The Annals of Statistics
1981, Vol. 9, No. 5, 1066–1078

WEAK CONVERGENCE AND EFFICIENT DENSITY ESTIMATION AT A POINT

BY A. M. KRIEGER AND J. PICKANDS, III

University of Pennsylvania

We consider estimators for a multivariate probability density at a point. Efficient choices require knowledge of the density and its second derivatives although these are not known. We use consistent, but not necessarily efficient, estimators for these and use them to replace the unknown values in the choices for an efficient estimator. Our second stage estimators and the unattainable efficient choices are asymptotically equivalent. This follows because we show that an entire class of estimators converges weakly to a limiting stochastic process. We find asymptotically efficient estimators of kernel type.

1. Introduction. Given a sequence of independent and identically distributed k-dimensional random vectors with probability density function $f(\mathbf{x})$, we consider the problem of estimating $f(\mathbf{x})$ at a point $\mathbf{x}$. We take $\mathbf{x} = \mathbf{0}$ without loss of generality. Many authors have studied the properties of estimators of the kernel type introduced by Rosenblatt (1956) and Parzen (1962). Necessary and sufficient conditions for consistency and asymptotic normality were given by Parzen (1962). Conditions for weak and strong consistency of these estimators were studied by Nadaraya (1965), Bhattacharya (1967), Schuster (1969), and most recently by Silverman (1978) among others. Cacoullos (1966) extended Parzen's work to the multivariate case. Optimum rates of convergence have been considered by Farrell (1972) and later by Sacks and Ylvisaker (1981). The latter authors treated rates of convergence in terms of minimaxity within classes specified by bounds on the density and its derivatives over a region.

We find asymptotically efficient estimator sequences from among this class. In the second section, we consider a specific kernel function and find the estimator which is asymptotically efficient in the mean squared error sense among all of those which use this function. Unfortunately, this estimator is a function of $f \equiv f(0)$ and $f''_{ii} \equiv \partial^2 f/\partial x_i^2|_{x=0}$, $i = 1, 2, \cdots, k$. We find consistent estimators for f and f''_{ii} in Section 3 and substitute them for f and f''_{ii} in the asymptotically efficient estimator. The proof that this sequence of two-stage estimators is asymptotically efficient makes use of weak convergence by the method of embedding used in Breiman (1968, pages 278–281). We extend our discussion to all estimators of kernel type in Section 4.

Bickel and Rosenblatt (1973) used weak convergence methods. Their stochastic processes had as argument the point at which the density is being estimated. In ours, the argument is a parameter of the estimation procedure, in particular the window width. Similar results involving bivariate density estimation were given by Rosenblatt (1975). He also used the technique of Poissonization of the sample size as we do in the present paper.

We are grateful to a referee for suggestions which led to a substantial improvement in this paper. We also thank the editor for allowing us to see a pre-publication copy of Sacks and Ylvisaker (1981).

2. Discussion of the results. We want to estimate a k-dimensional density function at a point $\mathbf{x}$, which we take, without loss of generality, to be $\mathbf{0}$. We consider estimators of

Received March 22, 1979; revised February 5, 1981.
AMS 1980 *subject classifications.* 60F05, 62G05, 62G20.
Key words and phrases. Weak convergence, multivariate density estimation, asymptotic efficiency, asymptotic normality, embedding.

the form

$$\hat{f}_n = N(\delta_n)/\{nA_k(\delta_n)\} = N(\delta_n)/(nC_k\delta_n^k), \tag{2.1}$$

where n is the sample size, $N(\delta_n)$ is the number of observations that lie in the sphere of radius δ_n, about $\mathbf{0}$, $A_k(\delta_n)$ is the volume of the k-dimensional sphere of radius δ_n, and C_k is the volume of the unit k-dimensional sphere (see Appendix A). Equivalently $\hat{f}_n$ can be written more conveniently in terms of the volume $v_n = C_k\delta_n^k$,

$$\hat{f}_n = M(v_n)/nv_n, \tag{2.1a}$$

where $M(v_n)$ is the number of observations that lie in the k-dimensional sphere of volume v_n, i.e., $M(v_n) = N(\delta_n)$.

In addition, we assume that the density function in the neighborhood of $\mathbf{0}$ is of the form

$$f(\mathbf{x}) = f(\mathbf{0}) + \{\mathbf{f}'(\mathbf{0})\}^\top\mathbf{x} + ½\,\mathbf{x}^\top\mathbf{f}''(\mathbf{0})\mathbf{x} + o(\mathbf{x}^\top\mathbf{x}) \quad \text{as} \quad \mathbf{x} \to \mathbf{0} \tag{2.2}$$

where the vector $\mathbf{f}'(\mathbf{0}) = \left(\frac{\partial f}{\partial x_1}, \cdots, \frac{\partial f}{\partial x_k}\right)\Big|_{x=0}$, the matrix $\mathbf{f}''(\mathbf{0}) = \left\{\frac{\partial^2 f}{\partial x_i \partial x_j}\Big|_{x=0};\ i = 1, \cdots, k,\ j = 1, \cdots, k\right\}$, and $\top$ denotes transposition. For brevity of notation we will eliminate reference to $\mathbf{0}$ and take f, $\mathbf{f}'$, and $\mathbf{f}''$ to be evaluated at $\mathbf{0}$. Then

$$E(\hat{f}_n) = H(v_n)/v_n, \tag{2.3}$$

where $\hat{f}_n$ is given by (2.1a) and $H(v_n)$ is the probability that a given observation lies within the sphere centered at $\mathbf{0}$ and having volume v_n. But

$$\begin{aligned} H(v_n) &= \int_{|\mathbf{x}|\le\delta_n} f(\mathbf{x})\,d\mathbf{x} \\ &= \int_{|\mathbf{x}|\le\delta_n} f\,d\mathbf{x} + \int_{|\mathbf{x}|\le\delta_n} \mathbf{f}'^\top\mathbf{x}\,d\mathbf{x} + \frac{1}{2}\int_{|\mathbf{x}|<\delta_n} \mathbf{x}^\top\mathbf{f}''\mathbf{x}\,d\mathbf{x}\{1 + o(1)\} \qquad \text{as } v_n \to 0. \end{aligned} \tag{2.4}$$

