The Annals of Statistics
1991, Vol. 19, No. 1, 329–337

INFORMATION INEQUALITY BOUNDS ON THE MINIMAX RISK (WITH AN APPLICATION TO NONPARAMETRIC REGRESSION)[1]

BY LAWRENCE D. BROWN AND MARK G. LOW

Cornell University and University of California, Berkeley

This paper compares three methods for producing lower bounds on the minimax risk under quadratic loss. The first uses the bounds from Brown and Gajek. The second method also uses the information inequality and results in bounds which are always at least as good as those from the first method. The third method is the hardest-linear-family method described by Donoho and Liu. These methods are applied in four examples, the last of which relates to a frequently considered problem in nonparametric regression.

0. Introduction. An earlier paper, Brown and Gajek (1990), describes lower bounds derivable from the information inequality on the Bayes risk under quadratic loss. There are two principle results of interest from that paper. The first is Corollary BG2.3, which is actually due to Borovkov and Sakhanienko (1980). [All references to numbered results in Brown and Gajek (1990) will be preceded by the letters BG.] Theorem BG2.7 then improves on that result at the cost of some algebraic complexity.

The supremum over all priors of the Bayes risk is a lower bound for the minimax value. [See, e.g., Lehmann (1983), page 256.] In fact, under mild conditions these two numbers are equal. [See, e.g., Le Cam (1986).] Thus the supremum of the Bayes risk bound of Corollary BG2.3 or of Theorem BG2.7 is also a lower bound for the minimax value. In some cases Corollary BG2.3 yields in this way a bound which is both easily obtained and easily expressed.

Either of these bounds can usually be considerably improved by a different but simple numerical procedure also based on the information inequality. This second procedure is also explained in the following examples. Comparison is also made to a method described recently in Donoho and Liu (1989) for problems involving Gaussian distributions, which extends ideas presented earlier in Ibragimov and Hasminskii (1981, 1984).

1. The prototypical Gaussian example. Let $X \sim N(\theta, 1)$ with $|\theta| \le L$, $L \le \infty$. Find the minimax value under ordinary squared error loss.

This is the prototypical Gaussian problem. It is the basis for Donoho and Liu's method. It has earlier been studied by Casella and Strawderman (1981),

Received November 1988; revised February 1990.

[1]Research supported in part by NSF DMS-85-06847 and NSF DMS-88-09016.

AMS 1980 *subject classifications.* Primary 62F10; secondary 62F15, 62C99, 60E15.

Key words and phrases. Information inequality (Cramér–Rao inequality), minimax risk, density estimation, nonparametric regression, estimating a bounded normal mean.

Bickel (1981) and Marazzi (1985), among others. In particular, Casella and Strawderman calculated the minimax value for some values of $L \le 2$. Recently, Kuks (1983), Donoho, Liu and MacGibbon (1990) and Feldman and Brown (1989) give much more complete tables of the minimax value. Because of the existence of these results, there really is no longer much need for explicit but nonexact lower bounds on the minimax value. The problem does, however, provide a good basis for presentation of the comparative methodologies; and because exact numerical results are known, it allows for an examination of the precision of the various bounds.

Bounds from Brown and Gajek (1990). The bound of Corollary BG2.3 is

$$\text{Bound(BG2.12)} = \frac{1}{1 + \int (g')^2/g} \le B(g), \tag{1.1}$$

where $B(g)$ denotes the Bayes risk for a prior density of g. It is thus maximized by choosing the (differentiable) density g which minimizes $\int (g')^2/g$. This is a well-known calculus-of-variations problem, whose solution is given by the density g defined in (BG3.3). [See Huber (1964) and Bickel (1981).] Thus from (BG3.4) the minimax value—call it $M(L)$—satisfies $M(L) \ge (1 + \pi^2/L^2)^{-1}$. This bound can be improved by applying Theorem BG3.7 to this same density, g. The resulting bound is the right side of (BG3.6). Of course, $B(g)$ itself is an even better bound, though finding it required much more extensive numerical calculations. Thus the three columns of Table BG3.1 provide successively better bounds on M, each of which requires successively more computations to evaluate.

