DEPARTMENT OF STATISTICS

TECHNICAL REPORT NO. 236

MAY 1970

A POLYNOMIAL ALGORITHM FOR DENSITY ESTIMATION

by

Grace Wahba

Typist / Lyla M. Yelk

This research was supported in part by the Air Force Office of Scientific Research under Grant AFOSR 69-1803 and by the U.S. Army Research Office under Grant No. DA-ARO-D-31-124-G917.

ABSTRACT

An algorithm for density estimation based on ordinary polynomial (Lagrange) interpolation is studied. Let $F_n(x)$ be $\frac{n}{n+1}$ times the sample c.d.f based on n order statistics, $t_0, t_1, \ldots t_{n-1}$, from a population with density $f(x)$. It is assumed that $f^{(\nu)}$ is continuous, $\nu = 0, 1, 2, \ldots, r$, $r = m-1$, and $f^{(m)} \varepsilon L_2(-\infty,\infty)$. $F_n(x)$ is first locally interpolated by the mth degree polynomial passing through $F_n(t_{ik_n})$, $F_n(t_{(i+1)k_n})$, $\ldots$ $F_n(t_{(i+m)k_n})$, where k_n is a suitably chosen number, depending on n. The density estimate is then, locally, the derivative of this interpolating polynomial. If

$$k_n = 0\left(n^{(\frac{2m-1}{2m})}\right) ,$$

then it is shown that the mean square convergence rate of the estimate to the true density is

$$0\,(n^{-(\frac{2m-1}{2m})}) .$$

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 continuous,

and

$$k_n = 0\,(n^{\frac{2m}{2m+1}}),$$

then it is shown that the mean square convergence rates are

$$0\left(n^{-\frac{2m}{2m+1}}\right),$$

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 $t_0, t_1, \ldots t_{n-1}$ 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 [2]) has the property that

$$E\left(f_n(x) - f(x)\right)^2 = 0\left(n^{-(\frac{2r}{2r+1})}\right), \quad r = 1, 2, \ldots \tag{1.1}$$

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

$$\hat{f}_{n,m}(x) = \begin{cases} 0 & , \quad x < t_{k_n} \\ \frac{d}{dx}\hat{F}_{n,m}(x), & t_{k_n} \leq x < t_{(\ell-m+1)k_n} \\ 0 & , \quad t_{(\ell-m+1)k_n} \leq x \end{cases} \tag{1.2}$$

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

For m=1,

$$\hat{F}_{n,1}(x) = F_n(t_{ik_n}) + x \frac{F_n(t_{(i+1)k_n}) - F_n(t_{ik_n})}{t_{(i+1)k_n} - t_{ik_n}}, \quad t_{ik_n} \leq x < t_{(i+1)k_n}$$

$i=1,2, \ldots, \ell-1.$

For $m \geq 2$, let $\hat{F}_{n,m,i}(x)$, $i=0,1,2, \ldots \ell-m-1$, be the mth degree polynomial which interpolates to $F_n(x)$ at the m+1 points $x = t_{ik_n}, t_{(i+1)k_n}, \ldots t_{(i+m)k_n}$. For $x \varepsilon [t_{(i+1)k_n}, t_{(i+2)k_n})$, define $\hat{F}_{n,m}(x)$ to coincide with $\hat{F}_{n,m,i}(x)$, $i=0,1,2,\ldots \ell-m-1$.

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 $x \notin [t_{k_n}, t_{(\ell-m+1)k_n})$ is arbitrarily chosen for notational convenience.

Under the assumption on f that

$$E \mid t_{(i+1)k_n} - t_{(i+2)k_n} \mid^p = 0 \left(\frac{k_n}{n+1} \right)^p, \quad |p| < 8 m \tag{*}$$

We prove

<u>Theorem 1:</u>

$$E \mid f(x) - \hat{f}_{n,m}(x) \mid^2 = 0 \left(n^{-\frac{2m-1}{2m}}\right) \qquad m = 1, 2, \ldots \tag{1.3}$$

Sufficient conditions for (*) are shown to be e.g. that f is supported on a closed interval [a,b] with $0 < \lambda \leq f(x) \leq \Lambda < \infty$, $x \varepsilon [a,b]$.

