The Annals of Statistics
1973, Vol. 1, No. 5, 838-850

ON SEQUENTIAL DISTINGUISHABILITY[1]

BY RASUL A. KHAN

Mathematics Research Center, University of Wisconsin

Let $X_1, X_2, \cdots$ be a sequence of independent and identically distributed random variables governed by an unknown member of a countable family $\mathscr{P} = \{P_\theta : \theta \in \Omega\}$ of probability measures. The family $\mathscr{P}$ is said to be sequentially distinguishable if for any ε $(0 < \varepsilon < 1)$ there exist a stopping time t and a terminal decision function $\delta(X_1, \cdots, X_t)$ such that $P_\theta\{t < \infty\} = 1$ $\forall \theta \in \Omega$ and $\sup_{\theta\in\Omega} P_\theta(\delta(X_1, \cdots, X_t) \neq \theta) \leqq \varepsilon$. Robbins [12] defined a general stopping time (see Section 2) as an approach to this problem. This paper is a study of this stopping time with applications to some exponential distributions.

1. Introduction. Let $\mathscr{P} = \{P_\theta : \theta \in \Omega\}$ be a countable family of probability measures defined on some fixed probability space. We are observing sequentially a sequence of random variables $X_1, X_2, \cdots$ asumed to be governed by some unknown member of the family $\mathscr{P}$. We want to stop at some finite stage and decide in favor of a member of the family $\mathscr{P}$ with a uniformly small probability of error. A sequential procedure (t, δ) consists of a stopping rule t and a terminal decision function $\delta(X_1, X_2, \cdots, X_t)$. The following definition is due to Robbins [12].

DEFINITION. The family $\mathscr{P}$ is said to be sequentially distinguishable if for any given ε $(0 < \varepsilon < 1)$ there exist a stopping rule t and a terminal decision function $\delta(X_1, \cdots, X_t)$ such that $P_\theta(t < \infty) = 1$ $\forall \theta \in \Omega$ and $P_\theta(\delta(X_1, \cdots, X_t) \neq \theta) \leqq \varepsilon$ uniformly in θ.

Motivated by Wald's sequential probability ratio test (SPRT) Robbins [12] defined a general stopping time (to be introduced in Section 2) which gives a uniform bound for the probabilities of error. He used this stopping time for estimating an integer mean of a normal distribution and gave several interesting results. This work is devoted to the study of this stopping time with emphasis on its applications to the sequential distinguishability problems for the exponential distributions.

2. A general stopping time of Robbins. With no loss of generality we can assume the existence of a countable family of probability densities $\{f_\theta(x) : \theta \in \Omega\}$ with respect to some σ-finite measure μ. Let μ_n denote the μ-measure in n

Received September 1971; revised November 1972.

[1] This paper is a part of the author's doctoral thesis completed at Columbia University under the guidance of Professor Herbert Robbins. It was sponsored by the United States Army under Contract No.: DA-31-124-ARO-D-462.

AMS 1970 *subject classifications*. Primary 62L10; Secondary 62L99.

Key words and phrases. Sequential distinguishability, countable family, stopping rule, sequential probability ratio test, optimality, Kullback-Leibler information measure, asymptotic optimality.

dimensions and $f_{\theta,n}$ the joint probability density function of $(X_1, X_2, \cdots, X_n)$ with respect to μ_n (for every $n \geqq 1$). For notational convenience we write $f_{i,n} = f_{\theta_i,n}$ and $P_i = P_{\theta_i}$, etc. In what follows we shall take a doubly indexed sequence of constants $\{a_{ij}\}$ such that $a_{ij} > 1$ and $\sum_{i\neq j} a_{ij}^{-1} \leqq \varepsilon$ for a given $0 < \varepsilon < 1$ and every j. Define the stopping time:

$$N = \inf\{n \geqq 1 : f_{i,n} \geqq \sup_{j\neq i} a_{ij} f_{j,n} \text{ for some } i\} \\ = \infty \quad \text{if no such } n, \tag{2.1}$$

and assert $\theta_i(P_i)$ if N stops with i, i.e. $\delta(X_1, \cdots, X_N) = \theta_i$ if N stops with i. Writing a_i = accept θ_i, and assuming N terminates, it follows that

$$\begin{aligned} P_j(e) = P_j(\text{error}) &= \textstyle\sum_{i\neq j} P_j(a_i) = \sum_{i\neq j} \sum_{n=1}^{\infty} \int_{\{N=n, a_i\}} f_{j,n}\, d\mu_n \\ &= \textstyle\sum_{i\neq j} \sum_{n=1}^{\infty} \int_{\{N=n, a_i\}} (f_{j,n}/f_{i,n}) f_{i,n}\, d\mu_n \\ &\leqq \textstyle\sum_{i\neq j} a_{ij}^{-1} P_i(a_i) \leqq \varepsilon \quad \forall j, \end{aligned}$$

and thus

$$\sup_i P_i(e) \leqq \varepsilon. \tag{2.2}$$

Thus the stopping time N does provide a uniform bound on the probabilities of error. But the first question at issue is: under what conditions is the stopping time N a bonafide stopping rule under all P_i? Sections 3 and 4 are devoted to this problem.

3. Preliminaries and necessary conditions for termination.

3.1. *Preliminaries*. There are certain measures of divergence between distributions which play an essential role throughout the paper. From now on it is assumed without explicit mention that we are dealing with i.i.d. (independent and identically distributed) sequence of random variables governed by a member of the family $\mathscr{P}$. Define the following:

(a) $$\lambda(i, j) = \int f_i(x) \log(f_i(x)/f_j(x))\, d\mu$$

(b) $$\rho(i, j) = \int [f_i(x) f_j(x)]^{\frac{1}{2}}\, d\mu \quad \text{and}$$

(c) $$D(i, j) = \int |f_i(x) - f_j(x)|\, d\mu.$$

The measure $\lambda(\cdot, \cdot)$ is usually called the Kullback–Leibler information measure. The measure $\rho(\cdot, \cdot)$ was introduced by Hellinger and is frequently used in probabilistic and statistical contexts (see [8], [2], and [9]). Finally, the measure $D(\cdot, \cdot)$ is a well-known metric. We refer to Kraft [9] and Kullback [10] for these information numbers.

It is obvious that $0 \leqq \rho(i, j) \leqq 1$, with equality at the respective extremes according as $P_i \perp P_j$ (orthogonality) or $P_i = P_j$. An application of Jensen's inequality shows that $\lambda(i, j) \geqq 0$, with equality only if $f_i(x) = f_j(x)$ a.s. μ. The relation between ρ, D and λ is given by the following lemma.

LEMMA 1. $2(1 - \rho^2(i, j))^{\frac{1}{2}} \geqq D(i, j) \geqq 2(1 - \rho(i, j))$. *Further*, $\exp(-\frac{1}{2}\lambda(i, j)) \leqq \rho(i, j) \leqq 1$, *so that* $\exp[-\frac{1}{2}\inf_{j\neq i} \lambda(i, j)] \leqq \sup_{j\neq i} \rho(i, j) \leqq 1$. *Moreover*, $\sup_{j\neq i} \rho(i, j) < 1$ *implies that* $\inf_{j\neq i} \lambda(i, j) > 0$.