A better information-inequality bound. A much better bound on M is available from the information inequality. That inequality implies that

$$M \ge (1 + b'(\theta))^2 + b^2(\theta) \quad \forall\, \theta, \tag{1.2}$$

where b is the bias function of any estimator. (Only estimators having everywhere finite risk and hence everywhere differentiable bias functions need be considered.) Let M_0 be the smallest constant for which there exists a differentiable function β on $(-L, L)$ satisfying

$$M_0 \ge (1 + \beta'(\theta))^2 + \beta^2(\theta). \tag{1.3}$$

Then, $M_0 \le M$; thus M_0 is a lower bound for M. It is easy to show that (1.3) has a solution if and only if there is a solution to

$$M_0 = (1 + \beta'(\theta))^2 + \beta^2(\theta). \tag{1.3'}$$

[See Brown and Farrell (1990).] Equation (1.3′) can be rewritten as

$$\beta'(\theta) = \left(M_0 - \beta^2(\theta)\right)^{1/2} - 1. \tag{1.4}$$

The method thus requires finding the smallest M_0 such that (1.4) has a solution on $(-L, L)$. (1.4) can be solved by separating variables. See, for

TABLE 1
Bounds for the minimax value $M(L)$ in Example 1

L	Bound($BG3.4$)	Bound(1.5)	M_0	$M(L)$
0.5	0.02470	0.160	0.126	0.199
1	0.09198	0.401	0.300	0.450
2	0.2884	0.642	0.547	0.645
3	0.4770	0.722	0.688	0.751
5	0.7170	0.771	0.829	0.857
10	0.9102	0.794	0.937	0.945
20	0.9759	0.800	0.980	

example, Kuks (1972). [A simple symmetry argument shows one need only see whether there exists a solution to (1.4) on $(0, L)$ with the initial condition $\beta(0) = 0$.] It should be clear that this method provides the best possible bound directly obtainable from the information inequality. See Table 1 for values of M_0.

The Ibragimov–Hasminskii constant. Donoho, Liu and MacGibbon (1990) establish (numerically) the bound

$$M \geq \frac{d}{1 + L^{-2}} \tag{1.5}$$

with $d \approx (1.247)^{-1}$. See also Feldman and Brown (1989). [The existence of the constant d was first established by Ibragimov and Hasminskii (1981).] Table 1 contains (for easy comparison) the bound $(1 + \pi^2/L^2)^{-1}$ from (BG3.4) as in Table BG3.1, the bound (1.5) with $d = (1.247)^{-1}$, the bound M_0 calculated as above and the actual value of $M(L)$ whenever available from Casella and Strawderman (1981) or Feldman and Brown (1989) or Kuks (1983). The value for $L = 10$ was supplied by R. C. Liu (private communication).

2. Estimation of the canonical binomial parameter. Let X be binomial(n, p) and consider the problem of estimating the canonical exponential family parameter $\theta = \ln(p/(1 - p))$ as in Example BG3.5. Maximization of the bound (BG2.12) of Corollary BG2.3 requires finding the density g which minimizes $D = \int (g'/I)^2/g$, where $I(\theta) = ne^\theta/(1 + e^\theta)^2$. Solution of this calculus-of-variations problem entails setting $g = w^2$ and solving the Euler equation $(w'/I)' - \lambda w = 0$, where λ is a Lagrange multiplier. The solution yields $g = 6e^{2\theta}/(1 + e^\theta)^4$, which is the prior density considered in Example BG3.5. Thus the first three rows of Table BG3.2 provide successively better lower bounds for M.

The method of Brown and Farrell (1990) explained in (1.2)–(1.4) can also be used here. This entails finding the smallest M_0 for which

$$\beta'(\theta) = \left(\left(M_0 - \beta^2(\theta)\right)I(\theta)\right)^{1/2} - 1 \tag{2.1}$$

TABLE 2
Values of the bound M_0 and of (BG2.16) *from Table* BG3.2

$n =$	1	3	5	10	25	100
$M_0 =$	0.211	0.400	0.505	0.646	0.800	0.934
(BG2.16) =	0.1314	0.3058	0.4186	0.5829	0.7707	0.9279

has a solution on $(-\infty, \infty)$. (Again, by symmetry it suffices to start with the initial condition $\beta(0) = 0$. Then it is necessary to solve (2.1) numerically on $[0, \infty)$.) Values of M_0 are given in Table 2 along with (for comparison) the value of the bound (BG2.16) from Table BG3.2. We do not know the values of M, except when $n = 1$. When $n = 1$ then $M = 0.221$ (as compared to $M_0 = 0.211$) corresponding to the procedure with $\delta(0) = -0.94$, $\delta(1) = +0.94$ which is Bayes for a prior concentrated on approximately the three points $\theta = 0, \pm 2.6$.