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

If, instead we assume $(r+1) = m$ continuous derivatives in a closed interval with x in the interior, and let

$k_n \sim \text{const.}\ (m,f)\ n^{\frac{2m}{2m+1}}$, we prove, assuming (*)

Theorem 2:

$$E \left| f(x) - \hat{f}_{n,m}(x) \right|^2 = 0 \left(n^{-\frac{2m}{2m+1}} \right) \tag{1.4}$$

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, which might be viewed as the sum of a squared bias and a 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 $x_0 < x_1 < \ldots x_m$ are m+1 real numbers, and $f^{(\nu)}, \nu = 0, 1, 2, \ldots r$ absolutely continuous on $[x_0, x_m]$, $f^{(m)} \varepsilon L_2[x_0, x_m]$. Let $\ell_\nu(x; x_0, x_1, \ldots x_m) = \ell_\nu(x)$ be the mth degree polynomials satisfying $\ell_\nu(x_\mu) = \delta_{\mu,\nu}$, $\mu, \nu = 0, 1, 2 \ldots m$. Then we have

Theorem 3

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

$$x \varepsilon [x_0, x_m]\ ,\ m = 1, 2$$

$$x \varepsilon [x_1, x_{m-1}]\ ,\ m \geq 3$$

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. (see [3], "unsymmetric case"). He obtained the interesting result that if $k_n = o(n^{2/3})$, and x is a point at which f' exists and is continuous, then

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

Van Ryzin's theorem tells us what happens if we proceed here as though f' was only square integrable (e.g. $k_n = 0(n^{1/2})$) but in fact f' exists and is continous at x.

We remind the reader that an extensive literature exists on density estimation. For a bibliography, see [4].

2. DESCRIPTION OF THE ALGORITHM AND THE MAIN THEOREMS

It is convenient to have some general formulae for interpolating polynomials. Let $x_0, x_1, \ldots x_m$ be m + 1 distinct real numbers . Let $\ell_\nu(x)$ be defined by

$$\ell_\nu(x) = \ell_\nu(x; x_0, x_1, \ldots x_m) = \frac{\prod_{\substack{\mu=0\\ \mu\neq\nu}}^{m} (x-x_\mu)}{\prod_{\substack{\mu=0\\ \mu\neq\nu}}^{m} (x_\nu - x_\mu)}, \nu = 0, 1, 2, \ldots m \tag{2.1}$$

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

$$\ell_\nu(x_\mu) = \begin{cases} 1, & \mu=\nu \\ 0, & \mu\neq\nu \end{cases} \tag{2.2}$$

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

$$\hat{\ell}_{i,\nu}(x) = \ell_\nu (x; t_{ik_n}, t_{(i+1)k_n}, \ldots t_{(i+m)k_n}) \tag{2.3}$$

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

$$\hat{f}_{n,m}(x) = \frac{d}{dx} \sum_{\nu=0}^{m} \hat{\ell}_{i,\nu}(x) \frac{(i+\nu)k_n+1}{(n+1)}, \quad i = i(x)$$

$$x \varepsilon [t_{k_n}, t_{(\ell-m+1)k_n}) \tag{2.4a}$$

$= 0$ otherwise

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

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

for $m \geq 2$, and by

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

when $m = 1$.

That is to say,

$$\sum_{\nu=0}^{m} \ell_{i,\nu}(x) \frac{(i+\nu)k_m+1}{(n+1)}$$

is the mth degree polynomial which interpolates to $F_n(t_{(i+\nu)k_n}), \nu = 0, 1, 2 \ldots m$. In view of the fact that

$$\sum_{\nu=0}^{m} \hat{\ell}_{i,\nu}(x) \equiv 1 \tag{2.5}$$

we may rewrite (2.4) as

$$\hat{f}_{n,m}(x) = \frac{d}{dx}\sum_{\nu=0}^{m}\hat{\ell}_{i,\nu}(x)\,\frac{\nu k_n}{(n+1)}\ ,\quad x\varepsilon[t_{k_n},\ t_{(\ell-m+1)k_n}) \quad i=i(x) \tag{2.6}$$