PROOF. The topmost inequalities follow from

$$[\int |f_i - f_j|\, d\mu]^2 \leqq \int |(f_i)^{\frac{1}{2}} - (f_j)^{\frac{1}{2}}|^2\, d\mu \int |(f_i)^{\frac{1}{2}} + (f_j)^{\frac{1}{2}}|^2\, d\mu = 2(1 - \rho^2(i, j)),$$

and

$$\int |f_i - f_j|\, d\mu \geqq \int |(f_i)^{\frac{1}{2}} - (f_j)^{\frac{1}{2}}|^2\, d\mu = 2(1 - \rho(i, j)).$$

The rest follows from

$$\begin{aligned}\rho(i, j) = E_i[(f_j(X)/f_i(X))]^{\frac{1}{2}} &= \exp\{\log E_i[(f_j(X)/f_i(X))]^{\frac{1}{2}} \\ &\geqq \exp\{E_i \log [(f_j(X)/f_i(X))]^{\frac{1}{2}}\} = \exp(-\tfrac{1}{2}\lambda(i, j)).\end{aligned}$$

3.2. *Necessary conditions for termination.* We define the following conditions:

(i) $\inf_{j\neq i} \lambda(i, j) > 0\ \forall i$.
(ii) $\sup_{j\neq i} \rho(i, j) < 1\ \forall i$.
(iii) $\inf_{j\neq i} D(i, j) > 0\ \forall i$.

It is easy to see that (ii) $\Leftrightarrow$ (iii) and (ii) $\Rightarrow$ (i). That these conditions are relevant to the stopping time (2.1) is given by the following lemma.

LEMMA 2. *If any of the conditions* (i), (ii) *or* (iii) *fails for all* i, *then* $P_i(N = \infty) = 1$, *so that these conditions are necessary for* $P_i(N < \infty) = 1$.

PROOF. It is enough to prove only for (ii). If (ii) fails for some i, then there exists an infinite subsequence $\{\rho(i, j_k)\}$ such that $j_k \neq i$ and $\rho(i, j_k) \to 1$ as $k \to \infty$. But this $\Rightarrow D(i, j_k) \to 0 \Rightarrow f_{j_k}(x) \to_{\mu} f_i(x) \Rightarrow f_{j_k}(X) \to_{P_i} f_i(X) \Rightarrow f_{j_{k_l}}(X) \to_{\text{a.s. } P_i} f_i(X)$. For notational convenience we denote by $f_{j_k}(X)$ the preceding resulting subsequence. Therefore it follows that

$$f_{j_k,n} = \prod_{\alpha=1}^{n} f_{j_k}(X_\alpha) \to_{\text{a.s. } P_i} f_{i,n} = \prod_{\alpha=1}^{n} f_i(X_\alpha) \qquad \forall n \geqq 1.$$

Hence we have

$$\sup_{j\neq i} (f_{j,n}/f_{i,n}) \geqq \sup_{j_k\neq i} (f_{j_k,n}/f_{i,n}) \geqq 1 \qquad \text{a.s.} \quad P_i \qquad \forall n \geqq 1.$$

But the fact that $a_{ij} > 1$ and $\sum_{i\neq j} a_{ij}^{-1} \leqq \varepsilon\ (0 < \varepsilon < 1) \Rightarrow \inf_{j\neq i} a_{ij} > 1$. Thus we have

$$\begin{aligned}\sup_{j\neq i} (a_{ij} f_{j,n}/f_{i,n}) &\geqq \sup_{j\neq i} (\inf_{j\neq i} a_{ij})(f_{j,n}/f_{i,n}) \\ &> \sup_{j\neq i}(f_{j,n}/f_{i,n}) \geqq 1 \qquad \text{a.s.} \quad P_i.\end{aligned}$$

Hence $\forall$ fixed $1 \leqq n < \infty$, we have

$$\sup_{j\neq i} (a_{ij} f_{j,n}/f_{i,n}) > 1 \qquad \text{a.s.} \quad P_i,$$

so that $P_i(N = \infty) = 1$ by the very definition of N.

REMARK. Lemma 2 entails that $\mathscr{P}$ is not sequentially distinguishable through N if the above conditions do not hold.

4. Sufficient conditions for termination. We recall the definition

$$\begin{aligned} N &= \inf\{n \geqq 1 : f_{i,n} \geqq \sup_{j\neq i} a_{ij} f_{j,n} \text{ for some } i\} \\ &= \infty \qquad \text{if no such } n,\end{aligned} \tag{4.1}$$

where $f_{i,n} = \prod_{\alpha=1}^{n} f_i(X_\alpha)$.

Clearly, the stopping time (4.1) terminates with probability one if

$$P_i\left\{\liminf_{n\to\infty} \sup_{j\neq i} a_{ij}\frac{f_{j,n}}{f_{i,n}} = 0\right\} = 1 \quad \forall i\,. \tag{4.2}$$

Two natural conditions which ensure (4.2) are given by the following theorem.

THEOREM 1. *If*

(a) $$P_i(\inf_{n\geqq 1}\sup_{j\neq i} a_{ij}f_{j,n}/f_{i,n} < \infty) = 1 \quad \forall i$$ *and*

(b) $$P_i(\inf_{n\geqq 1}\sup_{j\neq i} f_{j,n}/f_{i,n} < 1) = 1 \quad \forall i\,,$$

then (4.2) *holds, and hence*

$$P_i(N < \infty) = 1 \quad \forall i\,.$$

PROOF. Define the following stopping times:

$$\begin{aligned}
k_0 &= \inf\{n \geqq 1 : \sup_{j\neq i} a_{ij}\textstyle\prod_{\alpha=1}^{n} f_j(X_\alpha)/f_i(X_\alpha) < \infty\}\,,\\
k_1 &= \inf\{n \geqq 1 : \sup_{j\neq i}\textstyle\prod_{\alpha=k_0+1}^{n+k_0} f_j(X_\alpha)/f_i(X_\alpha) < 1\}\,,\\
k_2 &= \inf\{n \geqq 1 : \sup_{j\neq i}\textstyle\prod_{\alpha=k_1+k_0+1}^{n+k_1+k_0} f_j(X_\alpha)/f_i(X_\alpha) < 1\}, \cdots,\\
k_m &= \inf\{n \geqq 1 : \sup_{j\neq i}\textstyle\prod_{\alpha=k_{m-1}+\cdots+k_0+1}^{n+k_{m-1}+\cdots+k_0} f_j(X_\alpha)/f_i(X_\alpha) < 1\}\,.
\end{aligned}$$

