The Annals of Statistics
1995, Vol. 23, No. 2, 476–489

EFFICIENT ESTIMATION OF MONOTONE BOUNDARIES

BY A. P. KOROSTELEV, L. SIMAR AND A. B. TSYBAKOV

Academy of Sciences of Russia, Université Catholique de Louvain, and Academy of Sciences of Russia and Université Catholique de Louvain

Let $g: [0, 1] \to [0, 1]$ be a monotone nondecreasing function and let G be the closure of the set $\{(x, y) \in [0, 1] \times [0, 1]: 0 \le y \le g(x)\}$. We consider the problem of estimating the set G from a sample of i.i.d. observations uniformly distributed in G. The estimation error is measured in the Hausdorff metric. We propose the estimator which is asymptotically efficient in the minimax sense.

1. Introduction. Consider the estimation of support G of an unknown probability density based on a sample from this density. We assume that the boundary of G is defined by a monotone function and we mainly study the case where the underlying density is uniform in G.

Denote by $K = [0, 1] \times [0, 1]$ the unit square in the plane and let the points of the square K be denoted by $x = (x_1, x_2)$. Let $g(x_1)$, $0 \le x_1 \le 1$ be a monotone nondecreasing function such that

$$0 \le g(x_1) \le 1. \tag{1}$$

Consider the set

$$G^0 = \{x = (x_1, x_2): 0 \le x_1 \le 1, 0 \le x_2 \le g(x_1)\}.$$

Let G be the closure of G^0. Briefly, we say that G is *the set under* g. Note that $G \ne G^0$ if the corresponding function $g(x_1)$ is discontinuous. Let mes(G) be the Lebesgue measure of G. Denote by $\mathscr{G}_0$ the class of all sets G, such that mes(G) $\ne 0$ and G is under some monotone nondecreasing function g satisfying (1). We assume that the true density support G belongs to the class $\mathscr{G}_0$.

Let $\mathscr{X} = (X_1, \dots, X_n)$ be a sample of independent random variables uniformly distributed in G. We study the problem of estimation of G, given the observations $\mathscr{X}$. By an estimator $\hat{G}_n$ of G we mean an arbitrary closed set in K measurable with respect to $\mathscr{X}$.

The problem of estimating a monotone boundary with possible nonconvexity was first studied by Deprins, Simar and Tulkens (1984) in the context of measuring the efficiency of enterprises. They introduced the *free disposal hull* estimator of the set G, which is denoted F_n and defined as

$$F_n = \bigcup_{i=1}^{n} \mathrm{SE}(X_i).$$

Received December 1992; revised June 1993.
AMS 1991 *subject classifications.* 62G05, 62G20.
Key words and phrases. Monotone boundary, free disposal hull, Hausdorff distance, efficiency, minimax risk, estimation of support of a density.

Here and later SE(v) for a point $v = (v_1, v_2) \in K$ denotes the set of "southeastern" points with respect to v:

$$\mathrm{SE}(v) = \{x = (x_1, x_2) \in K: x_1 \ge v_1,\ x_2 \le v_2\}.$$

Graphically, the free disposal hull is the set under the "lowest" monotone step function g covering all the sample points X_i. In other words, the free disposal hull F_n has the minimal area among all the sets $G \in \mathscr{G}_0$ such that $G \supseteq \{X_1, \dots, X_n\}$. This shows that F_n is also the *maximum likelihood estimator* of G on the class $\mathscr{G}_0$. In fact, the maximum likelihood estimator for the uniform density on G is

$$G_n^{\mathrm{MLE}} = \arg\max_{G \in \mathscr{G}_0} \prod_{i=1}^{n} \mathrm{mes}(G)^{-1} I\{X_i \in G\} = \arg\min_{\substack{G \in \mathscr{G}_0: \\ G \supseteq \{X_1, \dots, X_n\}}} \mathrm{mes}(G) = F_n,$$

where $I\{\cdot\}$ is the indicator function.

In this paper we study the asymptotic behavior of the free disposal hull and some other related estimators. As an error criterion we use the Hausdorff distance between the estimator and the true set G. We show that a "blown-up" version of the free disposal hull is an efficient estimator in the asymptotical minimax sense. The main result of the paper is the exact asymptotics of minimax risk in the Hausdorff metric.

It is not surprising that one needs to blow up the free disposal hull in order to achieve efficiency: it is obviously downward biased since $F_n \subseteq G$. The asymptotics for F_n show that it is 50% as efficient as the optimal estimator (see Section 5 for the definition of efficiency).

In Section 2 we give some definitions and present the main result. Section 3 is devoted to the proof of the minimax lower bound. In Section 4 the optimal estimator is studied, and it is shown that the risk of this estimator attains asymptotically the lower bound. In Section 5 we discuss relative efficiencies of different monotone boundary estimators. Section 6 is concerned with some extensions. First, we give the extension of our main result to the multivariate case. Second, we consider the problem with nonuniform distributions of observations and show that the free disposal hull has the optimal rate of convergence to G.

We are not aware of earlier statistical study of monotone boundary estimators. However, there is some other work on density support estimation and on estimation of monotone functions which seems to be related to ours. Concerning the support estimation problem, we refer to the original papers of Geffroy (1964) and Rényi and Sulanke (1963, 1964). Rényi and Sulanke (1963, 1964) investigated the case of convex support G in two dimensions, and they proposed a natural estimator which is a convex hull of sample points. Note that the convex hull is also inward biased, as the free disposal hull. Ripley and Rasson (1977) consider a certain blown-up version of the convex hull in order to eliminate the bias. For further results on support estimation see, for example, Chevalier (1976), Devroye and Wise (1980), Korostelev and Tsybakov [(1993a, 1993b), Chap. 7] and Mammen and Tsybakov (1995).