$$= 0 \text{ otherwise}$$

We may now write

$$f(x) - \hat{f}_{n,m}(x) = \left\{f(x) - \sum_{\nu=1}^{m}\frac{d}{dx}\hat{\ell}_{i,\nu}(x)\int_{t_{ik_n}}^{t_{(i+\nu)k_n}} f(\xi)d\xi\right\}$$

$$+\left\{\sum_{\nu=1}^{m}\frac{d}{dx}\hat{\ell}_{i,\nu}(x)\ \psi_{i,\nu}\right\}\quad x\varepsilon[t_{k_n},\ t_{(\ell-m+1)k_n}) \tag{2.7}$$

$$= f(x) \qquad x\notin[t_{k_n},\ t_{(\ell-m+1)k_n})$$

where

$$i = i(x)$$

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

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

$$|f(x) - \hat{f}_{n,m}(x)|^2 \le 2\left|f(x) - \sum_{\nu=1}^{m}\frac{d}{dx}\hat{\ell}_{i,\nu}(x)\int_{t_{ik_n}}^{t_{(i+\nu)k_n}} f(\xi)d\xi\ \right|^2 \tag{2.9}$$

$$+ 2m\sum_{\nu=1}^{m}\left(\frac{d}{dx}\hat{\ell}_{i,\nu}(x)\right)^2\ \psi_{i,\nu}^2\ ,\quad x\varepsilon[t_{k_n}, t_{(\ell-m+1)k_n})$$

$$= f^2(x) \qquad x\notin[t_{k_n}, t_{(\ell-m+1)k_n})$$

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

Theorem 3 Let $x_0 < x_1 < \ldots x_m$ be m+1 real numbers and suppose f(x) satisfies $f^{(\nu)}(x)$ absolutely continuous on $[x_o, x_m]$, $f^{(m)}(x) \varepsilon L_2[x_o, x_m]$, m = r+1. Then

$$|f(x) - \sum_{\nu=1}^{m} \frac{d}{dx} \ell_\nu (x; x_0, x_1, \ldots x_m) \int_{x_0}^{x} f(\xi) d\xi|^2 \tag{2.10}$$

$$\leqslant a(m) \int_{x_0}^{x_m} [f^{(m)}(\xi)]^2 d\xi |x_m - x_0|^{2m-1}$$

with

$a(1) = 1$ $\qquad x \varepsilon [x_0, x_m]$, m = 1, 2

$a(2) = (5/2)^2$ $\qquad x \varepsilon [x_1, x_{m-1}]$, m = 3, 4, ... 1|

$a(m) = \left[\frac{2(m+3)}{(m-1)}\right]^2$, $m \geq 3$

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

$$|f(x) - \hat{f}_{n,m}(x)|^2 < 2a(m) \int_a^b [f^{(m)}(\xi)]^2 d\xi |t_{(i+m)k_n} - t_{ik_n}|^{2m-1} \tag{2.11}$$

$$+2m \sum_{\nu=1}^{m} [\frac{d}{dx} \hat{\ell}_{i,\nu}(x)]^2 \psi_{i,\nu}^2 , \quad i = i(x), \; x \varepsilon [t_{k_n}, t_{(\ell-m+1)k_n})$$

$$\leq f^2(x) \qquad x \notin [t_{k_n}, t_{(\ell-m+1)k_n})$$

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

In the case $|f^{(m)}(\xi)| \leq c$, $a \leq \xi \leq b$, we may write

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

$$= f^2(x), \qquad x \notin [t_{k_n}, t_{(\ell-m+1)k_n})$$

Therefore

$$E|f(x) - \hat{f}_{n,m}(x)|^2 \leq \begin{cases} \max_i \; 2a(m) \int_a^b [f^{(m)}(\xi)]^2 d\xi \; E|t_{(i+m)k_n} - t_{ik_n}|^{2m-1} \\ \quad + 2m \sum_{\nu=1}^{m} E^{1/2}\left[\frac{d}{dx}\hat{\ell}_{i,\nu}(x)\right]^4 E^{1/2}\psi_{i,\nu}^4 & (2.13) \\ \quad + f^2(x) \cdot P_r\{x \notin [t_{k_n}, t_{(\ell-m+1)k_n})\} \\ \\ \max_i \; 2a(m) \sup_{a \leq \xi \leq b} |f^{(m)}(\xi)|^2 \; E|t_{(i+m)k_n} - t_{ik_n}|^{2m} \\ \quad + 2m \sum_{\nu=1}^{m} E^{1/2}\left[\frac{d}{dx}\hat{\ell}_{i,\nu}(x)\right]^4 E^{1/2}\psi_{i,\nu}^4 & (2.14) \\ \quad + f^2(x) \cdot Pr\{x \notin [t_{k_n}, t_{(\ell-m+1)k_n})\} \end{cases}$$

