The Annals of Statistics
1984, Vol. 12, No. 3, 1090–1099

ON THE ESTIMATION OF A CONVEX SET[1]

BY MARC MOORE

Ecole Polytechnique, Montréal

Given independent observations $x_1, \cdots, x_n$ drawn uniformly from an unknown compact convex set D in $\mathbb{R}^p$ (p known) it is desired to estimate D from the observations. This problem was first considered, for $p = 2$, by Ripley and Rasson (1977). We consider a decision-theoretic approach where the loss function is $L(D, \hat{D}) = m(D \,\triangle\, \hat{D})$.We prove the completeness of the Bayes estimation rules. A form for the nonrandomized Bayes estimation rules is presented and applied, for an a priori law reflecting ignorance, to the cases $p = 1$ and where D is a rectangle in the plane; some comparisons are made with other estimation methods suggested in the literature. Finally, the consistency of the estimation rules is studied.

1. Introduction. Ripley and Rasson (1977) propose a solution to the following problem originally formulated by Professor D. G. Kendall: given a realization of a homogeneous planar Poisson process of unknown intensity within a compact convex set D, find D. Let $x_1, \cdots, x_N$ denote the points in D where realizations are observed. Conditionally on the value n of N, $x_1, \cdots, x_n$ are the values of n independent and uniformly distributed random vectors on D. When the value n of N is known, the problem is then reduced to the estimation of the contour of a compact convex set D given the position of n points uniformly drawn from it.

The solution analyzed by Ripley and Rasson (1977) (the R-R procedure) consists of a dilatation of the convex hull, $H(\mathbf{x})$, of $\mathbf{x} = (x_1, \cdots, x_n)$ about its centroid. More precisely, they suggest considering first the set $s(H(\mathbf{x})) = H(\mathbf{x}) - g(H(\mathbf{x}))$, where $g(H(\mathbf{x}))$ is the centroid of $H(\mathbf{x})$; when the Lebesgue measure of D, $m(D)$, is known, they argue that a translation of

$$[m(D)/m(H(\mathbf{x}))]^{1/2}s(H(\mathbf{x})), \tag{1.1}$$

is the maximum likelihood estimator of D. For $m(D)$ unknown, they find a constant c such that $E[m(cs(H(\mathbf{x})))] = m(D)$,

$$c = \{(n+1)/[n+1-E[V_{n+1}]]\}^{1/2},$$

where V_{n+1} is the number of vertices in the convex hull of $x_1, \cdots, x_{n+1}$ (i.e., if there would be one more point drawn from D), and propose to use it as the factor of $s(H(\mathbf{x}))$ in (1.1). Rasson (1979) obtains, under particular conditions, an expression for $E[V_{n+1}]$. Since the computation of $E[V_{n+1}]$ is difficult, we may choose, in practice, to replace it by an estimator; for example, the observed value v_n of V_n, i.e., to use $c = [n/(n - v_n)]^{1/2}$.

Received June 1982; revised September 1983.

[1] Research supported by the Natural Sciences and Engineering Research Council of Canada grant A-8211.

AMS 1980 *subject classifications.* Primary 62F99; secondary 62F15, 62C10, 60D05.

Key words and phrases. Estimation, convex set, decision theory, Bayesian, complete, convergence.

The final R-R estimator is

$$\hat{D} = g(H(\mathbf{x})) + cs(H(\mathbf{x})). \tag{1.2}$$

This solution is affine invariant, as is the problem itself.

In this paper, we first formulate this problem (still conditionally on the value of N) for convex sets in $\mathbb{R}^p$, in the context of statistical decision theory using an appropriate loss function. In Section 3 we prove the completeness of the class of Bayes estimation rules. The form of the nonrandomized Bayes procedures is derived in Section 4 and applied to the case $p = 1$ in Section 5, and to a class of rectangles in the plane in Section 6; some comparisons with other estimators are made. Finally, some observations about consistency are presented in Section 7. In particular, we show that under rather weak conditions the procedure suggested here and the R-R procedure are asymptotically equivalent.

