The Annals of Statistics
1985, Vol. 13, No. 1, 206–227

OPTIMAL DETECTION OF A CHANGE IN DISTRIBUTION[1]

By Moshe Pollak

The Hebrew University of Jerusalem

Suppose one is able to observe sequentially a series of independent observations $X_1, X_2, \cdots$ such that $X_1, X_2, \cdots, X_{\nu-1}$ are iid distributed according to a known distribution F_0 and $X_\nu, X_{\nu+1}, \cdots$ are iid distributed according to a known distribution F_1. Assume that ν is unknown and the problem is to raise an alarm as soon as possible after the distribution changes from F_0 to F_1. Formally, the problem is to find a stopping rule N which in some sense minimizes $E(N-\nu \mid N \geq \nu)$ subject to a restriction $E(N \mid \nu = \infty) \geq B$. A stopping rule that is a limit of Bayes rules is first derived. Then an almost minimax rule is presented; i.e. a stopping rule N^* is described which satisfies $E(N^* \mid \nu = \infty) = B$ for which

$$\sup_{1\leq\nu<\infty} E(N^* - \nu \mid N^* \geq \nu) - \inf_{\{\text{stopping rules } N \mid E(N\mid \nu=\infty)\geq B\}} \sup_{1\leq \nu<\infty} E(N-\nu \mid N \geq \nu) = o(1)$$

where $o(1) \to 0$ as $B \to \infty$.

1. Introduction. Suppose one is able to observe sequentially a series of independent observations whose distribution possibly changes from F_0 to F_1 at an unknown point in time. Formally, $X_1, X_2, \cdots$ are independent random variables such that $X_1, \cdots, X_{\nu-1}$ are each distributed according to a distribution F_0 and $X_\nu, X_{\nu+1}, \cdots$ are each distributed according to a distribution F_1, where $1 \leq \nu \leq \infty$ is unknown. The objective is to detect that a change took place "as soon as possible" after its occurrence, subject to a restriction on the rate of false detections.

The problem originally arose out of considerations of quality control. When a process is "in control," observations are distributed according to F_0. At the unknown point ν, the process jumps "out of control" and ensuing observations are distributed according to F_1. The aim is to raise an alarm "as soon as possible" after the process jumps "out of control".

The early literature on the problem deals mainly with a change in the mean of normal random variables having known, fixed variance, or a change in the probability of success in Bernoulli trials. Early solutions grouped observations and proposed testing each group individually for an indication of whether or not the process is "in control". Shewhart charts (Shewhart, 1931) were standard procedure for 20–30 years.

Gradually it became clear that much was being lost by regarding (groups of)

Received May 1982; revised June 1983.

[1] This research was supported by a grant from the United States-Israel Binational Foundation (BSF), Jerusalem, Israel.

AMS 1980 *subject classifications.* Primary 62L10, 62N10; secondary 62C10, 62F15, 60G40.

Key words and phrases. Quality control, control charts, stopping time, limit of Bayes rules, minimax rule, sequential analysis.

observations separately, thereby not enabling evidence to accumulate with time. Early attempts to rectify this include work by Dudding and Jennett (1944) and Weiler (1954). Their procedure involves constructing "warning lines" and "action lines," the proposal being that one would take action whenever a point passes an "action line" or whenever enough recent points pass a "warning line". (See also Page, 1955.)

Procedures that are in current use were initiated by Page (1954). In order to detect a change in a normal mean from μ_0 to $\mu_1 > \mu_0$, he proposed stopping and declaring the process to be "out of control" as soon as $S_n - \min_{1\le k\le n} S_k$ gets large, where $S_k = \sum_{i=1}^{k} (X_i - \mu^*)$ and $\mu_0 < \mu^* < \mu_1$ is suitably chosen. This and related procedures are known as CUSUM (cumulative sum) procedures (cf. Johnson and Leone, 1962, for a survey). Other methods were also proposed in related contexts (e.g., Girshick and Rubin, 1952; Chernoff and Zacks, 1964; Kander and Zacks, 1966). Nonparametric approaches have also been tried (cf. Bhattacharya and Frierson, 1981).

In order to evaluate and compare procedures, one needs to formalize a restriction on false detections as well as to formalize the objective of detecting a change "as soon as possible" after its occurrence. The restriction on false detections is usually formalized either as a rate restriction on stopping rule N (e.g. a requirement that $E(N \mid \nu = \infty) \ge B$) or a probability restriction (e.g. a requirement that $P(N \ge \nu) \le \alpha$ for all ν). The objective of detecting a change "as soon as possible" after its occurrence is usually formalized in terms of functionals of $N - \nu$. (See Aroian and Levene, 1950; Page, 1954; Shiryayev, 1964; Lorden, 1971; Pollak and Siegmund, 1975; Chen, 1978; Kenett and Pollak, 1983).

Shiryayev (1963, 1978) solved the problem in a Bayesian framework. He considered a loss function whereby one loses one unit if $N < \nu$ and one loses c units for each observation taken after ν if $N \ge \nu$. The prior on ν is assumed to be Geometric (p). Shiryayev showed that the Bayes solution prescribes stopping as soon as the posterior probability of a change having occurred exceeds a fixed level. He also solved a non-Bayesian version, minimizing the mean detection time after the onset of a stationary regime.

In the non-Bayesian setting of the problem, the only other optimality result is that of Lorden (1971). His results are based on the restriction that the stopping rules N must satisfy $E(N \mid \nu = \infty) \ge B$. The speed with which a stopping rule N detects a (true) change of distribution is evaluated by $\sup \operatorname{ess\,sup} E((N - \nu + 1)^+ \mid X_1, \cdots, X_{\nu-1})$. Lorden showed that a certain class of stopping rules is asymptotically ($B \to \infty$) optimal (in a first-order sense) and that Page's aforementioned procedure belongs to this class.

In this article, the problem is also treated (in a non-Bayesian setting) under the restriction that the permissible stopping rules N satisfy $E(N \mid \nu = \infty) \ge B$. We first study a stopping rule N_A considered by Roberts (1966), defined as the first time $n \ge 1$ that a certain (nonnegative) statistic R_n (defined in Section 2) exceeds the level A. (Shiryayev, 1963, considered a continuous-time analog.) Roberts (1966) compared this rule to other procedures using Monte Carlo methods, and found it to be very good. We show this rule to be a limit as $p \to 0$ of the Bayes solution of the aforementioned Bayesian problem considered by

Shiryayev. We also show this rule to be asymptotically ($p \to 0$) Bayes risk efficient. (See Theorem 1.)

Following Pollak and Siegmund (1975), the speed with which a stopping rule detects a (true) change of distribution is evaluated by $\sup_{1\le\nu<\infty} E(N - \nu \mid N \ge \nu, \nu)$. An almost minimax rule is presented (see Theorems 2 and 3); i.e. a stopping rule N_A^* is defined which satisfies $E(N_A^* \mid \nu = \infty) = B$ and, to within an $o(1)$ term, minimizes $\sup_{1\le\nu<\infty} E(N - \nu \mid N \ge \nu, \nu)$ over all stopping rules N which satisfy $E(N \mid \nu = \infty) \ge B$, where $o(1) \to 0$ as $B \to \infty$.

The construction of the rule and the proof of its almost minimaxity is based on the idea that "if an equalizer rule is extended Bayes, it is a minimax rule" (Ferguson, 1967, page 91). We first show that R_{n+1} is a time-independent function of R_n and the $(n + 1)$th observation X_{n+1}. We then let R_0^* have an arbitrary distribution and define R_{n+1}^* to be a function of R_n^* and X_{n+1} in the same manner as R_{n+1} depends on R_n and X_{n+1}. We show that there exists a choice ψ_A^* for the distribution of R_0^* which creates a sequence $\{R_n^*\}_{n=0}^{\infty}$ for which the distribution of R_n^* conditional on $\{N_A^* > n\}$ is the same for all $n \ge 0$ (when $\nu = \infty$) where $N_A^* = \min\{n \mid R_n^* \ge A\}$. Note that the distribution of $N_A^* - \nu$ conditional on $\{R_{\nu-1}^*, N_A^* \ge \nu\}$ is the same for all $\nu \ge 1$. Therefore, $E(N_A^* - \nu \mid N_A^* \ge \nu)$ is the same for all $\nu \ge 1$ (making N_A^* an "equalizer rule"). The proof is concluded by embedding this setup in a sequence of Bayesian problems which allow $\nu = 0$ as a prior possibility, showing N_A^* to be an asymptotically Bayes risk efficient limit of Bayes rules, and showing that the event $\{\nu = 0\}$ is asymptotically negligible.

Explicit examples are presented in Section 4.

2. A limit of Bayes rules. Consider the following Bayesian problem and denote it by $B(\gamma, p, c)$. Suppose ν has a prior distribution

$$P(\nu = 0) = \gamma$$

$$P(\nu = n) = (1 - \gamma)p(1 - p)^{n-1}, \quad n \ge 1$$

where p and γ are known constants, $0 < p \le 1$, $0 \le \gamma \le 1$. Conditional on ν, if $\nu > 1$ the observations $X_1, \cdots, X_{\nu-1}$ are independent with density f_0 with respect to a σ-finite measure μ and are independent of $X_\nu, X_{\nu+1}, \cdots$ which (conditional on ν) are independent with density f_1 ($\ne f_0$) with respect to μ. If $\nu = 0$ or $\nu = 1$ then $X_1, X_2, \cdots$ are independent with density f_1 with respect to μ. The aim is to choose an optimal stopping time N. The losses are 1 or $c(N - \nu)$ if $N < \nu$ or $N \ge \nu$, respectively, where $c > 0$ is a fixed constant.

This problem was solved by Shiryayev (1978), Section 4.3. Define

$$R_n = \sum_{k=1}^{n} \prod_{i=k}^{n} \frac{f_1(X_i)}{f_0(X_i)}, \quad R_{q,n} = \sum_{k=1}^{n} \prod_{i=k}^{n} \left(\frac{f_1(X_i)}{f_0(X_i)} \frac{1}{q}\right)$$

$$\Delta = \inf\left\{x \mid P_\infty\left(\frac{f_1(X_i)}{f_0(X_i)} \le x\right) > 0\right\} < 1.$$

$P_\nu(\cdot)$, $E_\nu(\cdot)$ will respectively denote probability and expectation conditional on ν when $1 \le \nu \le \infty$. (The case $\nu = \infty$ corresponds to the possibility that no change from f_0 will ever take place; i.e. $X_1, X_2, \cdots$ are iid with density f_0 with respect to μ.)