Since $\int_{|\mathbf{x}|\le\delta_n} x_i\,d\mathbf{x} = 0$ for all i, $\int_{|\mathbf{x}|\le\delta_n} x_ix_j\,d\mathbf{x} = 0$ for all $i \neq j$ and, by Lemma A.2 of Appendix A,

$$\begin{aligned} \int_{|\mathbf{x}|\le\delta_n} x_i^2\,d\mathbf{x} &= C_k\delta_n^{k+2}/(k+2) = C_k\delta_n^k\delta_n^2/(k+2) \\ &= v_n(v_n/C_k)^{2/k}/(k+2) = v_n^{(k+2)/k}/(k+2)C_k^{2/k}, \end{aligned}$$

it follows that

$$H(v_n) = fv_n + v_n^{(k+2)/k}\textstyle\sum_{i=1}^k f''_{ii}/\{2(k+2)C_k^{2/k}\} + o(v_n^{(k+2)/k}) \qquad \text{as } v_n \to 0. \tag{2.5}$$

By substituting (2.5) into (2.3) we see that

$$E(\hat{f}_n) = f + (v_n/C_k)^{2/k}\textstyle\sum_{i=1}^k f''_{ii}/\{2(k+2)\} + o(v_n^{2/k}) \qquad \text{as } v_n \to 0 \tag{2.6}$$

and so $E(\hat{f}_n) \to f$ as $v_n \to 0$ independently of the sample size n.

The variance of the estimator $\hat{f}_n$ is computed using the binomial distribution. Thus

$$\operatorname{var}(\hat{f}_n) = nH(v_n)\{1 - H(v_n)\}/n^2v_n^2 \sim H(v_n)/(nv_n^2) \qquad \text{as } v_n \to 0. \tag{2.7}$$

We mean by this that $\lim_{n\to\infty} \operatorname{Var}(\hat{f}_n)/\{H(v_n)/nv_n^2\} = 1$. Substituting (2.5) into (2.7),

$$\operatorname{Var}(\hat{f}_n) \sim (f/nv_n) + [(\textstyle\sum_{i=1}^k f''_{ii})(v_n/C_k)^{2/k}/\{2nv_n(k+2)\}](1 + o(1)) \sim (f/nv_n) \qquad \text{as } v_n \to 0. \tag{2.8}$$

For mean square consistency in general it is necessary and sufficient that

$$\lim_{n\to\infty} v_n = 0 \qquad \text{and} \qquad \lim_{n\to\infty} nv_n = \infty. \tag{2.9}$$

In order to obtain mean square error efficiency we want to find v_n that minimizes

$$\text{MSE}(v_n) \equiv E(\hat{f}_n - f)^2 = \text{Var}(\hat{f}_n) + \{E(\hat{f}_n) - f\}^2. \tag{2.10}$$

Substituting (2.6) and (2.8) into (2.10) we see that

$$\text{MSE}(v_n) \sim (f/nv_n) + [(v_n/C_k)^{4/k}(\textstyle\sum_{i=1}^k f''_{ii})^2/\{4(k+2)^2\}] \equiv \text{MSE}^*(v_n) \tag{2.11}$$

as $v_n \to 0$. To find the value of v_n that minimizes $\text{MSE}^*(v_n)$ for each n we differentiate $\text{MSE}^*(v_n)$ with respect to v_n and set the derivative equal to zero. Then

$$-(f/v_n^2) + [(v_n/C_k)^{(4-k)/k}(\textstyle\sum_{i=1}^k f''_{ii})^2/\{kC_k(k+2)^2\}] = 0. \tag{2.12}$$

Multiplying through by $(v_n/C_k)^2$ we can write $f/nC_k^2 = (v_n/C_k)^{(k+4)/k}A/C_k$, where $A \equiv (\sum_{i=1}^k f''_{ii})^2/\{k(k+2)^2\}$. So $(v_n/C_k) = (fC_k/nAC_k^2)^{k/(k+4)}$ and $v_n \equiv C_k^{4/(k+4)}(f/nA)^{k/(k+4)}$. In other words $\text{MSE}^*(v_n)$ is minimized when

$$v_n = t^* n^{-k/(k+4)}, \tag{2.13}$$

where

$$t^* = C_k^{4/(k+4)}\{(k+2)^2 kf/(\textstyle\sum_{i=1}^k f''_{ii})^2\}^{k/(k+4)}. \tag{2.14}$$

For every $t \in (0, \infty)$ define the estimators

$$\hat{f}_n(t) \equiv M(v_n)/nv_n, \tag{2.15}$$

where

$$v_n \equiv tn^{-k/(k+4)}.$$

Recall (2.1a). Here n is the sample size and $M(v_n)$ is the number of observations that lie within the sphere with volume v_n. By (2.11),

$$\text{MSE}^*(v_n) \equiv (f/tn^{4/(k+4)}) + n^{-4/(k+4)}(t/C_k)^{4/k}B,$$

where $B \equiv (\sum_{i=1}^k f''_{ii})^2/\{4(k+2)^2\}$. That is

$$\begin{aligned} n^{4/(k+4)}\text{MSE}^*(v_n) &= (f/t) + (t/C_k)^{4/k}(\textstyle\sum_{i=1}^k f''_{ii})^2/\{4(k+2)^2\} \\ &= \lim_{n\to\infty} n^{4/(k+4)}\text{MSE}(v_n) \end{aligned} \tag{2.16}$$

by (2.11). The minimum should occur at $t = t^*$, given by (2.14). We verify that it does. Differentiating (2.16) with respect to t gives

$$-f/t^2 + C_k^{-4/k}t^{-(k-4)/k}A,$$

where

$$A \equiv (\textstyle\sum_{i=1}^k f''_{ii})^2/k(k+2)^2, \quad t^{(k+4)/k} = fC_k^{4/k}/A \quad \text{and} \quad t^* = t = C_k^{4/(k+4)}(f/A)^{k/(k+4)}.$$

This is the same as (2.14).

We would like to use the estimators $\hat{f}_n(t^*)$ where t^* is given by (2.14). Unfortunately t^* depends upon the unknowns f and $\sum_{i=1}^k f''_{ii}$. So we consider an estimator of the form $\hat{f}_n(\tilde{t}^*)$ where $\tilde{t}^*$ is a consistent estimator for t^*. That is $\tilde{t}^* \to t^*$ in probability as $n \to \infty$. We will show that our estimator $\hat{f}_n(\tilde{t}^*)$ is asymptotically equivalent to the unattainable estimator $\hat{f}_n(t^*)$ in the sense that

$$n^{2/(k+4)}\{\hat{f}_n(t^*) - \hat{f}_n(\tilde{t}^*)\} \to 0 \tag{2.17}$$

in probability as $n \to \infty$.