3. A two-dimensional Gaussian example. The one-dimensional methods of this paper can yield useful minimax results in multidimensional problems. The basic principle is to consider one-dimensional subproblems as was done, for example, in Farrell (1972). See also Brown and Farrell (1990). Donoho and Liu (1989) have shown in Gaussian settings that it can be quite efficient to consider only linear subproblems.

This example and the next one present two multidimensional Gaussian examples which can be treated by these methods. It appears *that in examples of this type the Donoho and Liu hardest-linear-family methodology usually gives better numerical results than the information-inequality methodology.* However, in the two examples presented here the opposite is true, if only barely.

The advantage in the present method is that it allows consideration of one-dimensional subfamilies which are not linear. Indeed, it can be considered as a step along the path to consideration of two-dimensional (parallelogram) families. This intuitive idea is clarified by the following examples.

The hardest-linear-family method. Let X be a two-dimensional normal variable with mean $\theta = (\theta_1, \theta_2)$ and covariance I. Assume θ lies in the convex, balanced set

$$\Theta = \{(\theta_1, \theta_2) : |\theta_1 - \theta_2| \le 1\}.$$

The problem is to estimate θ_1 under squared error loss. Donoho and Liu propose considering linear subsets of Θ, of the form

$$\Theta_\alpha = \{(\theta_1, \theta_2) \in \Theta : \theta_2 = \alpha\theta_1\}.$$

The minimum variance unbiased estimator of θ_1 is normally distributed with variance $\sigma_\alpha^2 = (\alpha^2 + 1)^{-1}$. It is also a sufficient statistic. Hence the minimax value for estimating θ_1 in the family Θ_α is

$$\overline{M}_\alpha = \sigma_\alpha^2 M(L(\alpha)),$$

where $M(L)$ is given in Table 1 and

$$L^2(\alpha) = \frac{\sup\{\theta_1 \colon \theta \in \Theta_\alpha\}^2}{\sigma_\alpha^2} = \frac{1+\alpha^2}{(1-\alpha)^2}.$$

[The bound (1.5) easily yields

$$\begin{aligned}\sup \overline{M}_\alpha &\geq (1.247)^{-1}\sigma_\alpha^2\left(1 + L^{-2}(\alpha)\right)^{-1}\\ &= (1.247)^{-1}\left(1 - 2\alpha + 2\alpha^2\right)^{-1} \geq 0.5346,\end{aligned}$$

with equality at $\alpha = 1/2$, $L^2(\alpha) = 5$.] The tables of Donoho, Liu and MacGibbon (1990) yield $\overline{M} \doteq \sup \overline{M}_\alpha \approx 0.549 = M_\alpha$ for $\alpha = 0.70$.

The information-inequality method. Information-inequality methods are not restricted to linear families. Consequently, an intuitively natural family is

$$\begin{aligned}\Theta'_\alpha = \Theta_\alpha \cup \Big\{(\theta_1, \theta_2) \colon \theta_1 \geq (1-\alpha)^{-1}, \theta_2 = \theta_1 - 1; \text{ or}\\ \theta_1 \leq -(1-\alpha)^{-1}, \theta_2 = \theta_1 + 1\Big\}\end{aligned}$$

composed of three connecting line segments. For this family $I(\theta_1) = (\alpha^2 + 1)$ for $|\theta_1| < (1-\alpha)^{-1}$ and $I(\theta_1) = 2$ for $|\theta_1| > (1-\alpha)^{-1}$. The minimax bound is the least value M_0 for which a solution $b(\cdot)$ exists on $(-\infty, \infty)$ to

$$\beta'(\theta) = \left[\left(M_0 - \beta^2(\theta)\right)I(\theta)\right]^{1/2} - 1. \tag{3.1}$$

Note that a solution to (3.1) exists on $((1-\alpha)^{-1}, \infty)$ if and only if $M_0 - \beta^2((1-\alpha)^{-1}) \geq 1/2$ since $I(\theta_1) \equiv 2$ on this semi-infinite interval. From this fact, plus symmetry, it can be seen that a solution to (3.1) exists if and only if the solution to

$$\beta'(\theta) = \left[\left(M_0 - \beta^2(\theta)\right)(\alpha^2 + 1)\right]^{1/2} - 1, \qquad \beta(0) = 0, \tag{3.2}$$

satisfies $\beta^2((1-\alpha)^{-1}) \leq M_0 - 1/2$. It is relatively easy to numerically check whether this is so. For the choice $\alpha = 1/2$ (in fact, for any α, $0.4 \leq \alpha \leq 0.51$) this least value of M_0 is 0.599.

