The Annals of Mathematical Statistics
1971, Vol. 42, No. 4, 1339–1347

SEQUENTIAL DISCRIMINATION WITH LIKELIHOOD RATIOS

BY CHARLES YARBROUGH[1]
University of California, Berkeley

1. Introduction and summary. Let $X_1, X_2, \cdots$ be independent identically distributed random variables. You observe the X's sequentially, knowing that their distribution is one of countably many different probabilities. Within an arbitrary error level, can you decide which one? This is the general problem of sequential discrimination.

Freedman [4] showed that the discriminability of a family Θ is equivalent to a seemingly weaker condition. Namely, for any error level α and any particular $\theta \in \Theta$ there is a uniformly powerful fixed sample size test of $\{\theta\}$ versus $\Theta - \{\theta\}$ with error level uniformly as small as α. The proof is constructive. Given the fixed sample size tests there is a recipe for manufacturing a sequential procedure to decide among all the members of Θ.

The fixed sample size tests are, however, still required. Hoeffding and Wolfowitz [5] considered this problem at length. LeCam and Schwartz [7] also touched upon it briefly. Both papers considered separations in various topologies and structures.

Here we return to the original problem and ask whether likelihood ratios can be sensibly used. A rule is easy to specify. For each $\theta \in \Theta$ you pick a number greater than one. Now watch $X_1, X_2, \cdots$. At each step compute the likelihood ratio for every pair of probabilities. Eventually it may happen that for some θ all the ratios with θ in the numerator are as big as the pre-assigned number. If so, stop and declare θ to be the true distribution. This rule is the extension to the countable case of the general sequential probability ratio test proposed by Barnard [2] and detailed by Armitage [1]. It does require the computation of all the likelihood ratios; but since Θ is countable, there is always at least one base for calculating densities. Any one will do.

Likelihood ratio procedures have the advantage of being easy to formulate. Also, the comparison of densities seems to be a reasonably natural technique. However, it does not always work. An example will illustrate that it may fail spectacularly.

When do likelihood ratio procedures work? The principal result is a characterization of families which are likelihood ratio discriminable. Check each probability θ separately. There may be a number $K(\theta)$ bigger than one which will be eventually exceeded simultaneously by all the ratios with θ in the numerator. If not, likelihood ratios will not work. If so, then the values $K(\theta)$ may be chosen so as to limit the error to any desired level.

---

[1] Current address: Instituto de Nutrición de Centro América y Panamá, Guatemala.

Despite their failures, likelihood ratio procedures may work when other natural conditions fail. Freedman showed that if each $\theta \in \Theta$ is isolated in the topology of setwise convergence, then Θ is discriminable. The converse is false. There is a family which is likelihood ratio discriminable, but which has one element in the setwise closure of all the others. This is a direct consequence of a recent theorem by LeCam [8].

Finally, with many familiar families likelihood ratio procedures have finite expected stopping time. Cases vary, however, and there is a discriminable family which has infinite expected stopping time for sampling under one of its elements.

2. Preliminaries. $(\Omega, \mathscr{G})$ is a measurable space and Θ is a countable collection of probabilities defined on $(\Omega, \mathscr{G})$. For $\theta \in \Theta$ and $N = 1, 2, \cdots, \infty$ the probability space $(\Omega^N, \mathscr{G}^N, \theta^N)$ is the usual N-fold product and $\mathscr{G}^{(N)}$ is the sub-σ-field of $\mathscr{G}^\infty$ consisting of events depending only on the first N coordinates. For a point $\omega \in \Omega^\infty$ denote the Nth coordinate by ω_N. Let $X_N(\omega) = \omega_N$.

Since Θ is countable there exist σ-finite measures on $(\Omega, \mathscr{G})$ which dominate all the $\theta \in \Theta$. Fix one such measure μ. Let f_θ be the density of θ with respect to μ. Then f_θ is a real valued nonnegative $\mathscr{G}$ measurable function on Ω. Write $f_\theta(\omega_1, \cdots, \omega_N)$ as shorthand for $\prod_{i=1}^N f_\theta(\omega_i)$. Let $\theta^\infty \mid \mathscr{G}^{(N)}$ be the restriction of θ^∞ to $\mathscr{G}^{(N)}$. Then $(\partial\theta^\infty \mid \mathscr{G}^{(N)}/\partial\mu^\infty \mid \mathscr{G}^{(N)})\,(\omega) = f_\theta(\omega_1, \cdots, \omega_N)$ holds μ^∞-a.e. This is an easy consequence of the natural association between $\mathscr{G}^{(N)}$ and $\mathscr{G}^N$.