3. Main results. To prove (2.17) we treat $\hat{f}_n(t)$, after suitable normalization, as a stochastic process. Let

$$Y_n(t) = n^{2/(k+4)}\{\hat{f}_n(t) - f\} = Y_n^{(1)}(t) + Y_n^{(2)}(t), \tag{3.1}$$

where

$$Y_n^{(1)}(t) \equiv n^{2/(k+4)}[\hat{f}_n(t) - E\{\hat{f}_n(t)\}]$$

and

$$Y_n^{(2)}(t) \equiv n^{2/(k+4)}[E\{(\hat{f}_n(t)\} - f],$$

$o < \alpha \le t \le \beta < \infty$, α and β arbitrarily chosen.

It is easy to see that $Y_n^{(2)}(t)$ is non-random. Furthermore, from (2.6)

$$\lim_{n\to\infty} Y_n^{(2)}(t) = Y^{(2)}(t) \equiv (t/C_k)^{2/k} \textstyle\sum_{i=1}^k f''_{ii}/\{2(k+2)\} \tag{3.2}$$

uniformly for $t \in [\alpha, \beta]$.

We will show that

$$Y_n^{(1)}(t) \Rightarrow Y^{(1)}(t) \equiv f^{1/2}W(t)/t \tag{3.3}$$

as $n \to \infty$ where $W(t)$ is the standard Brownian motion and $\Rightarrow$ means weak convergence in L_∞ norm; i.e., $x_n(t) \Rightarrow x(t)$ means $\phi\{x_n(t)\} \to \phi\{x(t)\}$ in law for any L_∞ continuous functional ϕ. This implies, but is not implied by, convergence of finite dimensional distributions. From (3.3) it will follow, by (3.2), that

$$Y_n(t) \Rightarrow Y(t) \equiv \{f^{1/2}W(t)/t\} + (t/C_k)^{2/k} \textstyle\sum_{i=1}^k f''_{ii}/\{2(k+2)\} \tag{3.4}$$

as $n \to \infty$. If we can show that (3.3) holds we have our main result, since we will show that $EY^2(t)$ is minimized at $t = t^*$, where t^* is given by (2.14). By weak convergence it will follow that $n^{2/(k+4)}\{\hat{f}_n(\tilde{t}^*) - \hat{f}_n(t^*)] \to 0$ in probability as $n \to \infty$. Recall that we expect to find $\tilde{t}^*$ such that $\tilde{t}^* \to t^*$ in probability as $n \to \infty$.

Proving (3.3) is most easily accomplished by means of Breiman's (1968) method of embedding. We assume that the sample size n has a Poisson distribution with mean λ and consider the stochastic process

$$Y_\lambda^{(1)}(t) \equiv \lambda^{2/(k+4)}[\hat{f}_\lambda(t) - E\{\hat{f}_\lambda(t)\}] \tag{3.5}$$

for all λ and t, $0 < \lambda$, $t < \infty$, where

$$\hat{f}_\lambda(t) \equiv M(v)/\lambda v$$

and

$$v = v(t) \equiv t\lambda^{-k/(k+4)}.$$

Here, $M(v)$ is the number of points within a sphere of volume v. Notice that $M(v)$ is a homogeneous one-dimensional Poisson process with intensity 1 when the argument is $EM(v)$, which is nondecreasing as v increases.

In order to show that $Y_\lambda^{(1)}(t)$ converges to $Y^{(1)}(t)$ as $\lambda \to \infty$ we need the following lemma and theorem. Here and hereafter $a \vee b$ is the larger of a and b and $a \wedge b$ is the smaller.

LEMMA 3.1. *Let $X_1, \cdots, X_n$ be* i.i.d. *non-negative random variables with finite third moment, then*

$$\vee_{i=1}^n X_i/n^{1/2} \to 0$$

in probability as $n \to \infty$.

PROOF. We need to show that for any $\varepsilon > 0$,

$$\lim_{n\to\infty} P(\vee_{i=1}^n X_i/n^{1/2} > \varepsilon) = 0.$$

But

$$P(\vee_{i=1}^n X_i/n^{1/2} > \varepsilon) = P(\vee_{i=1}^n X_i > n^{1/2}\varepsilon) = P(\cup_{i=1}^n \{X_i > n^{1/2}\varepsilon\})$$
$$\le nP(X_i > n^{1/2}\varepsilon) = nP(X_i^3 > n^{3/2}\varepsilon^3) \le nE(X_i^3)/n^{3/2}\varepsilon^3.$$

The last inequality follows from the Markov inequality. It follows that

$$P(\vee_{i=1}^n X_i/n^{1/2} > \varepsilon) \le E(X_i^3)/n^{1/2}\varepsilon^3 \to 0 \quad \text{as} \quad n \to \infty.$$

LEMMA 3.2. *Let $N(t), 0 \le t < \infty$, be the one-dimensional homogeneous Poisson process with intensity* 1. *For all ω, $0 < \omega < \infty$, let*

$$X(t:\omega) \equiv \{N(\omega t) - \omega t\}/\omega^{1/2}.$$

Then

$$X(t:\omega) \Rightarrow W(t)$$

as $\omega \to \infty$.

PROOF. Let

$$X_-(t:\omega) = \{N([\omega t]) - \omega t\}/\omega^{1/2}$$

and

$$X_+(t:\omega) = \{N([\omega t] + 1) - \omega t\}/\omega^{1/2},$$

where $[x]$ is the greatest integer $\le x$. It follows that $X_-(t:\omega) \le X(t:\omega) \le X_+(t:\omega)$ for all t and ω. Therefore it is sufficient to show that as $\omega \to \infty$

$$X_-(t:\omega) \Rightarrow W(t) \quad \text{and} \quad X_+(t:\omega) \Rightarrow W(t).$$

We can write

$$X_-(t:\omega) = [\textstyle\sum_{j=1}^{[\omega t]} \{N_j - E(N_j)\} + [\omega t] - \omega t]([\omega]/\omega)^{1/2}/[\omega]^{1/2}$$

where N_j, $j = 1, 2, \cdots, [\omega t]$ are i.i.d. Poisson random variables with mean 1. If we let

$$X_-^*(t:\omega) = \textstyle\sum_{j=1}^{[\omega t]} \{N_j - E(N_j)\}/[\omega],$$

then

$$X_-(t:\omega) = ([\omega]/\omega)^{1/2} X_-^*(t:\omega) + ([\omega t] - \omega t)/\omega^{1/2}.$$