2. The problem. Let $\mathscr{D}$ be a uniformly bounded class of compact convex sets in $\mathbb{R}^p$ (p fixed and known), each of positive measure (there exists a compact convex set $K \subset \mathbb{R}^p$ such that $m(K) < \infty$ and $D \subseteq K$ for every $D \in \mathscr{D}$). Given the independent observations $x_1, \cdots, x_n$ drawn uniformly from an unknown element $D \in \mathscr{D}$, we wish to produce an estimator $\hat{D}$ of D. We consider a decision theoretic approach where: the set of "states of nature" is $\mathscr{D}$, the decision space is a class of compact convex sets, $\mathscr{D}'$, in $\mathbb{R}^p$ and the loss function is

$$L(D, \hat{D}) = m(D \,\triangle\, \hat{D}),$$

where $D \,\triangle\, \hat{D} = \{x\colon x \in D \cap \hat{D}^c \text{ or } x \in D^c \cap \hat{D}\}$. The function $L(\cdot, \cdot)$ defines a metric on $\mathscr{D} \times \mathscr{D}'$. The analogues to the metric L can be defined on $\mathscr{D} \times \mathscr{D}$ and $\mathscr{D}' \times \mathscr{D}'$; let $\mathscr{D}$ and $\mathscr{D}'$ denote the σ-algebras generated by the open sets for these metrics.

A decision rule δ is a function from the sample space, $\Omega = \prod^n D$ (the n-fold product of $D \in \mathscr{D}$), to the set of probability measures on $(\mathscr{D}', \mathscr{D}')$ such that, for every $B \in \mathscr{D}'$, the function

$$\mathbf{x} \to G_{\delta(\mathbf{x})}(B)$$

is $\mathscr{B}_p = \prod^n \mathscr{B}_p$-measureable, $G_{\delta(\mathbf{x})}(\cdot)$ being the probability measure on $(\mathscr{D}', \mathscr{D}')$ assigned to $\mathbf{x}$ by δ and $\mathscr{B}_p$ the class of Borel sets in $\mathbb{R}^p$.

The risk associated with δ in estimating D is

$$R(D, \delta) = \int_{\prod^n \mathbb{R}^p} \int_{\mathscr{D}'} L(D, D')\, dG_{\delta(\mathbf{x})}(D') \chi_\Omega(\mathbf{x})/[m(D)]^n \, d\mathbf{x}$$

$\chi_A(\cdot)$ being the characteristic function of the set A. Given an a priori probability measure λ on $(\mathscr{D}, \mathscr{D})$,

$$r(\lambda, \delta) = \int_{\mathscr{D}} R(D, \delta)\, d\lambda(D)$$

denotes the Bayes risk corresponding to δ relative to λ; the estimation rule δ_0 is Bayes with respect to λ, if

$$r(\lambda, \delta_0) = \inf_\delta r(\lambda, \delta).$$

3. Completeness of Bayes estimation rules. To study some properties of the class of Bayes estimation rules, we impose the following conditions on $\mathscr{D}$ and $\mathscr{D}'$:

(i) There exists a $\delta > 0$ such that $m(D) \geq \delta$ for every $D \in \mathscr{D}$.

(ii) $\mathscr{D}'$ is uniformly bounded.

(iii) There exists a $\delta' > 0$ such that $m(D') \geq \delta'$ for every $D' \in \mathscr{D}'$.

THEOREM 3.1. *Under the conditions* (i), (ii) *and* (iii), *the class of Bayes estimation rules is complete.*

The proof of this theorem is rather technical and essentially consists of a slight generalization of Theorem 10, page 385, in Berger (1980). The main steps may be summarized as follows:

1. Under the conditions imposed, the spaces $(\mathscr{D}, \mathscr{D})$ and $(\mathscr{D}', \mathscr{D}')$ are both compact. This follows from Blaschke's theorem (Valentine, 1964, page 37), using the fact that the metric considered here is equivalent to the Minkowski distance, (assumption (iii) could be removed if $\mathscr{D}'$ includes the empty set).