For any pair $\theta, \sigma \in \Theta$ define the following likelihood ratio functions on Ω^∞, with $0/0 = 1$.

$$R_{\theta\sigma N}(\omega) = f_\theta(\omega_1, \cdots, \omega_N)/f_\sigma(\omega_1, \cdots, \omega_N)$$

$$R_{\theta N}(\omega) = \inf_{\sigma \in \Theta - [\theta]} R_{\theta\sigma N}(\omega).$$

Of course $R_{\theta\sigma N}$ and $R_{\theta N}$ are $\mathscr{G}^{(N)}$ measurable. For $\omega \in \Omega^\infty$ and N a natural number, define the shift mapping S_N of Ω^∞ onto itself by $S_N(\omega) = (\omega_{1+N}, \omega_{2+N}, \cdots)$. Two easy facts:

(1) For fixed $\sigma \neq \theta$ $\qquad \theta^\infty[R_{\theta\sigma N} \uparrow \infty] = 1.$

(2) For all positive integers L, M, and $N = L + M$

$$R_{\theta L}(\omega) \cdot R_{\theta M}(S_L(\omega)) \leqq R_{\theta N}(\omega).$$

Fix $\theta \in \Theta$ and let Φ be Θ–$\{\theta\}$. Then to say that θ is in the weak closure of Φ means that for every real $\varepsilon > 0$, every natural number M, and every collection $A_1, A_2, \cdots, A_M$ of $\mathscr{G}$-sets there is a $\sigma \in \Phi$ such that $|\sigma[A_i] - \theta[A_i]| < \varepsilon$ for all $i = 1, 2, \cdots, M$.

3. Definition of likelihood ratio discrimination. The formal definition is motivated by the following idea. Given Θ and a maximum permissible error level α, choose a function K_α on Θ with $K_\alpha(\theta) > 1$ for all $\theta \in \Theta$. Sample until all the likelihood ratios $R_{\theta\sigma N}$ are at least $K_\alpha(\theta)$ for some θ against all its competitors $\sigma \in \Theta$–$\{\theta\}$. Stop and declare θ to be the true distribution.

Formally, let K be a real-valued function on Θ with $K(\theta) > 1$ for all $\theta \in \Theta$. Define on Ω^∞:

$$(3)\quad \tau_K(\omega) = \infty \qquad \text{if } R_{\theta N}(\omega) < K(\theta) \text{ for all } \theta \in \Theta \text{ and for all } N = 1, 2, \cdots$$
$$= \min [N \mid R_{\theta N}(\omega) \geqq K(\theta) \quad \text{for some} \quad \theta \in \Theta] \quad \text{otherwise}$$

$$(4)\quad D_K(\omega) = \theta_0 \text{ (arbitrary)} \qquad \text{if} \quad \tau_K(\omega) = \infty$$
$$= \theta \qquad \text{if} \quad \tau_K(\omega) < \infty \quad \text{and} \quad R_{\theta\tau_K(w)}(\omega) \geqq K(\theta).$$

Since the values of K are strictly larger than one, D_K is unambiguous. No two ratios $R_{\theta\sigma N}(\omega)$ and $R_{\sigma\theta N}(\omega)$ can both be greater than one. Both $[\tau_K = N]$ and $[\tau_K = N, D_K = \theta]$ are $\mathscr{G}^{(N)}$ measurable, so the obvious procedure of waiting until τ_K and then choosing D_K makes sense as a non-randomized sequential decision procedure.