For ease of exposition we will assume that $P_1(f_1(X_1)/f_0(X_1) = \infty) = 0$. Analogous results can be formulated when this probability is positive.

THEOREM 1. (I) *Suppose that the P_∞-distribution of $f_1(X_1)/f_0(X_1)$ has no atoms.*

(i) *For any $\Delta < A < \infty$ there exists a constant $0 < c^* < \infty$ and a sequence $\{p_i, c_i\}_{i=1}^{\infty}$ with $p_i \to_{i\to\infty} 0$, $c_i \to_{i\to\infty} c^*$ such that the stopping time*

$$(1) \qquad N_A = \min\{n \mid R_n \ge A, n \ge 1\}$$

is a limit as $i \to \infty$ of Bayes rules for $B(\gamma = 0, p = p_i, c = c_i)$.

(ii) *For any set of Bayesian problems $B(\gamma, p, c)$ with $\gamma = 0, p \to 0, c \to c^*$*

$$\limsup_{p\to 0, c\to c^*} \frac{1 - \{\text{Expected loss using a Bayes rule for problem } B(0, p, c)\}}{1 - \{\text{Expected loss using } N_A \text{ for problem } B(0, p, c)\}} = 1.$$

(iii) *For any $1 \le B < \infty$ there exists a unique $\Delta \le A < \infty$ such that $B = E_\infty N_A$.*

(II) *If the P_∞-distribution of $f_1(X_1)/f_0(X_1)$ has atoms, then for any $1 \le B < \infty$ there exists $\Delta \le A < \infty$ such that $B = E_\infty N'_A$, where N'_A is defined as N_A in* (1) *with the modification that there may be randomization* (*whether or not to continue sampling*) *if $R_n = A$. N'_A is also a limit to Bayes rules, and a result analogous to* (ii) *above holds.*

REMARK. Part (ii) of the theorem should be seen in light of the fact that as $p \to 0$ and $c \to c^*$, the Bayes rule stops before ν with probability tending to 1, so that 1 minus the expected loss is an appropriate quantity to consider asymptotically.

PROOF. The idea of the proof is to note that $R_{q,n} \to R_n$ as $q \to 1$, that, with $q = 1 - p$, $N_{q,A} = \min\{n \mid R_{q,n} \ge A\}$ is a Bayes rule for some $B(0, p, c)$ (and vice versa) and that for fixed c the threshold A (governing $N_{q,A}$) remains bounded as $p \to 0$. The details are spelled out in the following lemmas.

LEMMA 1. *For the Bayesian problem $B(\gamma, p, c)$ described in the beginning of this section, there exists a constant $\delta_{p,c}$ such that*

$$M_{p,c} = \min\{n \mid n \ge 0, \quad P(\nu \le n \mid \mathscr{F}_n) \ge \delta_{p,c}\}$$

is a Bayes rule (*where $\mathscr{F}_n = \mathscr{F}(X_1, \cdots, X_n)$*).

PROOF. This is merely a restatement of Theorem 7 of Shiryayev (1978), page 195.

LEMMA 2. *Let $M_{p,c}$, $\delta_{p,c}$ be as in Lemma 1 and suppose that $\gamma = 0$ in the Bayesian problem described in the beginning of this section. Then $M_{p,c} = N_{q,D}$ for $q = 1 - p$, $D = (1/p)(\delta_{p,c}/(1 - \delta_{p,c}))$.*

PROOF. Denoting $\prod_{i=1}^{0} f_0(X_i) = 1$, note that for $n \geq 1$

$$P(\nu = j \mid \mathscr{F}_n) = \begin{cases} \dfrac{(\prod_{i=1}^{j-1} f_0(X_i))(\prod_{i=j}^{n} f_1(X_i))pq^{j-1}}{\sum_{k=1}^{n} (\prod_{i=1}^{k-1} f_0(X_i))(\prod_{i=k}^{n} f_1(X_i))\, pq^{k-1} + (\prod_{i=1}^{n} f_0(X_i))q^n} & \text{if } 1 \leq j \leq n \\ \dfrac{(\prod_{i=1}^{n} f_0(X_i))pq^{j-1}}{\sum_{k=1}^{n} (\prod_{i=1}^{k-1} f_0(X_i))(\prod_{i=1}^{n} f_1(X_i))pq^{k-1} + (\prod_{i=1}^{n} f_0(X_i))q^n} & \text{if } n < j < \infty \end{cases}$$

and therefore, for $n \geq 1$

$$P(\nu \leq n \mid \mathscr{F}_n) = \frac{\sum_{k=1}^{n} \prod_{i=k}^{n} (f_1(X_i)/f_0(X_i))(1/q)}{\sum_{k=1}^{n} \prod_{i=k}^{n} ((f_1(X_i)/f_0(X_i))(1/q)) + 1/p} = \frac{R_{q,n}}{R_{q,n} + 1/p}.$$

Hence $P(\nu \leq n \mid \mathscr{F}_n) \geq \delta_{p,c}$ if and only if $R_{q,n} \geq (1/p)(\delta_{p,c}/(1 - \delta_{p,c})) = D$ and so $M_{p,c} = N_{q,D}$.

LEMMA 3. *There exists $0 < q_0 < 1$ such that for all $q_0 \leq q$, when $\nu = \infty$,*

$$H_q = \sum_{k=1}^{\infty} \prod_{i=1}^{k} (f_1(X_i)/f_0(X_i))(1/q)$$

is a.s. a finite-valued random variable, and $R_{q,n} \to_{n\to\infty}^{\text{dist}} H_q$.

PROOF.

$$0 = \log \int \frac{f_1(x)}{f_0(x)} f_0(x)\, d\mu(x) > \int \log(f_1(x)/f_0(x))\, f_0(x)\, d\mu(x).$$

Hence there exists $0 < q_0 < 1$ such that

$$\int \log \left(\frac{f_1(x)}{f_0(x)} \frac{1}{q}\right) f_0(x)\, d\mu(x) < 0$$

whenever $q_0 \leq q \leq 1$. Therefore, for such q there exists a constant $\eta < 0$ such that

$$\limsup_{k\to\infty} k^{-1} \sum_{i=1}^{k} \log\left(\frac{f_1(X_i)}{f_0(X_i)} \frac{1}{q}\right) \leq_{\text{a.s.}} \eta.$$

Let

$$H_{q,n} = \sum_{k=1}^{n} \prod_{i=1}^{k} \left(\frac{f_1(X_i)}{f_0(X_i)} \frac{1}{q}\right) = \sum_{k=1}^{n} \exp\left(\sum_{i=1}^{k} \log\left(\frac{f_1(X_i)}{f_0(X_i)} \frac{1}{q}\right)\right).$$

For a sequence $X_1, X_2, \cdots$ define

$$M = \max\left\{m \mid m^{-1} \sum_{i=1}^{m} \log\left(\frac{f_1(X_i)}{f_0(X_i)} \frac{1}{q_0}\right) > \tfrac{1}{2}\eta\right\}.$$

M is a.s. finite (since $\nu = \infty$). For $n > M$,

$$H_{q,n} \leq H_{q,M} + \sum_{k=M+1}^{n} e^{\eta k/2} < H_{q,M} + \frac{1}{1 - e^{\eta/2}}.$$

Since $H_{q,n}$ increases in n, it follows that $H_{q,n} \to^{\text{a.s.}}_{n\to\infty} H_q$ where H_q is a.s. finite. Since $H_{q,n}$ and $R_{q,n}$ have the same distribution, it follows that $R_{q,n} \to^{\text{dist}}_{n\to\infty} H_q$ and the proof of Lemma 3 is complete.

LEMMA 4. *Suppose* $\nu = \infty$. *Denote* $F_q(x) = \inf_n P(R_{q,n} \le x)$ *and let* q_0 *be as in Lemma* 3. *Then* $F_q(x)$ *is nondecreasing in* x *for fixed* q, $F_q(x)$ *is nondecreasing in* q *for fixed* x *and for* $q \ge q_0$ $\lim_{x\to\infty} F_q(x) = 1$.

PROOF. Clearly $F_q(x)$ is nondecreasing in q and in x. The rest of the lemma follows easily from Lemma 3.

LEMMA 5. Let $N_{q,A} = \min\{n \mid R_{q,n} \ge A\}$. *Suppose* $A > (\Delta/q)/(1 - \Delta/q) \ge 0$. *Then* $P_\infty(R_{q,n} \le x \mid N_{q,A} > n) \ge P_\infty(R_{q,n} \le x)$ *for* $n = 1, 2, \cdots$.

PROOF. Clearly $P_\infty(N_{q,A} < \infty) = 1$. Under the conditions stated above $P_\infty(N_{q,A} > n) > 0$ for $n = 1, 2, \cdots$.

The Lemma is trivial for $x \ge A$. For $x < A$ the proof is by induction on n. The statement of Lemma 5 is obviously correct for $n = 1$. Suppose that it is correct for $n = 1, 2, \cdots, m$. Let $Y_{q,n} = q^{-1}f_1(X_n)/f_0(X_n)$. Note that

$$R_{q,m+1} = Y_{q,m+1}(1 + R_{q,m}).$$

Now

$$\begin{aligned} P_\infty(R_{q,m+1} &\le x \mid N_{q,A} > m + 1) \\ &= P_\infty(R_{q,m} < (x/Y_{q,m+1}) - 1 \mid N_{q,A} > m,\ R_{q,m} < A/Y_{q,m+1} - 1) \\ &= \frac{P_\infty(R_{q,m} \le (x/Y_{q,m+1}) - 1 \mid N_{q,A} > m)}{P_\infty(R_{q,m} < (A/Y_{q,m+1}) - 1 \mid N_{q,A} > m)} \\ &\ge P_\infty(R_{q,m} \le (x/Y_{q,m+1}) - 1 \mid N_{q,A} > m) \\ &\ge P_\infty(R_{q,m} \le (x/Y_{q,m+1}) - 1) \\ &= P_\infty(R_{q,m+1} \le x) \end{aligned}$$

where the last inequality follows from the induction hypothesis.

This completes the proof of Lemma 5.