2. Since $(\mathscr{D}', \mathscr{D}')$ is compact, since $L(\cdot, \cdot)$ is bounded and continuous, and since the family of distributions is dominated, it can be shown that the set of risk functions is compact in the sense of pointwise convergence (applications of proposition 1 in Brown (1980)).

3. The joint density of the observations is, for almost every $\mathbf{x}$, continuous with respect to $D \in \mathscr{D}$. Then, since $L(\cdot, \cdot)$ is bounded, we may show that $R(D, \delta)$ is continuous with respect to D for every δ.

4. From 2. and 3., and the compactness of $(\mathscr{D}, \mathscr{D})$, we obtain the desired result (see Brown (1981), proposition 1.9).

To obtain Theorem 3.1 we have to consider all the decision rules (randomized and nonrandomized). From a practical point of view it is of interest to know if we can restrict the search of Bayes decision rules to nonrandomized procedures.

THEOREM 3.2. *Under the conditions* (i), (ii) *and* (iii), *for every* $\varepsilon > 0$ *and every decision rule* δ *there exists a nonrandomized decision rule* δ^* *such that* $|R(D, \delta) - R(D, \delta^*)| \leq \varepsilon$ *for every* $D \in \mathscr{D}$.

PROOF. From the conditions imposed on $\mathscr{D}$ the space $(\mathscr{D}, \mathscr{D})$ is compact for the metrics

$$\rho_1(D_1, D_2) = \sup_{S \in \mathscr{B}_p} |m(S \cap D_1)/m(D_1) - m(S \cap D_2)/m(D_2)|$$

and

$$\rho_2(D_1, D_2) = \sup_{D' \in \mathscr{D}'} |m(D_1 \Delta D') - m(D_2 \Delta D')| = m(D_1 \Delta D_2),$$

which are equivalent.

Also under the restrictions imposed on $\mathscr{D}'$, the space $(\mathscr{D}', \mathscr{D}')$ is compact for the metric

$$\rho'(D_1' D_2') = \sup_{D \in \mathscr{D}} | m(D \vartriangle D_1') - m(D \vartriangle D_2') | = m(D_1' \vartriangle D_2').$$

Then, from an easy adaptation of Theorem 8.1 in Dvoretzky, Wald and Wolfowitz (1951) we have the desired result. □

REMARK. For $p = 1$, the loss function presents a convexity property which permits us to prove that the nonrandomized procedures form an essentially complete class.

4. Nonrandomized Bayes estimation rules. The Bayes risk corresponding to a nonrandomized decision rule is

$$\begin{aligned} r(\lambda, \delta) &= \int_{\mathscr{D}} \int_{\prod^n \mathbb{R}^p} m(D \vartriangle \delta(\mathbf{x})) f(\mathbf{x} \mid D)\, d\mathbf{x}\, d\lambda(D) \\ &= \int_{\prod^n \mathbb{R}^p} \int_{\mathscr{D}} m(D \vartriangle \delta(\mathbf{x}))\, d\psi(D \mid \mathbf{x}) f(\mathbf{x})\, d\mathbf{x}, \end{aligned} \tag{4.1}$$

where

$$f(\mathbf{x} \mid D) = \chi_\Omega(\mathbf{x})/[m(D)]^n,$$

$$f(\mathbf{x}) = \int_{\mathscr{D}} f(\mathbf{x} \mid D)\, d\lambda(D),$$

and

$$d\psi(D \mid \mathbf{x}) = f(\mathbf{x} \mid D)\, d\lambda(D)/f(\mathbf{x})$$

is the posterior density of D given $\mathbf{x}$. From (4.1) it is clear that δ is a Bayes rule with respect to λ if, for each $\mathbf{x}$,

$$\int_{\mathscr{D}} m(D \vartriangle \delta(\mathbf{x}))\, d\psi(D \mid \mathbf{x}) \tag{4.2}$$

achieves its minimum at $\delta(\mathbf{x})$.