Definition. Suppose there is a function K_α on Θ for every real $\alpha \in (0, 1)$ such that for every $\theta \in \Theta$

$$(5)\qquad K_\alpha(\theta) > 1 \qquad \text{and}$$

$$(6)\qquad \theta^\infty[\tau_{K_\alpha} < \infty] = 1 \qquad \text{and}$$

$$(7)\qquad \theta^\infty[D_{K_\alpha} = \theta] \geqq 1 - \alpha.$$

Then likelihood ratio discrimination exists.

In this case we shall also say that the family Θ is likelihood ratio discriminable.

4. Theorems and discussion. The results announced in Section 1 are reiterated here in formal detail. The examples appear in the following section.

We are concerned with the properties of ratios of densities taken with respect to a fixed dominating measure μ. Is the choice of μ arbitrary? All procedures defined below involve only countable operations on sets defined by the countable family of measurable functions $f_\theta(\omega)/f_\sigma(\omega)$ for $\theta, \sigma \in \Theta$. So an affirmative answer is given by

(8) Fact. *Let μ and ρ be two measures which dominate all $\theta \in \Theta$. Let $f_\theta = \partial\theta/\partial\mu$ and $g_\theta = \partial\theta/\partial\rho$. Then for any pair $\theta, \sigma \in \Theta$ the ratios f_θ/f_σ and g_θ/g_σ are equal* θ-a.e.

To expose the main theorem we need a function $T_{\theta h}$ defined on Ω^∞ for each $\theta \in \Theta$ and each real h by

$$(9)\qquad T_{\theta h}(\omega) = \infty \qquad \text{if} \quad R_{\theta N}(\omega) < h \text{ for all } N = 1, 2, \cdots$$
$$= \min [N \mid R_{\theta N}(\omega) \geqq h] \qquad \text{otherwise.}$$

Useful facts about $T_{\theta h}$ are

$$(10)\qquad [T_{\theta h} = N] \in \mathscr{G}^{(N)}.$$

(11) $T_{\theta h}(\omega)$ is a monotone non-decreasing function of h.

(12) For K any function from Θ to $(1, \infty)$, for $N = 1, 2, \cdots$, and for all $\theta \in \Theta$

$$[\tau_K < N] \supset [T_{\theta K(\theta)} < N] \supset [\tau_K < N \quad \text{and} \quad D_K = \theta].$$

The event $[T_{\theta h} < \infty]$ is fundamental. It consists of those ω which will eventually identify θ at a likelihood ratio level of h. Suppose that there is a function K on Θ such that $\theta^\infty[T_{\theta K(\theta)} < \infty] = 1$ for all $\theta \in \Theta$. Then the likelihood ratio procedure surely stops, and it is necessary only to control the error. This is done by choosing new (and generally higher) values for K.

For all $\theta, \sigma \in \Theta$ $(R_{\sigma\theta N}; N = 1, 2, \cdots)$ is a nonnegative θ^∞-super-martingale with expectation $\leqq 1$. As such, it can be decomposed into a positive martingale with expectation one and a non-positive part. Applying the Kolmogorov inequality gives

$$\theta^\infty[R_{\sigma\theta N} \geqq K(\sigma) \quad \text{some } N = 1, 2, \cdots] \leqq 1/K(\sigma).$$

(13) $$\theta^\infty[D_K \neq \theta] \leqq \theta^\infty[R_{\sigma\theta N} \geqq K(\sigma) \text{ some } N \text{ and some } \sigma] \leqq \textstyle\sum_{\theta \in \Theta}[1/K(\sigma)]$$

which can be made arbitrarily small by a suitable choice of K.

If revising K does not alter the certainty of stopping, we are done. To check this we will have to examine $T_{\theta h}$ on $[1, \infty)$. Note that this includes $T_{\theta 1}$, even though all values of K are greater than one.

(14) LEMMA. *Fix* $\theta \in \Theta$ *and let* $h > g > 1$ *be real numbers. Then* $\theta^\infty[T_{\theta g} < \infty] = 0$ *if and only if* $\theta^\infty[T_{\theta h} < \infty] = 0$. *Also,* $\theta^\infty[T_{\theta g} < \infty] = 1$ *if and only if* $\theta^\infty[T_{\theta h} < \infty] = 1$. Informally: if you are sure to miss (or find) one of the $\theta \in \Theta$ by using some particular h, then you can use any other level. To prove (14) we need

(15) FACT. *Suppose* f *is a real-valued, non-increasing function on* $(1, \infty)$ *such that* $1 \geqq f(x) \geqq f(x^2) \geqq [f(x)]^2$ *holds for all* x.