The related work in the theory of monotone estimation is that of Grenander (1956), who showed that the maximum likelihood estimator for monotone density is a step function with jumps at the order statistics, the study of Grenander's estimator by Groeneboom (1985) and the paper of Birgé (1987), where explicit bounds are given on the minimax risks of monotone density estimators.

2. Some definitions and the main result. We use the Hausdorff metric to measure the difference between an estimator $\hat{G}_n$ and the true set G. Recall that the Hausdorff metric $d(G_1, G_2)$ for two closed sets G_1 and G_2 is defined by

$$d(G_1, G_2) = \max\left\{ \max_{x \in G_1} \rho(x, G_2);\ \max_{x \in G_2} \rho(x, G_1) \right\},$$

where $\rho(x, G) = \min_{y \in G}|x - y|$ is the Euclidean distance between a point x and a closed set G. We put by definition $d(G_1, G_2) = 0$ if either G_1 or G_2 is empty.

Let $w(t)$ be a *loss function*, that is, the function defined for nonnegative t and having the following properties: $w(t)$ is nonnegative, continuous, nondecreasing, $w(0) = 0$ and

$$w(t) \le \overline{w}(1 + t^{\alpha}), \qquad t \ge 0,$$

for some positive constants α and $\overline{w}$. For an arbitrary estimator $\hat{G}_n$ define the *risk function*

$$R(G, \hat{G}_n) = E_G\left[w\left(\psi_n^{-1} d(G, \hat{G}_n) \right) \right]. \tag{2}$$

Here E is the expectation with respect to the distribution P_G of observations and ψ_n is a normalizing factor, that is, a sequence of positive numbers.

Let $0 < \Delta < 1$ be fixed. Define the rectangles $K_1 = K \cap \{x_1 \ge 1 - \Delta\}$, $K_2 = K \cap \{x_2 \le \Delta\}$ and consider the following subset of $\mathscr{G}_0$:

$$\mathscr{G} = \{G \in \mathscr{G}_0 \colon G \supseteq K_1 \cup K_2\}.$$

We study the asymptotics of the *minimax* risk

$$r_n = \inf_{\hat{G}_n} \sup_{G \in \mathscr{G}} R(G, \hat{G}_n) \tag{3}$$

as $n \to \infty$. In what follows we show that the correct normalizing factor is $\psi_n = ((\log n)/n)^{1/2}$. Under this choice of ψ_n a finite nonzero limit exists for the minimax risk (3):

$$r = \lim_{n \to \infty} r_n,$$

and we find the exact value of r.

THEOREM 1. *If* $\psi_n = ((\log n)/n)^{1/2}$, *then for any loss function* $w(t)$ *the* *following equality holds*:

$$\lim_{n\to\infty} \inf_{\hat{G}_n} \sup_{G\in\mathscr{G}} R(G,\hat{G}_n) = w(1/\sqrt{2\pi}). \tag{4}$$

To prove this theorem it suffices to show that (i) for any estimator $\hat{G}_n$

$$\liminf_{n\to\infty} \sup_{G\in\mathscr{G}} E_G\left[w\left(\psi_n^{-1}d(G,\hat{G}_n)\right)\right] \geq w(1/\sqrt{2\pi}); \tag{5}$$

(ii) there exists an estimator G_n^* satisfying

$$\lim_{n\to\infty} R(G,G_n^*) \leq w(1/\sqrt{2\pi}) \tag{6}$$

uniformly in $G \in \mathscr{G}$.

The estimator G_n^* satisfying (6) is called *efficient*. In Section 4 we show that the $\psi_n/\sqrt{2\pi}$-neighbourhood of the free disposal hull, that is, $G_n^* = \{x \in K: \rho(x, F_n) \leq \psi_n/\sqrt{2\pi}\}$ is an efficient estimator.

REMARK. 1. One may consider a more general case when K is not a square but a rectangle $[0,a]\times[0,b]$, $a,b>0$. The result of Theorem 1 extends to this case, with $1/\sqrt{2\pi}$ replaced by $\sqrt{ab}/\sqrt{2\pi}$.

REMARK. 2. The reason why we deal with the restriction $\mathscr{G}$ of the class $\mathscr{G}_0$, and not with $\mathscr{G}_0$ itself, is to rule out the degenerate cases. For example, consider the sequence of monotone functions $g_n(x_1) = \frac{1}{2}I\{0 \leq x_1 < 1 - e^{-n}\} + I\{1 - e^{-n} \leq x_1 \leq 1\}$. Clearly, the sets G_n under the functions g_n, $n = 1,2,\dots,$ are in $\mathscr{G}_0$, but not in $\mathscr{G}$ for n large enough. The probability that there is at least one sample point in the slice $\{x = (x_1, x_2): 1 - e^{-n} \leq x_1 \leq 1, \frac{1}{2} \leq x_2 \leq 1\} \subseteq G_n$ is $O(e^{-n})$ as $n \to \infty$. Thus, one cannot "feel" from the data that g_n has a jump at $1 - e^{-n}$. However, the size of the jump is $\frac{1}{2}$, and this gives the main contribution to the Hausdorff distance between the set G_n and its estimator. This easily entails that the risks of F_n and G_n^* would explode as $n \to \infty$. The class $\mathscr{G}$ depends on Δ, but we omit it in the notation, since, as we show. the asymptotics of the minimax risk is independent of Δ.