For a fixed $\delta(\mathbf{x})$, (4.2) is the expectation with respect to the law $\psi(\cdot \mid \mathbf{x})$ of the measure of the random set $D \vartriangle \delta(\mathbf{x})$. As shown by Robbins (1944) Fubini's theorem yields that this expectation can be written as

$$\int_{\mathbb{R}^p} \psi(\{D: y \in D \vartriangle \delta(\mathbf{x})\} \mid \mathbf{x})\, dy. \tag{4.3}$$

For many classes $\mathscr{D}$ it is possible to obtain an explicit expression for (4.3) and hence the $\delta(\mathbf{x})$ giving it its minimum value. From (4.3) we can show that if

$$S(\mathbf{x}) = \{y \in \mathbb{R}^p: \psi(\{D: y \in D\} \mid \mathbf{x}) \geq 1/2\} \tag{4.4}$$

is convex (a.e.) then $\delta(\mathbf{x}) = S(\mathbf{x})$ is a Bayes procedure, a fact which may be useful in the computation of Bayes procedures.

5. The one-dimensional case. When $p = 1$, the problem considered reduces to the estimation of the parameters a and ℓ from observations $x_1, \cdots, x_n$ drawn uniformly on $[a - \ell, a + \ell]$.

Estimation of the upper limit of the interval $[0, \theta]$ from a sample drawn uniformly on this interval has been carefully studied, both from the fixed-sample and from the sequential points of view. For example, Craig (1943) obtains the best linear unbiased estimator based on the ordered sample, $\hat{\theta} = (n + 1)x_{(n)}/n$, Alvo (1978) gives an account of the sequential approach, Jones (1978) presents a review of the literature regarding the Bayesian approach, and Susarla and O'Bryan (1979) study the empirical Bayes approach to obtain an interval estimate for θ. Another form of the one-parameter uniform distribution which has deserved study is the one where the sample is uniformly distributed on $[\theta - \frac{1}{2}, \theta + \frac{1}{2}]$, $\theta > 0$; see, for example, Welch (1939), Wald (1950), Blyth (1951) and Fraser (1952).

The two-parameter uniform model, $[a - \ell, a + \ell]$ has received less attention. All the proposed estimators are based on the ordered sample $x_{(1)}, \cdots, x_{(n)}$; since every approach gives the same estimator for a, $(x_{(1)} + x_{(n)})/2$, we will consider only estimation of the length 2ℓ. Carlton (1946) obtains the maximum likelihood estimator: $x_{(n)} - x_{(1)}$, Lloyd (1952) derives the best linear unbiased (b.l.u.) estimator: $(n + 1)(x_{(n)} - x_{(1)})/(n - 1)$, while Ripley and Rasson (1977) give the maximum likelihood estimator based on $x_{(n)} - x_{(1)}$: $n(x_{(n)} - x_{(1)})/(n - 1)$. The least-squares affine invariant (l.s.a.i.) estimator, however, is $(n + 2)(x_{(n)} - x_{(1)})/n$.

Our approach here is decision-theoretic with loss function

$$L[(a, \ell), (\hat{a}, \hat{\ell})] = m([a - \ell, a + \ell] \,\Delta\, [\hat{a} - \hat{\ell}, \hat{a} + \hat{\ell}]). \tag{5.1}$$

In order to make comparisons with other methods of estimation for a and ℓ, we consider an a priori measure λ reflecting ignorance. The decision problem considered is not invariant under the affine group, but it becomes invariant under that group if the loss function is

$$L_1[(a, \ell), (\hat{a}, \hat{\ell})] = m([a - \ell, a + \ell] \,\Delta\, [\hat{a} - \hat{\ell}, \hat{a} + \hat{\ell}])/2\ell. \tag{5.2}$$

When the loss function L_1 is used, the a priori law reflecting ignorance is the right invariant Haar measure (Berger, 1980, page 262):

$$d\lambda(a, \ell) \propto 1/\ell. \tag{5.3}$$

It is easy to see that the generalized Bayes procedure for loss L_1 and prior (5.3), is the same as the generalized Bayes procedure for loss L and prior

$$d\lambda(a, \ell) \propto 1/\ell^2. \tag{5.4}$$

The prior measure (5.4) is the inner prior and (5.3) is the outer prior corresponding to affine transformation (Villegas, 1977).