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

Since

$$\frac{d}{dx}\hat{\ell}_{i,\nu}(x) = \sum_{\substack{\mu=0\\ \mu\neq\nu}}^{m} \frac{\prod_{\substack{\xi=0\\ \xi\neq\mu,\xi\neq\nu}}^{m} (x-t_{(i+\xi)k_n})}{\prod_{\substack{\xi=0\\ \xi\neq\nu}}^{m} (t_{(i+\nu)k_n}-t_{(i+\xi)k_n})} \tag{2.15}$$

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

$$\left| \frac{d}{dx}\hat{\ell}_{i,\nu}(x)\right| \le m(t_{(i+m)k_n}-t_{ik_n})^{m-1} \frac{1}{\min\limits_{\nu=0,1,..m-1} (t_{(i+\nu+1)k_n}-t_{(i+\nu)k_n})^m} \tag{2.16}$$

and

$$E^{1/2}\left| \frac{d}{dx}\hat{\ell}_{i,\nu}(x)\right|^4 \le m^2\, E^{1/4}(t_{(i+m)k_n}-t_{ik_n})^{8(m-1)} \times$$

$$E^{1/4}\left(\frac{1}{\min\limits_{\nu=0,1,...m-1} (t_{(i+\nu+1)k_n}-t_{(i+\nu)k_n})^{8m}} \right). \tag{2.17}$$

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

<u>Lemma 1</u> Let the support set of f(x) be [a,b] and suppose $0 < \lambda \le f(x) \le \Lambda$, $x\varepsilon[a,b]$, and let p, q $< m\, k_n$.

Then

$$E|t_{(i+\nu)k_n}-t_{ik_n}|^{p} \leq \frac{1}{\lambda p}\left[\frac{\nu k_n}{(n+1)}\left(1+0\left(\frac{1}{k_n}\right)\right)\right]^{p} \tag{2.18a}$$

$$E|t_{(i+\nu)k_n}-t_{ik_n}|^{-q} \leq \Lambda^{q}\left[\frac{n+1}{(\nu k_n)}\left(1+0\left(\frac{1}{k_n}\right)\right)\right]^{q} \tag{2.18b}$$

Thus, assuming the hypotheses of the Lemma,

$$E^{1/2}\left|\frac{d}{dx}\hat{\ell}_{i,\nu}(x)\right|^{4} \leq m^{2m}\cdot\frac{\Lambda^{2m}}{\lambda^{2(m-1)}}\left[\left(1+0\left(\frac{1}{k_n}\right)\right]\cdot\left(\frac{n+1}{k_n}\right)^{2} \tag{2.19}$$

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

$$\psi_{i,\nu} \sim \rho_\nu - \frac{\nu k_n}{n+1} \tag{2.20}$$

where

$$\rho_\nu \sim Be(\nu k_n, n-\nu k_n+1) \tag{2.21}$$

$$E\,\rho_\nu = \frac{\nu k_n}{(n+1)}$$

In the appendix we show the following

Lemma 2

$$E^{1/2}\,\psi^{4}_{i,\nu} \leq \frac{\sqrt{3}\nu k_n}{(n+1)^2}\left(1+0\left(\frac{\nu k_n}{n+2}\right)\right)^{\frac{1}{2}} \quad . \tag{2.22}$$

We next invoke Lemma 3, proved in the appendix:

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

$$\Pr\{x \notin [t_{k_n}, t_{(\ell-m+1)k_n})\} = 0\left(\frac{k_n}{n^2}\right) \tag{2.23}$$