Condition (a) implies $P_i(k_0 < \infty) = 1$. Now we have

$$\begin{aligned}
P_i(k_1 < \infty) &= P_i(\sup_{j\neq i}\textstyle\prod_{\alpha=k_0+1}^{n+k_0} f_j(X_\alpha)/f_i(X_\alpha) < 1 \quad\text{for some } n \geqq 1)\\
&= \textstyle\sum_{k'=1}^{\infty} P_i(\sup_{j\neq i}\prod_{\alpha=k'+1}^{n+k'} f_j(X_\alpha)/f_i(X_\alpha) < 1\\
&\qquad \text{for some } n \geqq 1 \mid k_0 = k')P_i(k_0 = k')\\
&= \textstyle\sum_{k'=1}^{\infty} P_i(\sup_{j\neq i}\prod_{\alpha=1}^{n} f_j(X_\alpha)/f_i(X_\alpha) < 1\\
&\qquad \text{for some } n \geqq 1)P_i(k_0 = k')\\
&= \textstyle\sum_{k'=1}^{\infty} P_i(k_0 = k') = P_i(k_0 < \infty) = 1\,.
\end{aligned}$$

In general, $P_i(k_m < \infty) = 1$, $m = 0, 1, 2, \cdots$ follows from the fact that $k_1, k_2, \cdots$ are i.i.d. Now set $N_m = k_0 + k_1 + \cdots + k_m$, and write

$$\begin{aligned}
&a_{ij}\textstyle\prod_{\alpha=1}^{N_m}(f_j(X_\alpha)/f_i(X_\alpha))\\
&\quad = (a_{ij}\textstyle\prod_{\alpha=1}^{k_0} f_j(X_\alpha)/f_i(X_\alpha)(\prod_{\alpha=k_0+1}^{k_1+k_0} f_j(X_\alpha)/f_i(X_\alpha))\cdots\\
&\qquad \times (\textstyle\prod_{\alpha=k_{m-1}+\cdots+k_0+1}^{k_m+\cdots+k_0} f_j(X_\alpha)/f_i(X_\alpha))\,.
\end{aligned}$$

Therefore,

$$\sup_{j\neq i} a_{ij}\textstyle\prod_{\alpha=1}^{N_m} f_j(X_\alpha)/f_i(X_\alpha) \leqq C(k_0)\prod_{r=1}^{m} Y_r$$

where

$$\begin{aligned}
C(k_0) &= \sup_{j\neq i} a_{ij}\textstyle\prod_{\alpha=1}^{k_0} f_j(X_\alpha)/f_i(X_\alpha)\,,\\
Y_1 &= \sup_{j\neq i}\textstyle\prod_{\alpha=k_0+1}^{k_1+k_0} f_j(X_\alpha)/f_i(X_\alpha), \cdots,\\
Y_m &= \sup_{j\neq i}\textstyle\prod_{\alpha=k_{m-1}+\cdots+k_0+1}^{k_m+\cdots+k_0} f_j(X_\alpha)/f_i(X_\alpha)\,.
\end{aligned}$$

It is easy to show that $Y_1, Y_2, \cdots$ are i.i.d. random variables, and $P_i(C(k_0) < \infty) = 1$ $\forall i$, $r = 1, 2, \cdots$. Moreover, $N_m \to \infty$ (a.s. P_i) as $m \to \infty$. To show that $\lim_{m\to\infty} \sup_{j\neq i} a_{ij} \prod_{\alpha=1}^{Nm} f_j(X_\alpha)/f_i(X_\alpha) = 0$ a.s. P_i, it is enough to show that

$$P_i(\lim_{m\to\infty} \sum_{r=1}^{m} \log Y_r = -\infty) = 1 \quad \forall i . \tag{4.3}$$

But since $E_i \log Y_1 < \log (E_i Y_1) < \log 1 = 0$, (4.3) follows from the strong law of large numbers. Thus we have

$$P_i\left(\liminf_{n\to\infty} \sup_{j\neq i} a_{ij} \frac{f_{j,n}}{f_{i,n}} = 0\right) = 1 \quad \forall i ,$$

which completes the proof of the theorem.

Condition (a) is necessary for termination since if it fails the rule never terminates with positive probability. We now give certain conditions which ensure condition (b) of Theorem 1. Define the stopping time:

$$\begin{aligned} t_i &= \inf \{n \geqq 1 : f_{i,n} > \sup_{j\neq i} f_{j,n}\} \\ &= \infty \qquad \text{if no such } n . \end{aligned}$$

Clearly, condition (b) holds if $P_i(t_i < \infty) = 1\ \forall i$. To this end, we define

$$\varphi_{i,r}(x) = \sup_{j:|j-i|>r} f_j(x) , \qquad r > 0 ,$$

and

$$\begin{aligned} \varphi_{i,r}^*(x) &= \varphi_{i,r}(x) \qquad \text{if} \quad \varphi_{i,r}(x) > 1 \\ &= 1 \qquad \text{otherwise} . \end{aligned}$$

We assume

M_1. For sufficiently large r, $E_i \log \varphi_{i,r}^*(X) < \infty\ \forall i$.
M_2. $f_i(x) \neq f_j(x)$ a.s. $\forall i \neq j$.
M_3. $\inf_{j\neq i} D(i, j) > 0\ \forall i$.
M_4. $\lim_{|j|\to\infty} f_j(x) = 0$ a.s.
M_5. $E_i|\log f_i(X)| < \infty\ \forall i$.

THEOREM 2. *Under the above assumptions,*

$$P_i\left(\lim_{n\to\infty} \sup_{j\neq i} \frac{f_{j,n}}{f_{i,n}} = 0\right) = 1 \quad \forall i ,$$

and hence

$$P_i(t_i < \infty) = 1 \quad \forall i .$$

PROOF. The proof is an adaptation of the ideas of Wald [14] and hence we omit the details. First, it is not hard to show that

$$\lim_{r\to\infty} E_i \log \varphi_{i,r}(X) = -\infty \quad \forall i . \tag{4.4}$$

Hence there exists a positive integer r such that

$$E_i \log \varphi_{i,r}(X) < E_i \log f_i(X) . \tag{4.5}$$

Since

$$\sup_{j\neq i} (f_{j,n}/f_{i,n}) \leqq \sum_{j\neq i:|j-i|\leqq r} \prod_{\alpha=1}^{n} f_j(X_\alpha) f_i^{-1}(X_\alpha) + \prod_{\alpha=1}^{n} \phi_{i,r}(X_\alpha) f_i^{-1}(X_\alpha) ,$$

hence the conclusion follows from the above inequality and the strong law of large numbers.

