The Annals of Statistics
1993, Vol. 21, No. 3, 1155–1165

DETECTING A CHANGE OF A NORMAL MEAN BY DYNAMIC SAMPLING WITH A PROBABILITY BOUND ON A FALSE ALARM

BY DAVID ASSAF[1], MOSHE POLLAK[2], YA'ACOV RITOV[1] AND BENJAMIN YAKIR

Hebrew University

We show that when dynamic sampling is feasible, there exist surveillance schemes for which the probability of a false alarm is bounded and which have a bounded expected delay when detecting a (true) change. In the case of detecting a change of a normal mean, we probe optimality and suggest procedures. These procedures compare favorably to those having a fixed sampling rate which have been developed for an expectation constraint on the average run length until a false alarm.

1. Introduction and summary. The classical change-point problem consists of minimizing the time to detect a change subject to a constraint on false alarms. The usual setup is one where independent observations $X_1, X_2, \dots, X_{\nu-1}$ are identically distributed according to some F_0, and ensuing observations $X_\nu, X_{\nu+1}, \dots$ are iid with distribution F_1, where the change-point ν is unknown.

A detection scheme consists of a stopping time T for the process $\{X_1, X_2, \dots\}$ at which one stops and declares a change to have occurred. The speed of detection of a scheme may be measured by the maximal expected delay $\sup_{1 \le \nu < \infty} E(T - \nu | T \ge \nu)$. The classical formulation for containing the rate of false alarms is an expectation constraint; that is, T is required to satisfy $E(T|\nu = \infty) \ge B$ for some prespecified (large) constant B.

A more restrictive, though appealing, condition on the rate of false alarms is a probability constraint: $P_{\nu=\infty}(T < \infty) \le \alpha$ where $0 \le \alpha < 1$ is a given constant. It turns out that all of the classical procedures for detecting a change (Shewhart, CUSUM, Shiryayev-Roberts etc.) have $P(T < \infty | \nu = \infty) = 1$. The reason for foregoing a probability constraint is that the price is too high; the expected delay is unbounded if $\alpha < 1$. [This was noticed by Pollak and Siegmund (1975). Formally, this can be derived from Theorem 2 of Pollak (1985).] In fact, perhaps also for this reason, apart from Pollak and Siegmund's 1975 work, no procedures have been formulated for a non-Bayesian setting of detecting a change under a probability constraint.

Received April 1989; revised June 1992.
[1]Supported by a grant from the Israel Foundation Trustees.
[2]Partially supported by the United States–Israel Binational Science Foundation.
AMS 1991 *subject classifications*. Primary 62L10; secondary 62F05.
Key words and phrases. Quality control, control charts, change-point detection, dynamic sampling.

The state of affairs for the continuous time version is similar.

Assaf (1988) introduced dynamic sampling to the surveillance context. In certain situations, one can control the amount of sampling. For example, in an ice cream production context, at each point in time one can sample an arbitrary amount of ice cream; or if sampling continuously, the sampling rate can be varied. It is reasonable to allow such variations, as long as the average sampling rate is fixed (or bounded). Assaf (1988) and Assaf and Ritov (1989) noticed that employing such dynamic sampling can result in a dramatic reduction of the expected delay $\sup_{1 \le \nu < \infty} E(T - \nu | T \ge \nu)$. In the non-Bayesian setting, their work considers only stopping rules T satisfying $P(T < \infty | \nu = \infty) = 1$.

Here we reconsider the surveillance problem under a probability constraint, when dynamic sampling is feasible. The main result of this paper is that when dynamic sampling is feasible, one can satisfy a probability constraint yet keep the expected delay bounded. In fact, the expected delay is significantly smaller than the corresponding delay for classical procedures with an expectation bound. This is not given free, since one still may need to sample an increasing (ad infinitum) amount as $\nu \to \infty$. The reason why this is feasible is that dynamic sampling allows sampling plans that do an unbounded amount of sampling in a fixed amount of time.

The details are given here in case the observations are Brownian motion and one is alert to a change in the mean parameter. We first study the case where average sampling is required to be bounded only as long as there is no change in mean (Section 2). In this case it is shown that it is possible to devise an optimal procedure under a probability bound $\alpha = 0$. An explicit formula is given for the resulting expected delay. Though not applicable in real circumstances, this case is of theoretical interest. A comparison is made to dynamic sampling schemes satisfying expectation constraints and to classical sampling schemes.