Note that Theorem 1 gives exact asymptotics of the minimax risk: not only the rate of convergence to 0, but also the asymptotic constant. Results of this kind were obtained recently in some other nonparametric function estimation problems—estimation of functions in L_2-norm with quadratic loss w on the classes of ellipsoids [Pinsker (1980), Efroimovich and Pinsker (1981, 1982), Golubev (1982) and Nussbaum (1985)], density estimation in L_2 with quadratic loss w on a class of entire functions [Ibragimov and Khas'minskii (1982)] and estimation of regression curves in L_∞-norm [Korostelev (1992) and Donoho (1992)].

Theorem 1 deals with the minimax risk in Hausdorff metric, which is closer to the L_∞-norm results. An advantage of Theorem 1 is that it is not restricted to the quadratic loss function, unlike the L_2-norm results of Pinsker's type. Another important feature is due to the marvelous properties of the class of monotone functions: the asymptotic constant in (4) is an absolute one. It can be calculated once and forever, without any additional knowledge of a priori constants, such as smoothness characteristics of functions. Finally, we note that Theorem 1 is the first exact asymptotical minimax result concerning nonparametric estimation of boundaries [see Korostelev and Tsybakov (1992, 1993b) for other problems of this kind].

3. The lower bound. Let us prove the lower bound (5). Choose an arbitrary small ε ($0 < \varepsilon < \min(1/3, (1-\Delta)/2)$) and put $\delta_n = (0.5 - \varepsilon)^{1/2}\psi_n$, $M = M_n = [\varepsilon\delta_n^{-1}/2]$. Assume that there are some monotone functions $g_0, g_1, \dots, g_M$ (possibly depending on n) such that

$$g_l(x_1) \le g_0(x_1), \qquad 0 \le x_1 \le 1, \text{ for all } l = 1, \dots, M.$$

Denote by $G_0, \dots, G_M$ the sets under the monotone functions $g_0, \dots, g_M$. Assume that

$$\int_0^1 (g_0 - g_l)\, dx_1 = \delta_n^2 \operatorname{mes}(G_0), \qquad l = 1, \dots, M, \tag{7}$$

$$d(G_i, G_j) \ge 2 d_0 \psi_n, \qquad i \ne j, \qquad i, j = 0, \dots, M, \tag{8}$$

for some positive d_0 and all n large enough. An example of such a sequence of sets is given in Lemma 2.

In the following we write for brevity P_l instead of P_{G_l}, $l = 0, \dots, M$.

LEMMA 1. *Let some monotone functions g_l and the corresponding sets G_l, $l = 0, \dots, M$, satisfy* (7) *and* (8). *Then for any estimator $\hat{G}_n$ the following equality holds*:

$$\liminf_{n\to\infty} \max_{0\le l\le M} P_l\left\{\psi_n^{-1} d(G_l, \hat{G}_n) \ge d_0\right\} = 1.$$

PROOF. It suffices to verify that for an arbitrarily small $p_0 \in (0,1)$ one has

$$\max_{0\le l\le M} P_l\left\{\psi_n^{-1} d(G_l, \hat{G}_n) \ge d_0\right\} \ge 1 - p_0 \tag{9}$$

if n is large enough. Assume that inequality (9) does not hold. Then there exist p_0 and an estimator $\hat{G}_n$ satisfying

$$P_0\left\{\psi_n^{-1} d(G_0, \hat{G}_n) \ge d_0\right\} < 1 - p_0, \tag{10}$$

$$\min_{1\le l\le M} P_l\left\{\psi_n^{-1} d(G_l, \hat{G}_n) < d_0\right\} > p_0 \tag{11}$$

for all n large enough.

The following remark is crucial for the proof: Under the assumption $G_l \subset G_0$, $l = 1, \dots, M$, the joint P_l-distribution of observations $X_1, \dots, X_n$ (i.e., with each X_i uniformly distributed in G_l) coincides with P_0-conditional distribution of these observations given $\mathscr{X} \in G_l^n$, where $G_l^n = G_l \times \cdots \times G_l$ is the direct n-product of G_l's. It means that for any Borel set $\mathscr{C}$ in K^n,

$$(12)\qquad P_l\{\mathscr{X}\in\mathscr{C}\} = P_0\{\mathscr{X}\in\mathscr{C}|\mathscr{X}\in G_l^n\} = \frac{P_0\{\mathscr{X}\in\mathscr{C}\cap G_l^n\}}{P_0\{\mathscr{X}\in G_l^n\}}.$$

It follows from the definition and assumption (7) that

$$(13)\qquad P_0\{\mathscr{X}\in G_l^n\} = \left(\mathrm{mes}(G_l)/\mathrm{mes}(G_0)\right)^n = \left(1-\delta_n^2\right)^n.$$

Thus, we have from (12) and (13) that for any $\mathscr{C}\subseteq K^n$,

$$(14)\qquad P_0\{\mathscr{X}\in\mathscr{C}\cap G_l^n\} = \left(1-\delta_n^2\right)^n P_l\{\mathscr{X}\in\mathscr{C}\},\qquad l=1,\dots,M.$$

Due to assumption (8) and the triangle inequality, the following inclusion takes place:

$$(15)\qquad \bigcup_{l=1}^{M}\left\{\psi_n^{-1}d\left(G_l,\hat{G}_n\right)<d_0\right\}\subseteq\left\{\psi_n^{-1}d\left(G_0,\hat{G}_n\right)\ge d_0\right\}.$$