LEMMA 6. *Let* $N_{q,A}$ *be as in Lemma* 5 *and let* q_0 *be as in Lemma* 3. *Suppose* ν *has a Geometric* (p) *prior distribution*, $0 < p < 1$. *Then for each* $c > 0$ *there exists a constant* D_c *such that for all* $q_0 \le q \le 1$, $0 < p < 1$, $E(N_{q,D} - \nu \mid N_{q,D} \ge \nu) > 2c^{-1}$ *if* $D \ge D_c$.

PROOF. Let $q_0 \le q \le 1$. $E(N_{q,D} - \nu \mid \nu = k, N_{q,D} \ge \nu) = EE(N_{q,D} - \nu \mid \nu = k, N_{q,D} \ge k, R_{q,k-1})$.

Let x_0 be such that $F_{q_0}(x_0) \ge 1/2$ (where F_q is as in Lemma 4). Suppose $D > x_0$.

For $k > 1$, on $\{R_{q,k-1} \le x_0\}$, since

$$R_{q,k-m} = \left(\prod_{i=k}^{k+m} \left(\frac{f_1(X_i)}{f_0(X_i)}\frac{1}{q}\right)\right) R_{q,k-1} + \sum_{j=k}^{k+m} \prod_{i=j}^{k+m} \left(\frac{f_1(X_i)}{f_0(X_i)}\frac{1}{q}\right)$$

it follows that

$$\begin{aligned} &E(N_{q,D} - \nu \mid \nu = k,\ N_{q,D} \ge k,\ R_{q,k-1}) \\ (2) \qquad &\ge E(N_{q,D} - \nu \mid \nu = k,\ N_{q,D} \ge k,\ R_{q,k-1} = x_0) \\ &\ge E(N_{q_0,D} - \nu \mid \nu = k,\ N_{q_0,D} \ge k,\ R_{q_0,k-1} = x_0) \end{aligned}$$

(where the last inequality follows from the fact that, given that one has not crossed D by time $k - 1$ and one is at state x_0 at time $k - 1$, one would cross D later if the process being observed (henceforth) is $R_{q,n}$ rather than if the process being observed (henceforth) is $R_{q_0,n}$). The right-hand side of expression (2) does not depend on p and k. It obviously is finite, increases with D and tends to ∞ as $D \to \infty$. Hence one can choose D_c (regardless of $0 < p < 1$, $q_0 \le q \le 1$) such that this expression is larger than $4c^{-1}$ (for all $k > 1$) if $D \ge D_c$. Hence, if $D \ge D_c$, for $k > 1$

$$\begin{aligned} &E(N_{q,D} - \nu \mid \nu = k,\ N_{q,D} \ge \nu) \\ &\ge \int_{\{r \le x_0\}} E(N_{q_0,D} - \nu \mid \nu = k,\ N_{q_0,D} \ge k,\ R_{q_0,k-1} = x_0) \times dP_{R_{q,k-1} \mid \nu = k, N_{q,D} \ge \nu}(r) \\ &\ge 4c^{-1} P_k\{R_{q,k-1} \le x_0 \mid \nu = k,\ N_{q,D} > k - 1\} \\ &= 4c^{-1} P_\infty\{R_{q,k-1} \le x_0 \mid N_{q,D} > k - 1\} \ge 4c^{-1} P_\infty\{R_{q,k-1} \le x_0\} \\ &\ge 2c^{-1} \end{aligned}$$

where the last inequality follows from Lemma 4 and the definition of x_0, and the inequality before it follows from Lemma 5. Clearly also $E(N_{q,D} - \nu \mid \nu = 1, N_{q,D} \ge \nu) \ge 2c^{-1}$. Hence $E(N_{q,D} - \nu \mid N_{q,D} \ge \nu) \ge 2c^{-1}$ for $D \ge D_c$, and the proof of Lemma 6 is complete.

LEMMA 7. *Under the conditions and notations used in Lemma* 2,

$$\limsup_{p \to 0} \delta_{p,c}/p \le D_c$$

where D_c is as defined in Lemma 6.

PROOF. By Lemma 6, $cE(M_{p,c} - \nu \mid M_{p,c} \ge \nu) \ge 2$ if $D = \delta_{p,c}/(p(1 - \delta_{p,c})) \ge D_c$ and p is small enough. This would contradict the fact that $M_{p,c}$ is a Bayes rule, for one would be doing better by stopping immediately (i.e. taking no observations at all). Hence $D_c \ge D \ge \delta_{p,c}/p$.

LEMMA 8. *Denote $D_c^* = \inf\{D_c \mid D_c$ as in Lemma 6$\}$. Then D_c^* is nonincreasing in c and $D_c^* \to \Delta/q_0$ as $c \to \infty$.*

PROOF. The proof follows straightforwardly from the definition of D_c.

LEMMA 9. *Suppose $\gamma = 0$ in the Bayesian problem described in the beginning of this section and N is a stopping time. Then*

$$\frac{P(N > \nu)}{p} \rightarrow_{p \to 0} E_\infty N.$$

PROOF.

$$p^{-1}P(N \geq \nu) = p^{-1} \sum_{k=1}^{\infty} P(N \geq k, \nu = k)$$
$$= \sum_{k=1}^{\infty} P_\infty(N \geq k)(1 - p)^{k-1} \rightarrow_{p \to 0} E_\infty N.$$

LEMMA 10. *Let $\delta_{p,c}$ be as in Lemma 1 and let $e(c) = \lim\inf_{p \to 0} \delta_{p,c}/p$. Then* $\lim\inf_{c \to 0} e(c) = \infty$.

PROOF. Without loss of generality let $\gamma = 0$ in the Bayesian problem described in the beginning of this section. Let $M_{p,c}$ be as in Lemma 1. Suppose $\lim\inf_{c \to 0} e(c) = e_\infty < \infty$. Then $\delta_{p_i,c_i}/(p_i(1 - \delta_{p_i,c_i})) < 1 + e_\infty$ for some subsequence $c_i \to 0$, $p_i \to 0$. Since

$$E(\text{Loss using } M_{p_i,c_i}) = P(M_{p_i,c_i} < \nu) + cP(M_{p_i,c_i} \geq \nu)E(M_{p_i,c_i} - \nu \mid M_{p_i,c_i} \geq \nu)$$

it follows from Lemma 9 that

$$\frac{1 - E(\text{Loss using } M_{p_i,c_i})}{p_i} = \frac{P(M_{p_i,c_i} \geq \nu)}{p_i} [1 - c_i E(M_{p_i,c_i} - \nu \mid M_{p_i,c_i} \geq \nu)]$$
$$\leq \frac{P(M_{p_i,c_i} \geq \nu)}{p_i} \leq \frac{P(N_{1+e_\infty} \geq \nu)}{p_i} \leq 1 + E_\infty N_{1+e_\infty}$$

for large enough i. Clearly $E_\infty N_{1+e_\infty} < \infty$. Hence one would obviously do better by applying a Page rule with large enough upper boundary (cf. Lorden, 1971), contradicting the fact that M_{p_i,c_i} is a Bayes rule.

This completes the proof of Lemma 10.

PROOF OF THEOREM 1. By virtue of Lemma 2, it can be seen from Shiryayev (1978), pages 195–198 that for any $A > \Delta$, $0 < p < 1$ there exists a value $c_{A,p} > 0$ such that $N_{q,A}$ with $q = 1 - p$ is a Bayes rule for $B(\gamma = 0, p, c = c_{A,p})$. From Lemma 8 and Lemma 10 it follows that $0 < \lim\inf_{p \to 0} c_{A,p} \leq \lim\sup_{p \to 0} c_{A,p} < \infty$. Choose a subsequence $p_i \searrow 0$ such that c_{A,p_i} converges to a finite constant $c^* \geq 0$ as $i \to \infty$. Since $R_{q,n} \searrow R_n$ as $p \searrow 0$ it follows that $M_{p_i, c_{A,p_i}} \rightarrow^{\text{a.s.}}_{i \to \infty} N_A$ and so N_A is a limit of Bayes rules.

To prove Theorem 1 (ii), regard the problems $B(\gamma = 0, p, c)$. For any stopping rule N:

$$\frac{1 - \{\text{Expected loss using } N \text{ for problem } B(0, p, c)\}}{p} \tag{3}$$
$$= \frac{P(N > \nu)}{p} [1 - cE(N - \nu \mid N \geq \nu)].$$

Denoting by $M_{p,c}$ the Bayes rule for $B(0, p, c)$ described in Lemma 1, it follows from Lemmas 1, 2, 7 and 8 that if c is bounded away from zero, then for sufficiently small p all the $M_{p,c}$'s are dominated by an N_D. Examination of the proof of Lemma 9 shows that $P(M_{p,c} \geq \nu)/p - E_\infty M_{p,c} \to_{p\to 0} 0$ uniformly for c bounded away from zero.

Consider the sequence of problems $B(\gamma = 0, p = p_i, c = c_{A,p_i})$. In problem $B(0, p_i, c_{A,p_i})$, $M_{p_i,c_{A,p_i}} = N_{q_i,A}$. Therefore

$$\lim_{i\to\infty}[P(M_{p_i,c_{A,p_i}} \geq \nu)/p_i]/[P(N_A \geq \nu)/p_i] = \lim_{i\to\infty} E_\infty M_{p_i,c_{A,p_i}}/E_\infty N_A = 1.$$

Also

$$P(\nu = k \mid M_{p_i,c_{A,p_i}} \geq \nu) = \frac{P(N_{q_i,A} \geq k \mid \nu = k)P(\nu = k)}{P(N_{q_i,A} \geq \nu)} \to_{i\to\infty} \frac{P_\infty(N_A \geq k)}{E_\infty N_A}.$$

In problem $B(0, p, c)$, if one uses N_A, then

$$P(\nu = k \mid N_A \geq \nu) = \frac{P(N_A \geq k \mid \nu = k)P(\nu = k)}{P(N_A \geq \nu)} \to_{p\to 0} \frac{P_\infty(N_A \geq k)}{E_\infty N_A}.$$

For fixed k, $E(N - \nu \mid N \geq \nu = k) = E_k(N - k \mid N \geq k)$. Since $M_{p_i,c_{A,p_i}} \to_{i\to\infty} N_A$ a.s. P_k, it follows that

$$\lim_{i\to\infty} E(M_{p_i,c_{A,p_i}} - \nu \mid M_{p_i,c_{A,p_i}} \geq \nu = k) = E_k(N_A - k \mid N_A \geq k)$$
$$= \lim_{i\to\infty} E(N_A - \nu \mid N_A \geq \nu = k).$$