Summary. In summary, the information-inequality method yields the statement, $M > 0.598$. The hardest-linear-family method yields only the statement, $M \geq 0.549$.

Postscript. David Donoho has pointed out that the set Θ is a semi-infinite rectangle, and that Donoho, Liu and MacGibbon (1990) contains relevant results for minimax problems with parameters restricted to lie in a rectangle. For the case at hand these results show that the precise minimax value is

$$M = \tfrac{1}{2}M\left(1/\sqrt{2}\right) + \tfrac{1}{2} = 0.663\dots. \tag{3.3}$$

Comparison of this value with the bounds 0.598 and 0.549, above, indicates the efficacy of these one-dimensional methods in this two-parameter problem.

4. Nonparametric regression. Consider a nonparametric regression problem with equally spaced predictor variables, y_i, on $(-1/2, 1/2)$. Thus observe $X_i \sim N(f(y_i), 1)$, indep., $y_i = -1/2 + i/(n+1)$, $i = 1, \ldots, n$.

Assume $f'(0)$ exists and

$$|f(x) - f(0) - f'(0)x| \le Bx^2/2 \tag{4.1}$$

so that $f''(0)$, if it exists, satisfies $|f''(0)| \le B$. It is desired to estimate $\theta = f(0)$ under squared error loss.

Of interest are asymptotic results, as $n \to \infty$, about M_n = the minimax risk for this problem. [This problem and the corresponding density estimation problem have been considered by many authors. See, for example, Rosenblatt (1956), Parzen (1962), Farrell (1972), Stone (1980), Sacks and Ylvisaker (1981), Sacks and Strawderman (1982) and Low (1989).]

The hardest-linear-family method. It is convenient, first, to review the hardest-linear-family analysis as applied to this problem. Let

$$g_\rho(t) = (\operatorname{sgn} \rho)(|\rho| - Bt^2/2)^+.$$

Consider linear families of the form

$$\Theta_\rho = \left\{ f \colon f(y) = \frac{\theta}{\rho} g_\rho(y),\ |\theta| \le \rho \right\}.$$

Note that $f(0) = \theta$ when $f = (\theta/\rho)g_\rho \in \Theta_\rho$. The minimum variance unbiased estimator of θ for the family Θ_ρ is

$$\delta_\rho(x) = c_{\rho,n} \sum r(y_i) x_i$$

with

$$c_{\rho,n} = \rho \left(\sum_{i=1}^{n} f_\rho^2(y_i) \right)^{-1} \sim \frac{\rho}{n} \left(\int g_\rho^2(t)\, dt \right)^{-1} = \frac{15\sqrt{B/2}}{16 n \rho^{3/2}}.$$

This has variance

$$\sigma_{\rho,n}^2 = c_{\rho,n}^2 \sum f_\rho^2(y_i) \sim \frac{15\sqrt{B/2}}{16 n \rho^{1/2}}. \tag{4.2}$$

The normalized length of this family is

$$L_{\rho,n} = \frac{\rho}{\sigma_{\rho,n}} \sim \frac{4\sqrt{n}\,\rho^{5/4}}{\sqrt{15}\,(B/2)^{1/4}} = L_\rho \tag{4.3}$$

since $\max\{|\theta| \colon \theta \in \Theta_\rho\} = \rho$. Hence the minimax risk for this family is

$$M_{\rho,n} = \sigma_{\rho,n}^2 M(L_{\rho,n}) \sim \left(\frac{15}{16\sqrt{2}} \right)^{4/5} \frac{B^{2/5}}{n^{4/5} L_\rho^{2/5}} M(L_\rho), \tag{4.4}$$

where (4.2) and (4.3) have been used to write the right side of (4.4) in terms of L. It follows that

$$(4.5) \qquad M^* \doteq \liminf_{n\to\infty}\left[\left(\frac{n^{4/5}}{B^{2/5}}\right)M_n\right] \ge \left(\frac{15}{16\sqrt{2}}\right)^{4/5} \sup L^{-2/5}M(L).$$