5. Moments and the concept of asymptotic optimality.

5.1. *A crude estimate of* $P_i(N > n)$. Recall that $\rho(i, j) = \int (f_i f_j)^{\frac{1}{2}} d\mu$, and $\rho_n(i, j) = \int_{R_n} (f_{i,n} f_{j,n})^{\frac{1}{2}} d\mu_n = \rho^n(i, j)$. It follows from (4.1) that

$$\begin{aligned} P_i(N > n) &\leqq P_i(\sup_{j \neq i} a_{ij} f_{j,n} > f_{i,n}) \\ &\leqq \textstyle\sum_{j \neq i} P_i\{(f_{j,n}/f_{i,n})^t > a_{ij}^{-t}\} \\ &\leqq \textstyle\sum_{j \neq i} a_{ij}^t \rho_t^n(i, j) , \qquad 0 < t < 1 , \end{aligned}$$

and hence

$$(5.1) \qquad P_i(N > n) \leqq \textstyle\sum_{j \neq i} \inf_{0<t<1} a_{ij}^t \rho_t^n(i, j) \leqq \sum_{j \neq i} \inf_{0<t<1} a_{ij}^t \rho_t(i, j) .$$

where $\rho_t(i, j) = \int f_i^{1-t}(x) f_j^t(x) \, d\mu$ $(0 < t < 1)$. In particular, we have

$$(5.2) \qquad P_i(N > n) \leqq \textstyle\sum_{j \neq i} (a_{ij})^{\frac{1}{2}} \rho^n(i, j) \leqq \sum_{j \neq i} (a_{ij})^{\frac{1}{2}} \rho(i, j) .$$

Inequalities (5.1) and (5.2) are crude. Some improvement is possible under two assumptions, namely (i) $\sup_{j \neq i} a_{ij} = C_i < \infty \ \forall i$, and (ii) $\overline{\Omega(i)}$ ('$\overline{}$' denoting closure) is bounded $\forall i$, where $\Omega(i) = \Omega - \{\theta_i\}$. We then have

$$P_{\theta_i}(N > n) \leqq P_{\theta_i}\{\sup_{\theta \in \overline{\Omega(i)}} (f_{\theta,n}/f_{\theta_i,n})^{\frac{1}{2}} \geqq C_i^{-\frac{1}{2}}\} .$$

Let $S(\theta_k, \varepsilon_k)$ denote open intervals with centers θ_k and radii $\varepsilon_k > 0$. Since $\overline{\Omega(i)}$ is compact, $\bigcup_{i=1}^h S(\theta_i, \varepsilon_i) \supset \overline{\Omega(i)}$ $(1 \leqq h < \infty)$. Thus

$$\begin{aligned} P_{\theta_i}(N > n) &\leqq P_{\theta_i}\{\sup_{\theta \in \bigcup_{k=1}^h S(\theta_k, \varepsilon_k)} (f_{\theta,n}/f_{\theta_i,n})^{\frac{1}{2}} > C_i^{-\frac{1}{2}}\} \\ &\leqq \textstyle\sum_{k=1}^h P_{\theta_i}\{\sup_{\theta \in S(\theta_k, \varepsilon_k)} f_{\theta,n}/f_{\theta_i,n})^{\frac{1}{2}} > C_i^{-\frac{1}{2}}\} ; \end{aligned}$$

$$(5.3) \qquad P_{\theta_i}(N > n) \leqq \textstyle\sum_{k=1}^h (C_i)^{\frac{1}{2}} \rho_{n,\varepsilon_k}(\theta_k, \theta_i) \leqq \sum_{k=1}^h (C_i)^{\frac{1}{2}} \rho_{1,\varepsilon_k}^n(\theta_k, \theta_i)$$

where $\rho_{n,\varepsilon_k}(\theta_k, \theta_i) = \int_{R^n} (\sup_{\theta \in S(\theta_k, \varepsilon_k)} f_{\theta,n} f_{\theta_i,n})^{\frac{1}{2}} d\mu_n$.

5.2. *Termination and the existence of moments of N.* It follows from (5.1) that if $\sum_{j \neq i} (a_{ij})^{\frac{1}{2}} \rho(i, j) < \infty \ \forall i$, then $P_i(N < \infty) = 1 \ \forall i$. Moreover, the convergence of $\sum_{j \neq i} (a_{ij})^{\frac{1}{2}} \rho(i, j)$ $(\forall i)$ entails that $\sup_{j \neq i} \rho(i, j) < 1 \ \forall i$. Using these facts we can prove the following.

THEOREM 3. *If* $\sum_{j \neq i} (a_{ij})^{\frac{1}{2}} \rho^n(i, j) < \infty \ \forall i$ *for some* $n \geqq 1$, *then* $P_i(N < \infty) = 1$ $\forall i$. *Moreover,* $E_i \exp(tN) < \infty$ *for some* $t > 0$ $(\forall i)$, *and hence* $E_i N^k < \infty$ $\forall k \geqq 1$.

PROOF. It is enough to show that there exist $C > 0$ and $0 \leqq \delta < 1$ (both may depend on i but not on n) such that

$$P_i(N > n) \leqq C\delta^n \quad \forall i .$$

Assuming the series converges for $n = m$, it follows that it converges for all $n \geqq m$. Hence for $n > m$, it follows from (5.2) that

$$\begin{aligned} P_i(N > n) &\leqq \textstyle\sum_{j \neq i} (a_{ij})^{\frac{1}{2}} \rho^n(i, j) \leqq \sum_{j \neq i} (a_{ij})^{\frac{1}{2}} \rho^m(i, j) (\sup_{j \neq i} \rho(i, j))^{n-m} \\ &\leqq \delta^{n-m} \textstyle\sum_{j \neq i} (a_{ij})^{\frac{1}{2}} \rho^m(i, j) , \quad \delta = \delta(i) = \sup_{j \neq i} \rho(i, j) \\ &\leqq \delta^n \delta^{-m} \textstyle\sum_{j \neq i} (a_{ij})^{\frac{1}{2}} \rho^m(i, j) , \qquad 0 < \delta < 1 , \end{aligned}$$

and hence

$$P_i(N > n) \leqq C\delta^n , \qquad 0 < C < \infty, 0 < \delta < 1 .$$

5.3. *Exponential class cases.* We do not have to depend upon the preceding crude inequalities for the exponential distributions. Let $f_\theta(x) = \exp(\theta x - b(\theta))$ be a probability function with respect to some σ-finite measure μ, and let Θ denote the natural parameter set. Further, let $\Omega = \{\theta_i : \cdots < \theta_{-1} < \theta_0 < \theta_1 < \cdots\} \subset \Theta$. Recall the notation $f_{i,n} = \exp(\theta_i S_n - nb(\theta_i))$, $S_n = X_1 + \cdots + X_n$, and $\lambda(i, j) = E_i \log(f_i(X)/f_j(X)) = (\theta_i - \theta_j)b'(\theta_i) - (b(\theta_i) - b(\theta_j))$. We have the following.

THEOREM 4. *Assume that* $\sup_{j\neq i}\{\log a_{ij}/\lambda(i, j)\} < \infty \ \forall i$, *and let* $m = k \sup_{j\neq i}\{\log a_{ij}/\lambda(i, j)\}$, $k > 1$. *Then there exists a number* ρ *such that*

$$P_i(N > m) \leqq 2\rho^m \quad \forall i , \qquad 0 \leqq \rho < 1 .$$

PROOF. Assume for simplicity that m is an integer. It follows from definition (4.1) that

$$(5.4) \qquad P_i(N > m) \leqq P_i(\sup_{j>i} a_{ij} f_{j,m} > f_{i,m}) + P_i(\sup_{j<i} a_{ij} f_{j,m} > f_{i,m}) .$$

Passing to logarithms and noting that $\lambda(i, j)/k > (\log a_{ij})/m$, (5.4) gives

$$(5.5) \qquad P_i(N > m) \leqq P_i\left(\sup_{j>i}\left\{k^{-1}\lambda(i, j) + m^{-1} \sum_{\alpha=1}^m \log \frac{f_j(X_\alpha)}{f_i(X_\alpha)}\right\} > 0\right) + P_i\left(\sup_{j<i}\left\{k^{-1}\lambda(i, j) + m^{-1} \sum_{\alpha=1}^m \log \frac{f_j(X_\alpha)}{f_i(X_\alpha)}\right\} > 0\right).$$