Therefore, in problem $B(0, p_i, c_{A,p_i})$,

$$\lim_{i\to\infty} E(M_{p_i,c_{A,p_i}} - \nu \mid M_{p_i,c_{A,p_i}} \geq \nu)/E(N_A - \nu \mid N_A \geq \nu) = 1.$$

It now follows from (3) that

$$\frac{1 - \{\text{Expected loss using } M_{p_i,c_{A,p_i}} \text{ for problem } B(0, p_i, c_{A,p_i})\}}{1 - \{\text{Expected loss using } N_A \text{ for problem } B(0, p_i, c_{A,p_i})\}} \to_{i\to\infty} 1.$$

Suppose now that Theorem 1 (ii) were not true. Then there would exist a sequence $\{p_j^*, c_j^*\}_{j=1}^\infty$ such that $p_j^* \to_{j\to\infty} 0$, $c_j^* \to_{j\to\infty} c^*$, $M_{p_j^*,c_j^*} \to_{j\to\infty} N_{A^*}$ a.s. P_∞ for some $\Delta \leq A^* < \infty$ and

$$1 < \lim_{j\to\infty} \frac{1 - \{\text{Expected loss using } M_{p_j^*,c_j^*} \text{ for problem } B(0, p_j^*, c_j^*)\}}{1 - \{\text{Expected loss using } N_A \text{ for problem } B(0, p_j^*, c_j^*)\}}.$$

$$= \lim_{j\to\infty} \frac{1 - \{\text{Expected loss using } N_A^* \text{ for problem } B(0, p_j^*, c_j^*)\}}{1 - \{\text{Expected loss using } N_A \text{ for problem } B(0, p_j^*, c_j^*)\}}.$$

From considerations similar to (3) above and its sequel, it would then follow that for large enough i, N_{A^*} would do better than N_A and hence better than $M_{p_i,c_{A,p_i}}$ for problem $B(\gamma = 0, p = p_i, c = c_{A,p_i})$, which contradicts the fact that $M_{p_i,c_{A,p_i}}$ is a Bayes rule.

Theorem 1 (iii) follows from the fact that if the P_∞-distribution of $f_1(X_1)/f_0(X_1)$ has no atoms then $E_\infty N_A$ increases continuously from 1 to ∞ as A

increases from Δ to ∞. Theorem 1 (II) follows from considerations similar to those of the proof of Theorem 3 below. The details are omitted.

3. Almost minimaxity. Let $\Delta/(1-\Delta) < A < \infty$ and let ψ be a probability measure on $[0, \infty)$. Let R_0^* be a random variable with distribution ψ and define

$$R_{n+1}^* = \frac{f_1(X_{n+1})}{f_0(X_{n+1})}(1 + R_n^*), \quad N_{A,\psi}^* = \min\{n \mid R_n^* \geq A,\ n \geq 0\}.$$

In other words, $N_{A,\psi}^*$ is a stopping time constructed by randomizing R_0^* according to the distribution ψ, defining the sequence $\{R_n^*\}_{n=1}^\infty$ to be a sequence $\{R_n\}_{n=1}^\infty$ started at R_0^* (instead of at $R_0^* = 0$, as is in effect the case in the previous section) and stopping the first time n that R_n^* attains the level A.

For ease of exposition we will again assume that $P_1(f_1(X_1)/f_0(X_1) = \infty) = 0$. We will first consider the case that the P_∞-distribution of $f_1(X_1)/f_0(X_1)$ has no atoms.

THEOREM 2. *Let f_0 and f_1 be such that*

$$E_\infty[(f_1(X_1)/f_0(X_1)) \times \log^+(f_1(X_1)/f_0(X_1))] < \infty.$$

Suppose that the P_∞-distribution of $f_1(X_1)/f_0(X_1)$ has no atoms. Then for every $1 < B < \infty$ there exists a value A, $\Delta/(1-\Delta) < A < \infty$, and a probability measure ψ_A^* such that $B = E_\infty N_{A,\psi_A^*}^*$ and such that if N is any stopping time which satisfies $E_\infty N \geq B$ then

$$\sup_{1\leq\nu<\infty} E_\nu(N - \nu \mid N \geq \nu) \geq \sup_{1\leq\nu<\infty} E_\nu(N_{A,\psi_A^*}^* - \nu \mid N_{A,\psi_A^*}^* \geq \nu) + o(1)$$

where $o(1) \to 0$ as $B \to \infty$. $E_\nu(N_{A,\psi_A^*}^* - \nu \mid N_{A,\psi_A^*}^* \geq \nu) + o(1)$ is constant in $1 \leq \nu < \infty$.

REMARK. One can obtain ψ_A^* by solving the functional equation (4) in the sequel. For examples see Section 4. See also Remarks 1 and 3 in Section 5.

PROOF. Let $(\Delta/q)(1 - \Delta/q) < A$, $0 < q < 1$, let ψ be a probability measure on $[0, \infty)$, let $R_{q,0}^*$ be a random variable with distribution ψ and define

$$R_{q,n+1}^* = \frac{f_1(X_{n+1})}{f_0(X_{n+1})}\frac{1}{q}(1 + R_{q,n}^*), \quad N_{q,A,\psi}^* = \min\{n \mid R_{q,n}^* > A,\ n \geq 0\}$$

$$F_n(x) = P_\infty(R_{q,n}^* \leq x \mid N_{q,A,\psi}^* > n)$$

$$\rho(t, x) = P_\infty(R_{q,n+1}^* \leq x \mid R_{q,n}^* = t,\ N_{q,A,\psi}^* > n + 1)$$

$$\xi(t) = P_\infty(N_{q,A,\psi}^* > n + 1 \mid R_{q,n}^* = t,\ N_{q,A,\psi}^* > n).$$

Note that $\rho(t, x)$, $\xi(t)$ depend on q, A but are independent of ψ, n. Since the P_∞-distribution of $f_1(X_1)/f_0(X_1)$ has no atoms, $\xi(t)$ is continuous in t and $\rho(t, x)$ is continuous in t, x for $0 \leq t, x < A$.

LEMMA 11. *Let $T(F)$ be a transformation mapping the set of probability* *measures on* $[0, A]$ *into itself defined by*

$$T(F)(x) = \frac{\int_0^A \rho(t, x)\xi(t)\, dF(t)}{\int_0^A \xi(t)\, dF(t)}.$$

Then

(i) $F_{n+1} = T(F_n)$.

(ii) *There exists a solution to the equation*

$$T(\phi) = \phi. \tag{4}$$

REMARK. For later purposes, note that $T(F)$ has no atoms.

PROOF. Since

$$\begin{aligned} (5)\quad & P_\infty(R^*_{q,n} \in dx \mid N^*_{q,A,\psi} > n + 1) = P_\infty(R^*_{q,n} \in dx \mid N^*_{q,A,\psi} > n,\ N^*_{q,A,\psi} > n + 1) \\ &= \frac{P_\infty(N^*_{q,A,\psi} > n + 1 \mid R^*_{q,n} = x,\ N^*_{q,A,\psi} > n)P_\infty(R^*_{q,n} \in dx \mid N^*_{q,A,\psi} > n)P_\infty(N^*_{q,A,\psi} > n)}{\int_0^A P(N^*_{q,A,\psi} > n + 1 \mid R^*_{q,n} = x,\ N^*_{q,A,\psi} > n)P_\infty(R^*_{q,n} \in dx \mid N^*_{q,A,\psi} > n)P_\infty(N^*_{q,A,\psi} > n)} \\ &= \frac{\xi(x)\, dF_n(x)}{\int_0^A \xi(x)\, dF_n(x)} \end{aligned}$$

it follows that

$$F_{n+1}(x) = \frac{\int_0^A \rho(t, x)\xi(t)\, dF_n(t)}{\int_0^A \xi(t)\, dF_n(t)} = T(F_n)(x).$$

This accounts for (i). Because of the continuity properties of ρ and ξ, clearly $T(F)(x)$ is a weak* continuous functional of F for any x and therefore $T(F)$ is a weak* continuous transformation of F. Hence, by the Schauder-Tychonoff fixed point theorem (cf. Dunford, Schwartz, 1958, page 456, Theorem V.10.5) there exists a probability measure ϕ on $[0, A]$ (actually on $((\Delta/q)/(1 - \Delta/q), A)$) such that $T(\phi) = \phi$, completing the proof of Lemma 11.

Denote by $B(\psi, p, c)$ the Bayesian problem analogous to the one described in Section 2 but in which γ is a random variable defined by $\gamma = R^*_{q,0}/(R^*_{q,0} + 1/p)$ where the distribution of $R^*_{q,0}$ is ψ. (It is assumed that one knows the value of γ before taking the first observation.)

Let $p = 1 - q$, let ϕ be a solution of (4), and choose $\psi = G$ where G is defined by

$$dG(x) = \frac{(1 + px)\, d\phi(x)}{\int (1 + pt)\, d\phi(t)}. \tag{6}$$

It is easy to see that the distribution of $R^*_{q,0}$ conditional on $\{\nu > 0\}$ is ϕ. Hence, by induction, it follows that the distribution of $R^*_{q,n}$ conditional on $\{N^*_{q,A,G} > n, \nu = k, k > n\}$ is ϕ. It follows (by conditioning on $R^*_{q,\nu-1}$) that $E(N^*_{q,A,G} - \nu \mid N^*_{q,A,G} \geq \nu, \nu)$ is constant in ν for $\nu \geq 1$.