(i). *If there is an* h *with* $f(h) = 1$ *then* $f \equiv 1$.

(ii). *If there is an* h *with* $f(h) = 0$ *then* $f \equiv 0$.

PROOF. *Claim* (i). By monotonicity $f(x) = 1$ for all $x \leqq h$. Fix $x > h$ and let N be an integer so large that $h^{2^N} \geqq x$. Then

$$f(x) \geqq f(h^{2^N}) \geqq [f(h^{2^{N-1}})]^2 \geqq \cdots \geqq [f(h)]^{2^N} = 1.$$

Claim (ii). By monotonicity $f(x) = 0$ for all $x > h$. Fix $x < h$ and let N be so large that $x^{2^N} \geqq h$. Then

$$0 = f(h) \geqq [f(x)]^{2^N}. \quad \square$$

PROOF OF (14). By (11) $[T_{\theta g} < \infty] \supset [T_{\theta h} < \infty]$. Moreover

$$[T_{\theta g^2} < \infty] \supset \textstyle\bigcup_{I=1}^\infty \bigcup_{J=1}^\infty [T_{\theta g} = I \text{ and } T_{\theta g}(S_I(\omega)) = J]$$

by (2). But the X's are independent and identically distributed, so for fixed I, $[R_{\theta J}(S_I(\omega)) \geqq g]$ is independent of $[T_{\theta g} = I]$ and has probability $\theta^\infty[T_{\theta g} = J]$. So

$$\theta^\infty[T_{\theta g} < \infty] \geqq \theta^\infty[T_{\theta g^2} < \infty] \geqq \textstyle\sum_{I=1}^\infty \sum_{J=1}^\infty \theta^\infty[T_{\theta g} = I] \cdot \theta^\infty[T_{\theta g} = J]$$
$$= \theta^\infty[T_{\theta g} < \infty]^2;$$

Fact (15) applies. □

For the next lemma we need the strong Markov property.

(16) FACT. *Let* τ *be a stopping time on* Ω^∞; *i.e.,* $[\tau = N] \in \mathscr{G}^{(N)}$. *Suppose* $\theta^\infty[\tau < \infty] = 1$ *and let* $\tau_0 = 0, \tau_1 = \tau$, *and* $\tau_{N+1} = \tau(S_{\tau_N})$. *Let*

$$Z_N = (\omega_{\tau_{N-1}(\omega)+1}, \cdots, \omega_{\tau_N(\omega)}).$$

Then $Z_1, Z_2 \cdots$ *are independently and identically distributed.*

(17) LEMMA. *Fix* $\theta \in \Theta$ *and let* $h > 1$ *be a real number. If*

$$\theta^\infty[T_{\theta 1} < \infty] = 1 \quad \textit{and} \quad \theta^\infty[T_{\theta h} < \infty] > 0 \quad \textit{then} \quad \theta^\infty[T_{\theta h} < \infty] = 1.$$

PROOF. For a fixed N (chosen below) define the following system on Ω^∞:

$$(18) \quad Y_0 = 0 \quad \text{and} \quad Y_K = \min[M \mid M > Y_{K-1} + N \quad \text{and} \quad R_{\theta M - Y_{K-1}}(S_{Y_{K-1}}) \geqq 1].$$

By (16) the increments $Y_{K+1} - Y_K$ are independent identically distributed. They are finite a.e. since $\theta^\infty[T_{\theta 1} < \infty] = 1$ implies that θ^∞-a.e $R_{\theta N}(\omega) \geqq 1$ infinitely often. So there is some N and p such that

$$\theta^\infty[T_{\theta h} \leqq N] \geqq p > 0.$$

Then $\theta^\infty[T_{\theta h} \leqq Y_1] \geqq p$, and Fact (2) implies

$$\theta^\infty[T_{\theta h} > Y_{K+1} \mid T_{\theta h} > Y_K] \leqq 1 - p.$$

So

$$(19) \qquad \theta^\infty[T_{\theta h} > Y_K] \leqq (1-p)^K$$

and $T_{\theta h}$ is finite. □

Lemmas (14) and (17) show that there are only three possibilities for $\theta^\infty[T_{\theta h} < \infty]$ as a function of h on $[1, \infty)$. It may be identically one, or zero on $(1, \infty)$ with any value at $h = 1$, or strictly positive and strictly less than 1. The main theorem will argue that likelihood ratios discriminate if and only if $\theta^\infty[T_{\theta h} < \infty]$ is identically one. First we rule out the other cases.