But $([\omega]/\omega)^{1/2} \to 1$ as $\omega \to \infty$, $([\omega t] - \omega t)/\omega^{1/2} \to 0$ uniformly in t as $\omega \to \infty$, and $X_-^*(t:\omega) \Rightarrow W(t)$ as $\omega \to \infty$. This is true because a random walk with finite incremental variance can always be placed on the same sample space with the standard Brownian motion $W(t)$ to which it converges uniformly in probability; see Breiman (1968, pages 278–281). Weak convergence is an immediate consequence of this. It follows that $X_-(t:\omega) \Rightarrow W(t)$ as $\omega \to \infty$.

To show that $X_+(t:\omega) \Rightarrow W(t)$ we write

$$X_+(t:\omega) = X_-(t:\omega) + \{N([\omega t] + 1) - N[\omega t]\}/\omega^{1/2}.$$

But $\{N([\omega t] + 1) - N[\omega t]\}/\omega^{1/2} \to 0$ in probability uniformly in $t \in (\alpha, \beta)$, for any $\alpha, \beta \in (0, \infty)$, as $\omega \to \infty$ by Lemma 3.1 and the fact that Poisson random variables are non-negative with finite third moments. Therefore $X_+(t:\omega) \Rightarrow W(t)$ and so $X(t:\omega) \Rightarrow W(t)$.

LEMMA 3.3.

$$Y_\lambda^{(1)}(t) \Rightarrow f^{1/2} W(t)/t \tag{3.6}$$

as $\lambda \to \infty$ where $Y_\lambda^{(1)}(t)$ is given by (3.5).

PROOF. Since $M(t\lambda^{-k/(k+4)})$ has a Poisson distribution, its mean equals its variance. But from (2.6) we see that

$$E\{M(t\lambda^{-k/(k+4)})\} \sim f\lambda t\lambda^{-k/(k+4)} \quad \text{as} \quad \lambda \to \infty.$$

Therefore,

$$[M(t\lambda^{-k/(k+4)}) - E\{M(t_\lambda^{-k/(k+4)})\}]/(f\lambda^{4/(k+4)})^{1/2} \Rightarrow W(t)$$

as $\lambda \to \infty$ by Lemma 3.2. Rewriting (3.5) it follows that

$$\begin{aligned}
Y_\lambda^{(1)}(t) &\equiv [\hat{f}_\lambda(t) - E\{\hat{f}_\lambda(t)\}]/\lambda^{-2/(k+4)} \\
&= [M(t\lambda^{-k/(k+4)}) - E\{M(t\lambda^{-k/(k+4)})\}]/\{(\lambda t\lambda^{-k/(k+4)})\lambda^{-2/(k+4)}\} \\
&= f^{1/2}[M(t\lambda^{-k/(k+4)}) - E\{M(t\lambda^{-k/(k+4)})\}]/t(f\lambda^{4/(k+4)})^{1/2} \\
&\Rightarrow f^{1/2}W(t)/t
\end{aligned}$$

as $\lambda \to \infty$.

THEOREM 3.4.

$$n^{-2/(k+4)}\{\hat{f}_n(t) - \hat{f}_\lambda(t)\} \Rightarrow 0$$

as $n \to \infty$ and so equation (3.4) *holds.*

PROOF. It is sufficient to prove that

$$\sup_{\alpha \le t \le \beta} |\hat{f}_n(t) - \hat{f}_\lambda(t)| \to 0$$

in probability as $n \to \infty$ faster than $n^{2/(k+4)}$. This follows as the difference depends on the deficit or surplus in the numbers of observations in the volumes $tn^{-k/(k+4)}$ and $t\lambda^{-k/(k+4)}$ which is $o(n^{-k/(k+4)})$ as $n \to \infty$ independently of t by the weak law of large numbers.

The only details that need to be considered are consistent estimators $\tilde{f}$ and $\tilde{f}''$ for f and f''_ι. Clearly then $\tilde{t}^*$ is a consistent estimator for t^*, given by (2.14), where

$$\tilde{t}^* \equiv C_k^{4/(k+4)}\{(k+2)^2k\tilde{f}/(\textstyle\sum_{\iota=1}^k \tilde{f}''_\iota)^2\}^{k/(k+4)}.$$

Although there is a sizable literature devoted to finding consistent estimators for derivatives of density functions at a point, (c.f., Singh (1977)), we will present our own.

It follows from (2.9) that if $\tilde{f} = N(v_n)/nv_n$ where $v_n = n^{-\delta}$ for any $0 < \delta < 1$, then $\tilde{f}$ is a consistent estimator of f. To find a consistent estimator for f''_ι we prove the following lemma.

LEMMA 3.5. *For each i, $i = 1, 2, \cdots, k$, let $\tilde{f}''_\iota = \{N_i(2a_n) - 2N_\iota(a_n)\}/2na_n^3$, where $N_\iota(a_n)$ equals the number of observations such that $|x_i| \le a_n$. If $na_n^5 \to \infty$ and $a_n \to 0$ as $n \to \infty$ then $\tilde{f}''_\iota$ is a consistent estimator of f''_ι.*