Let $0 < c < \infty$, $0 < p < 1$ be such that $N_{q,A}$ is the Bayes rule for $B(0, p, c)$. Regard now the Bayesian problem $B(G, p, c)$. Since the random value of $\gamma = P(\nu = 0)$ is assumed to be known prior to taking observations, the Bayes rule for the problem $B(G, p, c)$ is $N^*_{q,A,G}$. In a manner similar to the proof of Theorem 1, choose a subsequence $\{T_i, p_i, c_i, \phi_i, G_i\}_{i=1}^{\infty}$ such that, as $i \to \infty$, $p_i \to 0$, $c_i \to c^*$ and ϕ_i converges in distribution to a limit ψ^*, where ϕ_i is a fixed point of the $T = T_i$ defined by A and by $q_i = 1 - p_i$ as in (4), G_i is derived from ϕ_i as G is derived from ϕ in (6), and such that $N^*_{q_i,A,G_i}$ is the Bayes rule for the problem $B(G_i, p_i, c_i)$. Note (using, for instance, Prohorov metric on the set of probability measures on $[0, A]$) that $T_i(F) \to T(F)$ uniformly in F where T is defined by A, $q = 1$. Therefore $\phi_i = T_i(\phi_i) \to_{i\to\infty} T(\psi^*)$ and so $\psi^* = T(\psi^*)$. It follows that ψ^* has no atoms. Note that $G_i \to_{i\to\infty} \psi^*$, and therefore one can imagine the sequence of problems $B(G_i, p_i, c_i)$ with their respective $R^*_{q_i,0}$ to be embedded in a probability space in such a way that $R^*_{q_i,0} \to^{\text{a.s.}}_{p_i\to 0} R^*_0$, so that N^*_{1,A,ψ^*} is the limit as $i \to \infty$ of the Bayes rules $N^*_{q_i,A,G_i}$. Note that $P(N^*_{q_i,A,G_i} \geq 1) = 1$. Likewise, since ψ^* has no atoms, $N^*_{1,A,\psi^*} = N^*_{A,\psi^*}$ and so $P(N^*_{A,\psi^*} \geq 1) = 1$.

LEMMA 12. *Let*

$$m(k) = \begin{cases} \dfrac{\int_0^A x\, d\psi^*(x)}{\int_0^A x\, d\psi^*(x) + E_\infty N^*_{A,\psi^*}}, & k = 0 \\[2ex] \dfrac{P_\infty(N^*_{A,\psi^*} \geq k)}{\int_0^A x\, d\psi^*(x) + E_\infty N^*_{A,\psi^*}}, & k = 1, 2, \cdots. \end{cases}$$

If one uses $N^*_{q_i,A,G_i}$ *in the problem* $B(G_i, p_i, c_i)$ *then* $P(\nu = k \mid N^*_{q_i,A,G_i} \geq \nu) \to m(k)$ *as* $i \to \infty$. *Also, if one uses* N^*_{A,ψ^*} *instead of* $N^*_{q_i,A,G_i}$ *in problem* $B(G_i, p_i, c_i)$ *then* $P(\nu = k \mid N^*_{A,\psi^*} \geq \nu) \to m(k)$ *as* $i \to \infty$.

PROOF. Note that $P(\nu = 0)/p_i \to \int_0^A x\, d\psi^*(x)$ as $i \to \infty$. If one uses $N^*_{q_i,A,G_i}$ in problem $B(G_i, p_i, c_i)$ then (by using an argument similar to the proof of Lemma 9)

$$\frac{P(N^*_{q_i,A,G_i} \geq \nu)}{p_i} \to_{i\to\infty} \int_0^A x\, d\psi^*(x) + E_\infty N^*_{A,\psi^*}. \tag{7}$$

Therefore

$$P(\nu = k \mid N^*_{q_i,A,G_i} \geq \nu) = \frac{P(N^*_{q_i,A,G_i} \geq k \mid \nu = k)P(\nu = k)}{P(N^*_{q_i,A,G_i} \geq \nu)} \to_{i\to\infty} m(k).$$

A similar argument applies when using N^*_{A,ψ^*} instead of $N^*_{q_i,A,G_i}$.

LEMMA 13.

$$\lim_{i\to\infty} \frac{1 - \{\textit{Expected loss using } N^*_{A,\psi^*} \textit{ for } B(G_i, p_i, c_i)\}}{p_i}$$

$$= \lim_{i\to\infty} \frac{1 - \{\textit{Expected loss using } N^*_{q_i,A,G_i} \textit{ for } B(G_i, p_i, c_i)\}}{p_i}$$

$$\left[\int_0^A x\, d\psi^*(x) + E_\infty N^*_{A,\psi^*}\right]\left\{1 - c^*\left[E_1(N^*_{A,\psi^*} - 1)\frac{E_\infty N^*_{A,\psi^*}}{\int_0^A x\, d\psi^*(x) + E_\infty N^*_{A,\psi^*}} + \frac{\int_0^A x\, d\psi^*(x)}{\int_0^A x d\psi^*(x) + E_\infty N^*_{A,\psi^*}} \lim_{i\to\infty} E(N^*_{A,\psi^*} \mid \nu = 0)\right]\right\}.$$

PROOF. Note that

$$\frac{1 - \{\text{Expected loss using } N \text{ for problem } B(\psi, p, c)\}}{p} \tag{8}$$
$$= \frac{P(N \geq \nu)}{p}[1 - cE(N - \nu \mid N \geq \nu)].$$

For any n, the distribution of R_n^* conditional on $\{N_{A,\psi^*}^* > n, \nu > n\}$ is the limit of the distribution of $R_{q_i,n}^*$ conditional on $\{N_{q_i,A,G_i}^* > n, \nu > n\}$ as $i \to \infty$. The limits $\lim_{i\to\infty} E(N_{A,\psi^*}^* \mid \nu = 0)$, $\lim_{i\to\infty} E(N_{q_i,A,G_i}^* \mid \nu = 0)$ exist and are equal. One can compute $E(N_{q_i,A,G_i}^* \mid \nu = 0)$ by conditioning on $R_{q_i,0}^*$, and one can compute its limit (as $i \to \infty$) by noting (as can be shown by direct computation) that the limiting distribution of $R_{q_i,0}^*$ conditional on $\{\nu = 0\}$ has the density $x d\psi^*(x)/\int_0^A t\, d\psi^*(t)$. (This is also the limiting distribution of R_0^* conditional on $\{\nu = 0\}$.) Therefore, after computing $E(N_{A,\psi^*}^* - \nu \mid N_{A,\psi^*}^* \geq \nu)$ by conditioning on ν and on R_0^* or $R_{\nu-1}^*$ for $\nu = 0$ or $\nu \geq 1$, respectively, Lemma 13 follows from (8) and (7).

LEMMA 14. *$\int_0^A x\, d\psi^*(x)/[\int_0^A x\, d\psi^*(x) + E_\infty N_{A,\psi^*}^*] = O((\log A)/A)$ where $O((\log A)/A)/((\log A)/A)$ remains bounded as $A \to \infty$.*

PROOF. Since $R_{n+1}^* = (1 + R_n^*)f_1(X_{n+1})/f_0(X_{n+1})$, it follows that $E_\infty(R_{n+1}^* \mid F_n) = 1 + R_n^*$, and therefore $R_n^* - n$ is a P_∞-martingale with expectation $E_\infty(R_n^* - n) = E_\infty R_0^* = \int_0^A x\, d\psi^*(x)$.

From the optional sampling theorem it now follows that $\int_0^\infty x\, d\psi^*(x) = E_\infty R_{N_{A,\psi^*}^*}^* - E_\infty N_{A,\psi^*}^*$. Hence $\int_0^\infty x\, d\psi^*(x) + E_\infty N_{A,\psi^*}^* = E_\infty R_{N_{A,\psi^*}^*}^* \geq A$.

By a method analogous to the proof of Lemma 5, one can show that for all n

$$\psi^*(x) = P_\infty(R_n^* \leq x \mid N_{A,\psi^*}^* > n)$$
$$\geq P_\infty(R_n^* \geq x) \to_{n\to\infty} \lim_{n\to\infty} P_\infty(R_n \leq x) = P(H_1 \leq x)$$

where the limit follows from the fact that $R_n^* - R_n = R_0^* \exp\{\sum_{i=1}^n Z_i\} \to 0$ a.s. P_∞ as $n \to \infty$. Hence

$$\int_0^A x\, d\psi^*(x) = \int_0^A \left[1 - \int_0^x d\psi^*(t)\right] dx \leq \int_0^A P(H_1 > x)\, dx.$$

By Kesten (1973), Theorem 5, $xP(H_1 > x) \to 1$ as $x \to \infty$.

It follows that $\lim_{A\to\infty} \int_0^A x\, d\psi^*(x)/\log A \leq 1$, from which Lemma 14 follows.

To prove Theorem 2, note that for all $1 \leq \nu < \infty$

$$E_\nu(N_{A,\psi^*}^* - \nu \mid N_{A,\psi^*}^* \geq \nu) = E_1(N_{A,\psi^*}^* - 1).$$

Suppose there existed a sequence $\{A_j\}_{j=1}^\infty$, $A_j \to \infty$ as $j \to \infty$, and a sequence of stopping rules $\{M_j\}_{j=1}^\infty$ such that $P(M_j \geq 1) = 1$, $E_\infty M_j \geq E_\infty N_{A_j,\psi_j^*}^*$ (where ψ_j^* is the ψ^* of A_j) and that there existed $\eta > 0$ such that $\sup_{1\leq\nu<\infty} E_\infty(M_j - \nu \mid M_j \geq \nu) < E_1(N_{A_j,\psi_j^*}^* - 1) - \eta$. Without loss of generality, one can assume (by defining

R_0^{**} to be distributed as R_0^* and to be independent of R_0^*, $R_{q_i,0}^*$, $i = 1, 2, \cdots$, and applying M_j to the sequence R_0^{**}, X_1, X_2, $\cdots$) that M_j is independent of R_0^*, $R_{q_i,0}^*$, $i = 1, 2, \cdots$.