In some situations, one may be obligated to maintain a bound on the average sampling rate, both before and after a change. This situation is studied in Section 3 (where a typical example in which these constraints arise naturally is discussed). Dynamic sampling schemes satisfying a probability constraint of $0 < \alpha < 1$ are considered. A scheme is constructed for which the expected delay is for all $0 < \alpha < 1$ within a constant of the lowest possible (α-dependent) expected delay.

2. Prechange average sampling constraint.

2.1. *Continuous sampling at equidistant time periods.* We begin with a model in which sampling may be done on a continuous basis at discrete equally spaced points in time. (This may reasonably describe hourly inspection of an ice cream production process.) Denote the sampling points $n = 1, 2, \dots$. Suppose that, should one be sampling an infinite amount at each period, one would be observing independent Brownian motions with unit variance parameter, with zero mean before a change, and with mean parameter $\mu_1 > 0$ (of known value) after a change. Let ν denote the first sampling period for which the mean parameter is μ_1. Let $\{B_1(t)\colon 0 \le t < \infty\}, \{B_2(t)\colon 0 \le t < \infty\}, \dots$ be the observable Brownian motions at the sampling periods $1, 2, \dots,$ respectively. In the classical fixed sampling rate setting, the part of $\{B_i(t)\colon 0 \le t < \infty\}$ which will be observed is $B_i(s_0)$, where s_0 is a constant reflecting the fixed amount sampled at each period. In the notation of Section 1, $X_i = B_i(s_0)$. The dynamic sampling setup allows variation on the part of $\{B_i(t)\colon 0 \le t < \infty\}$ actually observed, resulting in an observed sample path $\{B_i(t)\colon 0 \le t \le t_i\}$. The interpretation of t_i is the sampling time of observation at the real time period i. Formally, t_i is a stopping time for $B_i(t)$ such that $\{t_i > t\}$ is in the σ-field generated by $\{B_j(s)\colon 0 \le s \le t_j,\ j = 1, \dots, i-1$ and $B_i(s),\ 0 \le s \le t\}$. Let $0 < \gamma < \infty$ be a bound on the average sampling rate, in the following sense: Allowable sampling schemes are such that as long as no change occurs, the expected sampling time until (the end of) sampling period n may not exceed γn, $n = 1, 2, \dots$. In other words, we require the stopping variable sequence $t_1, t_2, \dots$ to satisfy

$$E_\nu \sum_{i=1}^{n} t_i \le \gamma n, \qquad n = 1, 2, \dots, \nu - 1,\ \nu = 1, 2, \dots,$$

where E_ν (and P_ν) represents expectation (respectively, probability) when the first real time period at which the change is already in effect is ν. ($\nu = \infty$ denotes that no change ever occurs.) This requirement is equivalent to

$$(1) \qquad E_{\nu=\infty} \sum_{i=1}^{n} t_i \le \gamma n, \qquad n = 1, 2, \dots .$$

In certain situations one may be willing to sample at an enormous rate after a change takes place so as to detect the change quickly. A typical situation of this type occurs when production following the change is worthless (e.g., medicine failing to meet required standards). In this section, we model this by not requiring the rate of sampling after a change takes place to be bounded. For this case, consider first the constraint $\alpha = 0$. The case $\alpha > 0$ is discussed in subsection 2.3. Let N be an integer-valued stopping time with respect to the observed process; that is, $\{N > n\}$ is in the σ-field generated by $\{B_i(t)\colon 0 \le t \le t_i;\ 1 \le i \le n\}$. Denote

$$\Gamma = \{(N, \bar{t}) : \bar{t} \text{ satisfies (1) and } P_{\nu=\infty}(N < \infty) = 0\}.$$

THEOREM 1.