The same argument yields that the random events in the left-hand side of (15) are disjoint for different l. Hence

$$\begin{aligned}P_0\left\{\psi_n^{-1}d\left(G_0,\hat{G}_n\right)\ge d_0\right\} &\ge P_0\left\{\bigcup_{l=1}^{M}\left\{\psi_n^{-1}d\left(G_l,\hat{G}_n\right)<d_0\right\}\right\}\\ &=\sum_{l=1}^{M}P_0\left\{\psi_n^{-1}d\left(G_l,\hat{G}_n\right)<d_0\right\}\\ &\ge\sum_{l=1}^{M}P_0\left\{\psi_n^{-1}d\left(G_l,\hat{G}_n\right)<d_0;\mathscr{X}\in G_l^n\right\}.\end{aligned}$$

Applying (14) and assumption (11), we get

$$\begin{aligned}&P_0\left\{\psi_n^{-1}d\left(G_0,\hat{G}_n\right)\ge d_0\right\}\\ &\quad\ge\left(1-\delta_n^2\right)^n\sum_{l=1}^{M}P_l\left\{\psi_n^{-1}d\left(G_l,\hat{G}_n\right)<d_0\right\}>p_0M\left(1-\delta_n^2\right)^n\\ &\quad=p_0\left[\frac{\varepsilon}{2}\left((0.5-\varepsilon)\frac{\log n}{n}\right)^{-1/2}\right]\left(1-(0.5-\varepsilon)\frac{\log n}{n}\right)^n\\ &\quad\ge\frac{p_0\varepsilon}{2}\left(\frac{n^\varepsilon}{\log n}\right)^{1/2}>1-p_0\end{aligned}$$

for n large enough. The last inequality contradicts (10). This proves the lemma. □

LEMMA 2. *For any* $0 < \varepsilon < \min\{1/3, (1-\Delta)/2\}$ *there exist the sets* $G_l \in \mathscr{G}$, $l = 0, \dots, M$, *satisfying the assumptions of Lemma* 1 *with*

$$d_0 \geq \left(\frac{1-3\varepsilon}{2\pi}\right)^{1/2}, \tag{16}$$

whenever $n \geq n_0$, *where* n_0 *depends only on* ε.

PROOF. Consider the straight line segment with the endpoints $a_0 = (0, 1 - \varepsilon)$ and $a_1 = (\varepsilon, 1)$. Assume w.l.o.g. that $M = (\sqrt{2}\,\varepsilon/(2\sqrt{2}\,\delta_n)) = (\varepsilon/2)\delta_n^{-1}$ is an integer and divide the segment into equal subsegments of length $2\sqrt{2}\,\delta_n$. Denote the endpoints of the subsegments by $U_0 = a_0, U_1, \dots, U_{M-1}, U_M = a_1$.

Define the set $G_0 = \bigcup_{l=0}^{M} D_l$, where $D_l = \mathrm{SE}(U_l)$ and note that

$$\mathrm{mes}(G_0) = 1 - \varepsilon^2/2 - M\delta_n^2/2 \approx 1 - \varepsilon^2/2$$

for n large enough. For each l, $l = 1, \dots, M$, define W_l as the quarter of the circle of radius $\rho = (4\,\mathrm{mes}(G_0)/\pi)^{1/2}\delta_n$ with the center at U_l:

$$W_l = B(U_l, \rho) \cap D_l.$$

Here $B(u, \rho)$ denotes the Euclidean ball in $\mathbb{R}^2$ with center u and radius ρ.

The choice of radius ρ entails $\mathrm{mes}(W_l) = \pi\rho^2/4 = \mathrm{mes}(G_0)\delta_n^2$ and $\rho < 2\delta_n$. Define $G_l = G_0 \setminus W_l$. Since $\varepsilon < (1-\Delta)/2$, the sets $G_0, \dots, G_M$ belong to $\mathscr{G}$. It is easy to see that for the sets G_l, conditions (7) and (8) are satisfied with $2d_0 = \rho\psi_n^{-1}$. Thus

$$d_0 = \frac{1}{2}\left(\frac{4\,\mathrm{mes}(G_0)}{\pi}\right)^{1/2}(0.5 - \varepsilon)^{1/2} \geq \left(\frac{1-3\varepsilon}{2\pi}\right)^{1/2}$$

for $\varepsilon < \min\{1/3, (1-\Delta)/2\}$ and $n \geq n_0$, where n_0 depends on ε. □

PROOF OF THE LOWER BOUND (5). Lemmas 1 and 2 and the definition of the loss function imply

$$\begin{aligned}
&\liminf_{n\to\infty} \sup_{G\in\mathscr{G}} E_G\left[w\left(\psi_n^{-1} d\left(G, \hat{G}_n\right)\right)\right] \\
&\quad\geq \liminf_{n\to\infty} \max_{0\leq l\leq M} w(d_0) P_l\left\{\psi_n^{-1} d\left(G_l, \hat{G}_n\right) \geq d_0\right\} \\
&\quad= w(d_0) \geq w\left(\left(\frac{1-3\varepsilon}{2\pi}\right)^{1/2}\right).
\end{aligned}$$

since ε arbitrarily small, inequality (5) follows. □

4. Efficient estimator. Define the estimator G_n^* as the $\psi_n/\sqrt{2\pi}$-neighbourhood of the free disposal hull, that is, $G_n^* = \{x \in K\colon\ \rho(x, F_n) \leq \psi_n/\sqrt{2\pi}\}$. Clearly, $G_n^* \supseteq F_n$.