Now, $m^{-1}\sum_{\alpha=1}^m \log(f_j(X_\alpha)/f_i(X_\alpha)) = (\theta_j - \theta_i)m^{-1}S_m - (b(\theta_j) - b(\theta_i))$, and $\lambda(i, j) = (\theta_i - \theta_j)b'(\theta_i) - (b(\theta_i) - b(\theta_j))$. Further, note that $E_i X_1 = b'(\theta_i)$, $\sigma_{X_1}^2(\theta_i) = b''(\theta_i) > 0$ so that $b(\cdot)$ is a convex function with $b'(\cdot)$ strictly increasing. After making necessary computations, (5.5) gives

$$(5.6) \qquad P_i(N > m) \leqq P_i(S_m \geqq ma) + P_i(S_m \leqq cm)$$

where

$$a = k^{-1}b'(\theta_i) + (1 - k^{-1})(b(\theta_{i+1}) - b(\theta_i))/(\theta_{i+1} - \theta_i) ,$$

and

$$c = k^{-1}b'(\theta_i) + (1 - k^{-1})(b(\theta_i) - b(\theta_{i-1}))/(\theta_i - \theta_{i-1}) .$$

It is easy to show that $a > b'(\theta_i)$ and $c < b'(\theta_i)$. Hence from (5.6) and a theorem of Chernoff in [2] we obtain

$$P_i(N > m) \leqq [m(a)]^m + [m(c)]^m$$

where $m(\alpha) = \inf_t e^{-\alpha t}M(t)$, $M(t) = E_i e^{tX_1}$. It can easily be shown that $0 \leqq m(a)$, $m(c) < 1$, and hence

$$P_i(N > m) \leqq 2\rho^m , \quad \rho = \max(m(a), m(c)) .$$

COROLLARY. *If* $f_\theta(x) = \exp(\theta x - b(\theta))$, *then there is always a choice of* a_{ij} *for which the stopping time* N *is a bonafide stopping rule. Furthermore, the family*

$\mathscr{P} = \{P_\theta : \theta \in \Omega\}$ *where* P_θ *is the probability pertaining to* $f_\theta(x)$, *is sequentially distinguishable through* N iff *each* $\theta \in \Omega$ *is isolated in the usual topology of the real line* (*equivalently*, iff $\sup_{j \neq i} \rho(i, j) < 1 \ \forall i$).

5.4. *A lower bound for* $E_i N$ *and the concept of asymptotic optimality.*

LEMMA 3. *If* $0 < \inf_{j \neq i} \lambda(i, j) \leqq \lambda(i, j) < \infty \ \forall i$ *and* j, *then* $\forall i$

$$E_i N \geqq \sup_{j \neq i} \left[\frac{\log a_{ij}}{\lambda(i, j)} \right] \geqq \frac{-\log \varepsilon}{\inf_{j \neq i} \lambda(i, j)}, \tag{5.7}$$

where $\varepsilon (0 < \varepsilon < 1)$ *is the same as in* (2.2).

PROOF. We may assume that $E_i N < \infty \ \forall i$, since otherwise the lemma is trivial. It follows from the definition of N (4.1) that

$$\textstyle\sum_{\alpha=1}^{N} \log (f_i(X_\alpha)/f_j(X_\alpha)) \geqq \log a_{ij} \quad \forall j \neq i .$$

By Wald's lemma ([13], pages 170–171) we have

$$\lambda(i, j) E_i N \geqq \log a_{ij} \quad \forall j \neq i .$$

Hence

$$E_i N \geqq \sup_{j \neq i} \left[\frac{\log a_{ij}}{\lambda(i, j)} \right] \geqq \frac{-\log \varepsilon}{\inf_{j \neq i} \lambda(i, j)}, \qquad 0 < \varepsilon < 1 ,$$

where the last inequality follows from $\sum_{i \neq j} a_{ij}^{-1} \leqq \varepsilon$.

LEMMA 4. *Let* (t, δ) *be any procedure such that* $P_i(t < \infty) = 1 \ \forall i$, *and* $P_i(\text{error}) \leqq \varepsilon \ \forall i$. *If the assumption of Lemma* 3 *holds, then*

$$E_i t \geqq \frac{(1 - \varepsilon) \log (1 - \varepsilon)/\varepsilon + o(\log \varepsilon)}{\inf_{j \neq i} \lambda(i, j)} \quad \forall i .$$

PROOF. We may assume that $E_i t < \infty$. Set $S_t = \sum_{\alpha=1}^{t} \log (f_i(X_\alpha)/f_j(X_\alpha))$, $A_i = \{\text{accept } P_i\}$ and $B_i = \bigcup_{j \neq i} A_j$. Then $P_i(A_i) \geqq 1 - \varepsilon$, $P_i(B_i) \leqq \varepsilon$, and there is no loss in assuming $P_i(B_i) > 0$. Denoting by E_i^A the conditional expectation given A, using Wald's lemma and Jensen's inequality it follows that

$$\begin{aligned} \lambda(i, j) E_i t = E_i S_t &= P_i(A_i) E_i^{A_i} S_t + P_i(B_i) E_i^{B_i} S_t \\ &= P_i(A_i) E_i^{A_i} (-\log \textstyle\prod_{\alpha=1}^{t} (f_j(X_\alpha)/f_i(X_\alpha))) \\ &\quad + P_i(B_i) E_i^{B_i} (-\log \textstyle\prod_{\alpha=1}^{t} (f_j(X_\alpha)/f_i(X_\alpha))) \\ &\geqq P_i(A_i)(-\log E_i^{A_i} \textstyle\prod_{\alpha=1}^{t} (f_j(X_\alpha)/f_i(X_\alpha))) \\ &\quad + P_i(B_i)(-\log E_i^{B_i} \textstyle\prod_{\alpha=1}^{t} (f_j(X_\alpha)/f_i(X_\alpha))) . \end{aligned}$$

Now,

$$E_i^{A_i} \textstyle\prod_{\alpha=1}^{t} (f_j(X_\alpha)/f_i(X_\alpha)) = P_i^{-1}(A_i) \sum_{n=1}^{\infty} \int_{\{N=n, A_i\}} (f_{j,n}/f_{i,n}) f_{i,n} \, d\mu_n = \dfrac{P_j(A_i)}{P_i(A_i)} .$$

Similarly, $E_i^{B_i} \prod_{\alpha=1}^{t} (f_j(X_\alpha)/f_i(X_\alpha)) = P_j(B_i)/P_i(B_i)$. Thus

$$\begin{aligned} \lambda(i, j) E_i t &\geqq P_i(A_i) \log \{P_i(A_i)/P_j(A_i)\} + P_i(B_i) \log \{P_i(B_i)/P_j(B_i)\} \\ &\geqq (1 - \varepsilon) \log (1 - \varepsilon)/\varepsilon + P_i(B_i) \log \{P_i(B_i)/P_j(B_i)\} . \end{aligned}$$

Since $P_i(B_i) \leqq \varepsilon$, $P_j(B_i) \geqq P_j(A_j) \geqq 1-\varepsilon$, $P_i(B_i)\log\{P_i(B_i)/P_j(B_i)\} \leqq \varepsilon\log\varepsilon/(1-\varepsilon) \to 0$ as $\varepsilon \to 0$. The lemma now follows from the above inequalities.