(a) $$\inf_{(N,\bar{t})\in\Gamma} \ \sup_{1\le\nu<\infty} E_\nu(N - \nu | N \ge \nu) = e^{-(1/2)\mu_1^2\gamma}/(1 - e^{-(1/2)\mu_1^2\gamma}).$$

(b) *The infimum in* (a) *is attained by* $(N, \bar{t})$ *such that*

$$\begin{aligned} t_n &= \min\{t\colon B_n(t) \le \tfrac{1}{2}\mu_1 t - \tfrac{1}{2}\mu_1\gamma\} \\ &= \infty \quad \textit{if no such } t \textit{ exists}, \\ N &= \min\{n : t_n = \infty\} \\ &= \infty \quad \textit{if no such } n \textit{ exists}. \end{aligned}$$

PROOF. Clearly

$$\inf_{(N,\bar{t})\in\Gamma}\ \sup_{1\le\nu<\infty} E_\nu(N-\nu|N\ge\nu) \ge \inf_{(N,\bar{t})\in\Gamma} E_{\nu=1}N-1. \tag{2}$$

We will calculate the right-hand side of (2), and then show that the inequality in (2) is an equality. Let $(N,\bar{t})\in\Gamma$ and define

$$q_n \equiv P_{\nu=1}\Big(\max_{i=1,\dots,n} t_i<\infty\Big).$$

Note that

$$P_{\nu=\infty}\Big(N\le n,\ \max_{i=1,\dots,n} t_i<\infty\Big)=0 \quad\Leftrightarrow\quad P_{\nu=1}\Big(N\le n,\ \max_{i=1,\dots,n} t_i<\infty\Big)=0.$$

Since $P_{\nu=\infty}(N<\infty)=0$, it follows that

$$P_{\nu=1}(N>n)\ge P_{\nu=1}\Big(\max_{i=1,\dots,n} t_i<\infty\Big)=q_n.$$

Therefore

$$E_{\nu=1}N-1\ge\sum_{n=1}^{\infty} q_n.$$

Let $W(t)$ be a Brownian motion with drift parameter μ. Consider testing H_0: $\mu=\mu_1$ versus H_1: $\mu=0$. For fixed n, consider the test which applies t_j to $W(\Sigma_{i=1}^{j-1}t_i+t)-W(\Sigma_{i=1}^{j-1}t_i)$: $t\ge 0$, $j=1,\dots,n$, and which rejects H_0 when $\max_{i=1,\dots,n} t_i<\infty$. Because (1) is satisfied, it follows that this is a test of power one and of significance level q_n. When finite, the sampling time of this test is $\Sigma_{i=1}^{n}t_i$ and so its H_1-expected sampling time is $E_{\nu=\infty}\Sigma_{i=1}^{n}t_i$. Among all power one tests with this H_1-expected sampling time, the one with the lowest significance level is a power one sequential probability ratio test. This has as its stopping time

$$\begin{aligned} T_n &= \min\{t: W(t)-\tfrac{1}{2}\mu_1 t\le -c_n\}\\ &=\infty \quad \text{if no such } t \text{ exists,}\end{aligned}$$

where c_n is a constant such that $E_{H_1}T_n=E_{\nu=\infty}\Sigma_{i=1}^{n}t_i$.

Now

$$P_{H_0}(T_n<\infty)=e^{-\mu_1 c_n}$$

and

$$E_{H_1}T_n=2c_n/\mu_1$$

so that

$$P_{H_0}(T_n<\infty)=\exp\big[-(1/2)\mu_1^2 E_{H_1}T_n\big].$$

Hence

$$q_n\ge\exp\Big[-(1/2)\mu_1^2 E_{\nu=\infty}\sum_{i=1}^{n}t_i\Big]\ge\exp\big[-(1/2)\mu_1^2\gamma n\big]$$

and so

$$E_{\nu=1}N-1\ge\sum_{n=1}^{\infty}e^{-(1/2)\mu_1^2\gamma n}=\frac{e^{-(1/2)\mu_1^2\gamma}}{1-e^{-(1/2)\mu_1^2\gamma}}. \tag{3}$$

The $(N, \bar{t})$ described in Theorem 1(b) attains equality in (3). Furthermore for this $(N, \bar{t})$, $E_\nu(N - \nu | N \geq \nu)$ is constant in $1 \leq \nu < \infty$, so that equality is attained in (2). □

REMARK. The proof of Theorem 1 depends only on the fact that $W(t)$ is a process with stationary and independent increments which increases continuously in t. Therefore, if, for example, the observable process is a Poisson process instead of Brownian motion, and one is alert to a change of the intensity from λ_0 to $\lambda_1 < \lambda_0$, the same proof will yield an analog of Theorem 1.

