*

$$\left|f(x)-\sum_{\nu=1}^{m}\frac{d}{dx}l_\nu(x)\int_{x_0}^{x_\nu}f(\xi)\,d\xi\right|^2 \leqq a(m)\textstyle\int_{x_0}^{x_m}[f^{(m)}(\xi)]^2\,d\xi\,|x_m-x_0|^{2m-1} \qquad x\in[x_0,x_m], m=1,2, \quad x\in[x_1,x_{m-1}], m\geqq 3 \tag{3.1}$$

with

$$\begin{aligned} a(1) &= 1 \\ a(2) &= (\tfrac{5}{2})^2 \\ a(m) &= \left[\frac{2(m+3)}{(m-1)!}\right]^2, \qquad m \geqq 3. \end{aligned} \tag{3.2}$$

PROOF. The assumptions on f tell us that it has a Taylor series expansion in $[x_0, x_m]$ of the form

$$f(x) = \sum_{\nu=0}^{m-1} f^{(\nu)}(x_0)\frac{x^\nu}{\nu!}+\int_{x_0}^{x_m}\frac{(x-u)_+^{m-1}}{(m-1)!}f^{(m)}(u)\,du \qquad x_0 \leqq x \leqq x_m \tag{3.3}$$

where

$$\begin{aligned} (u)_+ &= u, & u &\geqq 0 \\ &= 0 & &\text{otherwise.} \end{aligned} \tag{3.4}$$

We may then write

$$(3.5)\quad f(x)-\tilde{f}(x)=\left\{\sum_{\nu=0}^{m-1} f^{(\nu)}(x_0)\frac{x^\nu}{\nu!}-\frac{d}{dx}\sum_{\mu=1}^{m} l_\mu(x)\sum_{\nu=0}^{m-1} f^{(\nu)}(x_0)\int_{x_0}^{x_\mu}\frac{\xi^\nu}{\nu!}d\xi\right\}$$
$$+\left\{\int_{x_0}^{x_m} f^{(m)}(u)\left[\frac{(x-u)_+^{m-1}}{(m-1)!}-\frac{d}{dx}\sum_{\mu=1}^{m} l_\mu(x)\int_{x_0}^{x_\mu}\frac{(\xi-u)_+^{m-1}}{(m-1)!}d\xi\right]du\right\},$$

where we are writing

$$(3.6)\qquad \tilde{f}(x)=\sum_{\nu=1}^{m}\frac{d}{dx}l_\nu(x)\int_{x_0}^{x_\nu}f(\xi)\,d\xi.$$

We first show that the first term in (3.5) is identically zero. By examining the coefficient of $f^{(\nu)}(x_0)$, $\nu = 0, 1, 2, \cdots m-1$, it is sufficient to show that

$$(3.7)\qquad \frac{x^\nu}{\nu!}=\frac{d}{dx}\sum_{\mu=1}^{m} l_\nu(x)\int_{x_0}^{x_\mu}\frac{\xi^\nu}{\nu!}d\xi.$$

Integrating both sides of (3.7) from x_0 to x, it is sufficient to show that

$$(3.8)\qquad \int_{x_0}^{x}\frac{\xi^\nu}{\nu!}d\xi=\sum_{\mu=1}^{m} l_\mu(x)\int_{x_0}^{x_\mu}\frac{\xi^\nu}{\nu!}d\xi.$$

Since both sides of this equation are polynomials of degree no greater than m, it is sufficient to show that they coincide at m points. But the right-hand side is exactly that polynomial which interpolates to

$$\int_{x_0}^{x}\frac{\xi^\nu}{\nu!}d\xi \quad \text{for} \quad x=x_0,x_1,\cdots,x_m.$$

We can now use (3.5) with the term in brackets set equal to zero, and the Cauchy-Schwarz inequality to write

$$(3.9)\quad |f(x)-\tilde{f}(x)|^2\leqq\int_{x_0}^{x_m}[f^{(m)}(u)]^2\,du\int_{x_0}^{x_m}$$
$$\cdot\left[\frac{(x-u)_+^{m-1}}{(m-1)!}-\frac{d}{dx}\sum_{\mu=1}^{m} l_\mu(x)\int_{x_0}^{x_\mu}\frac{(\xi-u)_+^{m-1}}{(m-1)!}d\xi\right]^2 du.$$