The problem of sequential distinguishability for $\mathscr{P} = \{P_\theta : \theta \in \Omega\}$ where $\Omega = \{\theta_i\}$, can be considered in relation to a sequence of SPRT's. Let t_{ij} be the stopping rule of an SPRT ([13]) for $H_i : \theta = \theta_i$ against $H_j : \theta = \theta_j$ when error probabilities are $\alpha = \beta = \varepsilon$. Note that $E_i N \geqq E_i t_{ij}$, so that $E_i N \geqq \sup_{j\neq i} E_i t_{ij}$. The following definition is due to Robbins [12].

DEFINITION. The stopping rule N is said to be asymptotically optimal (a.o.) if $\lim_{\varepsilon\to 0} E_i N/\sup_{j\neq i} E_i t_{ij} = 1\ \forall i$.

The definition is interesting because of the known optimality of SPRT (see [15]). That the stopping rule (4.1) is a.o. (if the lower bound (5.7) is an asymptotic upper bound) is given by the following.

THEOREM 5. *The stopping rule N is* a.o. *if* $\forall i$,

$$E_i N \sim (-\log\varepsilon)/\inf_{j\neq i}\lambda(i,j) \qquad \text{as} \quad \varepsilon \to 0 .$$

PROOF. It is well known ([13]) that

$$E_i t_{ij} \geqq [(1-\varepsilon)\log(1-\varepsilon)/\varepsilon + \varepsilon\log\varepsilon/(1-\varepsilon)]/\lambda(i,j) \quad \forall j \neq i .$$

Therefore, $\sup_{j\neq i} E_i t_{ij} \geqq n(\varepsilon)/\inf_{j\neq i}\lambda(i,j)\ \forall i$, where $n(\varepsilon) = (1-\varepsilon)\log(1-\varepsilon)/\varepsilon + \varepsilon\log\varepsilon/(1-\varepsilon) \sim -\log\varepsilon$ as $\varepsilon \to 0$. Thus, if $E_i N \sim (-\log\varepsilon)/\inf_{j\neq i}\lambda(i,j)\ \forall i$, then $\limsup_{\varepsilon\to 0} E_i N/\sup_{j\neq i} E_i t_{ij} \leqq 1$. Since $E_i N \geqq \sup_{j\neq i} E_i t_{ij}$, there is limiting equality and hence N is asymptotically optimal.

6. Applications to some exponential distributions.

6.1. *Sequential distinguishability for the Poisson family.* Let P_λ denote the probability under which $X_1, X_2, \cdots$ are i.i.d. Poisson random variables with mean $\lambda > 0$. Let the distribution of X_1 be governed by some unknown member of the family $\mathscr{P} = \{P_\lambda : \lambda \in \Lambda\}$ where $\Lambda = \{\lambda_i : 0 < \lambda_1 < \lambda_2 < \cdots\}$, $\lambda_0 = 0$. Assume that $\lambda_i - \lambda_{i-1} \geqq 1\ \forall i \geqq 1$. Later we modify our choice of a_{ij} to overcome this assumption.

We choose $a_{ij} = \alpha^{|\lambda_i - \lambda_j|}$, $\alpha > 1$. So, $\sum_{i\neq j} a_{ij}^{-1} \leqq \sum_{i\neq j}\alpha^{-|i-j|} \leqq 2/(\alpha-1)$. Setting $S_n = X_1 + \cdots + X_n$, after some computations we obtain

$$(6.1)\qquad N = \inf\left\{n \geqq \log\alpha : \left\{\frac{(\lambda_i - \lambda_{i-1})}{\log(\lambda_i/\lambda_{i-1})}\right\}(1 + n^{-1}\log\alpha) \leqq n^{-1}S_n \leqq \left\{\frac{(\lambda_{i+1} - \lambda_i)}{\log(\lambda_{i+1}/\lambda_1)}\right\}(1 - n^{-1}\log\alpha) \ \text{ for some } \ i\right\},$$

and assert λ_i if N stops with i. It is easy to see that

$$(6.2)\qquad P_{\lambda_i}(N < \infty) = 1 \quad \forall i, \qquad \text{and} \qquad \sup_{\lambda\in\Lambda} P_\lambda(\text{error}) \leqq 2/(\alpha - 1) .$$

Setting

$$q(i+1) = [(\lambda_{i+1} - \lambda_i)/\log(\lambda_{i+1}/\lambda_i)] > \lambda_i ,$$
$$q(i-1) = [(\lambda_i - \lambda_{i-1})/\log(\lambda_i/\lambda_{i-1})] < \lambda_i ,$$

from (5.7) we obtain

$$(6.3)\qquad E_{\lambda_i} N \geqq \frac{\log \alpha}{\min [\{1 - \lambda_i q^{-1}(i+1)\}, \{\lambda_i q^{-1}(i-1) - 1\}]} = K_i^{-1} \log \alpha ,$$

and thus

$$(6.4)\qquad \liminf_{\alpha\to\infty} (K_i E_i N)/\log \alpha \geqq 1 .$$

To compute an upper bound for $E_{\lambda_i} N$, recall that

$$(6.5)\qquad N = \inf\{n \geqq \log \alpha : q(i-1)(1 + n^{-1}\log \alpha) \leqq n^{-1} S_n \leqq q(i+1)(1 - n^{-1}\log \alpha) \text{ for some } i\} .$$

Set $r = R \log \alpha (R > K_i^{-1})$ and assume for simplicity that r is an integer. From (6.5) we have

$$(6.6)\qquad P_{\lambda_i}(N > r) \leqq P_{\lambda_i}(S_r > ar) + P_{\lambda_i}(S_r < br)$$

where $a = q(i+1)(1 - R^{-1}) > q(i+1)(1 - K_i) > \lambda_i$, and $b = q(i-1)(1 + R^{-1}) < q(i-1)(1 + K_i) < \lambda_i$. Since $E_{\lambda_i} X_1 = \lambda_i$, a theorem of Chernoff in [2] implies that

$$(6.7)\qquad P_{\lambda_i}(N > r) \leqq m^r(a) + m^r(b)$$

where $m(\zeta) = \inf_t e^{-\zeta t} E_{\lambda_i} e^{tX_1} < 1$ for $\zeta = a$ and $\zeta = b$. Thus, $N \sim K_i^{-1} \log \alpha$ in probability as $\alpha \to \infty$. Moreover, non-asymptotically, (6.7) implies: (i) $E_{\lambda_i} e^{tN} < \infty$ for some $t > 0$, and hence (ii) $E_{\lambda_i} N^k < \infty$ $\forall k \geqq 1$. It is easy to show that

$$E_{\lambda_i} N \leqq R \log \alpha + (R \log \alpha + 1)\{m(a)^{R(\log \alpha)} + m(b)^{R(\log \alpha)}\} + (m(a)^{R(\log \alpha)+1})/(1 - m(a)) + (m(b)^{R(\log \alpha)+1})/(1 - m(\mathrm{b})) ,$$

and hence

$$(6.8)\qquad \limsup_{\alpha\to\infty} (K_i E_{\lambda_i} N)/\log \alpha \leqq 1 ,$$

which combined with (6.4) gives

$$(6.9)\qquad E_{\lambda_i} N \sim K_i^{-1} \log \alpha \qquad \text{as} \quad \alpha \to \infty .$$

Moreover, if $\lambda_{i+1} - \lambda_i = 1$ $\forall i \geqq 1$, then the rule (6.1) is asymptotically optimal (Theorem 5).