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

$$E|f(x) - \hat{f}_{n,m}(x)|^2 \leq \left\{A\left(\frac{k_n}{(n+1)}\right)^{2m-1} + B\frac{1}{k_n}\right\} + 0\left(\frac{k_n}{n^2}\right) \tag{2.24}$$

$$\leqq \left\{C\left(\frac{k_n}{(n+1)}\right)^{2m} + B\frac{1}{k_n}\right\} + 0\left(\frac{k_n}{n^2}\right) \tag{2.25}$$

where

$$A = 2a(m)\int_a^b [f^{(m)}(\xi)]^2 d\xi \cdot \left(\frac{m}{\lambda}\right)^{2m-1}\left(1+0\left(\frac{1}{k_n}\right)\right) \tag{2.26a}$$

$$B = m^{2m+3}\frac{\Lambda^{2m}}{\lambda^{2(m-1)}}\sqrt{3}\left(1+0\left(\frac{1}{k_n}\right)+0\left(\frac{k_n}{n}\right)\right) \tag{2.26b}$$

$$C = 2a(m)\sup_{a\leq\xi\leq b}|f^{(m)}(\xi)|^2\left(\frac{m}{\lambda}\right)^{2m}\left(1+0\left(\frac{1}{k_n}\right)\right) \tag{2.26c}$$

A lemma given by Parzen (see [2], 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 †

---

† We assume $A, C \neq 0$. The dominant term of A and C equals 0 if f is a polynomial of degree < 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.

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

for (2.24), and

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

for (2.25).

We then have

$$E|f(x) - \hat{f}_{n,m}(x)|^2 \leq \begin{cases} Dn^{\left(-\frac{2m-1}{2m}\right)} + o\left(n^{-\frac{2m-1}{2m}}\right) & (2.29) \\ Gn^{\left(-\frac{2m}{2m+1}\right)} + o\left(n^{\frac{-2m}{2m+1}}\right) & (2.30) \end{cases}$$

where

$$D = \frac{2m}{(2m-1)^{2m-1}} \left(A\, B^{2m-1}\right)^{\frac{1}{2m}} \tag{2.31}$$

$$G = \frac{2m+1}{2m^{2m}} \left(C\, B^{2m}\right)^{\frac{1}{2m+1}} \tag{2.32}$$

We have thus proved:

<u>Theorem 1.</u> Let $f(x)$ be supported on $[a,b]$, with $0 < \lambda \leq f(x) \leq \Lambda$, $x \in [a,b]$, let $f^{(\nu)}$, $\nu = 0, 1, 2, \ldots r$ be continuous, let $f^{(m)} \varepsilon L_2 [a,b]$, $m = r+1$, and let the estimates $\hat{f}_{n,m}(x)$ be given by (2.4), with k_n chosen as in (2.27). Then

$$E|f(x) - f_{n,m}(x)|^2 \leq Dn^{-\frac{2m-1}{2m}} + o\left(n^{-\frac{2m-1}{2m}}\right) \tag{2.33}$$

where D is given by (2.31) .

Theorem 2. Let f(x) satisfy the assumptions of Theorem 1, and in addition suppose $\sup_{\xi\epsilon[a,b]} |f(\xi)|^2 < \infty$.

Then, if k_n is chosen as in (2.28),

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

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

$$|f(x) - \sum_{\nu=1}^{m} \frac{d}{dx} \ell_\nu(x) \int_{x_0}^{x_\nu} f(\xi)\, d\xi\,|^2 \tag{3.1}$$

$$\leq a(m) \int_{x_0}^{x_m} [f^{(m)}(\xi)]^2\, d\xi\, |x_m - x_0|^{2m-1}$$

$$x\epsilon [x_0,x_m],\ m = 1, 2$$

$$x\epsilon [x_1,x_{m-1}],\ m \geq 3$$

with

$$a(1) = 1 \qquad (3.2)$$

$$a(2) \quad (5/2)^2$$

$$a(m) = \left[\frac{2(m+3)}{(m-1)!}\right]^2 , \quad m \geq 3$$

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

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

where

$$(u)_+ = u, \ u \geq 0 \qquad (3.4)$$
$$= 0 \text{ otherwise}$$

We may then write

$$f(x) - \tilde{f}(x) = \left\{\sum_{\nu=0}^{m-1} f^{(\nu)}(x_0) \frac{x^\nu}{\nu!} - \frac{d}{dx} \sum_{\mu=1}^{m} \ell_\mu(x) \sum_{\nu=0}^{m-1} f^{(\nu)}(x_0) \int_{x_0}^{x_\mu} \frac{\xi^\nu}{\nu!}\, d\xi\right\} \qquad (3.5)$$