It is our purpose to examine the integrand

$$(3.10)\qquad \left[\frac{(x-u)_+^{m-1}}{(m-1)!}-\frac{d}{dx}\sum_{\mu=1}^{m} l_\mu(x)\int_{x_0}^{x_\mu}\frac{(\xi-u)_+^{m-1}}{(m-1)!}d\xi\right]^2.$$

Let $h_u(x)$ be defined, for u, $x\in[x_0, x_m]$ by

$$(3.11)\qquad h_u(x)=\int_{x_0}^{x}\frac{(\xi-u)_+^{m-1}}{(m-1)!}d\xi=\frac{(x-u)_+^m}{m!}$$

and $p_u(x)$ by

$$(3.12)\quad p_u(x) = \sum_{\nu=1}^{m} l_\nu(x) \int_{x_0}^{x_\nu} \frac{(\xi-u)_+^{m-1}}{(m-1)!}\, d\xi = \sum_{\nu=0}^{m} l_\nu(x) h_u(x_\nu) = \sum_{\nu=1}^{m} l_\nu(x) h_u(x_\nu),$$

thus $p_u(x)$ is the mth degree polynomial which interpolates to $h_u(x)$ at the points $x_0, x_1, \cdots x_m$.

Thus (3.9) may be written

$$(3.13)\quad |f(x)-\tilde{f}(x)|^2 \leqq \int_{x_0}^{x_m} [f^{(m)}(u)]^2\, du \int_{x_0}^{x_m} \left[\frac{d}{dx}(h_u(x)-p_u(x))\right]^2 du.$$

We calculate directly a bound on $|(d/dx)(h_u(x)-p_u(x))|$ for $m = 1, 2$, and then give a general bound good for $m \geqq 3$.

For $m = 1$

$$h_u(x)-p_u(x) = (x-u)_+ - \frac{(x-x_0)}{(x_1-x_0)}(x_1-u)$$

and

$$(3.14)\quad \left|\frac{d}{dx}(h_u(x)-p_u(x))\right| = \left|(x-u)_+^0 - \frac{(x_1-u)}{(x_1-x_0)}\right| \leqq 1.$$

For $m = 2$

$$(3.15)\quad h_u(x)-p_u(x) = \frac{(x-u)_+^2}{2!} - \left\{\frac{(x-x_0)(x-x_2)}{(x_1-x_0)(x_1-x_2)}\frac{(x_1-u)_+^2}{2!} + \frac{(x-x_0)(x-x_1)}{(x_2-x_1)(x_2-x_0)}\frac{(x_2-u)^2}{2!}\right\}.$$

We have

$$(3.16)\quad \left|\frac{d}{dx}h_u(x)\right| = |(x-u)_+| \leqq |x_2-x_0|.$$

The maximum of $|(d/dx)p_u(x)|$ clearly occurs at $x = x_2$. We have

$$(3.17)\quad \frac{d}{dx}p_u(x)\Big|_{x=x_2} = \frac{(x_2-x_0)}{(x_1-x_0)(x_1-x_2)}\frac{(x_1-u)_+^2}{2!} + \frac{(x_2-x_1)+(x_2-x_0)}{(x_2-x_1)(x_2-x_0)}\frac{(x_2-u)^2}{2!}.$$

For $u \geqq x_1$, the first term is zero, and since $(x_2-u)^2 \leqq (x_2-x_1)^2$, the second term is clearly bounded in absolute value by $|x_2-x_0|$. For $x_0 \leqq u \leqq x_1$, a rearrangement of terms gives

$$(3.18)\quad \frac{d}{dx}p_u(x)\Big|_{x=x_2} = \frac{1}{2!}\left\{\frac{(x_2-u)^2}{(x_2-x_0)} - \frac{(x_1-u)^2}{(x_1-x_0)} + (x_2-u)+(x_1-u)\right\} \quad x_0 \leqq u < x_1,$$

which is clearly bounded in absolute value by $\frac{3}{2}\,|x_2-x_0|$. Hence

$$(3.19)\quad \left|\frac{d}{dx}[(h_u(x)-p_u(x)]\right| \leqq \tfrac{5}{2}|x_2-x_0| \qquad m = 2.$$

We now assume $m \geqq 3$.

