The Annals of Statistics
1974, Vol. 2, No. 2, 359-361

ESTIMATION OF THE kth DERIVATIVE OF A DISTRIBUTION FUNCTION[1]

By Carl Maltz

California State University, Long Beach

Estimation of the kth derivative of a df by means of the kth-order difference quotients of the empiric df is investigated. In particular, consistency conditions are given, the asymptotic bias, variance, and mean-squared error of the estimator are computed, and means of minimizing the latter are discussed.

1. Introduction. Let $X_1, X_2 \cdots, X_n$ be a random sample distributed according to a df F. Suppose that F possesses a kth derivative $F^{(k)}$ at a point x. In this paper we discuss the estimation of $F^{(k)}(x)$ through use of the difference quotient

$$F_n^{(k)}(x; h_n) = (2h_n)^{-k} \textstyle\sum_{j=0}^{k} (-1)^j F_n(\tilde{x}_j)\binom{k}{j}, \tag{1}$$

where $\tilde{x}_j = x + (k-2j)h_n$. F_n denotes the empiric df based on $X_1, \cdots, X_n$ and $\{h_n\}$ is a suitably chosen sequence of positive numbers converging to zero. When there is no danger of confusion we will omit the subscript on h. We assume throughout that $k \geqq 1$.

We investigate the consistency, asymptotic bias, variance, and mean square error of this estimator, and discuss the minimization of the latter through judicious choice of the sequence $\{h_n\}$. This generalizes results of Rosenblatt (1956) who treated the case $k = 1$. Gaffey (1959) made use of the estimator (1) and essentially proved Theorem 2 of this paper. Schuster (1969) considered a different estimator for $F^{(k)}(x)$ for which he proved a.s. uniform convergence subject to certain regularity conditions.

2. Asymptotic bias, variance and mean square error.

Theorem 1. *Assume that F' exists at x. Then*

$$\operatorname{Var}(F_n^{(k)}(x; h)) = n^{-1}(2h)^{1-2k}F'(x)\binom{2k-2}{k-1} + o(n^{-1}h^{1-2k}). \tag{2}$$

Proof. We have

$$\begin{aligned}(2h)^{2k} \operatorname{Var}(F_n^{(k)}(x; h)) &= \operatorname{Var}\left(\textstyle\sum_{j=0}^{k} (-1)^j F_n(\tilde{x}_j)\binom{k}{j}\right) \\ &= \textstyle\sum_{i,j=0}^{k} (-1)^{i+j}\binom{k}{i}\binom{k}{j} \operatorname{Cov}(F_n(\tilde{x}_i), F_n(\tilde{x}_j)) \\ &= \frac{1}{n} \textstyle\sum_{i,j=0}^{k} (-1)^{i+j}\binom{k}{i}\binom{k}{j}[F(\tilde{x}_j \wedge \tilde{x}_i) - F(\tilde{x}_i)F(\tilde{x}_j)],\end{aligned} \tag{3}$$

Received November 1972; revised April 1973.

[1] This paper is based on part of the author's dissertation at the University of California, Los Angeles. Part of this research was supported under NSF GP-5224.

AMS 1970 *subject classifications*. Primary 62G05; Secondary 62F10.

Key words and phrases. Estimation, derivatives, distribution function, difference quotients, asymptotic variance, asymptotic mean-square error, asymptotic bias.

where $\tilde{x}_i \wedge \tilde{x}_j$ denotes the minimum of $\tilde{x}_i$ and $\tilde{x}_j$. By the definition of the derivative,

$$F(\tilde{x}_j) = F(x + (k - 2j)h) = F(x) + (k - 2j)hF'(x) + o(h)$$

as $h \to 0$. Inserting this, together with the analogous formula for $F(\tilde{x}_i)$ in (3) we obtain:

$$\begin{aligned}(2h)^{2k} \operatorname{Var}(F_n^{(k)}(x; h)) \\ &= \frac{1}{n} \textstyle\sum_{i,j=0}^{k} (-1)^{i+j}\binom{k}{j}\binom{k}{i} \cdot [F(x) + \{(k - 2j) \wedge (k - 2i)\}F'(x)h \\ &\quad - F^2(x) - h(2k - 2j - 2i)F(x)F'(x) + o(h)] .\end{aligned} \tag{4}$$

Using the equations $\sum_{i=0}^{k} (-1)^i \binom{k}{i} = \sum_{i=0}^{k} (-1)^i \binom{k}{i} i = 0$ it is easily seen that (4) reduces to

$$\begin{aligned}(2h)^{2k} \operatorname{Var}(F_n^{(k)}(x; h)) \\ &= -2F'(x)hn^{-1} \textstyle\sum_{i,j=0}^{k} (-1)^{i+j}\binom{k}{i}\binom{k}{j}(i \vee j) + o(n^{-1}h)\end{aligned} \tag{5}$$

where $i \vee j$ denotes the larger of i and j. The sum in (5) is equal to the constant term in the expansion in powers of z of the function

$$z(1 - z^{-1})^k \frac{d}{dz} [(1 - z)^k](2 \textstyle\sum_{m=0}^{\infty} z^{-m} - 1) = (-1)^k k(z + 1)(1 - z)^{2k-2} z^{1-k} .$$

The binomial theorem yields the value $-\binom{2k-2}{k-1}$, and (2) follows immediately.