2.2. *Continuous sampling in continuous real time.* Suppose one may sample continuously, being permitted to vary the sampling rate subject to a bound γx on the P_∞-expected sampling time during the first x time units, $0 \leq x < \infty$. [A rigorous treatment involves expressing variations in sampling rate in terms of the instantaneous variance; cf. Assaf (1988).]

Regard three statisticians: Statistician A may sample continuously as above, and receives the information as it arrives. Statistician B may view the information gathered by Statistician A in a given time interval $((n-1)\delta, n\delta]$ (in the chronological sequence at which it was gathered) only at the time point $n\delta$, $n = 1, 2, \dots$. Statistician C may sample (and view) only at times $n\delta$, subject to a bound $n\gamma/\delta$ on the P_∞-expected sampling time during the first n sampling points $(\delta, 2\delta, \dots, n\delta)$. To emphasize the differences between the statisticians, note that Statisticians A and B have the option of behaving like Statistician C. This can be done by sampling only at times $x = n\delta$ (subject to the appropriate constraints on the expected total amount sampled).

Note that the optimal expected delay of Statistician B does not exceed that of Statistician A by more than δ. Argue that since Statistician B should make decisions only at times $n\delta$, the best information for a decision is the most up-to-date information; hence the optimal behavior for Statistician B is to gather information only at points $n\delta$. More formally, sufficiency considerations imply that Statistician B pays no penalty for viewing the process only at points $n\delta$. In other words, Statisticians B and C should behave the same way. By virtue of the results of Section 2.1 (with time scale $n\delta$ instead of n), the optimal procedure for Statistician C is to make a power one SPRT (of H_0: $\mu = \mu_1$ vs. H_1: $\mu = 0$) every δ time units, such that the expected ($\nu = \infty$) sampling time per observation period is $\delta\gamma$. Let $\delta \to 0$. By Theorem 1, it follows that the infimal expected delay for Statistician A is

$$\lim_{\delta \to 0} \delta \frac{e^{-(1/2)\mu_1^2\gamma\delta}}{1 - e^{-(1/2)\mu_1^2\gamma\delta}} = \frac{2}{\mu_1^2\gamma}. \tag{4}$$

2.3. *Comments and comparisons to other schemes.* Assaf and Ritov (1989) study the non-Bayesian continuous time model with an expectation constraint of the type $E_{\nu=\infty}T \geq B$. The rules suggested in their paper are of the (A, c, δ) type: Perform an SPRT every $\delta > 0$ time units with lower stopping level $-c\delta$ and upper stopping level A. The stopping time T is defined as the first time that the upper level A is hit. The "best" procedure is attained in the limit as $\delta \to 0$, and is denoted as $(A, c, 0)$.

The (A, c, δ) procedures yield a (scaled) geometric distribution for T, thus forcing $\alpha = P_{\nu=\infty}(T < \infty)$ to equal 1. The procedure suggested in Section 2 above may be thought of as an (A, c, δ) procedure with $A = \infty$, yielding $\alpha = P_{\nu=\infty}(T < \infty) = 0$.

The expected delay with finite A will evidently be shorter. However, since B is usually set to be fairly high, the saving in expected delay is far from being dramatic. As an example, consider a typical case where $\gamma = 1$, $B = 793$ and $\mu_1 = 1$. By virtue of (3), the procedure suggested in this paper yields an expected delay of 2 with $\alpha = 0$. The $(A, c, 0)$ procedure has an expected delay of 1.96 and $\alpha = 1$ [Assaf and Ritov (1989)]—a saving of only 2% at the price of increasing α from 0 to 1.