(20) THEOREM. *Suppose there is a* $\theta \in \Theta$ *such that* $\lim_{h \to 1} \theta^\infty[T_{\theta h} < \infty] < 1$. *Then* Θ *is not likelihood ratio discriminable.*

PROOF. Let $\lim_{h \to 1} \theta^\infty[T_{\theta h} < \infty] = 1 - \delta$. The monotonicity of $T_{\theta h}$ (11) implies $\theta^\infty[T_{\theta K(\theta)} < \infty] \leqq 1 - \delta$ for all K on Θ with $K(\theta) > 1$ for all $\theta \in \Theta$. But by (12)

$$\theta^\infty[\tau_K < \infty, D_K = \theta] \leqq \theta^\infty[T_{\theta K(\theta)} < \infty] \leqq 1 - \delta.$$

So either $\theta^\infty[\tau_K < \infty] < 1$ or $\theta^\infty[D_K = \theta] \leqq 1 - \delta$ contrary to definitions (6) and (7). □

(21) COROLLARY. *If* $\theta^\infty[T_{\theta h} = \infty] = 1$ *for some real* $h > 1$ *and some* $\theta \in \Theta$ *then there is no likelihood ratio discrimination.*

The main result is a characterization theorem.

(22) THEOREM. Θ *is likelihood ratio discriminable if and only if for every* $\theta \in \Theta$ *there is a real* $K(\theta) > 1$ *such that* $\theta^\infty[T_{\theta K(\theta)} < \infty] = 1$.

PROOF. Lemmas (14), (17), and Theorem (20) demonstrate the necessity; Lemma (14) and inequality (13) the sufficiency. □

REMARK. Of course, if some function K works, then any other with values greater than one will too. It is an open question, however, whether likelihood ratio procedures which surely stop for all $\theta \in \Theta$ imply the discriminability of Θ.

Here are three direct consequences of (22).

(23) COROLLARY. *Any finite* Θ *is likelihood ratio discriminable.*

PROOF. By (1) $R_{\theta\sigma N} \uparrow \infty$ with θ^∞-probability one. The finiteness of Θ implies $\min_{\sigma \in \Theta - \{\theta\}} R_{\theta\sigma N} \uparrow \infty$ also. □

REMARK. As is well known, this result can be obtained by a direct appeal to the law of large numbers.

(24) COROLLARY. Θ *is likelihood ratio discriminable if for every* $\theta \in \Theta$ *there is a natural number* N *and a set* $A \in \mathscr{G}^{(N)}$ *with* $\theta^\infty[A] > 0$ *and* $\sigma^\infty[A] = 0$ *for all* $\sigma \in \Theta - \{\theta\}$.

PROOF. Fix θ, N, A. For all $\sigma \neq \theta$, $f_\sigma(\omega_1, \cdots, \omega_N) = 0$ on A, which implies $R_{\theta N}(\omega) = \infty$ on A. The events $[S_K(\omega) \in A]$ for $K = 0, N, 2N, \cdots$ are independent and equally probable. They have positive probability. One such event is sure to occur, so

$$\theta^\infty[T_{\theta_2} < \infty] \geqq \theta^\infty[R_{\theta K} = \infty \quad \text{some} \quad K = 1, 2, \cdots] = 1. \ \square$$

(25) COROLLARY. *Fix* $\theta \in \Theta$. *Let* C_N *be the weak closure of* $[\sigma^N \mid \sigma \in \Theta$ and $\sigma \neq \theta]$. *If* $\theta^N \in C_N$ *for all* $N = 1, 2, \cdots$ *then likelihood ratio discrimination does not exist.*