By the Newton form of the remainder for Lagrange interpolation (see, for example, Isaacson and Keller (1966, page 248), we have, that

$$h_u(x) - \sum_{v=0}^m l_v(x) h_u(x_v) = \prod_{v=0}^m (x - x_v) h_u[x_0, x_1, \cdots, x_m, x] \tag{3.20}$$

where $h\,[x_0, x_1, \cdots x_m, x]$ is the $(m+1)$st order divided difference of h_u at the points $x_0, x_1, \cdots x_m, x$. It will be convenient to use identities relating the $(m+1)$st to the mth and $(m-1)$st order divided differences, in particular

$$\begin{aligned} h_u[x_0, x_1, \cdots, x_m, x] &= \frac{h_u[x_1, \cdots, x_m, x] - h_u[x_0, \cdots, x_{m-1}, x]}{(x_m - x_0)} \\ &= \frac{1}{(x_m - x_0)} \left\{ \frac{h_u[x_2, \cdots, x_m, x] - h_u[x_0, \cdots, x_{m-1}, x]}{(x_m - x_1)} \right. \\ &\quad \left. - \frac{h_u[x_1, \cdots, x_{m-1}, x] - h_u[x_0, \cdots, x_{m-2}, x]}{(x_{m-1} - x_0)} \right\}. \end{aligned} \tag{3.21}$$

Thus we may combine (3.20) and (3.21) to write

$$\begin{aligned} &\frac{d}{dx}(h_u(x) - p_u(x)) \\ &= \sum_{v=0}^m \left(\prod_{j \neq v} (x - x_j) \right) \left\{ \frac{h_u[x_1, x_2, \cdots, x_m, x] - h_u[x_0, x_1, \cdots, x_{m-1}, x]}{(x_m - x_0)} \right\} \\ &\quad + \frac{\prod_{v=0}^m (x - x_v)}{(x_m - x_0)} \left\{ \frac{d}{dx} \left[\frac{h_u[x_2, \cdots, x_m, x] - h_u[x_1, \cdots, x_{m-1}, x]}{(x_m - x_1)} \right. \right. \\ &\quad \left. \left. - \frac{h_u[x_1, \cdots, x_{m-1}, x] - h_u[x_0, \cdots, x_{m-2}, x]}{(x_{m-1} - x_0)} \right] \right\}. \end{aligned} \tag{3.22}$$

Now if $y_0 < y_1 < \cdots < y_m$ are any $m+1$ points in the interval $[x_0, x_m]$, we show that

$$|h_u[y_0, y_1, \cdots, y_m]| \leqq \sup_{x_0 \leqq \xi \leqq x_m} \frac{1}{(m-1)!} |h_u^{(m)}(\xi)|. \tag{3.23}$$

This follows by writing

$$|h_u[y_0, y_1, \cdots, y_m]| = \left| \frac{h_u[y_1, y_2, \cdots, y_m] - h_u[y_0, y_1, \cdots, y_{m-1}]}{(y_m - y_0)} \right|. \tag{3.24}$$