Evidently, any procedure satisfying a probability constraint of $0 < \alpha < 1$ cannot yield a better expected delay than one with an expectation constraint $E_{\nu=\infty}T \geq B$, $B \to \infty$, which will result in the same expected delay of the procedure suggested in this paper which has $\alpha = 0$. It is however theoretically possible to devise procedures having any $0 < \alpha < 1$ by taking a time dependent upper limit $A(t)$ and using an obvious $(A(t), c, 0)$ procedure. Of course, $\alpha = 0$ cannot be attained by any practical procedure. A suitable function $A(t)$ [necessarily $\lim_{t\to\infty} A(t) = \infty$] may be chosen to obtain any specified value of $0 < \alpha < 1$. These procedures will yield improved expected delays for some value of ν but not for the supremum.

The reader should note that the fixed rate sampling procedures such as CUSUM and Shiryayev-Roberts yield expected delays of approximately 10 for the same problem [Pollak and Siegmund (1985)] and all of them have $\alpha = 1$.

3. Pre- and post-change average sampling constraints. In this section, we consider the situation in which the expected average sampling rate is required to be bounded all the time, including the period after the change. We envision a somewhat flexible sampling budget, so that the statistician may guardedly borrow against the future, as long as on the average the expenses remain within limits. We regard the situation in which surveillance will resume after an alarm is raised and the process is (putatively) reset to its in-control state, that is, $\mu = 0$. Therefore, we formally regard pairs $(N, \bar{t})$ as in the previous section, but with the added proviso that if $N = n$, then $(t_{n+1}, t_{n+2}, \dots)$ will have the same structure as $(t_1, t_2, \dots)$, and N will next be applied to $(t_{n+1}, t_{n+2}, \dots)$. We will call this a renewal structure. We replace (1) with the more restrictive set of constraints

$$(5)\qquad \begin{aligned} &E_{\nu=\infty} \sum_{i=1}^{n} t_i \leq \gamma n, \qquad n = 1, 2, \dots, \\ &E_{\nu} \sum_{i=1}^{N} t_i \leq \gamma E_{\nu} N, \qquad \nu = 1, 2, \dots. \end{aligned}$$

(Of course, one can contemplate even more restrictive constraints.) Since the repeated surveillance is in a sense a renewal process, the average sampling rate per observation will be bounded by γ in the long run if (5) is satisfied. [We stipulate a renewal structure because it seems reasonable from a practical point of view. Formally, Theorem 2 below and its proof will be valid if we drop the renewal structure requirement. However, the second requirement of (5) seems to make no sense unless we consider a renewal structure. If we assume a renewal structure, then it means that the long run average sampling rate is, indeed, *gamma*.]

Practical situations in which this scenario arises are conceivable. A typical case is when a false alarm has severe consequences, whereas sampling past the change-point is relatively tolerable. As a more specific example, consider a clinic conducting a trial using a new and delicate instrument to measure certain characteristics in a blood sample. The instrument is difficult to calibrate, and comparison of measurements made after different calibrations is difficult. Therefore, the instrument is calibrated once, and measurements on blood samples are made consecutively, until the instrument goes out of calibration. (As in all instruments, measurement has variance.) The clinic prepares a very large quantity of identical copies of a blood sample with known characteristics. To monitor the calibration level, the clinic may send any number of samples from the known lot to the laboratory for measurement. Clearly, there is a great premium for not stopping the measurement process in order to recalibrate when the process is in control. A probability bound on a false alarm would be an appropriate constraint. Dynamic sampling is clearly feasible. (While the sampling units in this scenario are discrete, if the measurement variance is large then continuous sampling would be a good approximation.) Since measurements made after the change-point will be discarded, one would want to stop and recalibrate as soon as possible after a change.

It is clearly not possible to satisfy an $\alpha = 0$ probability constraint and have N stop with positive P_ν-probability. It is, however, possible to meet any $0 < \alpha < 1$ level. Formally, we add the constraints

$$P_{\nu=\infty}(N < \infty) \le \alpha, \qquad P_\nu(N < \infty) = 1 \quad \text{for } 1 \le \nu < \infty. \tag{6}$$

So, let

$$\Omega = \{(N, \bar{t}) : (N, \bar{t}) \text{ has a renewal structure and satisfies (5) and (6)}\}.$$