PROOF. Fix real $h > 1$ and positive integer N. $[T_{\theta h} = N] \subset [R_{\theta\sigma N} \geqq h]$ for all $\sigma \neq \theta$. So

$$h \cdot \sigma^\infty[T_{\theta h} = N] = \int_{[T_{\theta h} = N]} h \cdot f_\sigma(\omega_1, \cdots, \omega_N)\, \partial\mu^\infty \mid \mathscr{G}^{(N)}$$
$$\leqq \int_{[T_{\theta h} = N]} f_\theta(\omega_1, \cdots, \omega_N)\, \partial\mu^\infty \mid \mathscr{G}^{(N)} = \theta^\infty[T_{\theta h} = N].$$

Weak closure implies $\sup_{\sigma \neq \theta} \sigma^\infty[A] \geqq \theta^\infty[A]$ for all $A \in \mathscr{G}^{(N)}$. So

$$\theta^\infty[T_{\theta h} = N] = 0 \quad \text{and} \quad \theta^\infty[T_{\theta h} < \infty] = 0. \quad \square$$

REMARK. In this last case a stronger result is true: no discrimination is possible by any means. This fact is a generalization of Theorem 4 in Freedman to rules measurable on the σ-fields of $X_1, \cdots, X_N$ for $N = 1, 2, \cdots$.

Can every discriminable family be separated by likelihood ratios? Example 1 shows that the answer is no. In fact $\theta^\infty[T_{\theta h} = \infty] = 1$ for all $\theta \in \Theta$ and $h > 1$.

Example 2 is based on a theorem by LeCam. It exhibits a likelihood ratio discriminable Θ which has one element in the weak closure of the rest. The key is that in two dimensions all the probabilities are weakly separated. This peculiar situation is not well understood, but it does show that separation in the weak topology is not equivalent to discriminability.

When likelihood ratios do discriminate, what can be said about the stopping time? Since $\tau_K \leqq T_{\theta K(\theta)}$ by (11), we first look to $T_{\theta h}$ and show

(26) LEMMA. *For any real $h > g > 1$, $E_\theta[T_{\theta g}] < \infty$ if and only if $E_\theta[T_{\theta h}] < \infty$.*

PROOF. $T_{\theta x}$ is monotone in x which shows the necessity. So suppose $E_\theta[T_{\theta g}] < \infty$ for some g. Then $E_\theta[T_{\theta 1}] < \infty$ and there is an N and p with $\theta^\infty[T_{\theta h} < N] = p > 0$. Define $Y_0, Y_1, \cdots$ by (18). Then by (19) and (16)

$$E_\theta[T_{\theta h}] \leqq E_\theta Y_1/(1-p) \leqq (E_\theta[T_{\theta 1}+N])/(1-p) < \infty. \square$$

Thus if $T_{\theta h}$ has finite expectation, you may change the test levels. This lemma coupled with Wald's identity gives a result which covers many common families.

(27) THEOREM. *Suppose there is an N such that $\infty > E_{\theta N}[\log R_{\theta N}] > 0$. Then $E_\theta[T_{\theta h}] < \infty$ for all $h > 1$. In particular this applies to the normal family $\Theta = [N(u_I, 1), I = 1, 2, \cdots]$ in the discriminable case that all u_I are isolated.*

PROOF. By (2) and Lemma (26) it is sufficient to show that if $Y_1, Y_2, \cdots$ are independent and distributed as $\log R_{\theta N}$, then for some $h, \tau_h = \inf[M \mid \sum_{I=1}^{M} Y_I > h]$ has finite expectation. But if $E\tau_0 < \infty$ then $E\tau_h < \infty$; and for any independent and identically distributed $Y_1, Y_2 \cdots$, if $\infty > EY_1 > 0$, then $E\tau_0 < \infty$ ([3] page 380). $\square$

REMARK. The finiteness of $E_{\theta N}[\log R_{\theta N}]$ is not necessary.

Any more general result must take into account Example 3. There is a family which is likelihood ratio discriminable and which has infinite expected stopping time under sampling from one of its members.