Special case. If $\lambda_i = i$, i.e. $\Lambda = \{1, 2, \cdots\}$, then

$$N = \inf\{n \geqq \log \alpha : (\log^{-1} i/(i-1))(1 + n^{-1} \log \alpha) \leqq n^{-1} S_n \leqq (\log^{-1} (i+1)/i)(1 - n^{-1} \log \alpha) \text{ for some } i\} .$$

Also,

$$E_i N \sim \frac{\log \alpha}{\min [\{1 - i \log (1 + 1/i)\}, \{i \log i/(i-1) - 1\}]} \qquad \text{as} \quad \alpha \to \infty .$$

Since $\min [\{1 - i \log (1 + 1/i)\}, \{i \log i/(i-1) - 1\}] = 1 - i \log (1 + i^{-1})$, $i \geqq 1$, hence $E_i N \sim (\log \alpha)/\{1 - i \log (1 + i^{-1})\}$ as $\alpha \to \infty$. This is the asymptotic expression also obtained by McCabe [11].

6.2. *Sequential distinguishability for the normal family.* Let P_θ denote the probability under which $X_1, X_2, \cdots$ are i.i.d. $N(\theta, 1)$, $-\infty < \theta < \infty$. Here we take $\Omega = \{\theta_i : -\infty = \theta_{-\infty} < \cdots < \theta_{-1} < \theta_0 < \theta_1 < \cdots < \theta_\infty = \infty\}$. As in the Poisson case, assume that $\theta_i - \theta_{i-1} \geqq 1$ $\forall i$, so that Ω is an ordered set in the usual direction with at least unit spacing. Choose $a_{ij} = \alpha^{|\theta_i - \theta_j|}$, $\alpha > 1$. Setting $S_n = X_1 + \cdots + X_n$, some computations yield

$$N = \inf\{n \geqq 1 : n^{-1}\log\alpha - \tfrac{1}{2}(\theta_i - \theta_{i-1}) \leqq (n^{-1}S_n - \theta_i) \leqq \tfrac{1}{2}(\theta_{i+1} - \theta_i) - n^{-1}\log\alpha \quad \text{for some } i\}, \tag{6.10}$$

and assert θ_i if N stops with i. It follows that (i) $P_{\theta_i}(N < \infty) = 1$ $\forall i$, and (ii) $\sup_i P_{\theta_i}(\text{error}) \leqq 2/(\alpha - 1)$.

We omit the details of exact lower and upper bounds for $E_{\theta_i}N$. These bounds are easily obtained in a fashion similar to that in Section 6.1. However, it is easy to show that $E_{\theta_i}N \sim (2\log\alpha)/\min[(\theta_{i+1} - \theta_i), (\theta_i - \theta_{i-1})]$ as $\alpha \to \infty$. If $\theta_{i+1} - \theta_i = 1$ $\forall i$, then the rule (6.10) is asymptotically optimal. In particular, setting $\theta_i = i$, we get Robbins' procedure (see [12]) where some other details are also given.

6.3. *Asymptotically optimal rules without uniform spacing.* In the preceding two sections we obtained a.o. (asymptotically optimal) rules for sequences which are uniformly spaced. We now give a.o. rules without uniform spacing.

6.3.1. *The normal case.* For simplicity, consider $\Omega = \{\theta_i : \theta_1 < \theta_2 < \cdots\}$. Modify the previous choice of a_{ij} as follows:

$$\begin{aligned} a_{ij} &= \alpha^{(\theta_j - \theta_i)/(\theta_{i+1} - \theta_i)} && \text{if } j > i, \\ &= \alpha^{(\theta_i - \theta_j)/(\theta_i - \theta_{i-1})} && \text{if } j < i, \qquad \alpha > 1. \end{aligned}$$

From (6.10) it follows that

$$N = \inf\{n \geqq 1 : (\log\alpha)/n(\theta_i - \theta_{i-1}) - (\theta_i - \theta_{i-1})/2 \leqq (n^{-1}S_n - \theta_i) \leqq (\theta_{i+1} - \theta_i)/2 - (\log\alpha)/n(\theta_{i+1} - \theta_i) \quad \text{for some } i\}. \tag{6.11}$$

It can easily be shown that

$$E_{\theta_i}N \sim (\log\alpha)/\inf_{j \neq i} I(\theta_i, \theta_j) \quad \text{as} \quad \alpha \to \infty, \tag{6.12}$$

where $I(\theta_i, \theta_j) = (\theta_i - \theta_j)^2/2$ is the Kullback-Leibler measure. Thus, if our modified choice of a_{ij} satisfy

(1) $\sum_{i \neq j} a_{ij}^{-1} \leqq \psi(\alpha) \downarrow 0$ as $\alpha \to \infty$, and

(2) $\sum_{i \neq j} a_{ij}^{-1} \lesssim 2/\alpha$ as $\alpha \to \infty$,

then by Theorem 5 the stopping rule (6.11) is asymptotically optimal. Here are some conditions under which (1) and (2) hold.

a. Assume: (i) there exists $\delta > 0$ such that $\inf_i \inf_{j \neq i} |\theta_i - \theta_j| = \delta$, and (ii) $\theta_{i+1} - \theta_i \leqq \Delta$ $\forall i \geqq 1$, then

$$\sum_{i \neq j} a_{ij}^{-1} \leqq 2/\alpha + o(a^{-1}).$$

To see this we note that (i) for $i \leqq j-2$,

$$(\theta_j-\theta_i)/(\theta_{i+1}-\theta_i) \geqq 1+\frac{\delta}{\Delta}+\cdots+\left(\frac{\delta}{\Delta}\right)^{j-i-1},$$

and (ii) for $i \geqq j+2$,

$$(\theta_i-\theta_j)/(\theta_i-\theta_{i-1}) \geqq 1+\frac{\delta}{\Delta}+\cdots+\left(\frac{\delta}{\Delta}\right)^{i-j-1}.$$

Therefore,

$$\textstyle\sum_{i\neq j} a_{ij}^{-1} \leqq 2/\alpha + 2\alpha^{-1}\sum_{k=1}^{\infty}\alpha^{-k\delta/\Delta} = 2/\alpha + 2/\alpha(\alpha^{\delta/\Delta}-1).$$

b. Assume: (i) Successive differences are non-increasing and (ii) either there exists a $\delta > 0$ such that $\inf_i \inf_{j\neq i}|\theta_i-\theta_j| = \delta$, or

$$\textstyle\sum_{i=1}^{j-2}\alpha^{-(\theta_j-\theta_i)/(\theta_{i+1}-\theta_i)} \leqq g(\alpha)\downarrow 0 \qquad \text{as} \quad \alpha\to\infty,$$

then $\sum_{i\neq j} a_{ij}^{-1} \leqq 2/\alpha + o(\alpha^{-1})$.