We prove that G_n^* is an efficient estimator. First, we need some definitions. Let G be the set under the monotone function g. Define the *edge of* G as the graph of g considered as a monotone set-valued function. Denote the edge of

G by ∂G. Let $u \in \partial G$ be an arbitrary point at the edge of G. For any radius $h > 0$ denote by $W(u, h)$ the intersection

$$W(u,h) = B(u,h) \cap G.$$

Fix some small $\varepsilon > 0$ and define the radii $\rho_k = \rho_k(u)$ such that

$$\mathrm{mes}(W(u,\rho_k)) = \mathrm{mes}(G)(0.5 + k\varepsilon)\psi_n^2, \qquad k = 1, 2, \dots, \sqrt{n}$$

($\sqrt{n}$ is assumed to be a large integer). Since the edge of G is defined by a monotone function, it is easy to see that $\pi\rho_k^2/4 \le \mathrm{mes}(W(u, \rho_k)) \le 3\pi\rho_k^2/4$, whenever $u \in \partial G$. This entails

$$\rho_k \le \sqrt{4\,\mathrm{mes}(W(u,\rho_k))/\pi} \le 2\psi_n\sqrt{(1 + 2k\varepsilon)/(2\pi)}\,, \qquad k = 1, 2, \dots, \sqrt{n}\,.$$

Let the set of points $\{u_1, \dots, u_N\}$ be a minimal $\varepsilon\psi_n$-net on ∂G w.r.t. the Euclidean distance in $\mathbb{R}^2$. Thus, $\rho(u_{l-1}, u_l) \le \varepsilon\psi_n$, $l = 2, \dots, N$. Since the length of ∂G is bounded from above by 2, we have $N = N_n \le (2/\varepsilon)\psi_n^{-1} \le 2\sqrt{n}/\varepsilon$ uniformly in $G \in \mathscr{G}$. Introduce the random events

$$A_{k,n} = \bigcap_{l=1}^{N} \{\text{at least one of } X_1, \dots, X_n \text{ belongs to } W(u_l, \rho_k(u_l))\},$$
$$k = 1, 2, \dots, \sqrt{n}\,.$$

LEMMA 3. *For any $G \in \mathscr{G}$ and for all n large enough,*

$$P_G\{A_{k,n}\} \ge 1 - (2/\varepsilon)n^{-k\varepsilon}, \qquad k = 1, 2, \dots, \sqrt{n}\,. \tag{17}$$

PROOF. For n large,

$$\begin{aligned} P_G\{A_{k,n}\} &= 1 - P_G\left\{\bigcup_{l=1}^{N} (\text{no observations in } W(u_l, \rho_k(u_l)))\right\} \\ &\ge 1 - N\left(1 - (0.5 + k\varepsilon)\psi_n^2\right)^n \ge 1 - \left(2\sqrt{n}/\varepsilon\right)n^{-0.5-k\varepsilon} \\ &= 1 - (2/\varepsilon)n^{-k\varepsilon}. \end{aligned}$$ □

PROOF OF INEQUALITY (6). We have

$$\max_{x \in G_n^* \setminus G} \rho(x, G) \le \psi_n/\sqrt{2\pi}$$

since $F_n \subseteq G$ almost surely and G_n^* is the $\psi_n/\sqrt{2\pi}$-neighbourhood of F_n. Now, if $A_{k,n}$ holds,

$$\begin{aligned} \max_{1 \le l \le N} \rho(u_l, G_n^* \cap G) &\le \max_{1 \le l \le N} \rho_k(u_l) - \psi_n/\sqrt{2\pi} \\ &\le \left(2\sqrt{1 + 2k\varepsilon} - 1\right)\psi_n/\sqrt{2\pi} \le (1 + 2k\varepsilon)\psi_n/\sqrt{2\pi}\,. \end{aligned}$$

Since the set $\{u_1, \dots, u_N\}$ is an $\varepsilon\psi_n$-net on ∂G, the last inequality implies that

$$\max_{x \in G} \rho(x, G_n^* \cap G) \le (1 + 2k\varepsilon)\psi_n/\sqrt{2\pi} + \varepsilon\psi_n \le (1 + 5k\varepsilon)\psi_n/\sqrt{2\pi}$$

when $A_{k,n}$ holds. Hence

$$A_{k,n} \subseteq \{d(G, G_n^*) \le (1 + 5k\varepsilon)\psi_n/\sqrt{2\pi}\}, \qquad k = 1, 2, \dots, \sqrt{n}.$$

Using the obvious property $d(G, G_n^*) \le \sqrt{2}$ for any $G, G_n^* \subseteq K$, and applying Lemma 3, we have