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

where we are writing

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

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

$$\frac{x^\nu}{\nu!} = \frac{d}{dx} \sum_{\mu=1}^{m} \ell_\nu(x) \int_{x_o}^{x_\mu} \frac{\xi^\nu}{\nu!} \, d\,\xi \qquad (3.7)$$

Integrating both sides of (3.6) from x_0 to x, it is sufficient to show that

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

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

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

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

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

It is our purpose to examine the integrand

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

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

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

and $p_u\ (x)$ by

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

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

Thus (3.9) may be written

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

We calculate directly a bound on $\left|\frac{d}{dx} (h_u(x) - p_u(x))\right|$ for m = 1, 2, and then give a general bound good for $m \geq 3$.

For m = 1

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

and

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

For m = 2

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

$$\left. + \frac{(x-x_0)\ (x-x_1)}{(x_2-x_1)(x_2-x_0)} \frac{(x_2-u)^2}{2!}\right\} \qquad (3.15)$$

We have

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

The maximum of $\left|\frac{d}{dx} p_u(x)\right|$ clearly occurs at $x = x_2$. We have

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

$$+ \frac{(x_2-x_1)+(x_2-x_0)}{(x_2-x_1)(x_2-x_0)} \frac{(x_2-u)^2}{2!} \qquad (3.17)$$

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

$$\frac{d}{dx} \; p_u(x) \Big|_{x=x_2} = \frac{1}{2!} \left\{ \frac{(x_2-u)^2}{(x_2-x_0)} - \frac{(x_1-u)^2}{(x_1-x_0)} \right.$$

$$\left. + (x_2-u) + (x_1-u) \right\} \qquad x_o \leq u < x_1, \tag{3.18}$$

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

$$\left| \frac{d}{dx} \; [(h_u(x) - p_u(x)] \right| \leq \frac{5}{2} |x_2 - x_0| \qquad m=2 \tag{3.19}$$

We now assume $m \geq 3$.

By the Newton form of the remainder for Lagrange interpolation (see, for example, Isaacson and Keller [1], p. 248), we have, that

$$h_u(x) - \sum_{\nu=0}^{m} \ell_\nu(x) h_u(x_\nu)$$

$$= \prod_{\nu=0}^{m} (x-x_\nu) \; h_u \; [x_0, x_1, \ldots x_m, x] \tag{3.20}$$

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

$$h_u[x_0, x_1, \ldots x_m, x] = \frac{h_u[x_1, \ldots x_m, x] - h_u[x_0, \ldots x_{m-1}, x]}{(x_m - x_0)} \tag{3.21}$$

$$= \frac{1}{(x_m - x_0)} \left\{ \frac{h_u[x_2 \ldots x_{m-1}, x] - h_u[x_0, \ldots x_{m-1}, x]}{(x_m - x_1)} - \frac{h_u[x_1, \ldots x_{m-1}, x] - h_u[x_0, \ldots x_{m-2}, x]}{(x_{m-1} - x_0)} \right\}$$

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

$$\frac{d}{dx}\big(h_u(x) - p_u(x)\big) =$$

$$\sum_{\nu=0}^{m} \left(\prod_{j \neq \nu} (x - x_j) \right) \left\{ \frac{h_u[x_1, x_2, \ldots, x_m, x] - h_u[x_0, x_1, \ldots, x_{m-1}, x]}{(x_m - x_0)} \right\}$$

$$+ \frac{\prod_{\nu=0}^{m} (x - x_\nu)}{(x_m - x_0)} \left\{ \frac{d}{dx} \left[\frac{h_u[x_2, \ldots x_m, x] - h_u[x_1, \ldots x_{m-1}, x]}{(x_m - x_1)} - \frac{h_u[x_1, \ldots x_{m-1}, x] - h_u[x_0, \ldots x_{m-2}, x]}{(x_{m-1} - x_0)} \right] \right\} \tag{3.22}$$

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

$$|h_u[y_0, y_1, \ldots y_m]| \leq \sup_{x_0 \leq \xi \leq x_m} \frac{1}{(m-1)!} \left| h_u^{(m)}(\xi) \right| \tag{3.23}$$