PROOF. We apply (2.6) with dimension $k = 1$ and $a_n = v_n/2$. Here v_n is length, which is volume in one dimension. Now

$$E\{N_\iota(a_n)\}/2na_n = f^{(\iota)} + a_n^2 f''_\iota/6 + o(a_n^2)$$

as $a_n \to 0$, where $f^{(\iota)}$ is the marginal density for X_ι. Substituting $2a_n$ for a_n, it transpires that

$$E\{N_i(2a_n)\}/4na_n = f^{(\iota)'} + 4a_n^2 f''_\iota/6 + o(a_n^2)$$

as $a_n \to 0$. Now

$$\begin{aligned}
E(\tilde{f}''_\iota) &= E\{N_\iota(2a_n) - 2N_\iota(a_n)\}/2na_n^3 = (2/a_n^2)[E\{N_\iota(2a_n)/4na_n\} - 2E\{N_\iota(a_n)/4na_n\}] \\
&= (2/a_n^2)\{f^{(\iota)} + 4a_n^2 f''_\iota/6 - f^{(\iota)} - a_n^2 f''_\iota/6 + o(a_n^2)\} \\
&= (2/a_n^2)(3a_n^2 f''_\iota/6) + o(1) = f''_\iota + o(1)
\end{aligned}$$

as $a_n \to 0$. We have shown our estimator sequence to be asymptotically unbiased as $a_n \to 0$, hence as $n \to \infty$. Now we will show that the variance $V(\tilde{f}''_\iota) \to 0$ as $a_n \to \infty$. By definition

$$\begin{aligned}
(4n^2a_n^6)V(\tilde{f}''_\iota) &= V\{N_\iota(2a_n) - 2N_\iota(a_n)\} \\
&= V\{N_\iota(2a_n)\} - 4\,\mathrm{Cov}\{N_i(2a_n), N_\iota(a_n)\} + 4V\{N_\iota(a_n)\} \sim 4na_n f^{(\iota)}
\end{aligned}$$

as $n \to \infty$ by (2.8.) So

$$V(\tilde{f}''_{\iota}) \sim (4f^{(\iota)}na_n)/4n^2a_n^6 = f^{(\iota)}/na_n^5$$

as $a_n \to 0$ and so as $n \to \infty$. If $na_n^5 \to \infty$, then $V(\tilde{f}''_{\iota}) \to 0$ as $n \to \infty$.

4. A more general class of estimators. In this section we consider estimators of the form

$$\hat{f}_n^{(h)} \equiv \sum_{\iota=1}^n h\left(\frac{C_k \| X_\iota \|^k}{a_n}\right)/na_n, \tag{4.1}$$

where $h(x)$ is a nonnegative, nonincreasing function on the real line such that $\int_0^\infty h(x)\,dx = 1$, and $a_n = n^{-k/(k+4)}$ so that $\hat{f}_n^{(h)}$ is not degenerate in the limit. Such a function is used as a "window function." Up to this point we have considered only window functions of the form

$$h(x \mid t) \equiv 1/t,\ 0 \le x \le t, \tag{4.2}$$

and 0, otherwise. As previously, $\hat{f}_n(t)$ is, for each t, $0 < t < \infty$, the estimator using the window function $h(x \mid t)$.

Consider an estimator of the form

$$\hat{f}_n \equiv E\hat{f}_n(T) \tag{4.3}$$

where T is a nonnegative random variable independent of $\hat{f}_n(t)$ for all t. Let U be a hypothetical random variable uniformly distributed on the unit interval and let $T \equiv \phi(U)$ where ϕ is a monotone non-decreasing function. Any random variable T may be represented, uniquely, in this way. Notice that the window functions (4.2) correspond to hypothetical random variables T having degenerate distributions. The corresponding function $\phi(u) \equiv t$ for all $u \in (0, 1)$.

Now we show that any estimator of the form (4.3) has a representation of the form (4.1). Indeed,

$$h(x) \equiv Eh(x \mid T) = \int_x^\infty t^{-1}\,dF(t) = \int_0^1 h(x \mid \phi(u))\,du, \tag{4.4}$$

where $F(t)$ is the cumulative distribution function for T.

We consider the asymptotic behavior of

$$\hat{f}_n^{(\phi)} = E\hat{f}_n(T = \phi(U)), \tag{4.5}$$

where U is uniformly distributed on the unit interval and independent of $\hat{f}_n(t)$ for all t. By Theorem 3.4,

$$n^{2/(k+4)}\{\hat{f}_n(t) - f\} \Rightarrow Y(t) \equiv \beta\frac{W(t)}{t} + \gamma t^\alpha, \tag{4.6}$$

as $n \to \infty$ where

$$\alpha \equiv 2/k, \qquad \beta \equiv f^{1/2} \quad \text{and} \quad \gamma \equiv C_k^{-2/k}\sum_{\iota=1}^k f''_{\iota\iota}/\{2(k+2)\}.$$

LEMMA 4.1. *Let $\hat{f}_n^{(\phi)}$ be an estimator sequence of the form* (4.5). *Then*

$$n^{2/(k+4)}(\hat{f}_n^{(\phi)} - f) \Rightarrow Y^{(\phi)} \tag{4.7}$$

as $n \to \infty$, where the random variable

$$\begin{aligned} Y^{(\phi)} &= EY(T = \phi(U)) = \int_0^1 Y(\phi(u))\,du \\ &= \beta\int_0^1 W(\phi(u))\,du/\phi(u) + \gamma\int_0^1 \phi^\alpha(u)\,du. \end{aligned} \tag{4.8}$$

This is an immediate consequence of (4.5) and (4.6).

LEMMA 4.2. *The expectation*

$$E(Y^{(\phi)})^2 = \beta^2 I_0 + \gamma^2 I_1^2, \tag{4.9}$$

where

$$I_0 \equiv E(1/T_0 \vee T_1) = 2\int_0^1 u\,du/\phi(u), \qquad I_1 \equiv ET^\alpha = \int_0^1 \phi^\alpha(u)\,du, \tag{4.10}$$

where the hypothetical random variables T, T_0 and T_1 are mutually independent and identically distributed as $\phi(U)$ where U is uniformly distributed on the unit interval.

PROOF. By (4.8),

$$E(Y^{(\phi)})^2 \equiv E\int_0^1\int_0^1 Y(\phi(u))Y(\phi(v))\,du\,dv$$

$$= \beta^2 E\int_0^1\int_0^1 W(\phi(u))W(\phi(v))\,du\,dv/\phi(u)\phi(v) \tag{4.11}$$

$$+ 2\beta\gamma E\int_0^1\int_0^1 W(\phi(u))\phi^\alpha(v)\,du\,dv/\phi(u) + \gamma^2\int_0^1\int_0^1 \phi^\alpha(u)\phi^\alpha(v)\,du\,dv.$$

Consider the first term on the right side of (4.11). Clearly

$$E\int_0^1\int_0^1 W(\phi(u))W(\phi(v))\,du\,dv/\phi(u)\phi(v) = \int_0^1\int_0^1 EW(\phi(u))W(\phi(v))\,du\,dv/\phi(u)\,\phi(v)$$

$$= \int_0^1\int_0^1 \phi(u) \wedge \phi(v)\,du\,dv/\phi(u)\phi(v)$$

$$= \int_0^1\int_0^1 du\,dv/\phi(u) \vee \phi(v) = E(1/T_0 \vee T_1).$$

Now $T_0 \vee T_1 \equiv \phi(U_0) \vee \phi(U_1) = \phi(U_0 \vee U_1)$ where $U_0 \vee U_1$ has cumulative distribution function u^2. So the first term on the right side of (4.11) is $\beta^2 I_0$ where I_0 is given by (4.10). Consider the second term. Apparently

$$E\int_0^1\int_0^1 W(\phi(u))\phi^\alpha(v)\,du\,dv = \int_0^1\int_0^1 EW(\phi(u))\phi^\alpha(v)\,du\,dv = 0.$$

The last term

$$\gamma^2\int_0^1\int_0^1 \phi^\alpha(u)\phi^\alpha(v)\,du\,dv = \gamma^2 I_1^2,$$

where I_1 is given by (4.10). The conclusion (4.9) follows.