$$\begin{aligned}
&E_G\big[w\big(\psi_n^{-1} d(G, G_n^*)\big)\big] \\
&\quad \le w\left(\frac{1+5\varepsilon}{\sqrt{2\pi}}\right) P_G\left\{\psi_n^{-1} d(G, G_n^*) \le \frac{1+5\varepsilon}{\sqrt{2\pi}}\right\} \\
&\qquad + \sum_{k=1}^{\sqrt{n}-1} w\left(\frac{1+5(k+1)\varepsilon}{\sqrt{2\pi}}\right) \\
&\qquad \times P_G\left\{\frac{1+5k\varepsilon}{\sqrt{2\pi}} < \psi_n^{-1} d(G, G_n^*) \le \frac{1+5(k+1)\varepsilon}{\sqrt{2\pi}}\right\} \\
&\qquad + w\big(\sqrt{2}\,\psi_n^{-1}\big) P_G\left\{\frac{1+5\sqrt{n}\,\varepsilon}{\sqrt{2\pi}} < \psi_n^{-1} d(G, G_n^*)\right\} \\
&\quad \le w\left(\frac{1+5\varepsilon}{\sqrt{2\pi}}\right) + \sum_{k=1}^{\sqrt{n}-1} w\left(\frac{1+5(k+1)\varepsilon}{\sqrt{2\pi}}\right) P_G\{\bar{A}_{k,n}\} \\
&\qquad + w\big(\sqrt{2n}\big) P_G\{\bar{A}_{\sqrt{n},n}\} \\
&\quad \le w\left(\frac{1+5\varepsilon}{\sqrt{2\pi}}\right) + \frac{2}{\varepsilon}\bar{w}\left[\sum_{k=1}^{\sqrt{n}-1}\left(1 + \left(\frac{1+5(k+1)\varepsilon}{\sqrt{2\pi}}\right)^{\alpha}\right)\right. \\
&\qquad\qquad \left. \times n^{-k\varepsilon} + \big(1 + (2n)^{\alpha/2}\big) n^{-\sqrt{n}\,\varepsilon}\right] \\
&\quad \le w\left(\frac{1+5\varepsilon}{\sqrt{2\pi}}\right) + C_1\left[\sum_{k=1}^{\infty} k^{\alpha} n^{-k\varepsilon} + n^{-(\sqrt{n}\,\varepsilon - \alpha)}\right],
\end{aligned}$$

where C_1 is a positive constant. A routine analysis shows that the last expression in square brackets vanishes as $n \to \infty$. Thus, uniformly in $G \in \mathscr{G}$,

$$\lim_{n\to\infty} E_G\big[w\big(\psi_n^{-1} d(G, G_n^*)\big)\big] \le w\left(\frac{1+5\varepsilon}{\sqrt{2\pi}}\right)$$

for an arbitrarily small $\varepsilon > 0$. This proves (6). □

5. Relative efficiency of boundary estimators. As we mentioned already, the free disposal hull is a downward biased estimator of G. The following theorem shows that the free disposal hull itself gives the constant $2/\sqrt{2\pi}$ instead of $1/\sqrt{2\pi}$.

THEOREM 2. *Under the assumptions of Theorem* 1, *we have*

$$\lim_{n\to\infty} \sup_{G\in\mathscr{G}} R(G, F_n) = w\big(2/\sqrt{2\pi}\big).$$

PROOF. Note that

$$\begin{aligned} d(G, F_n) &= \max_{x\in G} \rho(x, F_n) \le \max_{1\le l\le N} \rho_k(u_l) + \varepsilon\psi_n \\ &\le \big(2\sqrt{1+2k\varepsilon}\big)\psi_n/\sqrt{2\pi} + \varepsilon\psi_n \end{aligned}$$

when $A_{k,n}$ holds. Hence, the inequality $\lim_{n\to\infty}\sup_{G\in\mathscr{G}} R(G, F_n) \le w(2/\sqrt{2\pi})$ for any loss function $w(t)$ is proved by the method of Section 4. To verify the inverse inequality, let us return to the construction used in Lemma 2. Let, again, W_l be the quarters of circles with the centers at U_l, $l = 1,\dots, M$ and of the radius

$$\rho = (4\,\mathrm{mes}(G_0)/\pi)^{1/2}\delta_n = 2((1-2\varepsilon)\mathrm{mes}(G_0)/(2\pi))^{1/2}\psi_n.$$

Introduce the random events

$$B_l = \{\text{there are no observations in } W_l\}, \qquad l = 1,\dots, M.$$

Note that the mean number of observations X_i, $i = 1,\dots, n$, belonging to at least one of W_l is equal to

$$nM\,\mathrm{mes}(G_0)\,\delta_n^2 = (0.5-\varepsilon)M(\log n) = (1-2\varepsilon+o(1))M(\log M),$$

where $o(1)\to 0$ as $n\to\infty$, and due to the law of large numbers,

$$\begin{aligned} \lim_{n\to\infty} P_{G_0}\{&\text{number of observations belonging to} \\ &\text{at least one of } W_l \ge (1-\varepsilon)M(\log M)\} = 0. \end{aligned}$$

Using the asymptotics for urn models [Johnson and Kotz (1977), page 318] we find that if one throws $(1-\varepsilon)M(\log M)$ balls independently in M urns, then there is at least one empty urn with probability close to 1 as M tends to infinity. Thus,

$$\lim_{n\to\infty} P_{G_0}\left\{\bigcup_{l=1}^{M} B_l\right\} = 1.$$

This equality shows that the free disposal hull does not cover at least one of W_l and, consequently,

$$\psi_n^{-1} d(G_0, F_n) \ge \rho \ge 2((1-3\varepsilon)/(2\pi))^{1/2}$$

with the probability tending to 1 as $n\to\infty$. Since ε is arbitrarily small and $w(t)$ is continuous, this implies

$$\lim_{n\to\infty} R(G_0, F_n) \ge w\big(2/\sqrt{2\pi}\big). \qquad \square$$

Other improvements of F_n (besides G_n^*) are also possible. For example, consider the parallel shift of F_n in the "northwestern" direction:

$$G_n^{**} = \bigcup_{i=1}^{n} \mathrm{SE}(X_i + \nu),$$

where $\nu = (-\sqrt{2}\,\psi_n/4, \sqrt{2}\,\psi_n/4)$. Following the lines of Section 4 and Theorem 2 [but with the sets $W(u, \rho)$ chosen as intersections of G with small squares rather than circles], one can prove that

$$\lim_{n\to\infty} \sup_{G\in\mathscr{G}} E_G\left[w\left(\psi_n^{-1} d(G, G_n^{**})\right)\right] = w(1/2).$$