This follows by writing

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

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

$$h_u [y_1, y_2, \ldots y_m] = \frac{1}{(m-1)!} h_u^{(m-1)} (\xi_2) \tag{3.25}$$

$$h_u[y_0, y_1, \ldots y_{m-1}] = \frac{1}{(m-1)!} h_u^{(m-1)}(\xi_1)$$

and

$$h_u[y_0, y_1, \ldots y_m] = \frac{1}{(m-1)!} \left| \frac{h_u^{(m-1)}(\xi_2) - h_u^{(m-1)}(\xi_1)}{y_m - y_0} \right| \leq \sup_{x_0 \leq \xi \leq x_1} \frac{1}{(m-1)!} \left| h_u^{(m)}(\xi) \right| \tag{3.26}$$

Similarly, it can be shown that

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

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

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

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

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

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

$$\frac{d}{dx}\left(h_u(x) - p_u(x)\right) \tag{3.28}$$

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

$$\le 2\ \frac{(m+3)}{(m-1)!}\ |x_m-x_1|^{m-1}$$

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

APPENDIX

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

Lemma 1. Let t_ν, and $t_{\nu+k}$ be the νth and the $\nu+k$ th order statistics from a random sample of size n from a population with density $f(x)$ where ν itself may be a random variable ($\nu \leq n-k$), and where $f(x)$ is supported on a closed interval $[a,b]$ with $0 < \lambda \leq f(x) \leq \Lambda$, $x\varepsilon[a,b]$. Then, for p, q < k,

$$E|\, t_{\nu+k}-t_\nu|^p \leq \frac{1}{\lambda p} \quad \frac{(k+p-1)(k+p-2)\ldots(k)}{(n+p)(n+p-1)\ \ldots\ (n+1)} = (\frac{1}{\lambda})^p (\frac{k}{n+1})^p \left(1+0\ (\frac{1}{k}\ )\right) \qquad (A.2)$$

$$E|\, t_{\nu+k}-t_\nu|^{-q} \leq \Lambda^q \quad \frac{n(n-1)\ \ldots\ (n-p+1)}{(k-1)(k-2)\ldots\ (k-p)} = \Lambda^q\ (\frac{n+1}{k})^q \left(1+0\ (\frac{1}{k})\right) \qquad (A.2)$$

Proof. The proof is effected, if we can show that the inequalities hold for any fixed $\nu \leq n-k$.

Assuming ν fixed now, the joint density $g(x,y)$ of t_ν and $t_{\nu+k}$ is

$$g(x,y) = \frac{n!}{(\nu-1)!(k-1)!(n-\nu-k)!} \quad F^{\nu-1}(x)[F(y)-F(x)]^{k-1}[1-F(y)]^{n-\nu-k} \quad f(x)f(y) \qquad (A.3)$$

$$x < y\ ,$$

$$= 0 \text{ otherwise}$$

Therefore $E|t_{\nu+k}-t_\nu|^p$ is given by

$$E|t_{\nu+k}-t_{\nu}|^{p} = \frac{n!}{(\nu-1)!(k-1)!(n-\nu-k)!} \quad x$$

$$\iint_{x<y} F^{\nu-1}(x)[F(y)-F(x)]^{k-1+p}\,\frac{[y-x]^{p}}{[F(y)-F(x)]^{p}}\,[1-F(y)]^{n-\nu-k}\quad f(x)f(y)\;dxdy$$

$$\leq \frac{\left[\dfrac{n!}{(\nu-1)!(k-1)!(n-\nu-k)!}\right]}{\left[\dfrac{(n+p)!}{(\nu-1)!(k+p-1)!(n-\nu-k)!}\right]}\left[\frac{(n+p)!}{(\nu-1)!(k+p-1)!(n-\nu-k)!}\right] \quad x$$

$$\iint_{x<y} F^{\nu-1}(x)[F(y)-F(x)]^{k-1+p}\,\frac{1}{\min\limits_{u}|f(u)|^{p}}\,[1-F(y)]^{n-\nu-k}\;f(x)f(y)dxdy$$

$$\leq \frac{n!}{(n+p)!}\;\frac{(k-1)!}{(k+p-1)!}\qquad\frac{1}{\min\limits_{u}|f(u)|^{p}} \tag{A.4}$$