LEMMA 4.3. *For any fixed ϕ, $E(Y^{(a\phi)2})$ is minimized when*

$$a \equiv (\beta^2 I_0/2\alpha\gamma^2 I_1^2)^{1/(2\alpha+1)}. \tag{4.12}$$

The actual minimum is of the form

$$J^{(\phi)} \equiv \inf_a E(Y^{(a\phi)})^2 \equiv \psi(\alpha, \beta, \gamma)(I_0^\alpha I_1)^{2/(2\alpha+1)}, \tag{4.13}$$

where

$$\psi(\alpha, \beta, \gamma) = \beta^{4\alpha/(2\alpha+1)}\gamma^{2/(2\alpha+1)}((2\alpha)^{1/(2\alpha+1)} + (2\alpha)^{-(2\alpha)/(2\alpha+1)}). \tag{4.14}$$

and α, β and γ are given by (4.6).

REMARK. Notice that α, β and γ do not depend upon the shape of the window since they do not depend upon I_0 and I_1, given by (4.10). Therefore $\psi(\alpha, \beta, \gamma)$, given by (4.14), does not depend upon the shape, either. The result of Lemma 4.3 depends upon the shape only through I_0 and I_1.

PROOF. By Lemma 4.2,

$$E(Y^{(a\phi)})^2 \equiv \beta^2 a^{-1} I_0 + \gamma^2 a^{2\alpha} I_1^2. \tag{4.15}$$

Now

$$\partial E(Y^{(a\phi)})^2/\partial a = 0 = -\beta^2 a^{-2} I_0 + 2\alpha\gamma^2 a^{2\alpha-1} I_1^2.$$

We multiply through by a^2. Now $\beta^2 I_0 = 2\alpha\gamma^2 a^{2\alpha+1} I_1^2$ and so $a = (\beta^2 I_0/2\alpha\gamma^2 I_1^2)^{1/(2\alpha+1)}$. Thus (4.12) is true. Now we substitute this minimizing value of a in (4.15). First

$$\beta^2 a^{-1} I_0 = (\beta^2 I_0)^{1-1/(2\alpha+1)}(2\alpha\gamma^2 I_1^2)^{1/(2\alpha+1)} = \psi_1(\alpha, \beta, \gamma)(I_0^\alpha I_1)^{2/(2\alpha+1)},$$

where

$$\psi_1(\alpha, \beta, \gamma) = \beta^{4\alpha/(2\alpha+1)}(2\alpha\gamma^2)^{1/(2\alpha+1)}.$$

Second

$$\begin{aligned}\gamma^2 a^{2\alpha} I_1^2 &= \gamma^2 I_1^2(\beta^2 I_0/2\alpha\gamma^2 I_1^2)^{2\alpha/(2\alpha+1)} = (\gamma^2 I_1^2)^{1-2\alpha/(2\alpha+1)}(\beta^2 I_0)^{2\alpha/(2\alpha+1)}(2\alpha)^{-2\alpha/(2\alpha+1)}\\ &= \beta^{4\alpha/(2\alpha+1)}\gamma^{2/(2\alpha+1)}(2\alpha)^{-2\alpha/(2\alpha+1)}(I_0^\alpha I_1)^{2/(2\alpha+1)} \equiv \psi_2(\alpha, \beta, \gamma)(I_0^\alpha I_1)^{2/(2\alpha+1)}.\end{aligned}$$

Thus (4.13) is true with (4.14).

LEMMA 4.4. *The scale invariant functional $J^{(\phi)}$, given by* (4.13), *is minimzed when*

$$\phi(u) = u^{1/(\alpha+1)} \tag{4.16}$$

or, equivalently, by any positive multiple of (4.16).

PROOF. By Lemma 4.3, we choose $\phi(u)$ so as to minimize $I_0^\alpha I_1$ where I_0^α and I_1 are given by (4.10). Differentiating with respect to $\phi(u)$ for each u,

$$\partial(I_0^\alpha I_1)/\partial\phi = 0 = -(a_1 u/\phi^2) + a_2\phi^{\alpha-1}, \tag{4.17}$$

where a_1 and a_2 are positive constants irrelevant by scale invariance. Solving (4.17) for ϕ as a function of u, the conclusion of Lemma 4.4 follows.

REMARK. To verify (4.17) we can minimize the functional $I_0^\alpha I_1$ over functions of the form $\phi(\mathrm{u}) = \mathrm{u}^m$, $0 < m < \infty$. Now

$$I_0 \equiv 2\int_0^1 u\,du/\phi(u) = 2\int_0^1 u^{1-m}\,du = 2(2-m)^{-1}u^{2-m}|_0^1 = 2(2-m)^{-1},\ 0 < m < 2.$$

Notice on the other hand that if $m \geq 2$, then $I_0 = 2\int_0^1 u^{1-m}\,du = \infty$. In any case

$$I_1 = \int_0^1 \phi^\alpha(u)\,du = \int_0^1 u^{\alpha m}\,du = [u^{\alpha m+1}/(\alpha m+1)]|_0^1 = 1/(\alpha m+1).$$

So

$I_0^\alpha I_1 = 2^\alpha(2-m)^{-\alpha}(\alpha m+1)^{-1}$, $\log I_0^\alpha I_1 = \alpha\log 2 - \alpha\log(2-m) - \log(\alpha m+1)$,

$$\partial\log I_0^\alpha I_1/\partial m = 0 = \frac{\alpha}{2-m} - \frac{\alpha}{\alpha m+1}$$

and $2-m = \alpha m + 1$ and so $m(\alpha+1) = 1$. Thus $m = 1/(\alpha+1)$ and $u^m = u^{1/(\alpha+1)}$.