Note that

$$N_{A,\psi^*}^* \le N_A \le \min\left\{n \,\middle|\, \max_{1\le k\le n} \frac{f_1(X_1), \cdots, f_1(X_n)}{f_0(X_1), \cdots, f_0(X_n)} \ge A\right\}$$

which is Page's stopping rule. Hence (cf. Lorden, 1971) $E_1 N_{A,\psi^*}^* = O(\log A)$, where $O(\log A)/\log A$ remains bounded and bounded away from zero as $A \to \infty$. By Lemma 14, $\int_0^A x\, d\psi^*(x)/[\int_0^A x\, d\psi^*(x) + E_\infty N_{A,\psi^*}^*] = O((\log A)/A)$, and so, for large enough j,

$$\int_0^{A_j} x\, d\psi_j^*(x) \bigg/ \left[\int_0^{A_j} x\, d\psi_j^*(x) + E_\infty N_{A_j,\psi_j^*}^*\right] < \frac{\eta}{E_1 N_{A_j,\psi_j^*}^*}.$$

Letting $\{p_i, G_i, c_i\}_{i=1}^\infty$ be as in Lemma 12 for $A = A_j$, note that when using M_j in problem $B(G_i, p_i, c_i)$

$$\begin{aligned}
&E(M_j - \nu \mid M_j \ge \nu)\\
&\quad = E(M_j - \nu \mid M_j \ge \nu > 0)P(M_j \ge \nu > 0) + E(M_j \mid \nu = 0)P(\nu = 0)\\
&\to_{i\to\infty} \lim_{i\to\infty} E(M_j - \nu \mid M_j \ge \nu > 0)\, \frac{E_\infty M_j}{\int_0^{A_j} x\, d\psi_j^*(x) + E_\infty M_j}\\
&\quad + [1 + E(M_j - 1 \mid \nu = 1)]\, \frac{\int_0^{A_j} x\, d\psi_j^*(x)}{\int_0^{A_j} x\, d\psi_j^*(x) + E_\infty M_j}\\
&\quad < E_1(N_{A_j,\psi_j^*}^* - 1) - \eta + \frac{\int_0^{A_j} x\, d\psi_j^*(x)}{\int_0^{A_j} x\, d\psi_j^*(x) + E_\infty M_j}\\
&\quad \le E_1(N_{A_j,\psi_j^*}^* - 1) - \eta + \frac{\int_0^{A_j} x\, d\psi_j^*(x)}{\int_0^{A_j} x\, d\psi_j^*(x) + E_\infty N_{A_j,\psi_j^*}^*}\\
&\quad \le E_1(N_{A_j,\psi_j^*}^* - 1) - E_1 N_{A_j,\psi_j^*}^*\, \frac{\int_0^{A_j} x\, d\psi_j^*(x)}{\int_0^{A_j} x\, d\psi_j^*(x) + E_\infty N_{A_j,\psi_j^*}^*}\\
&\quad + \frac{\int_0^{A_j} x\, d\psi_j^*(x)}{\int_0^{A_j} x\, d\psi_j^*(x) + E_\infty N_{A_j,\psi_j^*}^*}\\
&\quad = E_1(N_{A_j,\psi_j^*}^* - 1)\, \frac{E_\infty N_{A_j,\psi_j^*}^*}{\int_0^{A_j} x\, d\psi_j^*(x) + E_\infty N_{A_j,\psi_j^*}^*}\\
&\quad \le \lim_{i\to\infty} E(N_{A_j,\psi_j^*}^* - \nu \mid N_{A_j,\psi_j^*}^* \ge \nu).
\end{aligned}$$

It would follow that if j and i are large enough, M_j would do better than the Bayes rule N_{q_i,A,G_i}^* for problem $B(G_i, p_i, c_i)$, which obviously cannot be the case.

Therefore

$$\lim_{A\to\infty}[\sup_{1\le\nu<\infty}E(N^*_{A,\psi^*}\ge\nu) - \inf_{\{N \mid E_\infty N\ge E_\infty N^*_{A,\psi^*}\}}\sup_{1\le\nu<\infty}E(N-\nu \mid N\ge\nu)] = 0.$$

To complete the proof of Theorem 2, it is only left to establish a connection between B and A.

LEMMA 15. *Let $\delta = \Delta/(1-\Delta)$ and let P^x denote the P_∞-distribution of the sequence $R^*_0, R^*_1, R^*_2, \cdots$ given that $R^*_0 = x$. There exists $\delta < x < A$ and a constant $0 < w$ such that for all $0 \le n < \infty$*

$$P^x(N^*_A > n) \ge wP^\delta(N^*_A > n)$$

where $N^*_A = \min\{n \mid n \ge 0, R^*_n \ge A\}$.

PROOF. There exists $\delta < x < A$, $0 < w$ such that

$$P^A\{R_1 < A\} - P^\delta(R_1 \le x)/P^\delta(R_1 > x) = w.$$

Now

$$\frac{P^x(N^*_A > n+1)}{P^x(N^*_A > n)} \tag{9}$$
$$= \int_{[0,A)} \frac{P^x(N^*_A > n,\, R^*_n \in dz)P^z(R^*_1 < A)}{P^x(N^*_A > n)} \ge P^A(R^*_1 < A) > 0.$$

Also,

$$P^\delta(N^*_A \le n+1) \ge P^\delta(R^*_1 \ge x)P^x(N^*_A \le n) + P^\delta(R^*_1 < x)P^\delta(N^*_A \le n)$$
$$= P^\delta(R^*_1 \ge x)P^x(N^*_A \le n) + P^\delta(R^*_1 < x)$$
$$[P^\delta(N^*_A \le n+1) - P^\delta(N^*_A = n+1)].$$

Therefore

$$P^\delta(N^*_A \le n+1) \ge P^x(N^*_A \le n) - \frac{P^\delta(R^*_1 < x)}{P^\delta(R^*_1 \ge x)} P^\delta(N^*_A = n+1)$$

or

$$P^\delta(N^*_A > n+1) \le P^x(N^*_A > n) + \frac{P^\delta(R^*_1 < x)}{P^\delta(R^*_1 \ge x)} P^\delta(N^*_A = n+1).$$

Thus by (9)

$$\frac{P^x(N^*_A > n)}{P^\delta(N^*_A > n)} \ge \frac{P^\delta(N^*_A > n+1)}{P^\delta(N^*_A > n)} - \frac{P^\delta(R^*_1 < x)}{P^\delta(R^*_1 \ge x)}\frac{P^\delta(N^*_A = n+1)}{P^\delta(N^*_A > n)}$$
$$\ge P^A(R^*_1 < A) - \frac{P^\delta(R^*_1 < x)}{P^\delta(R^*_1 \ge x)} = w.$$

This concludes the proof of Lemma 15. (This proof is due to David Siegmund.) Note that the proof utilized only the monotonicity properties of R_n^*.

LEMMA 16. *If* $T(\phi_1) = \phi_1$ *and* $T(\phi_2) = \phi_2$ *(for* T *defined as in Lemma* 11) *then* $\int_0^A \xi(t)\, d\phi_1(t) = \int_0^A \xi(t)\, d\phi_2(t)$.

PROOF. Suppose without loss of generality that $\int_0^A \xi(t)\, d\phi_1(t) > \int_0^A \xi(t)\, d\phi_2(t)$. Note that the P_∞-distributions of N_{A,ϕ_i}^* are geometric ($p_i = 1 - \int_0^A \xi(t)\, d\phi_i(t)$) respectively, $i = 1, 2$. Using the notation of Lemma 15, letting x and w be as in Lemma 15, it follows that for all $n \geq 1$,

$$\begin{aligned}\left(\int_0^A \xi(t)\, d\phi_2(2)\right)^n &= P_\infty(N_{A,\phi_2}^* > n) \geq \phi_2(x)P^x(N_A^* > n - 1)\\ &\geq \phi_2(x)P^x(N_A^* > n) \geq \phi_2(x)wP^\delta(N_A^* > n)\\ &\geq \phi_2(x)wP_\infty(N_{A,\phi_1}^* > n) = \phi_2(x)w\left(\int_0^A \xi(t)\, d\phi_1(t)\right)^n.\end{aligned}$$

Since clearly $\phi_2(x) > 0$, this will be violated for large enough n if $\int_0^A \xi(t)\, d\phi_1(t) > \int_0^A \xi(t)\, d\phi_2(t)$ as supposed above.

The following lemma will conclude the proof of Theorem 2.

LEMMA 17. (i) $E_\infty N_{A,\phi}^*$ *is the same for all measures* ϕ *satisfying* $T(\phi) = \phi$.

(ii) *Let* $\{\phi_A\}$, $\Delta/(1 - \Delta) < A < \infty$, *be a set of measures satisfying* $T_A(\phi_A) = \phi_A$ (*where* T_A *is the* T *of Lemma* 11 *for the value* A). *Then* $E_\infty N_{A,\phi_A}^*$ *is a continuous function of* A, *ranging from* 1 *to* ∞ *as* A *ranges from* $\Delta/(1 - \Delta)$ *to* ∞.

PROOF. Since the P_∞-distribution of $N_{A,\phi}^*$ is geometric with parameter $p = 1 - \int_0^A \xi(t)\, d\phi(t)$, (i) follows from Lemma 16.

Let $\Delta/(1 - \Delta) < A_0 < \infty$, and suppose that A_i, $i = 1, 2, \cdots$ is a sequence converging to A_0. Let ϕ_{i_j}, $j = 1, 2, \cdots$ be a converging sequence of probability measures on $[0, A_{i_j}]$ satisfying $\phi_{i_j} = T_{A_{i_j}}(\phi_{i_j})$, where $\{A_{i_j}\}_{j=1}^\infty$ is a subsequence of $\{A_i\}_{i=1}^\infty$. Denote the limiting measure by Γ. It is easy to see (again by employing Prohorov metric) that there exists $A^* \geq A_0$ such that $T_{A_{i_j}}(F) \to_{j\to\infty} T_{A_0}(F)$ uniformly in F whose support is contained in $[0, A^*]$. It follows that $T_{A_0}(\Gamma) = \Gamma$ and that

$$\int_0^{A_{i_j}} \xi_{A_{i_j}}(t)\, d\phi_{i_j}(t) \to \int_0^{A_0} \xi_{A_0}(t)\, d\Gamma(t)$$

($\xi_A(t)$ is the function $\xi(t)$ indexed by the A to which it corresponds), so that $E_\infty N_{A_{i_j},\phi_{i_j}}^* \to_{j\to\infty} E_\infty N_{A_0,\Gamma}^*$. Therefore (by virtue of (i)) $E_\infty N_{A,\phi_A}^*$ is continuous in A.

Clearly, $E_\infty N_{A,\phi_A}^* \to 1$ as $A \to \Delta/(1 - \Delta)$, and (by Lemma 14 and its proof) $E_\infty N_{A,\phi_A} \to \infty$ as $A \to \infty$.

This completes the proof of Theorem 2.

If the P_∞-distribution of $f_1(X_1)/f_0(X_1)$ has atoms, a similar result holds. What remains to be shown (so that the proof of Theorem 2 could be carried through) is basically that there exists a solution to (4). This can be shown by a continuity argument. Assume first that there are at least 2 atoms. Since the number of atoms is countable, one can assume without loss of generality that these atoms are at the x_1-points $1, 2, \cdots, n$ or $1, 2, \cdots$, where all the atoms are different, and all of the rest of the mass is concentrated on $(-\infty, 0]$. Define $\{U_{1i}, U_{2i}\}_{i=1}^{\infty}$ to be independent among themselves and independent of $X_1, X_2, \cdots$, where U_{1i} are Normal$(0, \sigma^2)$ variables conditioned on being in $[0, 1]$ and U_{2i} are Normal$(0, \sigma^2)$ variables conditioned on being in $[-1, 1]$.