Then, since h has $m-1$ continuous derivatives, we may write, by the mean value theorem, that for some $\xi_2 \in [y_1, y_m]$, $\xi_1 \in [y_0, y_{m-1}]$,

$$\begin{aligned} h_u[y_1, y_2, \cdots, y_m] &= \frac{1}{(m-1)!} h_u^{(m-1)}(\xi_2) \\ h_u[y_0, y_1, \cdots, y_{m-1}] &= \frac{1}{(m-1)!} h_u^{(m-1)}(\xi_1) \end{aligned} \tag{3.25}$$

and

$$(3.26)\quad h_u[y_0, y_1, \cdots, y_m]$$

$$= \frac{1}{(m-1)!}\left|\frac{h_u^{(m-1)}(\xi_2) - h_u^{(m-1)}(\xi_1)}{y_m - y_0}\right| \leqq \sup_{x_0 \leqq \xi \leqq x_m} \frac{1}{(m-1)!}\left|h_u^{(m)}(\xi)\right|.$$

Similarly, it can be shown that

$$(3.27)\quad \frac{d}{dx} h_u[y_0, y_1, \cdots, y_{m-2}, x] = \lim_{\Delta \to 0} h_u[y_0, y_1, \cdots, y_{m-2}, x, x+\Delta]$$

$$\leqq \sup_{x_0 \leqq \xi \leqq x_m} \frac{1}{(m-1)!}\left|h_u^{(m)}(\xi)\right|.$$

Now, for $x_0 \leqq u \leqq x_m$, we have

$$h_u^{(m)}(x) = 1 \qquad x > u$$

$$h_u^{(m)}(x) = 0 \qquad x < u.$$

Thus, combining (3.22), (3.26) and (3.27) results, for $x_1 \leqq x \leqq x_{m-1}$, in

$$(3.28)\quad \frac{d}{dx}(h_u(x) - p_u(x))$$

$$\leqq \frac{2}{(m-1)!}\left\{\sum_{v=0}^{m}\left|\frac{\prod_{j=0,\, j\neq v}^{m}(x - x_j)}{(x_m - x_0)}\right| + \left|\frac{\prod_{j=0}^{m}(x - x_j)}{(x_m - x_0)}\left(\frac{1}{(x_m - x_1)} + \frac{1}{(x_{m-1} - x_0)}\right)\right|\right\}$$

$$\leqq 2\frac{(m+3)}{(m-1)!}\left|x_m - x_1\right|^{m-1}.$$

Substituting (3.14), (3.19) and (3.28) into (3.13) gives the theorem.

4. Concluding remarks. This work was prompted partly by the results of an investigation into the use of polynomial spline interpolation for density estimation. Suppose we are given $l+1$ ordinates y_i, $i = 0, 1, 2, \cdots l$, and $l+1$ abscissae $F(y_i) = i/l$, $i = 0, 1, 2, \cdots l$, of the function $F(x)$. Let m be an odd positive integer and let $F(x)$ be that function in the class of all functions defined on $[y_0, y_l]$ with absolutely continuous $(m-1)/2$th derivative and square integrable $(m+1)/2$th derivative, which minimizes

$$\int_{y_0}^{y_l} [\tilde{F}^{(m+1)/2}(u)]^2\, du$$

subject to the conditions $\tilde{F}(y_i) = F(y_i)$, $i = 0, 1, 2, \cdots l$. The solution $\tilde{F}(x)$ is a polynomial spline on $[y_0, y_l]$; it is a polynomial of degree m in each interval $[y_i, y_{i+1}]$, $i = 0, 1, 2, \cdots l-1$, and it possesses $m-1$ continuous derivatives on

$[y_0, y_l]$. Suppose $f(x) = F'(x)$, $x \in [y_0, y_l]$, and $f(x)$ has $m-1$ continuous and m square integrable derivatives. Then theorems of the form

$$\left| f(x) - \frac{d}{dx}\tilde{F}(x) \right|^2 = O(\Delta^{2m-1}) \tag{4.1}$$

where $\Delta = \sup_i |y_{i+1} - y_i|$ are well known.

By replacing (3.1) by (4.1) it is not hard to show that the bias term in a density estimate using odd degree spline interpolation goes to zero at the same rate as the bias term using Lagrange interpolation of the same order. It is conjectured that the variance term using spline interpolation also behaves as the variance term using Lagrange interpolation, but a proof was not found. If this conjecture is true, spline interpolation for density estimation is appealing because of the convergence rates as well as the continuity properties of the estimate. Boneva, Kendall and Stefanov (1971) have recently deduced some appealing kernels to use in a kernel-type estimate, by using splines.