Similarly

$$E|t_{\nu+k}-t_\nu|^{-q} = \frac{n!}{(\nu-1)!(k-1)!(n-\nu-k)!} \times$$

$$\iint_{x<y} F^{\nu-1}(x)[F(y)-F(x)]^{k-1-q}\frac{[F(y)-F(x)]^q}{|y-x|^q}[1-F(y)]^{n-\nu-k} f(x)f(y)dxdy$$

$$\leq \frac{\dfrac{n!}{(\nu-1)!(k-1)!(n-\nu-k)!}}{\dfrac{(n-q)!}{(\nu-1)!(k-q-1)!(n-\nu-k)!}} \cdot \frac{(n-q)!}{(\nu-1)!(k-q-1)!(n-\nu-k)!} \times$$

$$\iint_{x<y} F^{\nu-1}(x)[F(y)-F(x)]^{k-1-q} \max_u f^q(u)[1-F(y)]^{n-\nu-k} f(x)f(y)dxdy$$

$$= \frac{n!\,(k-q-1)!}{(n-q)!(k-1)!} \max_u f^q(u) \qquad \text{(A.5)}$$

<u>Lemma 2</u> Let $\psi = \rho-\frac{k}{n+1}$, 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

$$E\psi^4 = \mu_4 - 4\mu_3\mu_1 + 6\mu_2\mu_1^2 - 4\mu_1\mu_1^3 + \mu^4 .$$

$$= \frac{(k+3)(k+2)(k+1)\ k}{(n+4)(n+3)(n+3)(n+1)}$$

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

$$+ \frac{6(k+1)\ k^3}{(n+2)\ (n+1)^3}$$

$$- \frac{3\ k^4}{(n+1)^4}$$

Letting

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

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

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

$$f_4 = 0$$

we have

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

$$- \frac{k}{(n+1)} \left\{ \frac{1}{(n+1)^3} \left[f_1(k+3)(k+2)(k+1) - f_2 4(k+2)(k+1)k + f_3 6(k+1)\ k^2 \right] \right\}$$

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

<u>Lemma 3</u> 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 $\frac{\nu}{m} \to 0$. Then

$$P_r\left\{t_\nu > x\right\} \leq \frac{\nu}{(n+1)^2(n+2)} \quad \frac{1}{\left(F^{-1}(x) - \frac{\nu}{n+1}\right)^2} = 0\left(\frac{\nu}{n^2}\right)$$

ii) If $\frac{n-\nu}{n} \to 0$, then

$$P_r\{t_\nu < x\} \leq \frac{\nu(n-\nu+1)}{(n+1)^2(n+2)} \quad \frac{1}{\left(\frac{\nu}{n+1} - F^{-1}(x)\right)^2} = 0\left(\frac{n-\nu}{n^2}\right)$$

Proof: i) $P_r\{t_\nu > x\} = P_r\{\rho_\nu > F^{-1}(x)\}$, where

$$\rho_\nu \sim Be(\nu, n-\nu-1)$$

But, since var $\rho_\nu = \frac{\nu(n-\nu+1)}{(n+1)^2(n+2)}$, Chebychev's inequality gives, for $\frac{\nu}{n+1} < F(x)$,

$$P_r\{\rho_\nu > F^{-1}(x)\} \leq P_r\left\{\left|\rho_\nu - \frac{\nu}{n+1}\right| \geq F^{-1}(x) - \frac{\nu}{n+1}\right\}$$

$$\leq \frac{\nu(n-\nu+1)}{(n+1)^2(n+2)} \quad \frac{1}{\left(F^{-1}(x) - \frac{\nu}{n+1}\right)^2}$$

A similar equation is written for ii) .

REFERENCES

[1] Isaacson, Eugene, and Keller, Herbert, (1966) Analysis of Numerical Methods, John Wiley & Sons, New York.

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

[3] Van Ryzin, J. (1970). On a histogram method of density estimation. University of Wisconsin, Department of Statistics T.R. No. 226.

[4] Wegman, E.J. (1970) Nonparametric probability density estimation. University of North Carolina at Chapel Hill Institute of Statistics Mimeo Series No. 638.