LEMMA 4.5. *The optimal shape for the window function is*

$$h(x) = (\alpha+1)(1-x^\alpha)/\alpha,\ 0 \le x \le 1,$$

and 0 *for* $x \ge 1$, *where* $\alpha = 2/k$.

REMARK. Recall that x denotes volume. Thus $x \equiv C_k r^k$ where $r \equiv \|\mathbf{x}\|$ is the radius. Furthermore, $\alpha \equiv 2/k$ and so

$$h(C_k r^k) \equiv (k+2)(1 - C_k^{2/k} r^2)/2 = (k+2)(1 - C_k^{2/k}\|\mathbf{x}\|^2)/2.$$

The actual optimal shape is quadratic in the radius regardless of the dimension.

PROOF OF LEMMA 4.5. By definition $T \equiv U^{1/(\alpha+1)}$, where U is uniformly distributed on the unit interval. Therefore

$$\begin{aligned} P(T > t) &= p(U^{1/(\alpha+1)} > t) \\ &= p(U > t^{\alpha+1}) = 1 - t^{\alpha+1},\ 0 \le t \le 1. \end{aligned}$$

Now $h(x) = 0$, for $x \ge 1$ by (4.4), and for $x \in [0, 1]$,

$$h(x) = \int_x^1 t^{-1}\, d(t^{\alpha+1}) = (\alpha+1)\int_x^1 t^{\alpha-1}\, dt = (\alpha+1)t^\alpha\,|_x^1/\alpha = (\alpha+1)(1-x^\alpha)/\alpha.$$

The lemma is proved.

REMARK. Note that $\int_0^\infty h(x)\,dx = 1$ as required.

How efficient are the estimators based upon truncation? By (4.13) the asymptotic efficiency is

$$E \equiv (I_0^\alpha I_1/\bar{I}_0^\alpha \bar{I}_1)^{2/(2\alpha+1)},$$

where I_0 and I_1 are given by (4.10) with $\phi(u)$ given by (4.16) and $\bar{I}_0$ and $\bar{I}_1$ are the same but with $\phi(u) \equiv 1$, $0 \le u \le 1$. Now

$$I_0 \equiv 2\int_0^1 u^{1-(1/(\alpha+1))}\, du = 2(\alpha+1)/(2\alpha+1), \tag{4.18a}$$

$$I_1 \equiv \int_0^1 (u^{1/(\alpha+1)})^\alpha\, du = (\alpha+1)/(2\alpha+1), \tag{4.18b}$$

and

$$\bar{I}_0 \equiv 2\int_0^1 u\, du = 1 = \int_0^1 1^\alpha\, du = \bar{I}_1.$$

Thus

$$E_k \equiv 2^{2\alpha/(2\alpha+1)}\{(\alpha+1)/(2\alpha+1)\}^{2(\alpha+1)/(2\alpha+1)} = 2^{4/(k+4)}\{(k+2)/(k+4)\}^{2(k+2)/(k+4)}.$$

Notice that

$$\lim_{k\to\infty} E_k = 1.$$

TABLE 4.1
Relative efficiency of optimal estimators from Section 3 and Section 4

Dimension	Efficiency
1	.9432
2	.9245
3	.9189
4	.9186
5	.9205
6	.9233
7	.9265
8	.9298
9	.9329
10	.9359
50	.9789
100	.9886

Table 4.1 illustrates the relative efficiency of the optimal estimator found in Section 3 with respect to the optimal estimator found in this section for selected values of the dimension k of the density function.

To obtain the optimal estimator we first found a in (4.12). Substituting I_0 and I_1 from (4.18a) and (4.18b) into (4.12), we found that

$$a = (\beta^2 I_0/2\alpha\gamma^2 I_1^2)^{1/(2\alpha+1)} = (\beta^2/2\alpha\gamma^2)^{1/(2\alpha+1)}(I_0/I_1^2)^{1/(2\alpha+1)}. \tag{4.19}$$

Now, by (4.6) $\alpha = 2/k$, $\beta = f^{1/2}$ and $\gamma = C_k^{-2/k}\sum_{i=1}^k f_{ii}''/\{2(k+2)\}$ and so

$$\begin{aligned}(\beta^2/2\alpha\gamma^2)^{1/(2\alpha+1)} &= [4(k+2)^2 kf/4C_k^{-4/k}(\textstyle\sum_{i=1}^k f_{ii}'')^2\}]^{k/(k+4)}\\ &= C_k^{4/(k+4)}\{(k+2)^2 kf/(\textstyle\sum_{i=1}^k f_{ii}'')^2\}^{k/(k+4)} = t^*\end{aligned}$$

given by (2.14). Thus

$$\begin{aligned} a \equiv t^*(I_0/I_1^2)^{1/(2\alpha+1)} &= 2^{1/(2\alpha+1)}t^*\{(\alpha+1)/(2\alpha+1)\}^{-1/(2\alpha+1)}\\ &= 2^{k/(k+4)}t^*\{(k+2)/(k+4)\}^{-k/(k+4)}\\ &= t^*\{2(k+4)/(k+2)\}^{k/(k+4)}.\end{aligned} \tag{4.20}$$

By Lemma 4.5, the optimal window function is

$$h(x) \equiv \{(\alpha+1)/\alpha\}\{1-(x/a)\}^{\alpha}/a = (k+2)\{1-(x/a)\}^{2/k}/2a,$$

where a is given by (4.20) with t^* given by'(2.14).

REMARK. For the class of estimators studied in Section 3, $I_0 = I_1 = 1$ and so $a = t^*$. This implies that $h(x) = 1/t^*$ for $0 \le x \le t^*$, and 0, otherwise. Substituting this function $h(x)$ into (4.1) we obtain the optimal estimator for this smaller class.