APPENDIX

This appendix is given over to the proofs of Lemmas 1, 2 and 3 in Section 2.

LEMMA 1. (a) *Suppose* $f(u) > \lambda$ *for* $u \in [x-\varepsilon, x+\varepsilon]$, *and* $\max_{u \geqq x+\varepsilon} |u(1-F(u))| \leqq M$, $\max_{u \leqq x-\varepsilon} |uF(u)| \leqq M$. *Let* $i = i(x)$ *be defined, for* $x \in [t_{2k_n}, t_{(l-m+1)k_n})$ *as that value* i *which satisfies* $t_{(i+1)k_n} \leqq x < t_{(i+2)k_n}$ *for* $m \geqq 2$, *and by that value* i *which satisfies* $t_{ik_n} \leqq x < t_{(i+1)k_n}$ *when* $m = 1$. *Let* $i = i(x) = -(m+1)$, *for* $x \notin [t_{2k_n}, t_{(l-m+1)k_n})$ *and let* $t_\xi \equiv 0$ *for* $\xi < 0$. *Then, for fixed* p $(1 \leqq p \leqq k_n/2)$

$$E(t_{(i+m)k_n} - t_{ik_n})^p \leqq \frac{1}{\lambda^p}\left(\frac{mk_n}{n+1}\right)^p (m)\left(1 + O\left(\frac{1}{k_n}\right)\right). \tag{A.1}$$

(b) *Suppose* $f(u) \leqq \Lambda$. *Then, for any* j, *and fixed* q $(1 \leqq q \leqq k_n)$, *and* $m < k_n$,

$$E|t_{(j+m)k_n} - t_{jk_n}|^p \leqq \Lambda^q \left(\frac{n+1}{mk_n}\right)^q \left(1 + O\left(\frac{1}{k_n}\right)\right). \tag{A.2}$$

PROOF OF (*a*). Define the events

$$\begin{aligned}
B_j&: \quad t_{(j+1)k_n} \leqq x < t_{(j+2)k_n}\\
B_j^*&: \quad t_{jk_n} \leqq x < t_{(j+m)k_n}\\
A_j(u,v)&: \quad t_{jk_n} \in (u, u+du) \cap t_{(j+m)k_n} \in (v, v+dv) \cap B_j\\
A_j^*(u,v)&: \quad t_{jk_n} \in (u, u+du) \cap t_{(j+m)k_n} \in (v, v+dv) \cap B_j^*
\end{aligned}$$

for $j = 1, 2, \cdots, l-m-1$. Since for $m \geqq 2$, $B_j \subset B_j^*$, then $A_j \subset A_j^*$, and we have

$$\begin{aligned}
E|t_{(i+m)k_n} - t_{ik_n}|^p &= \textstyle\sum_{j=1}^{l-m-1} E\{|t_{(j+m)k_n} - t_{jk_n}|^p \mid B_j\} P(B_j)\\
&= \textstyle\sum_{j=1}^{l-m-1} \int_{u \leqq x} \int_{v \geqq x} |v-u|^p P(A_j(u,v))\\
&\leqq \textstyle\sum_{j=1}^{l-m-1} \int_{u \leqq x} \int_{v \geqq x} |v-u|^p P(A_j^*(u,v)).
\end{aligned} \tag{A.3}$$

For $m = 1$, (A.3) is valid with B_j and A_j replaced by $B_j{}^*$ and $A_j{}^*$.