THEOREM 2. *There exists a constant C (which depends on μ_1 and γ) such that*

$$\frac{2}{\gamma\mu_1^2}\log\left(\frac{1}{\alpha}\right) - 1 \le \inf_{(N,\bar{t})\in\Omega}\ \sup_{1\le\nu<\infty} E_\nu(N-\nu \mid N \ge \nu) \tag{7}$$

$$\le \frac{2}{\gamma\mu_1^2}\log\left(\frac{1}{\alpha}\right) + C. \tag{8}$$

REMARK. In the fixed sampling rate case, among all stopping times satisfying $E_\infty T \ge B$, the least upper bound of the expected delay is of the order

$(2/(\mu_1^2\gamma))\log B$ as $B \to \infty$ [Lorden (1971)]. Theorem 2 implies that there exists a dynamic sampling scheme satisfying $\alpha = 1/B$ with comparable expected delay.

PROOF. Note that if $(N, \bar{t}) \in \Omega$, then N can be used to define a power one test of H_0: $\nu = \infty$ versus H_1: $\nu = 1$ with significance level α. Therefore, (by the optimality of the SPRT) the H_1-expected sampling time until stopping is at least $(2/\mu_1^2)\log(1/\alpha)$. From (5) we therefore obtain $E_{\nu=1}N \geq (2/(\gamma\mu_1^2))\log(1/\alpha)$, from which the inequality on the left of (7) follows.

We proceed to prove the inequality on the right of (7) by construction. Note that it suffices to show the inequality for all $0 < \alpha \leq \alpha_0$, where $0 < \alpha_0 < 1$ is arbitrary. Set $0 < \varepsilon < 1$, whose exact value will be specified later. Let $b = (1/\mu_1)\log(1/\alpha)$, $b_n = b + (n-1)\varepsilon$ for $n = 1, 2, \ldots$. Let $T_{n,i} = \inf\{t\colon B_n(t) - (1/2)\mu_1 t \notin (-\varepsilon, b_i)\}$, where B_n is as in Section 2. Now construct $(N, \bar{t})$, as follows. Let $t_1 = T_{1,1}$. If $B_1(T_{1,1}) - (1/2)\mu_1 T_{1,1} = -\varepsilon$, define $t_2 = T_{2,2}$; otherwise define $t_2 = T_{2,1}$. Continue this recursively: Formally, denoting $J(n) = \max\{j\colon 1 \leq j < n,\ B_j(t_j) - (1/2)\mu_1 t_j \neq -\varepsilon\}$, $J(n) = 0$ if no such j exists, define $t_n = T_{n, n-J(n)}$. Let $X(t)$ be the process obtained by concatenating $\{B_n(t) - (1/2)\mu_1 t\colon 0 \leq t \leq t_n\}_{n=1}^{\infty}$; in other words, letting $n_t = \max\{j\colon j \geq 0,\ \sum_{r=1}^{j} t_r \leq t\}$, write for $0 \leq t < \infty$

$$X(t) = \sum_{j=1}^{n_t}\left[B_j(t_j) - \tfrac{1}{2}\mu_1 t_j\right] + B_{n_t+1}\left(t - \sum_{j=1}^{n_t} t_j\right) - \tfrac{1}{2}\mu_1\left(t - \sum_{j=1}^{n_t} t_j\right).$$

Define $\tau = \min\{t\colon X(t) \geq (1/\mu_1)\log(1/\alpha)\}$, $\tau = \infty$ if no such t exists. Clearly $P_{\nu=\infty}(\tau < \infty) = \alpha$. Define $N = n_\tau$. (Note that if $N = n$, then $\sum_{i=1}^{n} t_i$ will equal τ.) Now modify the sequence $\bar{t}$ so that if $N = n$, then $(t_{n+1}, t_{n+2}, \ldots)$ will have the same structure as $(t_1, t_2, \ldots)$. We will show that for appropriate ε, this (modified) pair $(N, \bar{t})$ is contained in Ω and that, for all $1 \leq \nu < \infty$, $E_\nu(N - \nu | N \geq \nu) \leq (2/(\gamma\mu_1^2))\log(1/\alpha) + C$ for some constant $C < \infty$.