THEOREM 2.

$$\begin{aligned}EF_n^{(k)}(x; h) &= F^{(k)}(x) + o(1) \\ &= F^{(k)}(x) + \frac{k}{6} h^2 F^{(k+2)}(x) + o(h^2)\end{aligned}$$

provided $F^{(k)}$ (*or* $F^{(k+2)}$) *exists at* x.

PROOF. The theorem follows immediately from (1) and the equations

$$EF_n(\tilde{x}_j) = \sum_{r=0}^{m} \frac{h^r F^{(r)}(x)(k - 2j)^r}{r!} + o(h^m) , \qquad m = k, k + 2 \tag{6}$$

and

$$\sum_{j=0}^{k} (-1)^j \binom{k}{j}(k - 2j)^r \begin{cases} = 0 & r = 0, 1, \cdots, k - 1, k + 1 \\ = 2^k k! & r = k \\ = 2^k k! \binom{k+2}{3}! & r = k + 2 .\end{cases} \tag{7}$$

The following are immediate consequences of Theorems 1 and 2:

COROLLARY 1. *If* $F^{(k)}$ *exists at* x, *and the sequence* $\{h_n\}$ *satisfies the conditions* $h_n \to 0$ *and* $n(h_n)^{2k-1} \to \infty$ *as* $n \to \infty$, *then* $F_n^{(k)}(x; h_n) \to_P F^{(k)}(x)$.

COROLLARY 2. *If* $F^{(k+2)}$ *exists at* x, *then the mean square error* (MSE) *of* $F_n^{(k)}(x; h)$

is given by:

$$\text{MSE}\,(F_n^{(k)}(x; h)) = \frac{k^2h^4}{36}(F^{(k+2)}(x))^2 + n^{-1}(2h)^{1-2k}F'(x)\binom{2k-2}{k-1} + o(h^4 + n^{-1}(2h)^{1-2k})\,. \tag{8}$$

3. Minimization of asymptotic mean square error. Let us now consider a sequence h_n of the form $Kn^{-\alpha}$. It is clear that the asymptotically optimal choice of α is such that the first two terms on the right in (8) are of the same order. That is, choose α so that

$$n^{-4\alpha} = n^{-1+\alpha(2k-1)}\,, \qquad \text{or} \qquad \alpha = (2k+3)^{-1}\,.$$

With this choice we have

$$\text{MSE}\,(F_n^{(k)}(x; h_n)) = \left[\frac{k^2}{36}K^4(F^{(k+2)}(x))^2 + (2K)^{1-2k}F'(x)\binom{2k-2}{k-1}\right]n^{-(4/(2k+3))} + o(n^{-(4/(2k+3))})\,. \tag{9}$$

The optimal choice of K is the one minimizing the right side of (9). Assuming that $F^{(k+2)}(x)$ and $F'(x)$ are not zero, this optimal choice K_0 is easily found to be:

$$K_0 = \left[\frac{9\binom{2k}{k}F'(x)}{k2^{2k}\{F^{(k+2)}(x)\}^2}\right]^{1/(2k+3)}\,. \tag{10}$$

Unfortunately, K_0 depends on the unknown values $F'(x)$ and $F^{(k+2)}(x)$. It seems reasonable, therefore, to replace $F'(x)$ and $F^{(k+2)}(x)$ in (10) by the estimates $F_n^{(1)}(x; h_n^{(1)})$ and $F_n^{(k+2)}(x; \bar{h}_n)$ where $h_n^{(1)} = n^{-\frac{1}{5}}$ and $\bar{h}_n = n^{-(1/(2k+7))}$, if both estimates are $\neq 0$. Otherwise set $K_n = 1$.

Acknowledgment. The author gratefully acknowledges several significant improvements suggested by the referee.

REFERENCES

GAFFEY, WILLIAM R. (1959). A consistent estimator of a component of a convolution. *Ann. Math. Statist.* **30** 198–205.

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

SCHUSTER, EUGENE F. (1969). Estimation of a probability density function and its derivatives. *Ann. Math. Statist.* **40** 1187–1195.

DEPARTMENT OF MATHEMATICS
CALIFORNIA STATE UNIVERSITY
LONG BEACH, CALIFORNIA 90815