Now, suppose that some estimator $\hat{G}_n$ satisfies

$$\lim_{n\to\infty} \sup_{G\in\mathscr{G}} E_G\left[w\left(\psi_n^{-1} d\left(G, \hat{G}_n\right)\right)\right] = w(\gamma)$$

with a constant $\gamma = \gamma(\hat{G}_n) > 0$. Define the *relative efficiency* of $\hat{G}_n$ by the formula

$$\mathrm{Eff}\left(\hat{G}_n\right) = \left(\sqrt{2\pi}\,\gamma\left(\hat{G}_n\right)\right)^{-1}. \tag{18}$$

It follows from (5) that $\mathrm{Eff}(\hat{G}_n) \le 1$ for any $\hat{G}_n$. The relative efficiencies of the estimators F_n, G_n^{**} and G_n^{*} are

$$\mathrm{Eff}(F_n) = 0.5, \qquad \mathrm{Eff}(G_n^{**}) = 2/\sqrt{2\pi} \approx 0.7979, \qquad \mathrm{Eff}(G_n^{*}) = 1.$$

Note that the efficiencies calculated by means of (18) are independent of the loss function w. Since the normalizing factor has the form $\psi_n = ((\log n)/n)^{1/2}$, we easily see that for any loss function the precision of G_n^* is the same as that of $\hat{G}_{N(n)}$, where $N(n) = n/(\mathrm{Eff}(\hat{G}_n))^{-2}$. The number $N(n)$ may be therefore called "equivalent number of observations for $\hat{G}_n$."

6. Extensions.

Multidimensional case. Let $g(x_1, \dots, x_s)$ be a monotone function in each of the arguments $x_1, \dots, x_s$ and let $0 \le g(x_1, \dots, x_s) \le 1$, $s \ge 1$. Define G, the *set under* g, as the closure of the $(s+1)$-dimensional set

$$G^0 = \left\{x = (x_1, \dots, x_s, x_{s+1}):\right.$$
$$\left. 0 \le x_{s+1} \le g(x_1, \dots, x_s), (x_1, \dots, x_s) \in [0,1]^s\right\}.$$

Let $\mathscr{G}_s$ be the class of sets G under the monotone functions $g(x_1, \dots, x_s)$ such that $G \supseteq K_1 \cup \dots \cup K_{s+1}$, where

$$K_j = \left\{x = (x_1, \dots, x_{s+1}) \in [0,1]^{s+1}: x_j \ge 1 - \Delta\right\}, \qquad j = 1, \dots, s,$$

$$K_{s+1} = \left\{x = (x_1, \dots, x_{s+1}) \in [0,1]^{s+1}: x_{s+1} \le \Delta\right\}.$$

The problem of estimating G from independent uniformly distributed in G observations $X_1, \dots, X_n$ has the same solution as above. However, now the correct normalizing factor is

$$\psi_n = ((\log n)/n)^{1/(s+1)}.$$

The constant $1/\sqrt{2\pi}$ in (4) must be replaced in the solution d_s of the equation

$$2^{-(s+1)}\operatorname{mes}(B_{s+1}(2d_s)) = s/(s+1).$$

Here $B_{s+1}(\rho)$ is Euclidean ball of radius ρ in R^{s+1}. Solving this equation, one gets the following result.

THEOREM 3. *For any $s \geq 1$ and any loss function $w(t)$ the following equality holds*:

$$\lim_{n\to\infty} \inf_{\hat{G}_n} \sup_{G\in\mathscr{G}_s} E_G\Big[w\Big((n/(\log n))^{1/(s+1)} d\big(G, \hat{G}_n\big)\Big)\Big] = w(d_s),$$

where

$$d_s = \Big(s(s+1)^{-1}\big(\operatorname{mes}(B_{s+1}(1))\big)^{-1}\Big)^{1/(s+1)}$$
$$= \left(\frac{1}{\sqrt{\pi}}\right)\left(s(s+1)^{-1}\Gamma\left(\frac{s+3}{2}\right)\right)^{1/(s+1)}$$

Nonuniform distributions. Consider again the two-dimensional case ($s = 1$). An important question is whether one can drop the assumption of uniformity. Let $Q(x)$ be a continuous strictly positive function on $K = [0,1] \times [0,1]$. Assume that the P_G-distribution of each independent observation X_i, $i = 1, \dots, n$, has density

$$q_G(x) = Q(x) I\{x \in G\} \Big/ \int_G Q(x)\, dx. \tag{19}$$

It is difficult to find the exact minimax constant in this model because the lower and upper bounds are expressed in terms of minimal and maximal values of $q_G(x)$ over the edge of G. These values do not coincide, except for some special examples of Q. However, it can be proved that the free disposal hull F_n has the optimal rate of convergence $\psi_n = ((\log n)/n)^{1/2}$.

THEOREM 4. *Let X_i, $i = 1, \dots, n$, be independent observations having density* (19). *Then for any loss function $w(t)$ there exists a constant C such that*

$$\lim_{n\to\infty} \sup_{G\in\mathscr{G}_s} E_G\Big[w\Big((n/(\log n))^{1/2} d(G, F_n)\Big)\Big] \leq C.$$

PROOF. Use the notation of Section 4. Let the radii $\tilde{\rho}_k$ be such that

$$\operatorname{mes}(W(u, \tilde{\rho}_k)) \geq (Q_{\max}/Q_{\min})\operatorname{mes}(G)(0.5 + k\varepsilon)\psi_n^2,$$

where $Q_{\max}$ and $Q_{\min}$ are the maximal and minimal values of $Q(x)$ in K.