For simplicity, write $v_j = jk_n, j = 1, 2, \cdots, l-m-1, k = mk_n$, and define $g^n_{v,k}(u, v)$ by

(A.4) $$g^n_{v,k}(u,v) = \frac{n!}{(v-1)!(k-1)!(n-v-k)!} F^{v-1}(u)[F(v)-F(u)]^{k-1}[1-F(v)]^{n-v-k}.$$

Then,

(A.5) $$\begin{aligned} P(A_j{}^*(u,v)) &= g^n_{v_j,k}(u,v)\,du\,dv, && u \leqq x, v \geqq x \\ &= 0 && \text{otherwise.} \end{aligned}$$

Now, note that, for $x-\varepsilon \leqq u \leqq x \leqq v \leqq x+\varepsilon$,

(A.6) $$\frac{|v-u|}{F(v)-F(u)} \leqq \frac{1}{\min_{x-\varepsilon \leqq \xi \leqq x+\varepsilon} f(\xi)} \leqq \frac{1}{\lambda}.$$

Also, for $u \leqq x-\varepsilon$, $v \geqq x+\varepsilon$,

(A.7) $$F(v)-F(u) \geqq F(x+\varepsilon)-F(x-\varepsilon) \geqq 2\lambda\varepsilon,$$

and so, using $|v-u|^p \leqq 2^{p-1}[|v|^p+|u|^p]$, and the assumptions on the tail behavior of F one obtains

(A.8) $$\begin{aligned} &|v-u|^p F^{v_j-1}(u)[F(v)-F(u)]^{k-1}|1-F(v)|^{n-v_j-k} \\ &\leqq 2^{p-1}\frac{M^p}{(2\lambda\varepsilon)^{2p}} F^{v_j-1-p}(u)[F(v)-F(u)]^{k+2p-1}|1-F(v)|^{n-v_j-k} \\ &\quad + 2^{p-1}\frac{M^p}{(2\lambda\varepsilon)^{2p}} F^{v_j-1}(u)[F(v)-F(u)]^{k+2p-1}|1-F(v)|^{n-v_j-k-p}. \end{aligned}$$

Therefore

(A.9) $$\begin{aligned} &\sum_{j=1}^{l-m-1} \int_{x-\varepsilon \leqq u \leqq x \leqq v \leqq x+\varepsilon}\int |v-u|^p P(A_j{}^*(u,v)) \\ &\quad + \sum_{j=1}^{l-m-1} \int_{u \leqq x-\varepsilon}\int_{v \geqq x+\varepsilon} |v-u|^p P(A_j{}^*(u,v)) \\ &\leqq \frac{1}{\lambda^p} \sum_{j=1}^{l-m-1} \frac{\dfrac{n!}{(v_j-1)!(k-1)!(n-v_j-k)!}}{\dfrac{(n+p)!}{(v_j-1)!(k+p-1)!(n-v_j-k)!}} \\ &\quad \cdot \int_{x-\varepsilon \leqq u \leqq x \leqq v \leqq x+\varepsilon}\int g^{n+p}_{v_j,k+p}(u,v)\,du\,dv \\ &\quad + \frac{M^p}{2^{p+1}(\lambda\varepsilon)^{2p}} \sum_{j=1}^{l-m-1} \frac{\dfrac{n!}{(v_j-1)!(k-1)!(n-v_j-k)!}}{\dfrac{(n+p)!}{(v_j-1-p)!(k+2p-1)!(n-v_j-k)!}} \\ &\quad \cdot \int_{u \leqq x-\varepsilon}\int_{v \geqq x+\varepsilon} g^{n+p}_{v_j-p,k+2p}(u,v)\,du\,dv \end{aligned}$$

$$+\frac{M^p}{2^{p+1}(\lambda\varepsilon)^{2p}}\sum_{j=1}^{l-m-1}\frac{\overline{\dfrac{n!}{(v_j-1)!(k-1)!(n-v_j-k)!}}}{\overline{\dfrac{(n+p)!}{(v_j-1)!(k+2p-1)!(n-v_j-k-p)!}}}$$

$$\cdot\int_{u\leqq x-\varepsilon}\int_{v\geqq x+\varepsilon}g^{n+p}_{v_j,k+2p}(u,v)\,du\,dv.$$