Define (where possibly $n = \infty$):

$$Y_i = \begin{cases} X_i & \text{if } X_i \le 0 \\ X_i + U_{1i} & \text{if } X_i = 1 \\ X_i + U_{2i} & \text{if } 1 < X_i < n \\ X_i - U_{1i} & \text{if } X_i = n. \end{cases}$$

The P_∞-distribution of the likelihood ratio of Y_i has no atoms. Clearly, $Y_i \to X_i$ a.s. as $\sigma \to 0$. It is not difficult to see that one can choose a sequence $\sigma_i \searrow 0$ such that the respective $N^*_{A_i}$-type rules with expected P_∞-stopping time B converges a.s. P_ν, $\nu \ge 0$, to a rule of a form similar to that described in Theorem 2. From this convergence one gets the existence of a solution ϕ to (4). The property that possibly does not hold using the above argument is the fact that ϕ may have atoms, so the rule may randomize in case $R_n = A$. A continuity argument shows the resulting rule to be a Bayes risk-efficient limit of Bayes rules. The details are omitted.

If the P_∞-distribution of $f_1(X_1)/f_0(X_1)$ has exactly one atom (and is of mass less than one), assume, without loss of generality, that the atom is at $x_1 = 1$ and all of the rest of the mass is concentrated on $(-\infty, 0]$. Let $\{U_{1i}\}_{i=1}^{\infty}$ be independent and independent of $X_1, X_2, \cdots$; U_{1i} being Normal$(0, \sigma^2)$ variables conditioned in $[0, 1]$. Let z be such that $P_\infty(f_1(X_1)/f_0(X_1) \in (z - \varepsilon, z + \varepsilon)) > 0$ for all $\varepsilon > 0$, where $z \neq f_1(1)/f_0(1)$.

Define

$$Y_i = \begin{cases} 1 - U_1 & \text{if } X_i = 1 \\ U_{1i} & \text{if } X_i \le 0,\ (f_1(X_i)/f_0(X_i)) \in (z - \varepsilon, z + \varepsilon) \\ X_i & \text{if } X_i \le 0,\ (f_1(X_i)/f_0(X_i)) \notin (z - \varepsilon, z + \varepsilon). \end{cases}$$

An argument similar to the above can be carried through by decreasing ε and σ.

As a consequence, we get Theorem 3, where N^{**}_{A,ψ^*} is a stopping rule defined like N^*_{A,ψ^*} of Theorem 2, with the added possibility of randomizing as to whether to stop or continue sampling when $R^*_n = A$.

THEOREM 3. *Let f_0 and f_1 be such that*

$$E_\infty[(f_1(X_1)/f_0(X_1))\log^+(f_1(X_1)/f_0(X_1))] < \infty.$$

For every $1 < B < \infty$ there exists a value A, $\Delta/(1 - \Delta) \le A < \infty$, and a measure ψ^* such that $B = E_\infty N^{**}_{A,\psi^*}$ and such that if N is any stopping time which satisfies

$E_\infty N \geq B$ then

$$\sup_{1\le\nu<\infty} E_\nu(N - \nu \mid N \geq \nu) \geq \sup_{1\le\nu<\infty} E_\nu(N^{**}_{A,\psi^*} - \nu \mid N^{**}_{A,\psi^*} \geq \nu) + o(1)$$

where $o(1) \to 0$ as $B \to \infty$.

REMARKS. Under additional technical assumptions, the solution of (4) can be shown to be unique, and for any probability F on $[0, A]$, $T^n(F)$ converges as $n \to \infty$ to this solution. The methods required are those of Theorem 6.3 of Krein and Rutman (1948) and its proof. These assumptions hold, for instance, if f_0 and f_1 are $N(0, 1)$ and $N(\theta, 1)$ densities, respectively.

The requirement that

$$E_\infty[(f_1(X_1)/f_0(X_1))\log^+(f_1(X_1)/f_0(X_1))] < \infty$$

was used only in the proof of Lemma 14, enabling application of Theorem 5 of Kesten (1973).

4. Examples. It is generally difficult to find explicitly the measures ψ^* of Theorem 2 and Theorem 3. We discuss here two examples, one a discrete case (Bernoulli random variables) and one continuous (exponential), in which some solutions are presented and difficulties in obtaining a general solution are made apparent. For most other examples—even when the random variables are normal—an explicit solution is not apparent.

EXAMPLE 1. Suppose X_i are Bernoulli variables, $f_0(x) = p_0^x q_0^{1-x}$, $f_1(x) = p_1^x q_1^{1-x}$, $x = 0, 1$. Suppose $p_1 > p_0$. Then

$$\frac{f_1(x)}{f_0(x)} = \frac{p_1^x q_1^{1-x}}{p_0^x q_0^{1-x}} = \frac{q_1}{q_0}\left(\frac{p_1 q_0}{p_0 q_1}\right)^x$$

$$\delta = \Delta/(1 - \Delta) = \textstyle\sum_{n=1}^{\infty} (q_1/q_0)^n = q_1/(q_0 - q_1).$$

Let $A \geq \delta$. Regard $N^*_{1,A,\psi}$. For $\delta \leq t, x \leq A$

$$\rho(t, x)\xi(t) = P_\infty\left[\frac{q_1}{q_0}\left(\frac{p_1 q_0}{p_0 q_1}\right)^{X_1}(1 + t) \leq x\right] = P_\infty\left[X_1 \leq \log\left(\frac{q_0}{q_1}\frac{x}{1 + t}\right)\Big/\log\left(\frac{p_1 q_0}{p_0 q_1}\right)\right]$$

$$= \begin{cases} 1 & \text{if } p_1/p_0 \leq x/(1 + t) \\ q_0 & \text{if } q_1/q_0 \leq x/(1 + t) < (p_1/p_0) \\ 0 & \text{if } x/(1 + t) < q_1/q_0 \end{cases}$$

$$\xi(t) = \begin{cases} 1 & \text{if } p_1/p_0 \leq A/(1 + t) \\ q_0 & \text{if } q_1/q_0 \leq A/(1 + t) < p_1/p_0 \\ 0 & \text{if } A/(1 + t) < q_1/q_0 \end{cases}$$

$$T(F)(x) = \frac{q_0 F\left\{\left(x\frac{p_0}{p_1} - 1, x\frac{q_0}{q_1} - 1\right]\right\} + F\left(x\frac{p_0}{p_1} - 1\right)}{q_0 F\left\{\left(A\frac{p_0}{p_1} - 1, A\frac{q_0}{q_1} - 1\right]\right\} + F\left(A\frac{p_0}{p_1} - 1\right)}.$$

We want to solve $F = T(F)$. Clearly the solution must be fully concentrated on $\{\delta\}$ if $A < (p_1/p_0)(q_0/(q_0 - q_1))$.

Suppose now that

(10) $$(p_1/p_0)(q_0/(q_0 - q_1)) < A < (p_1q_0/p_0q_1)(q_0/(q_0 - q_1)) - 1.$$

Let w denote the denominator of $T(F)$. For

(11) $$\delta \le x \le (q_1/q_0)(A + 1) \qquad (\le (p_1/p_0)(q_0/(q_0 - q_1)))$$

solving for $F = T(F)$ yields

$$(w/q_0)F(x) = F(x(q_0/q_1) - 1)$$

(for x satisfying (11)). Hence $F(\delta) = 0$ (unless $w = q_0$, which cannot be the case for A satisfying (10)). Set $y = xq_0/q_1 - 1$; i.e. $x = (y + 1)q_1/q_0$. It follows that for $\delta \le y \le A$

(12) $$F((q_1/q_0)(y + 1)) = (q_0/w)F(y).$$

For

(13) $$(q_1/q_0)(A + 1) \le x \le (p_1/p_0)(q_0/(q_0 - q_1))$$

solving for $F = T(F)$ yields

(14) $$F(x) = q_0/w$$

(for x satisfying (13)). Let $a_1 = (A + 1)q_1/q_0$, $b_1 = (p_1/p_0)q_0/(q_0 - q_1)$, and define recursively for $i = 2, 3, \cdots$ $a_i = (a_{i-1} + 1)q_1/q_0$, $b_i = (b_{i-1} + 1)q_1/q_0$. Applying (12) to (14) recursively yields

$$F(x) = (q_0/w)^i$$

for $a_i \le x \le b_i$ (see Figure 1). For

(15) $$(p_1/p_0)(q_0/(q_0 - q_1)) \le x < A$$

solving for $F = T(F)$ yields

(16) $$F(x) = [q_0 + p_0F(x(p_0/p_1) - 1)]/w$$

(for x satisfying (15)). Applying (16), one is able to partially fill the void in Figure 1 for x satisfying (15) by a series of horizontal lines analogous to those in Figure 1. One similarly partially fills the void for $b_{i+1} < x < a_i$, $i = 1, 2, \cdots$ by employing (12). The voids still remaining are partially filled, recursively, in an analogous manner. Thus one arrives at an expression for F in terms of w. In particular, one arrives at an expression for $F(Ap_0/p_1 - 1)$ in terms of w, so that setting $x = A$ in (16) yields a solution for w. (For sample, if $p_0 = 1/3$, $p_1 = 2/3$,

$$A = [(p_1q_0)/(p_0q_1)]q_0/(q_0 - q_1) - 1,$$

a short calculation shows that $Ap_0/p_1 - 1 = a_2 = b_2$, so that $F(Ap_0/p_1 - 1) = F(a_2) = (q_0/w)^2$. Applying this to (16) gets w to be the positive solution of $w^3 - 2/3 w^2 - 4/27 = 0$.) One can show that the solution ψ^* of $F = T(F)$ is continuous,

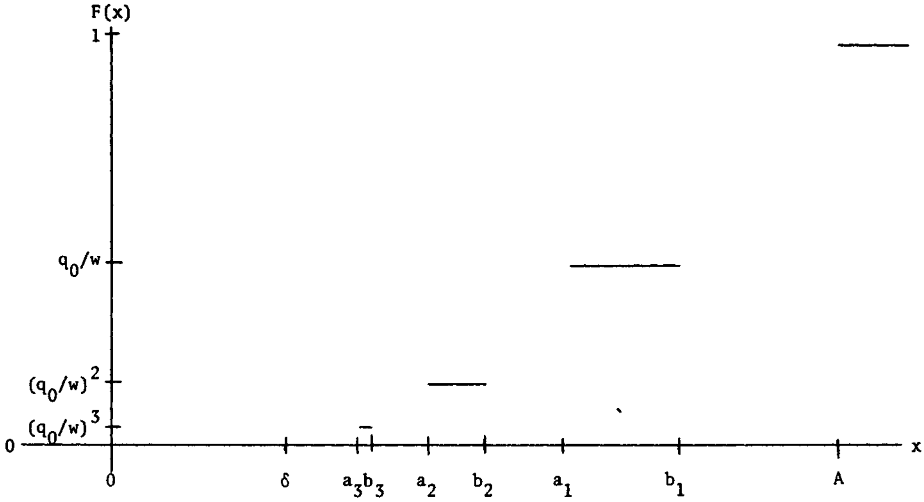

FIG. 1.

and so $N^*_{1,A,\psi^*} = N^*_{A,\psi^*}$ and the rule requires no randomization (whether to stop or continue sampling) when $R^*_n = A$. If $A = (p_1/p_0)q_0/(q_0 - q_1)$ then $q_0/w = 1$ so a solution ψ^* of $F = T(F)$ must be concentrated on $\{\delta\}$. (Note that this ψ^* is actually not a solution when regarding N^*_{1,A,ψ^*}, but is a solution when regarding N^*_{A,ψ^*}.) So again the rule does not randomize on $\{R^*_n = A\}$.