Note that

$$P_G(X_i \in W(u_l, \tilde{\rho}_k)) = \int_{W(u_l, \tilde{\rho}_k)} q_G(x)\, dx$$
$$\geq \int_{W(u_l, \tilde{\rho}_k)} (Q_{\min}/(Q_{\max} \operatorname{mes}(G)))\, dx \geq (0.5 + k\varepsilon)\psi_n^2.$$

Since Lemma 3 remains unchanged, the rest of proof is the same as that in Section 4. □

Acknowledgment. We would like to thank Mats Rudemo for the remark concerning the definition of $\mathscr{G}$.

REFERENCES

BIRGÉ, L. (1987). On the risk of histograms for estimating decreasing densities. *Ann. Statist.* **15** 1013–1022.

CHEVALIER, J. (1976). Estimation du support et du contenu du support d'une loi de probabilité. *Ann. Inst. H. Poincaré Sec. B* **12** 339–364.

DEPRINS, D., SIMAR, L. and TULKENS, H. (1984). Measuring labor efficiency in post offices. In *The Performance of Public Enterprises: Concepts and Measurements* (M. Marchand, P. Pestieau and H. Tulkens, eds.) 243–267. North-Holland, Amsterdam.

DEVROYE, L. and WISE, G. L. (1980). Detection of abnormal behavior via nonparametric estimation of the support. *SIAM J. Appl. Math.* **38** 480–488.

DONOHO, D. (1992). Asymptotic minimax risk (for sup-norm): solution via optimal recovery. Unpublished manuscript.

EFROIMOVICH, S. Y. and PINSKER, M. S. (1981). Estimation of square-integrable spectral density based on a sequence of observations. *Problems Inform. Transmission* **17** 182–196.

EFROIMOVICH, S. Y. and PINSKER, M. S. (1982). Estimation of square-integrable probability density of a random variable. *Problems Inform. Transmission* **18** 175–189.

GEFFROY, J. (1964). Sur un problème d'estimation géométrique. *Publ. Inst. Statist. Univ. Paris* **13** 191–120.

GOLUBEV, G. K. (1982). On minimax filtering of functions in L_2. *Problems Inform. Transmission* **18** 67–75.

GRENANDER, U. (1956). On the theory of mortality measurement. *Skand. Aktuarietidskr.* **39** 125–153.

GROENEBOOM, P. (1985). Estimating the monotone density. In *Proceedings of the Berkeley Conference in Honor of Jerzy Neyman and Jack Kiefer* (L. Le Cam and R. A. Olshen, eds.) **2** 539–555. Wadsworth, Belmont, CA.

IBRAGIMOV, I. A. and KHAS'MINSKII, R. Z. (1982). Estimation of a distribution density belonging to a class of entire functions. *Theory Probab. Appl.* **27** 551–562.

JOHNSON, N. L. and KOTZ, S. (1977). *Urn Models and Their Applications*. Wiley, New York.

KOROSTELEV, A. P. (1992). Exact asymptotic minimax estimate of a nonparametric regression in the uniform norm. *Theory Probab. Appl.* **38** 875–882 (in Russian).

KOROSTELEV, A. P. and TSYBAKOV, A. B. (1992). Asymptotically minimax image reconstruction problems. In *Topics in Nonparametric Estimation. Adv. Soviet Math.* (R. Z. Khasminskii, ed.) **12** 45–86. Amer. Math. Soc., Providence, RI.

KOROSTELEV, A. P. and TSYBAKOV, A. B. (1993a). Estimation of the density support and its functionals. *Problems Inform. Transmission* **29** 1–15.

KOROSTELEV, A. P. and TSYBAKOV, A. B. (1993b). *Minimax Theory of Image Reconstruction*. Springer, New York.

MAMMEN, E. and TSYBAKOV, A. B. (1995). Asymptotical minimax recovery of sets with smooth boundaries. *Ann. Statist.* **23** 502–524.

NUSSBAUM, M. (1985). Spline smoothing and asymptotic efficiency in L_2. *Ann. Statist.* **13** 984–997.

PINSKER, M. S. (1980). Optimal filtering of square integrable signals in Gaussian white noise. *Problems Inform. Transmission* **16** 120–133.

RÉNYI, A. and SULANKE, R. (1963). Über die konvexe Hülle von n zufällig gewählten Punkten. *Z. Wahrsch. Verw. Gebiete* **2** 75–84.

RÉNYI, A. and SULANKE, R. (1964). Über die konvexe Hülle von n zufällig gewählten Punkten. II. *Z. Wahrsch. Verw. Gebiete* **3** 138–147.

RIPLEY, B. D. and RASSON, J. P. (1977). Finding the edge of a Poisson forest. *J. Appl. Probab.* **14** 483–491.

A. P. KOROSTELEV
INSTITUTE OF SYSTEM ANALYSIS
ACADEMY OF SCIENCES OF RUSSIA
PROSPEKT 60-LET OKTYABRYA 9
MOSCOW 117312
RUSSIA

L. SIMAR
INSTITUT DE STATISTIQUE
UNIVERSITÉ CATHOLIQUE DE LOUVAIN
34, VOIE DU ROMAN PAYS
1348, LOUVAINE-LA-NEUVE
BELGIUM

A. B. TSYBAKOV
UNIVERSITÉ PARIS 6
4, PLACE JUSSIEU
LSTA, B.P. 158
75252 PARIS CEDEX 05
FRANCE