Let $C^{n+p}_{v,k}$ be the event $s_{v,n+p}\leqq x<s_{v+k,n+p}$ where $s_{v,n+p}$ is the vth order statistic from a sample of size $n+p$. Then

$$P(C^{n+p}_{v,k})=\int_{u\leqq x}\int_{v\geqq x}g^{n+p}_{v,k}(u,v)\,du\,dv. \tag{A.10}$$

If $2p\leqq k_n$, then

$$C^{n+p}_{v_j,k+p}\cap C^{n+p}_{v_l,k+p}=\varnothing \tag{A.11}$$

$$C^{n+p}_{v_j-p,k+2p}\cap C^{n+p}_{v_l-p,1+2p}=\varnothing \tag{A.12}$$

$$C^{n+p}_{v_j,k+2p}\cap C^{n+p}_{v_l,k+2p}=\varnothing \tag{A.13}$$

whenever $|j-l|\geqq m+1$.

Therefore, by (A.10) and (A.11), the first term on the right of (A.9) is bounded by

$$\frac{1}{\lambda^p}\frac{n!}{(n+p)!}\cdot\frac{(k+p-1)!}{(k-1)!}(m)=\frac{1}{\lambda^p}\left(\frac{mk_n}{n+1}\right)^p(m)\left(1+O\left(\frac{1}{k_n}\right)\right). \tag{A.14}$$

Furthermore $\sum_{j=1}^{l-m-1}\int_{u\leqq x-\varepsilon}\int_{v\geqq x+\varepsilon}g^{n+p}_{v_j-p,k+2p}(u,v)\,du\,dv$ and $\sum_{j=1}^{l-m-1}\int_{u\leqq x-\varepsilon}\int_{v\geqq x+\varepsilon}g_{v_j,k+2p}(u,v)\,du\,dv$ are both bounded by $(m+1)\times P_r$ {less than $k+2p-1$ observations out of $n+p$ in $[x-\varepsilon,\ x+\varepsilon]$}. Since $F(x+\varepsilon)-F(x-\varepsilon)\geqq 2\varepsilon\lambda$, the above probability is $O(1/n)$, (for $k_n/n\to 0$) by Chebychev's Theorem.

Thus the sum of the second and third terms in (A.9) is bounded by

$$\max_j\frac{M^p}{2^{p+1}(\lambda\varepsilon)^{2p}}\frac{n!}{(n+p)!}\frac{(k+2p-1)!}{(k-1)!}\left[\frac{(v_j-1-p)!}{(v_j-1)!}+\frac{(n-v_j-k-p)!}{(n-v_j-k)!}\right]\times O\left(\frac{1}{n}\right)$$
$$=O\left(\left(\frac{k_n}{n}\right)^p\times\frac{1}{n}\right). \tag{A.15}$$

PROOF OF (b).

$$E\left|t_{(j+m)k_n}-t_{jk_n}\right|^{-q}=\int_{u<v}\int|v-u|^{-q}g^n_{v_j,k}(u,v)\,du\,dv. \tag{A.16}$$

Using the fact that

$$\left|\frac{F(v)-F(u)}{v-u}\right|\leqq\max_u f(u)\leqq\Lambda,$$

the right-hand side of (A.16) is bounded by

$$\text{(A.17)}\quad \frac{\dfrac{n!}{(v_j-1)!(k-1)!(n-v_j-k)!}}{\dfrac{(n-q)!}{(v_j-1)!(k-q-1)!(n-v_j-k)!}}\Lambda^q\int\cdots\int_{u<v} g^{n-q}_{v_j,k-q}(u,v)\,du\,dv$$

$$=\Lambda^q\left(\frac{n+1}{mk_n}\right)^q\left(1+O\left(\frac{1}{k_n}\right)\right).$$

LEMMA 2. *Let* $\psi=\rho-[k/(n+1)]$, *where* $\rho\sim Be(k,n-k+1)$, *then*

$$E\psi^4=\frac{3k^2}{(n+1)^4}\left(1-\frac{2k}{(n+2)}+o\left(\frac{k}{n}\right)\right).$$