We first show that (5) is satisfied. Consider the first constraint in (5). Clearly $T_{n,i} \leq \inf\{t\colon B_n(t) - (1/2)\mu_1 t \leq -\varepsilon\}$. Hence $E_\infty T_{n,i} \leq 2\varepsilon/\mu_1$, implying $E_\infty \sum_{i=1}^{n} t_i \leq 2\varepsilon n/\mu_1$, $n = 1, 2, \ldots$. Therefore, the first constraint in (5) is satisfied if we take $\varepsilon \leq \gamma\mu_1/2$.

We now turn to the second constraint in (5). Denote

$$\begin{aligned} p_n &= P_{\nu=1}\left(B_n(T_{n,n}) - \frac{1}{2}\mu_1 T_{n,n} = b_n\right) \\ &= \frac{e^{\mu_1\varepsilon} - 1}{e^{\mu_1\varepsilon} - e^{-\mu_1 b_n}} \\ &\leq \frac{e^{\mu_1\varepsilon} - 1}{e^{\mu_1\varepsilon} - e^{\mu_1 b}}, \\ q_n &= 1 - p_n \\ &= \frac{1 - e^{-\mu_1 b_n}}{e^{\mu_1\varepsilon} - e^{-\mu_1 b_n}} \leq e^{-\mu_1\varepsilon}. \end{aligned}$$

The sequence $\{q_n\}_{n=1}^{\infty}$ is increasing. Note that

$$\begin{aligned} E_\nu(N-(\nu-1)|N\geq\nu) &= \sum_{n=1}^{\infty} P_{\nu=1}(N-(\nu-1)\geq n|N\geq\nu) \\ &= 1+\sum_{n=0}^{\infty}\prod_{i=\nu}^{\nu+n} q_i \\ &\geq \sum_{n=0}^{\infty} q_\nu^n \\ &= \frac{1}{1-q_\nu} \\ &= \frac{e^{\mu_1\varepsilon}-e^{-\mu_1[b+(\nu-1)\varepsilon]}}{e^{\mu_1\varepsilon}-1} \\ &\geq \frac{1-e^{-\mu_1 b}}{\mu_1\varepsilon} \end{aligned}$$

and that

$$\begin{aligned} E_\infty\left(\sum_{i=1}^{N\wedge(m+1)} t_i - \sum_{i=1}^{N\wedge m} t_i\right) &= E_\infty t_{m+1} 1(N\geq m+1) \\ &= P_\infty(N\geq m+1)\, E_\infty T_{m+1,m+1} \\ &\leq P_\infty(N\geq m+1)\frac{2\varepsilon}{\mu_1} \end{aligned}$$

so that

$$E_\infty \sum_{i=1}^{N\wedge(\nu-1)} t_i \leq \frac{2\varepsilon}{\mu_1}\sum_{n=1}^{\nu-1} P_\infty(N\geq n). \tag{9}$$

Hence, since $P_\infty(N>n)\geq 1-\alpha$,

$$\begin{aligned} \frac{E_\nu\sum_{i=1}^N t_i}{E_\nu N} &= \frac{E_\infty\sum_{i=1}^{N\wedge(\nu-1)} t_i + E_\nu\big(1(N\geq\nu)\sum_{i=\nu}^N t_i\big)}{E_\infty(N\wedge(\nu-1)) + E_\nu(1(N\geq\nu)(N-(\nu-1)))} \\ &\leq \frac{(2\varepsilon/\mu_1)\sum_{n=1}^{\nu-1}P_\infty(N\geq n) + (2/\mu_1)(b+(\nu-1)\varepsilon)P_\infty(N\geq\nu)}{\sum_{n=1}^{\nu-1}P_\infty(N\geq n) + \big((1-e^{-\mu_1 b})/\mu_1\varepsilon\big)P_\infty(N\geq\nu)} \\ &\leq \frac{2\varepsilon}{\mu_1} + \frac{2\varepsilon b}{1-e^{-\mu_1 b}} + \frac{2(\nu-1)\varepsilon P_\infty(N\geq\nu)/\mu_1}{(1-\alpha)(\nu-1)} \\ &\leq \varepsilon\left(\frac{2b}{1-e^{-\mu_1 b}} + \frac{2(2-\alpha)}{\mu_1(1-\alpha)}\right). \end{aligned}$$