Now, $\sup L^{-2/5}M(L)$ occurs near that value of L which maximizes $L^{-2/5}/(1+L^{-2})$ since locally $M(L)$ is approximately proportional to $(1+L^{-2})^{-1}$. That value of L is $L=2$; hence

$$(4.6) \qquad \sup L^{-2/5}M(L) \ge 2^{-2/5}M(2) = 0.4889,$$

with 0.4889 being very near $\sup L^{-2/5}M(L)$. From (4.5) and (4.6),

$$(4.7) \qquad M^* \ge 0.352$$

[Table 2 of Donoho and Liu (1989) gives the bound $2^{-2/5}L^{-2/5}M(L) \ge 0.370$. (The factor $2^{2/5}$ results from the fact that their table gives values for $\sup(2L)^{2q-2}M(L)$ with, here, $q = 4/5$.) This also yields (4.6) and (4.7).]

One obvious one-parameter subfamily to which to apply the information inequality is

$$(4.8) \qquad \Theta' = \{f\colon f = g_\theta(t)\colon |\theta| \le B/8\}.$$

For this family the information function is

$$(4.9) \qquad I_n'(\theta) = \#\{i\colon By_i^2/2 \le \theta\} \sim 2n(2|\theta|/B)^{1/2}.$$

Hence the information inequality implies the existence of a function β such that

$$(4.10) \qquad M_n \ge (1+\beta'(\theta))^2 I_n'^{-1}(\theta) + \beta^2(\theta).$$

Now, let $\zeta = (n^2/B)^{1/5}\theta$ and $\gamma(\zeta) = (n^2/B)^{1/5}\beta((B/n^2)^{1/5}\theta)$. Let $n \to \infty$ and $M^* = \liminf(n^{4/5}M_n/B^{2/5})$, as above. Then (4.10) yields

$$M^* \ge (1+\gamma'(\zeta))^2\frac{1}{\sqrt{8\zeta}} + \gamma^2(\zeta).$$

Consequently, a lower bound for M^* is the smallest value m^* for which the equation

$$(4.11) \qquad \gamma'(\zeta) = \left((m^* - \gamma^2(\zeta))\sqrt{8\zeta}\right)^{1/2} - 1$$

has a solution on $(-\infty,\infty)$. It is enough to investigate solubility of (4.11) on $(0,\infty)$ subject to $\gamma(0) = 0$. This yields only the bound

$$(4.12) \qquad M^* \ge 0.340$$

as opposed to the slightly better bound in (4.7).

There is, however, a way to improve (4.12). Motivated by (4.3) and an analogy with the example of Section 3, choose $l > 0$ and define

$$\rho_n = \left(\frac{\sqrt{15}\,(B/2)^{1/4} l}{4\sqrt{n}} \right)^{4/5} = \left(\frac{\sqrt{B}}{n} \right)^{2/5} 0.8484 l^{4/5}$$

and

$$\Theta'' = \left\{ f\colon f = \frac{\theta}{\rho_n} g\rho_n;\ |\theta| \le \rho_n \text{ or } f = g_\theta \rho_n \le |\theta| \le \frac{B}{8} \right\}.$$

Θ'' is composed of the linear family Θ_{ρ_n}, having asymptotic normalized half-length $L = l$, plus the part of the nonlinear family Θ' corresponding to larger $|\theta|$. The information function is now asymptotic to

$$I_n''(\theta) = \begin{cases} 1.3894 l^{2/5} \left(\dfrac{n}{\sqrt{B}} \right)^{4/5}, & |\theta| < \rho_n, \\ I_n'(\theta), & |\theta| > \rho_n, \end{cases}$$

since $I_n''(\theta) = (\sigma^2_{\rho_n, n})^{-1}$ for $|\theta| < \rho_n$, where $\sigma^2_{\rho_n, n}$ is defined by (4.2). Thus as in (4.11) a lower bound for M^* is the smallest value m^* for which

$$\gamma'(\zeta) = \left((m^* - \gamma^2(\zeta)) i(\zeta) \right)^{1/2} - 1 \tag{4.13}$$

has a solution, where

$$i(\zeta) = \begin{cases} 1.3894 l^{2/5}, & |\zeta| < 0.8484 l^{4/5}, \\ \sqrt{8\zeta}, & |\zeta| > 0.8484 l^{4/5}. \end{cases}$$

Values of l in (approximately) the range (1.4, 1.6) yield

$$M^* \ge 0.361. \tag{4.14}$$

[The value $l = 2$, motivated by (4.6), yields only $M^* \ge 0.359$.]