5. Examples.

EXAMPLE 1. *A discriminable family which is not likelihood ratio discriminable.* All probabilities in $\Theta = [\theta_{IM}; I = 2, 3, \cdots; M = 1, 2, \cdots, 2^I]$ are defined on [0, 1] with respect to Lebesgue measure λ. Let $A_{IM} = [(M-1)/2^I, M/2^I]$. Then

$$f_{\theta_{IM}}(x) = \tfrac{1}{2} + 2^{I-1} \qquad \text{for } x \in A_{IM}$$
$$= \tfrac{1}{2} \qquad \text{otherwise.}$$

To prove discriminability we show each element of Θ is weakly isolated. Fix I and M. Let

$$\beta = [A_{IK}, K = 1, 2, \cdots, 2^I \quad \text{and} \quad K \neq M;\ A_{I+1,2M}; \quad \text{and} \quad A_{I+1,2M+1}].$$

For any $\sigma \neq \theta_{IM}$ a direct computation yields $\sup_{B \in \beta} |\theta_{IM}^\infty[B] - \sigma^\infty[B]| \geqq \frac{1}{4}$.

However, Θ is not likelihood ratio discriminable. Fix N. Now fix K and consider $R_{\theta_{21}\theta_{KL}N}$ for $L = 1, \cdots, 2^K$. Since all densities are at least $\frac{1}{2}$ everywhere, each factor in the ratio is at most three. But for any value of X_1 and any K there is an L for which $f_{\theta_{KL}}$ is higher than $f_{\theta_{21}}$. In fact as $K \to \infty$ $\inf_{L=1,2,\cdots,2^K}[R_{\theta_{21}\theta_{KL}N}] \to 0$. So $R_{\theta_{21}N} = 0$ with θ_{21}^{∞}-probability one.

Here the failure of likelihood ratios is complete. For any N, I, and M we have $\theta_{IM}^{\infty}[R_{\theta_{IM}N} = 0] = 1$.

REMARK. The family of densities given by Kraft ([6] page 132) provides another example in which likelihood ratios fail completely.

EXAMPLE 2. *A likelihood ratio discriminable family with one element in the weak closure of the rest.* This example rests on a little understood property of weak closure. LeCam [8] proves the existence of a countable family $V = [v_N]$ with continuous densities defined on [0, 1] whose weak closure includes λ. On the square λ^2 is separated from $[v^2 \mid v \in V]$. V need not be discriminable. The following construction copes with this possibility.

Let $B_N = [(\frac{1}{2})^{N+1}, 1]$ and let θ_0 be λ. For $N = 1, 2, \cdots$ define θ_N to be zero on $[0, (\frac{1}{2})^{N+1})$ except for mass $(\frac{1}{2})^{N+1}$ on the midpoint $(\frac{1}{2})^{N+2}$. On B_N let

$$\theta_N[A] = v_N[A \cap B_N] \cdot \lambda[B_N]/v_N[B_N].$$

That θ_0 is in the weak closure of $(\theta_1, \theta_2, \cdots)$ is a direct consequence of the fact that by construction $\theta_N - v_N$ converges uniformly to zero on all measurable sets.

Corollary (24) shows likelihood ratio discriminability for $\Theta = [\theta_N, N = 1, 2, \cdots]$. For $N = 1, 2 \cdots$ θ_N puts positive mass on the point $(\frac{1}{2})^{N+2}$ while the others, having continuous densities there, do not. For θ_0 we must look for a separation on the unit square. Here, by construction [8], there is an A with $\theta_0{}^2[A] \geqq \frac{3}{4}$ but $v_N{}^2[A] = 0$ for all other N. Let $D = A \cap [\frac{2}{3}, 1]^2$. Then $\theta_N{}^2[D] = 0$ for $N = 1, 2, \cdots$ but $\lambda^2[D] > 0$.