From (4.4) we verify that the Bayes procedure for loss L and prior (5.4) is

$$\hat{a}_0 = (x_{(1)} + x_{(n)})/2, \quad \hat{\ell}_0 = (2^{1/n} - \tfrac{1}{2})(x_{(n)} - x_{(1)}). \tag{5.5}$$

Here

$$\psi(\{[a - \ell, a + \ell]: y \in [a - \ell, a + \ell]\} \mid x_{(1)}, \cdots, x_{(n)})$$
$$= \begin{cases} [(x_{(n)} - x_{(1)})/(x_{(n)} - y)]^n & \text{if } y < x_{(1)} \\ [(x_{(n)} - x_{(1)})/(y - x_{(1)})]^n & \text{if } y > x_{(n)}. \end{cases}$$

If the prior (5.4) is used with the quadratic loss function $(a - \hat{a})^2 + (\ell - \hat{\ell})^2$ we obtain $(n + 1)(x_{(n)} - x_{(1)})/(n - 1)$ as the Bayes estimator for 2ℓ; if the prior (5.3) is used with the same loss function we obtain $n(x_{(n)} - x_{(1)})/(n - 2)$.

The Bayes estimator (5.5) for 2ℓ, has smaller bias and m.s.e. than the maximum likelihood estimators. However it is not better (more bias and m.s.e.) than the b.l.u. estimator and the l.s.a.i. estimator. The comparison of the m.s.e. between (5.5) and the l.s.a.i. estimator (also the b.l.u. estimator) is in a sense unfair since the estimators were developed for different loss functions; it may, however, give information on the robustness of a procedure relative to loss function.

Since $(n + 2)/n \geq 2^{(n+1)/n} - 1$ for every n, the length of the estimated interval is smaller with (5.5) than with the l.s.a.i. estimator, which itself produces a smaller length than the b.l.u. estimator.

6. Example: a class of rectangles. To illustrate the application of our approach to convex sets in the plane, we consider the situation where both $\mathscr{D}$ and $\mathscr{D}'$ are the same set of rectangles with sides parallel to given axes. This set is characterized by the coordinates of the center, (t, u), and the half-lengths of the sides, r_1 and r_2, i.e., an element of $\mathscr{D}$ is $[t - r_1, t + r_1] \times [u - r_2, u + r_2]$. We use the inner prior (Villegas, 1977) corresponding to the transformation $(x, y) \rightarrow (c_1 x + b_1, c_2 y + b_2)$; c_1, c_2, b_1, b_2 are some constants:

$$d\lambda(t, u, r_1, r_2) \propto r_1^{-2} r_2^{-2}.$$

The Bayes procedure may be computed from (4.3) using the fact that

$$\psi(\{D: y \in D\} \mid x_1, \cdots, x_n) = \left[\frac{M_1(x_1, \cdots, x_n) M_2(x_1, \cdots, x_n)}{M_1(x_1, \cdots, x_n, y) M_2(x_1, \cdots, x_n, y)}\right]^n,$$

where $M_i(z_1, \cdots, z_k) = \max_{1 \leq j \leq k}\{z_{ji}\} - \min_{1 \leq j \leq k}\{z_{ji}\}$, $i = 1, 2$. We obtain,

$$\hat{t} = \min_{1 \leq j \leq n}\{x_{j1}\} + M_1(x_1, \cdots, x_n)/2,$$

$$\hat{u} = \min_{1 \leq j \leq n}\{x_{j2}\} + M_2(x_1, \cdots, x_n)/2,$$

and $\hat{r}_1$, $\hat{r}_2$ the solutions of the equations:

$$4\hat{r}_2 + \frac{M_1M_2}{(n-1)}\left\{-2(2n+2)M_1^{n-1}\left(\frac{M_1}{2}+\hat{r}_1\right)^{-n} + 8(M_1M_2)^{n-1}\left(\frac{M_2}{2}+\hat{r}_2\right)^{-n+1}\left(\frac{M_1}{2}+\hat{r}_1\right)^{-n}\right\} = 0,$$

$$4\hat{r}_1 + \frac{M_1M_2}{(n-1)}\left\{-2(2n+2)M_2^{n-1}\left(\frac{M_2}{2}+\hat{r}_2\right)^{-n} + 8(M_1M_2)^{n-1}\left(\frac{M_2}{2}+\hat{r}_2\right)^{-n}\left(\frac{M_1}{2}+\hat{r}_1\right)^{-n+1}\right\} = 0$$

where the arguments $x_1, \cdots, x_n$ have been deleted in $M_1(x_1, \cdots, x_n)$ and $M_2(x_1, \cdots, x_n)$.

As an illustration we consider a case where $n = 6$ and the points are (0, 5), (3, 0), (6, 3), (3, 8), (2, 3) and (4, 5); we obtain $\hat{t} = 3$, $\hat{u} = 4$, $\hat{r}_1 = 3{,}6785$, $\hat{r}_2 = 4{,}9046$. Figure 1 illustrates the estimation obtained and the one given by (1.2) with $c = \sqrt{3}$.

Here a generalized prior has been utilized because we wanted to represent ignorance. It is to be noted that the decision-theoretic approach permits, using an appropriate probability measure as prior, to incorporate prior information.

7. Convergence. In this section we study the consistency of the Bayes and R-R procedures and their asymptotic equivalence.

DEFINITION. Given a compact convex set D, and an estimator $\hat{D}_n$ produced by a rule δ based on n points drawn uniformly from D, then δ is said to be convergent if $m(D \vartriangle \hat{D}_n)$ converges in probability to zero.

THEOREM 7.1. *The* R-R *procedure used with* $c = [n/(n - v_n)]^{1/2}$ *is convergent.*

PROOF. From (1.2) it is sufficient to show that $m(H_n(\mathbf{x}) \vartriangle D)$ and V_n/n both converge to zero in probability ($H_n(\mathbf{x})$ is the convex hull based on n sample points).

Consider first the case where $D = [0, 1] \times [0, 1]$; that $m(H_n(\mathbf{x}) \vartriangle D)$ converges to zero almost surely then follows from the Borel-Cantelli theorem. The same result for an arbitrary $D \in \mathscr{D}$ follows since there exists a homeomorphism between the interior of D and $S = (0, 1) \times (0, 1)$.

To prove that V_n/n converges to zero in probability it is sufficient to note that