PROOF. Using the formula for the moments of a $Be(k,n-k+1)$ random variable

$$\mu_r=\frac{\Gamma(n+1)\Gamma(k+r)}{\Gamma(n+1+r)\Gamma(k)}$$

gives

$$E\psi^4=\mu_4-4\mu_3\mu_1+6\mu_2{\mu_1}^2-4\mu_1{\mu_1}^3+{\mu_1}^4$$
$$=\frac{(k+3)(k+2)(k+1)k}{(n+4)(n+3)(n+2)(n+1)}-\frac{4(k+2)(k+1)k^2}{(n+3)(n+2)(n+1)^2}+\frac{6(k+1)k^3}{(n+2)(n+1)^3}-\frac{3k^4}{(n+1)^4}.$$

Letting

$$f_1=1-\frac{(n+1)^3}{(n+4)(n+3)(n+2)}=\frac{6}{(n+2)}-\frac{19}{(n+2)(n+3)}+\frac{27}{(n+4)(n+3)(n+2)}$$

$$f_2=1-\frac{(n+1)^2}{(n+3)(n+2)}=\frac{3}{(n+2)}-\frac{4}{(n+3)(n+2)}$$

$$f_3=1-\frac{(n+1)}{(n+2)}=\frac{1}{(n+2)}$$

$$f_4=0$$

we have

$$E\psi^4=\frac{k}{(n+1)}\left\{\frac{1}{(n+1)^3}[(k+3)(k+2)(k+1)-4(k+2)(k+1)k+6(k+1)k^2-3k^3]\right\}$$
$$-\frac{k}{(n+1)}\left\{\frac{1}{(n+1)^3}[f_1(k+3)(k+2)(k+1)-f_24(k+2)(k+1)k+f_36(k+1)k^2]\right\}$$
$$=\frac{3k^2}{(n+1)^4}\left(1-\frac{2k}{(n+2)}+o\left(\frac{k}{n}\right)\right).$$

LEMMA 3. *Let t_ν be the* νth *order statistic of a sample of size n from a population with* c.d.f. *F. Suppose* $F(x) > 0$. (i) *Let* $\nu/m \to 0$. *Then*

$$P_r\{t_\nu > x\} \leqq \frac{\nu}{(n+1)^2(n+2)} \frac{1}{\left(F(x) - \frac{\nu}{n+1}\right)^2} = O\left(\frac{\nu}{n^2}\right).$$

(ii) *If* $(n-\nu)/n \to 0$, *then*

$$P_r\{t_\nu < x\} \leqq \frac{\nu(n-\nu+1)}{(n+1)^2(n+2)} \frac{1}{\left(\frac{\nu}{n+1} - F(x)\right)^2} = O\left(\frac{n-\nu}{n^2}\right).$$

PROOF. (i) $P_r\{t_\nu > x\} = P_r\{\rho_\nu > F(x)\}$, where $\rho_\nu \sim Be(\nu, n-\nu-1)$. But, since $Var\ \rho_\nu = \nu(n-\nu+1)/[(n+1)^2(n+2)]$, Chebychev's inequality gives for $\nu/(n+1) < F(x)$,

$$P_r\{\rho_\nu > F(x)\} \leqq P_r\left\{\left|\rho_\nu - \frac{\nu}{n+1}\right| \geqq F(x) - \frac{\nu}{n+1}\right\}$$
$$\leqq \frac{\nu(n-\nu+1)}{(n+1)^2(n+2)} \frac{1}{\left(F(x) - \frac{\nu}{n+1}\right)^2}.$$

A similar equation is written for (ii).

REFERENCES

[1] 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.

[2] ISAACSON, EUGENE and KELLER, HERBERT (1966). *Analysis of Numerical Methods*. Wiley, New York.

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

[4] VAN RYZIN, J. (1970). On a histogram method of density estimation. Univ. of Wisconsin, Department of Statistics T. R. No. 226.

[5] WEGMAN, E. J. (1970). Nonparametric probability density estimation. Univ. of North Carolina Institute of Statistics, Mimeo Series No. 638.