APPENDIX

LEMMA A.1.

$$\int_{\mathbf{x}^T\mathbf{x}\le\delta^2} d\mathbf{x} = \delta^k C_k, \tag{A.1}$$

where $C_k = \prod_{j=1}^{k} I_j$ *and*

$$I_j = \int_{-\pi/2}^{\pi/2} \cos^j \theta \, d\,\theta = \begin{cases} \pi \prod_{i=1}^{j/2} (2i-1)/2i & \text{if } j \text{ is even} \\ 2 \prod_{i=1}^{(j-1)/2} 2i/(2i+1) & \text{if } j \text{ is odd.} \end{cases}$$

PROOF. If we let $\mathbf{y} = \mathbf{x}/\delta$ it follows that

$$A_k(\delta) \equiv \int_{\mathbf{x}^\top \mathbf{x} \le \delta^2} d\mathbf{x} = \delta^k \int_{\mathbf{y}^\top \mathbf{y} \le 1} d\mathbf{y}. \tag{A.2}$$

All we need to find is C_k, the area of a k-dimensional unit sphere

$$C_k = \int_{y_k=-1}^{1} \left\{ \int_{y_{k-1}=-(1-y_k^2)^{1/2}}^{+(1-y_k^2)^{1/2}} \cdots \int_{y_1=-(1-y_k^2 \cdots y_2^2)^{1/2}}^{(1-\sum_{i=2}^{k} y_i^2)^{1/2}} dy_1 \cdots dy_{k-1} \right\} dy_k. \tag{A.3}$$

But the term in brackets in (A.3) equals

$$A_{k-1}\{(1-y_k^2)^{1/2}\} = (1-y_k^2)^{(k-1)/2} C_{k-1}.$$

This implies that

$$C_k = C_{k-1} \int_{-1}^{1} (1-y_k^2)^{(k-1)/2} \, dy_k. \tag{A.4}$$

If we let $y_k = \sin\theta$, then (A.4) becomes

$$C_k = C_{k-1} I_k = C_{k-1} \int_{\theta=-\pi/2}^{\pi/2} \cos^k\theta \, d\theta. \tag{A.5}$$

It is clear that $I_0 = \pi$ and $I_1 = 2$. Furthermore if we use integration by parts, letting $u = \cos^{j-1}\theta$ and $dv = \cos\theta \, d\theta$, then for $j \ge 2$ we find that

$$\int \cos^j\theta \, d\theta = \cos^{j-1}\theta \sin\theta + (j-1) \int \cos^{j-2}\theta \, d\theta - (j-1) \int \cos^j\theta \, d\theta.$$

It follows that

$$I_j = \int_{-\pi/2}^{\pi/2} \cos^j\theta \, d\theta = (j-1) I_{j-2} - (j-1) I_j$$

since

$$\cos(-\pi/2) = \cos(\pi/2) = 0.$$

Finally

$$I_j = (j-1) I_{j-2}/j; \tag{A.6}$$

c.f. Thomas (1965, page 367). The relation (A.6) and the fact that $I_0 = \pi$ and $I_1 = 2$ give the values of I_j as indicated in the statement of the lemma.

LEMMA A.2

(i) $\int_{\mathbf{x}^\top \mathbf{x} \le \delta^2} x_i \, d\mathbf{x} = 0$
(ii) $\int_{\mathbf{x}^\top \mathbf{x} \le \delta^2} x_i x_j \, d\mathbf{x} = 0$ if $i \ne j$
(iii) $\int_{\mathbf{x}^\top \mathbf{x} \le \delta^2} x_i^2 \, d\mathbf{x} = \delta^{k+2} C_k/(k+2)$

PROOF. Clearly (i) and (ii) are true by symmetry. Now

$$\int_{\mathbf{x}^\top \mathbf{x} \le \delta^2} x_k^2 \, d\mathbf{x} = \delta^{k+2} \int_{\mathbf{y}^\top \mathbf{y} \le 1} y_k^2 \, d\mathbf{y}. \tag{7}$$

But

$$\int_{\mathbf{y}^\top \mathbf{y} \le 1} y_k^2 \, d\mathbf{y} = C_{k-1} \int_{y_k=-1}^{1} y_k^2 (1 - y_k^2)^{(k-1)/2} \, dy_k.$$

If we let $y_k = \sin \theta$ then we find that

$$\int_{\mathbf{y}^\top \mathbf{y} \le 1} y_k^2 (1 - y_k^2)^{(k-1)2} \, d\mathbf{y} = C_{k-1} \int_{-\pi/2}^{\pi/2} \sin^2\theta \cos^k\theta \, d\theta$$

$$= C_{k-1} \int_{-\pi/2}^{\pi/2} (\cos^k\theta - \cos^{k+2}\theta) \, d\theta$$

$$= C_{k-1}(I_k - I_{k+2}) = C_{k-1} I_k/(k+2) = C_k/(k+2)$$

by (A.6)

REFERENCES

BHATTACHARYA, P. K. (1967). Estimation of a probability density function and its derivatives. *Sankhyā Ser. A* **29** 373–382.

BICKEL, P. J. and ROSENBLATT, M. (1973). On some global measures of the deviations of density function estimates. *Ann. Statist.* **1** 1071–1095.

BREIMAN, L. (1968). *Probability.* Addison-Wesley, Reading, Mass.

CACOULLOS, T. (1966). Estimation of a multivariate density. *Ann. Inst. Statist. Math.* **18** 178–189.

FARRELL, R. H. (1972). On the best obtainable asymptotic rates of convergence in estimation of a density function at a point. *Ann. Math. Statist.* **43** 170–180.

NADARAYA, E. A. (1965). On nonparametric estimates of density function and regression curves. *Theor. Probab. Appl.* **10** 186–190.

PARZEN, E. (1962). On estimates of a probability density function. *Ann. Math. Statist.* **40** 854–864.

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

ROSENBLATT, M. (1975). A quadratic measure of deviation of two-dimensional density estimates and a test of independence. *Ann. Statist.* **3** 1–14.

SACKS, J. and YLVISAKER, D. (1981). Asymptotically optimum kernels for density estimates at a point. *Ann. Statist.* **9** 334–346.

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

SILVERMAN, B. W. (1978). Weak and strong uniform consistency of the kernel estimate of a density and its derivatives. *Ann. Statist.* **6** 177–184.

SINGH, R. S. (1977). Improvement of some known nonparametric uniformly consistent estimates of derivatives of a density. *Ann. Statist.* **5** 394–399.

THOMAS, G. B. JR. (1965). *Calculus and Analytic Geometry (3rd Edition)*, Addison-Wesley, Reading, Mass.

DEPARTMENT OF STATISTICS
THE WHARTON SCHOOL
UNIVERSITY OF PENNSYLVANIA
PHILADELPHIA, PA 19104