$$E[V_n/n] = 1 - E[A_{n-1}]/m(D),$$

$$\operatorname{Var}[V_n/n] \le (1 - E[A_{n-1}]/m(D))(E[A_{n-1}]/m(D)),$$

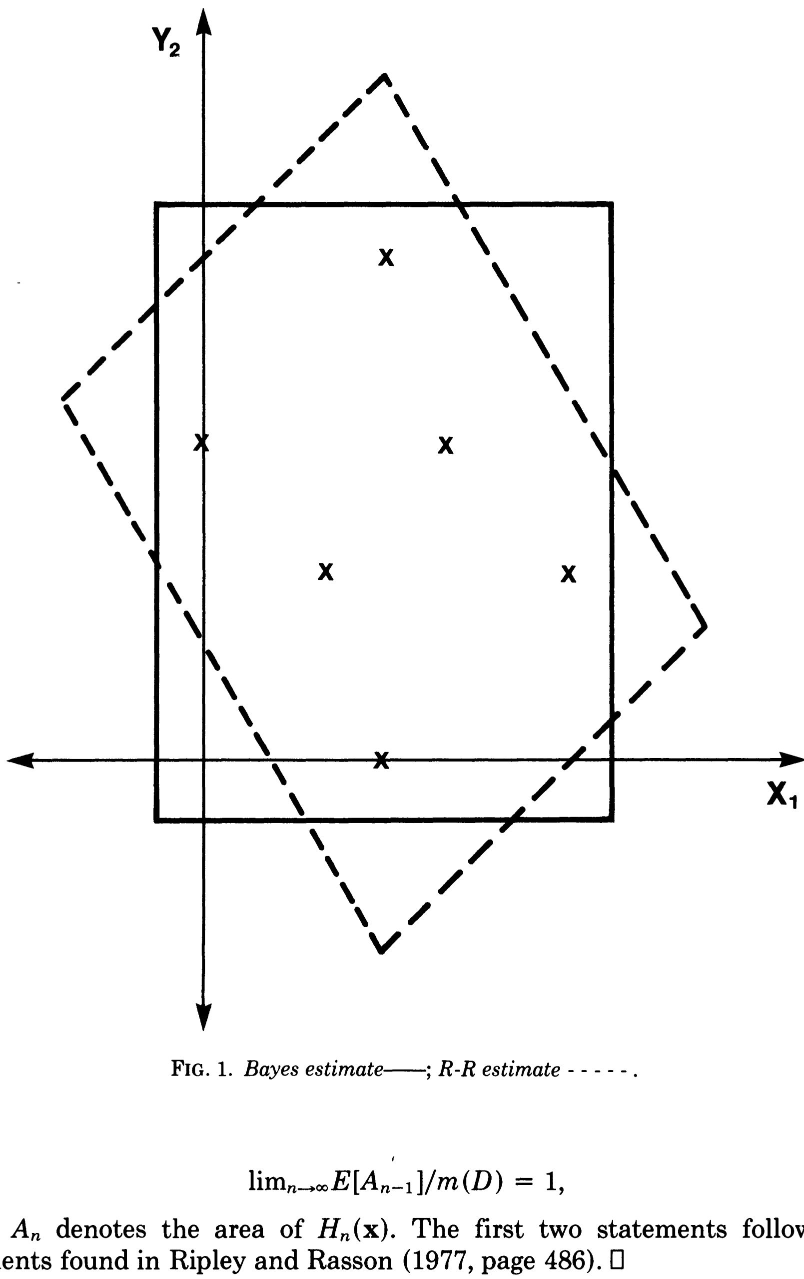

FIG. 1. *Bayes estimate——; R-R estimate - - - - -.*

and

$$\lim_{n\to\infty} E[A_{n-1}]/m(D) = 1,$$

where A_n denotes the area of $H_n(\mathbf{x})$. The first two statements follow from arguments found in Ripley and Rasson (1977, page 486). □

We now consider the nonrandomized Bayes procedures defined in Section 4. The following theorem is easy to prove.

THEOREM 7.2. *If the a priori measure λ is such that the posterior density*

$d\,\psi(D \mid \mathbf{x})$ converges in probability to the degenerate density at D_0, the true value of D, then the nonrandomized Bayes procedure, if it exists, is convergent.

COROLLARY. *If $\hat{D}_{1n}$ denotes the estimator obtained with the* R-R *procedure based on n sample points, and if $\hat{D}_{2n}$ is the Bayes estimator, based on the same sample points and corresponding to an a priori measure for which the condition of Theorem* 7.2 *is satisfied, then $m(\hat{D}_{1n} \Delta \hat{D}_{2n})$ converges in probability to zero.*

REMARK 1. In the case where the class $\mathscr{D}$ is characterized by a vector parameter $(\theta_1, \cdots, \theta_p)$, and where λ is an a priori probability measure admitting a density, we see, from Doob (1949), that the condition of Theorem 7.2 is satisfied.