Summary. In this example compare (4.14) and (4.7) to see that the two methods yield comparable bounds for M^*, with the information-inequality method yielding a bound about 2-1/2% better than the hardest-linear-family method.

The above problem is "Hölderian with exponent $q = 4/5$" as defined in Donoho and Liu (1989). This evidence and further results in Brown and Farrell (1990) and Low (1989) indicate that the former method should perform better than the latter for $q \ge 4/5$ (approximately).

It is plausible that the best rectangle method implicit in Donoho, Liu and MacGibbon (1990) (see also the postscript to the preceding section) could be used to improve on the above bounds. This conjecture remains to be investigated.

Acknowledgment. We wish to thank David Donoho for sending us preprint copies of his work, and for very useful additional personal correspondence.

REFERENCES

BICKEL, P. J. (1981). Minimax estimation of the mean of a normal distribution when the parameter space is restricted. *Ann. Statist.* **9** 1301–1309.

BOROVKOV, A. A. and SAKHANIENKO, A. U. (1980). On estimates of the expected quadratic risk. *Probab. Math. Statist.* **1** 185–195. (In Russian.)

BROWN, L. D. and FARRELL, R. H. (1990). A lower bound for the risk in estimating the value of a probability density. *J. Amer. Statist. Assoc.* **85** 1147–1153.

BROWN, L. D. and GAJEK, L. (1990). Information inequalities for the Bayes risk. *Ann. Statist.* **18** 1578–1594.

CASELLA, G. and STRAWDERMAN, W. E. (1981). Estimating a bounded normal mean. *Ann. Statist.* **9** 870–878.

DONOHO, D. L. and LIU, R. C. (1989). The hardest one-dimensional subproblem. Technical Report, Univ. California, Berkeley.

DONOHO, D. L., LIU, R. C. and MACGIBBON, B. (1990). Minimax risks over hyperrectangles and implications. *Ann. Statist.* **18** 1416–1437.

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

FELDMAN, I. and BROWN, L. D. (1989). The minimax risk for estimating a bounded normal mean. Cornell Statist. Center Preprint Series.

HUBER, P. J. (1964). Robust estimation of a location parameter. *Ann. Math. Statist.* **35** 73–101.

IBRAGIMOV, I. A. and HASMINSKII, R. Z. (1981). On some problems of nonparametric estimation of the values of a linear function. *Dokl. Akad. Nauk USSR* **256** 781–783. (In Russian.)

IBRAGIMOV, I. A. and HASMINSKII, R. Z. (1984). On nonparametric estimation of the values of a linear functional in Gaussian white noise. *Theory Probab. Appl.* **29** 18–32.

KUKS, YA. (1972). A lower bound for the minimax risk. *Comm. Estonian Acad. Sci.* **21** 303–306 (in Russian).

KUKS, J. (1983). An approximately minimax estimator of the mean of a one-dimensional normal distribution. *Izv. Akad. Nauk Estonskoj SSR.* **32** 286–293. (In Russian.)

LE CAM, L. (1986). *Asymptotic Methods in Statistical Decision Theory*. Springer, Berlin.

LEHMANN, E. L. (1983). *Theory of Point Estimation*. Wiley, New York.

LEVIT, B. YA. (1986). On bounds for the minimax risk. *Probab. Theory Math. Statist.* **2** 203–216.

LOW, M. G. (1989). A unified asymptotic minimax theory for nonparametric density estimation and nonparametric regression. Ph.D. dissertation, Cornell Univ.

MARAZZI, A. (1985). On constrained minimization of the Bayes risk for the linear model. *Statist. Decisions* **3** 277–297.

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

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

SACKS, J. and STRAWDERMAN, W. E. (1982). Improvements on linear minimax estimates. In *Statistical Decision Theory and Related Topics III* (S. S. Gupta and J. O. Berger, eds.) **2** 287–304. Academic, New York.

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

STONE, C. J. (1980). Optimal rates of convergence for nonparametric estimators. *Ann. Statist.* **8** 1348–1360.

MATHEMATICS DEPARTMENT
CORNELL UNIVERSITY
ITHACA, NEW YORK 14853

STATISTICS DEPARTMENT
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720