Therefore, (5) will be satisfied if we take

$$\varepsilon = \frac{\gamma}{2b/(1 - e^{-\mu_1 b}) + 2(2 - \alpha)/(\mu_1(1 - \alpha))}. \tag{10}$$

(Note that this also satisfies the previously derived condition $\varepsilon \le \gamma\mu_1/2$.) Now

$$\begin{aligned} E_\nu(N - (\nu - 1)|N \ge \nu) &= 1 + \sum_{n=0}^{\infty} \prod_{i=\nu}^{\nu+n} q_i \\ &\le \sum_{j=0}^{\infty} e^{-\mu_1 \varepsilon j} \\ &= \frac{1}{1 - e^{-\mu_1 \varepsilon}}. \end{aligned}$$

Since $(1 - e^{-x})^{-1} \le 1 + 1/x$ for positive x close enough to zero and since $b \to \infty$ as $\alpha \to 0$, it follows from (10) that there exists $0 < \alpha_0 < 1$ such that if $\alpha \le \alpha_0$ then

$$E_\nu(N - (\nu - 1)|N \ge \nu) \le \frac{2}{\mu_1\gamma}\left[\frac{b}{1 - e^{-\mu_1 b}} + \frac{2 - \alpha_0}{\mu_1(1 - \alpha_0)}\right] + 1,$$

so that by the definition of b

$$\begin{aligned} E_\nu(N - \nu|N \ge \nu) &\le \frac{2}{\mu_1\gamma}\left[b + \frac{be^{-\mu_1 b}}{1 - e^{-\mu_1 b}} + \frac{2 - \alpha_0}{\mu_1(1 - \alpha_0)}\right] \\ &\le \frac{2}{\mu_1^2\gamma}\log\left(\frac{1}{\alpha}\right) + C, \end{aligned}$$

where C is an upper bound of

$$\frac{2}{\mu_1\gamma}\left[\frac{be^{-\mu_1 b}}{1 - e^{-\mu_1 b}} + \frac{2 - \alpha_0}{\mu_1(1 - \alpha_0)}\right]$$

over b's for which $\alpha \le \alpha_0$. □

4. Additional comments. In this paper, we dealt with the case of known μ_1. While one can come up with examples where this may be realistic (detection of the presence of a signal of known frequency from an airplane or a submarine), the case usually encountered in practice is when μ_1 is unknown. The methods proposed in this paper depend heavily on the knowledge of μ_1. The primary problem arising when μ_1 is unknown is that of sampling design; the expected sampling rate must be bounded under all μ_1. Nevertheless, extensions of ideas contained in this article enable a treatment of the unknown μ_1 case. Also, the whole problem may be thought of anew by reconsidering the constraints (5). A possible alternative would be to require that $\sum_{i=1}^{n} t_i \le \gamma n$, $n = 1, 2, \dots,$ a.s. (P_ν), $\nu = 1, 2, \dots, \infty$. These questions are under study by

Benjamin Yakir in his Ph.D. dissertation and partial results have already been obtained.

Acknowledgments. The authors would like to thank an Associate Editor and the referee for their constructive criticism and comments.

REFERENCES

ASSAF, D. (1988). A dynamic sampling approach for detecting a change in distribution. *Ann. Statist.* **16** 236–253.

ASSAF, D. and RITOV, Y. (1989). Dynamic sampling procedures for detecting a change in the drift of Brownian motion: A non-Bayesian model. *Ann. Statist.* **17** 793–800.

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

POLLAK, M. (1985). Optimal detection of a change in distribution. *Ann. Statist.* **13** 206–227.

POLLAK, M. and SIEGMUND, D. (1975). Approximations to the expected sample size of certain sequential tests. *Ann. Statist.* **3** 1267–1282.

POLLAK, M. and SIEGMUND, D. (1985). A diffusion process and its application to detecting a change in the drift of Brownian motion. *Biometrika* **72** 267–280.

DAVID ASSAF
MOSHE POLLAK
YA'ACOV RITOV
DEPARTMENT OF STATISTICS
HEBREW UNIVERSITY
JERUSALEM

BENJAMIN YAKIR
DEPARTMENT OF BIOSTATISTICS
UNIVERSITY ROCHESTER
ROCHESTER, NEW YORK 14620