REMARK 2. In the case of a more general space $\mathscr{D}$ we may use a result given by Strasser (1981) to verify the condition of Theorem 7.2. We can show that an approximate maximum likelihood estimator, $\hat{D}_n$, as defined by Strasser (1981), is here such that

$$f(\mathbf{x} \mid \hat{D}_n)/f(\mathbf{x} \mid D_0) \geq 1$$

for almost every $\mathbf{x} = x_1, \cdots$. Then, from Wald (1949), $\hat{D}_n$ is almost surely convergent to estimate D_0, the true value of the parameter. So from Theorem 2.5 of Strasser (1981), if $\mathscr{D}$ and the a priori measure satisfy the conditions imposed there, then the posterior density converges in probability to the degenerate density in D_0.

Acknowledgements. The author wishes to thank the associate editor for the numerous and very constructive suggestions made on a first version of this work. They led to a shorter and improved paper.

REFERENCES

ALVO, M. (1978). Sequential estimation of a truncation parameter. *J. Amer. Statist. Assoc.* **73** 404–407.

BERGER, J. O. (1980). *Statistical Decision Theory, Foundations, Concepts and Methods.* Springer-Verlag, New York.

BLYTH, C. R. (1951). On minimax statistical decision procedures and their admissiblity. *Ann. Math. Statist.* **22** 22–43.

BROWN, L. D. (1980). A necessary condition for admissibility. *Ann. Statist.* **8** 540–544.

BROWN, L. D. (1981). A complete class theorem for statistical problems with finite sample spaces. *Ann. Statist.* **9** 1289–1301.

CARLTON, A. G. (1946). Estimating the parameters of a rectangular distribution. *Ann. Math. Statist.* **17** 355–358.

CRAIG, A. T. (1943). A note on the best linear estimate. *Ann. Math. Statist.* **14** 88–90.

DAVID, H. A. (1981). *Order Statistics*, 2nd ed. Wiley, New York.

DOOB, J. L. (1949). Application of the theory of martingales. *Colloq. Internat. du CNRS* **13** 22–28.

DVORETZKY, A., WALD, A., and WOLFOWITZ, J. (1951). Elimination of randomization in certain statistical decision procedures and zero-sum two-person games. *Ann. Math. Statist.* **22** 1–21.

EVANS, E. (1983). Estimating events. *Ann. Statist.* **11** 1218–1224.

FRASER, D. A. S. (1952). Sufficient statistics and selection depending on the parameter. *Ann. Math. Statist.* **23** 417–426.

JONES, P. W. (1978). Bayesian point estimation of the unknown upper limit of a uniform distribution. *Biometrical J.* **20** 619–622.

LLOYD, E. H. (1952). Least squares estimation of location and scale parameter using order statistics. *Biometrika* **39** 88–95.

RASSON, J. P. (1979). Estimation de formes convexes du plan. *Statistiques et Analyse des données* **1** 31–46.

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

ROBBINS, H. (1944). On the measure of a random set I. *Ann. Math. Statist.* **15** 70–74.

STRASSER, H. (1981). Consistency of maximum likelihood and Bayes estimates. *Ann. Statist.* **9** 1107–1113.

SUSARLA, V., and O'BRYAN, T. (1979). Empirical Bayes interval estimates involving uniform distributions. *Comm. Statist. A-Theor. Methods* **8** 385–397.

VALENTINE, F. A. (1964). *Convex Sets.* McGraw-Hill, New York.

VILLEGAS, C. (1977). Inner statistical inference. *J. Amer. Statist. Assoc.* **72** 453–458.

WALD, A. (1949). Note on the consistency of the maximum likelihood estimate. *Ann. Math. Statist.* **20** 595–601.

WELCH, B. L. (1939). On confidence limits and sufficiency with particular reference to parameters of location. *Ann. Math. Statist.* **10** 58–69.

DÉPARTEMENT DE MATHÉMATIQUES APPLIQUÉES
ECOLE POLYTECHNIQUE
C.P. 6079, SUCC. "A"
MONTRÉAL, QUÉBEC
CANADA H3C 3A7