EXAMPLE 1. Let $\theta_i = \log i$, $i \geqq 1$. Successive differences are non-increasing but $\delta = 0$. However,

$$\begin{aligned}\textstyle\sum_{i=1}^{j-2}\alpha^{-(\theta_j-\theta_i)/(\theta_{i+1}-\theta_i)} &\leqq \textstyle\sum_{i=1}^{j-2}\alpha^{-i(j-i)/j}\\ &\leqq \textstyle\sum_{i<j:i\leqq j/2}\alpha^{-i(j-i)/j} + \sum_{i<j:i>j/2}\alpha^{-i(j-i)/j}\\ &\leqq 2\textstyle\sum_{k>0}\alpha^{-k/2} = 2/((\alpha)^{\frac{1}{2}}-1).\end{aligned}$$

c. Assume: (i) Successive differences are non-decreasing and (ii) there exists a Δ such that for $i \geqq j+2$, $(\theta_i-\theta_j)/(\theta_i-\theta_{i-1}) \geqq \Delta(i-j)$. Then $\sum_{i\neq j} a_{ij}^{-1} \leqq (2\alpha-1)/\alpha(\alpha-1) + 1/\alpha^{\Delta}(\alpha^{\Delta}-1)$. However, $\sum_{i\neq j} a_{ij}^{-1} \leqq 2/\alpha + o(\alpha^{-1})$.

EXAMPLE 2. Let $\theta_i = i^2$, $i \geqq 1$. For $i \geqq j+2$,

$$\begin{aligned}(\theta_i-\theta_j)/(\theta_i-\theta_{i-1}) &= (i+j)(i-j)/(2i-1)\\ &\geqq (i-j)\inf_{i\geq 3}\frac{(i+1)}{(2i-1)} = \tfrac{4}{5}(i-j).\end{aligned}$$

d. If $\Omega = (\theta_1 > \theta_2 > \cdots)$, then the same modified choice of α_{ij} gives an a.o. rule provided that $\sum_{i\neq j} a_{ij}^{-1} \leqq \psi(\alpha)\downarrow 0$ as $\alpha\to\infty$, and $\sum_{i\neq j} a_{ij}^{-1} \leqq 2/\alpha + o(\alpha^{-1})$.

EXAMPLE 3. Let $\theta_i = i^{-1}$, i.e. $\Omega = (1, \frac{1}{2}, \frac{1}{3}, \cdots)$. In this case

$$\begin{aligned}a_{ij} &= \alpha^{(i+1)(j-i)/j} &&\text{if } j > i,\\ &= \alpha^{(i-1)(i-j)/j} &&\text{if } j < i.\end{aligned}$$

Thus,

$$\begin{aligned}\textstyle\sum_{i\neq j} a_{ij}^{-1} &= \textstyle\sum_{i<j:i\leqq j/2}\alpha^{-(i+1)(j-i)/j} + \sum_{i<j:i>j/2}\alpha^{-(j-i)(i+1)/j}\\ &\quad + \textstyle\sum_{i=j+1}^{\infty}\alpha^{-(i-j)(i-1)/j}\\ &\leqq \textstyle\sum_{k=1}^{\infty}\alpha^{-(k+1)/2} + \sum_{k=1}^{\infty}\alpha^{-k/2} + \sum_{k=1}^{\infty}\alpha^{-k} \leqq (3+\alpha^{\frac{1}{2}}+\alpha^{-\frac{1}{2}})/(\alpha-1).\end{aligned}$$

Nevertheless, $\sum_{i\neq j} a_{ij}^{-1} \leqq 2/\alpha + o(\alpha^{-1})$.

6.3.2. *The Poisson case*. Let $\Lambda = \{0 < \lambda_1 < \lambda_2 < \cdots\}$. Modifying the choice of a_{ij} as in Section 6.3.1, (6.1) now takes the form

$$N = \inf\left\{n \geqq \log\alpha : \left\{\frac{(\lambda_i - \lambda_{i-1})}{\log(\lambda_i/\lambda_{i-1})}\right\}(1 + (\log\alpha)/n(\lambda_i - \lambda_{i-1})) \leqq S_n/n \right.$$
$$\left. \leqq \left\{\frac{(\lambda_{i+1} - \lambda_i)}{\log(\lambda_{i+1}/\lambda_i)}\right\}(1 - (\log\alpha)/n(\lambda_{i+1} - \lambda_i)) \text{ for some } i\right\}.$$

It is easy to see that the stopping rule N is asymptotically optimal provided that $\sum_{i\neq j} a_{ij}^{-1} \leqq 2/\alpha + o(\alpha^{-1})$. The treatment of a general exponential model is deferred to a subsequent publication.

Acknowledgment. I wish to express my thanks to Professor Herbert Robbins for his encouragement, inspiration and help throughout this work. I also wish to thank Professor Robert Berk for his critical reading and for several helpful suggestions.

REFERENCES

[1] BLACKWELL, DAVID and GIRSCHICK, M. A. (1954). *Theory of Games and Statistical Decisions*. Wiley, New York.

[2] CHERNOFF, H. (1952). A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations. *Ann. Math. Statist.* **23** 493–507.

[3] DOOB, J. L. (1953). *Stochastic Processes*. Wiley, New York.

[4] FISHER, LLOYD and VAN NESS, JOHN W. (1969). Distinguishability of probability measures. *Ann. Math. Statist.* **40** 381–392.

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

[6] HAMMERSLEY, J. M. (1950). On estimating restricted parameters. *J. Roy. Statist. Soc. Ser. B* **12** 192–240.

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

[8] KAKUTANI, S. (1948). On equivalence of infinite product measures. *Ann. Math.* **49** 214–224.

[9] KRAFT, CHARLES (1955). Some conditions for consistency and uniform consistency of statistical procedures. California Univ. Publications in Statistics **2**, 125–141.

[10] KULLBACK, SOLOMON (1959). *Information Theory and Statistics*. Wiley, New York.

[11] MCCABE, GEORGE P. (1970). Some problems in sequential discrimination. Ph. D. thesis, Columbia Univ.

[12] ROBBINS, H. (1970). Sequential estimation of an integer mean. *Scientists at Work, Festschrift in honor of Herman Wold*. Almqvist and Wiksells, Uppsala.

[13] WALD, A. (1947). *Sequential Analysis*. Wiley, New York.

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

[15] WALD, A. and WOLFOWITZ, J. (1948). Optimum character of the sequential probability ratio test. *Ann. Math. Statist.* **19** 326–329.

209 FRANKLIN HALL
KENT STATE UNIVERSITY
KENT, OHIO 44240