When B is such that to satisfy $B = E_\infty N^{**}_{A,\psi^*}$ one has to choose $A = \delta$, one may have to randomize on $\{R^*_n = A = \delta\}$.

EXAMPLE 2. $f_0(x) = e^{-x}1(x > 0)$, $f_1(x) = \theta e^{-\theta x}1(x > 0)$, $\theta > 1$. Here $f_1(x)/f_0(x) = \theta \exp\{(1-\theta)x\}$ for $x > 0$. Hence, $\Delta = 0$. So:

$$\rho(t, x)\xi(t) = P_\infty(\theta e^{(1-\theta)X_1}(1 + t) \leq x) = P_\infty\left(X_1 \geq \frac{1}{1-\theta} \log \frac{1}{\theta} \frac{x}{1+t}\right)$$

$$= \left(\frac{1}{\theta} \frac{x}{1+t}\right)^{1/(\theta-1)} \wedge 1$$

$$\xi(t) = \left(\frac{1}{\theta} \frac{A}{1+t}\right)^{1/(\theta-1)} \wedge 1.$$

Hence, for F concentrated on $[0, A]$,

$$T(F)(x) = \frac{((1/\theta)x)^{1/(\theta-1)} \int_{((1/\theta)x-1)\vee 0}^{A} (1/(1+t))^{1/(\theta-1)}\, dF(t) + F((1/\theta)x - 1)}{((1/\theta)A)^{1/(\theta-1)} \int_{((1/\theta)A-1)\vee 0}^{A} (1/(1+t))^{1/(\theta-1)}\, dF(t) + F((1/\theta)A - 1)}.$$

Thus, if $0 < A < \theta$ and $0 < x < A$ then

$$T(F)(x) = ((x/A))^{1/(\theta-1)}$$

and so

$$\psi^*(x) = \begin{cases} 0 & \text{if } x \le 0 \\ (x/A)^{1/(\theta-1)} & \text{if } 0 \le x \le A \\ 1 & \text{if } A \le x. \end{cases}$$

For larger values of A, a method similar in nature to that employed in Example 1 may be employed. For $\theta < 1$ an explicit solution is not apparent.

5. Remarks.

1. Theorem 2 and Theorem 3 are an improvement of the result of Lorden (1971), who computed the minmax value up to a $o(\log B)$ term, where $o(\log B)/\log B \to 0$ as $B \to \infty$. See also Remark 3 in the sequel.

2. The analogs of the statistics R_n and the stopping rule N_A presented in this article can be applied to much more general situations. In general, suppose P_∞ and $\{P_\nu\}_{\nu=1}^{\infty}$ are probability measures on the sample space of $X_1, X_2, X_3, \cdots$ such that the marginal P_ν-distribution of $X_1, \cdots, X_{\nu-1}$ is equal to the marginal P_∞-distribution of $X_1, \cdots, X_{\nu-1}$ for all $1 \le \nu < \infty$. Let $R^{\#}_{q,n} = \sum_{k=1}^{n} q^{-(n-k+1)}\, dP_k(X_1, \cdots, X_n)/dP_\infty(X_1, \cdots, X_n)$, let $R^{\#}_n = R^{\#}_{1,n}$ and suppose $\nu \sim$ Geometric (p), $p = 1 - q$. Then it is still true that $P(\nu \le n \mid \mathscr{F}_n) = R^{\#}_{q,n}/(R^{\#}_{q,n} + 1/p)$. Hence $N^{\#}_A = \min\{n \mid R^{\#}_n \ge A\}$ is the limit as $p \to \infty$ of $\min\{n \mid P(\nu \le n \mid \mathscr{F}_n) \ge pA\}$. While this in general may not be a limit of Bayes rules, under certain circumstances it may still be a good procedure.

As an example, consider the problem considered in Section 2 in the more practical case that f_0 is known but f_1 is not. Suppose it can be assumed that f_1 belongs to a subset of a parametric family $\{f_\theta\}_{\theta\in\Theta}$, and suppose that J is a prior on Θ. Then $R^{\#}_n$ becomes a mixture-type analog of R_n. It can be shown that the resulting rules $N^{\#}_A$ are optimal in the sense of Pollak (1978).

3. It is of practical importance to evaluate the operating characteristics of the proposed rules. As in the proof of Lemma 14, it can be shown that $R^{\#}_n - n$ is a P_∞-martingale with zero expectation, so that $E_\infty N^{\#}_A = E_\infty R^{\#}_{N_A} \approx A$ may hold. A better approximation may be obtained by evaluating the overshoot $R^{\#}_{N_A}/A$. This is done in Pollak (1983) for the stopping rules N_A considered in Section 2 and their mixture-type analogs. Operating characteristics of the average run length after a change occurs are also derived in Pollak (1983), and these can be used to obtain explicit expressions for the asymptotic minmax value of N^*_A of Section 2.

4. A theory for Brownian motion is presented in Pollak and Siegmund (1984).

Acknowledgements. The author would like to thank Professors Peter Bickel, Abba Krieger, David Siegmund, Benjamin Weiss, and a referee for a number of very helpful discussions and suggestions.

REFERENCES

AROIAN, L. A. and LEVENE, H. (1950). The effectiveness of quality control charts. *J. Amer. Statist. Assoc.* **45** 520–529.

BHATTACHARYA, P. K. and FRIERSON, D. (1981). A nonparametric chart for detecting small disorders. *Ann. Statist.* **9** 544–554.

CHEN, R. (1978). A surveillance system for congenital malformations. *J. Amer. Statist. Assoc.* **73** 323–327.

CHERNOFF, H. and ZACKS, S. (1964). Estimating the current mean of a normal distribution which is subjected to a change in time. *Ann. Math. Statist.* **35** 999–1018.

DUDDING, B. P. and JENNETT, W. J. (1944). *Quality Control Chart Technique.* General Electric, London.

DUNFORD, N. and SCHWARTZ, J. T. (1958). *Linear Operators, Part 1.* Interscience, New York.

FERGUSON, T. S. (1967). *Mathematical Statistics: A Decision Theoretic Approach.* Academic, New York.

GIRSHICK, M. A. and RUBIN, H. (1952). A Bayes approach to a quality control model. *Ann. Math. Statist.* **23** 114–125.

JOHNSON, N. L. and LEONE, F. C. (1962). Cumulative sum control charts: mathematical principles applied to their construction and use. Parts I, II, III. *Industrial Quality Control* **18** 15–21; **19** 29–36; **20** 22–28.

KANDER, Z. and ZACKS, S. (1966). Test procedures for possible changes in parameters of statistical distributions occurring at unknown time points. *Ann. Math. Statist.* **37** 1196–1210.

KENETT, R. and POLLAK, M. (1983). On sequential detection of a shift in the probability of a rare event. *J. Amer. Statist. Assoc.* **78** 389–395.

KESTEN, H. (1973). Random difference equations and renewal theory for products of random matrices. *Acta Math.* **131** 207–248.

KREIN, M. D. and RUTMAN, M. A. (1948). Linear operators leaving invariant a cone in Banach space. *Trans. Amer. Math. Soc.*, Ser. 1, **10** 199–325.

LORDEN, G. (1971). Procedures for reacting to a change in distribution. *Ann. Math. Statist.* **42** 1897–1908.

PAGE, E. S. (1954). Continuous inspection schemes. *Biometrika* **41** 100–115.

PAGE, E. S. (1955). Control charts with warning lines. *Biometrika* **42** 243–257.

POLLAK, M. and SIEGMUND, D. (1975). Approximations to the expected sample size of certain sequential tests. *Ann. Statist.* **6** 1267–1282.

POLLAK, M. (1978). Optimality and almost optimality of mixture stopping rules. *Ann. Statist.* **9** 910–916.

POLLAK, M. (1984). Average run lengths of an optimal method of detecting a change in distribution. Department of Statistics, Stanford University Technical Report.

POLLAK, M. and SIEGMUND, D. (1984). A diffusion process and its application to detecting a change in the drift of a Brownian motion process. Department of Statistics, Stanford University Technical Report.

ROBERTS, S. W. (1966). A comparison of some control chart procedures. *Technometrics* **8** 411–430.

SHEWHART, W. A. (1931). *The Economic Control of the Quality of Manufactured Product.* Macmillan, New York.

SHIRYAYEV, A. N. (1963). On optimum methods in quickest detection problems. *Theory Probab. Appl.* **8** 22–46.

SHIRYAYEV, A. N. (1964). On Markov sufficient statistics in non-additive Bayes problems of sequential analysis. *Theory Probab. Appl.* **9** 604–618.

SHIRYAYEV, A. N. (1978). *Optimal Stopping Rules.* Springer-Verlag, New York.

WEILER, K. (1954). A new type of control chart limits for means, ranges and sequential runs. *J. Amer. Statist. Assoc.* **49** 298–314.

DEPARTMENT OF STATISTICS
HEBREW UNIVERSITY
JERUSALEM 91905
ISRAEL