EXAMPLE 3. *A likelihood ratio discriminable family with $E_{\theta_0}\tau_K = \infty$.* All elements of $\Theta = [\theta_0, \theta_1, \cdots]$ have densities with respect to Lebesgue measure. The first, θ_0, is λ on [0, 1]. Let p, a, and b be real numbers to be defined below. Let δ_N be the Nth binary fractional digit of the real number x. Define θ_N, $N \geqq 1$ by

$$\begin{aligned} f_{\theta_N}(x) &= a \cdot \delta_N(x) & & x \in [0, 1-p) \\ &= 1/a & & x \in [1-p, 1] \\ &= b_N & & x \in [N, N+1] \\ &= 0 & & \text{otherwise.} \end{aligned}$$

For every $p \in (0, 1)$ there are real numbers $a > 1$ and $b_N > b > 0$ for $N = 1, 2, \cdots$ such that the θ's are probabilities.

First we must show that Θ is likelihood ratio discriminable when $p \geqq \frac{1}{2}$. For $M \neq N$, $\theta_M{}^{\infty}[(N, N+1) = 0$ while $\theta_N{}^{\infty}[(N, N+1)] = b_N > 0$. The proof used for Corollary (24) gives $\theta_N{}^{\infty}[T_{\theta_N 2} < \infty] = 1$ for $N = 1, 2, \cdots$. To handle θ_0 we need

FACT. *Under* θ_0^∞ *the sequence* $\log R_{\theta_0 N}$, $N = 1, 2, \cdots$ *is a random walk with jumps of* $\pm$ log(a) *having respective probabilities* p *and* $1-p$.

PROOF. Let Y_i equal a when $0 \leqq X_i < 1-p$ and equal $1/a$ when $1-p \leqq X_i \leqq 1$. Let J be any finite set of integers and suppose $0 \leqq X_i < 1-p$ for all $i \in J$. With θ_0^∞-probability one there is an $M = M(X_i, i \in J)$ with $f_{\theta_M}(X_i) = a$ for all $i \in J$. For each possible J, condition on $[X_i < 1-p$ if and only if $i \in J]$ to get $\theta_0^\infty[R_{\theta_0 N} = \prod_{i=1}^N Y_i] = 1$ for $N = 1, 2, \cdots$. □

For $p \geqq \frac{1}{2}$ the random walk eventually reaches any positive height, so $\theta_0^\infty[T_{\theta_0 2} < \infty] = 1$ and Theorem (22) says that Θ is likelihood ratio discriminable.

Now we show that for $p = \frac{1}{2}$ and any function K from Θ to $(1, \infty)$ we have $E_{\theta_0}\tau_{\theta_0 K} = \infty$. For $M = 1, 2, \cdots$ $\theta_0^\infty[R_{\theta_M N} = 1$ for $N = 1, 2, \cdots] = 1$ so $\theta_0^\infty[T_{\theta_N 2} < \infty] = 0$ which implies $\theta_0^\infty[D_K \neq \theta_0] = 0$ and $\tau_K = T_{\theta_0 K(\theta_0)}$ with θ_0^∞-probability one. But for $p = \frac{1}{2}$ the expected time for a coin walk to become positive is infinite.

Acknowledgment. I owe special thanks to my thesis advisor David Freedman for patient and careful criticisms and also to Lucien LeCam for many useful conversations.

REFERENCES

[1] ARMITAGE, P. (1950). Sequential analysis with more than two alternative hypotheses, and its relation to discriminant function analysis. *J. Roy. Statist. Soc. Ser. B.* **12** 137–144.

[2] BARNARD, G. A. (1947). Book review. *J. Amer. Statist. Assoc.* **42** 658–664.

[3] FELLER, W. (1966). *An Introduction to Probability Theory and its Applications* **2**. Wiley, New York.

[4] FREEDMAN, D. A. (1967). A remark on sequential discrimination. *Ann. Math. Statist.* **38** 1666–1670.

[5] HOEFFDING, W. and WOLFOWITZ, J. (1958). Distinguishability of sets of distributions. *Ann. Math. Statist.* **29** 700–718.

[6] KRAFT, C. (1955). Some conditions for consistency and uniform consistency of statistical procedures. *Univ. California Publ. Statist.* **2** 125–142.

[7] LECAM, L. and SCHWARTZ, L. (1960). A necessary and sufficient condition for the existence of consitent estimates. *Ann. Math. Statist.* **31** 140–150.

[8] LECAM, L. (1970). On the weak convergence of probability measures. *Ann. Math. Statist.* **41** 621–625.