Quickest Sequential Opportunity Search in Multichannel Systems

Lifeng Lai[1], H. Vincent Poor[2], Yan Xin[3], and Georgios Georgiadis[4]

[1] University of Arkansas, Little Rock
Little Rock, AR 72204, USA
(e-mail: lxlai@ualr.edu)

[2] Princeton University
Princeton, NJ 08544, USA
(e-mail: poor@princeton.edu)

[3] NEC Laboratories America, INC.
Princeton, NJ 08540, USA
(e-mail: yanxin@nec-labs.com)

[4] Stanford University
Stanford, CA 94305, USA
(e-mail: georgios@stanford.edu)

Abstract. The problem of sequentially finding an independent and identically distributed (i.i.d.) sequence that is drawn from a probability distribution Q_1 by searching over multiple sequences, some of which are drawn from Q_1 and the others of which are drawn from a different distribution Q_0, is considered. Within a Bayesian formulation, a sequential decision rule is derived that optimizes a tradeoff between the probability of false alarm and the number of samples needed for the decision. In the case in which one can observe one sequence at a time, surprisingly, it is shown that the cumulative sum (CUSUM) test, which is well-known to be optimal for a non-Bayesian statistical change-point detection formulation, is optimal for the problem under study. Specifically, the CUSUM test is run on the first sequence. If a reset event occurs in the CUSUM test, then the sequence under examination is abandoned and the rule switches to the next sequence. If the CUSUM test stops, then the rule declares that the sequence under examination when the test stops is generated by Q_1.

Keywords. CUSUM test, optimal stopping, quickest detection, renewal theory, sequential testing.

1 Introduction

In the classical sequential testing problem, first studied by Wald (1945), one sequentially observes an independent and identically distributed (i.i.d.) sequence generated by one of two distributions Q_0 or Q_1, and wishes to test hypothesis H_1 that the sequence is generated by Q_1 against hypothesis H_0 that the sequence is generated by Q_0. The goal is to find a decision rule that uses a minimal number of samples, on average, while satisfying certain error probability constraints, or that optimizes some other tradeoff between error probabilities and the average number of samples. Under this model, the sequential probability ratio test (SPRT) was shown to be optimal by Wald and Wolfowitz (1948). Poor (2009) provides a comprehensive review of this topic.

In this paper, we consider a generalization of the sequential testing problem: sequential search over multiple sequences. In particular, we consider N sequences $\{Y_k^i; k = 1, 2, \cdots\}, i = 1, \cdots, N$, where for each i, $\{Y_k^i; k = 1, 2, \cdots\}$ are i.i.d. observations taking values in a set Ω endowed with a σ-field $\mathcal{F}$ of events, that obey one of the two hypotheses:

$$H_0: \quad Y_k^i \sim Q_0, \quad k = 1, 2, \cdots$$

versus

$$H_1: \quad Y_k^i \sim Q_1, \quad k = 1, 2, \cdots$$

where Q_0 and Q_1 are two distinct, but equivalent, distributions on $(\Omega, \mathcal{F})$. We use q_0 and q_1 to denote densities of Q_0 and Q_1, respectively, with respect to some common dominating measure. The sequences for different values of i are independent. Moreover, whether the i^{th} sequence $\{Y_k^i; k = 1, 2, \cdots\}$ is generated by Q_0 or Q_1 is independent of all other sequences. Here, we assume that for each i, hypothesis H_1 occurs with prior probability π_0 and H_0 with prior probability $1 - \pi_0$. Assuming that one can observe only one sequence at a time, our goal is to find a sequence that is generated by Q_1 in a way that minimizes an appropriate measure of error probability and sampling cost. This model is motivated by many applications. For example, in so called "cognitive radio" systems, wireless communication devices need to find unoccupied frequency bands before they can transmit information. Hence, a wireless device should listen to each possible frequency band to determine whether it is free or not. In this scenario, the observations from one frequency band consist of one sequence, Q_0 corresponds to the distribution of the received signal when there are other transmissions in the band, and Q_1 corresponds to the distribution of the received signal when the frequency band is free. The task of finding a free frequency channel clearly can be modelled as that of finding a sequence generated by Q_1. It is of interest to do so with minimal delay, in order to make optimal use of spectral resources. However, the device can typically examine only one band at a time due to hardware limitations. Thus this problem fits the above model very well.

To proceed with the above test, at each time, we select a sequence, say sequence j, and make an observation from this sequence. After making each observation, we can choose from one of the following three actions: 1) stop sampling and claim that the sequence we are currently observing is generated by Q_1; 2) continue to the next observation from the same sequence to gather more evidence about its statistical behavior; or 3) abandon the sequence that we are currently observing and switch to another sequence. Hence if a sequence is abandoned, we will not come back and test it again. Without loss of generality, we start taking samples from the first sequence, and switch to the second sequence if we decide to abandon the first sequence. Similarly, we will switch to the $(i+1)^{th}$ sequence if we decide to abandon the ith sequence. To ensure that there is always a sequence to switch to, we consider the case $N = \infty$.

We use s_k to denote the index of the sequence that we are observing at time k. Hence, we observe $\{Y_k^{s_k}; k = 1, 2, \cdots\}$ sequentially. The observations generate the filtration $\{\mathcal{F}_k; k = 1, 2, \cdots\}$ with $\mathcal{F}_k = \sigma(Y_1^{s_1}, Y_2^{s_2}, \cdots, Y_k^{s_k})$. We use ϕ_k to denote the $\mathcal{F}_k$-measurable switching function at time k. Here, $\phi_k(\mathcal{F}_k) = 1$ if we decide to abandon sequence s_k and switch to the next sequence, that is $s_{k+1} = 1 + s_k$. On the other hand $\phi_k(\mathcal{F}_k) = 0$ if we decide to continue observing sequence s_k, that is $s_{k+1} = s_k$. Let $\mathcal{T}$ denote the set of all stopping times with respect to the filtration $\mathcal{F}_k$. Note that the sequence $s_1, s_2, \cdots$, and hence the filtration $\mathcal{F}_1, \mathcal{F}_2, \cdots$, depends on the sequence $\phi_1, \phi_2, \cdots$ of switching functions. A stopping time $\tau \in \mathcal{T}$ will decide when we should stop sampling and declare that the sequence we are currently observing is generated by Q_1. More specifically, if $\tau = k$, we should stop sampling at time k, and declare that sequence s_k is generated by Q_1. There are two performance indices: 1) the error probability that sequence s_τ is generated by Q_0, that is $P(H^{s_\tau} = H_0)$, where H^j is the true hypothesis satisfied by sequence j; and 2) the average number of samples we take to make a decision, that is $\mathbb{E}\{\tau\}$.

Our goal is to determine the stopping time τ and the switching rules $\phi = \{\phi_1, \phi_2, \cdots\}$ to solve the following optimization problem:

$$\inf_{\tau \in \mathcal{T}, \phi} \left[P(H^{s_\tau} = H_0) + c\mathbb{E}\{\tau\} \right]. \tag{1}$$

Here $c > 0$ is a constant that represents the cost of taking one sample. We assume $c < 1 - \pi_0$, as the case $c \geq 1 - \pi_0$ is trivial: we simply do not take any observations and choose a sequence at random as being generated by Q_1.

2 Solution

We use $\pi_k = P(H^{s_k} = H_1 | \mathcal{F}_k)$ to denote the posterior probability that sequence s_k is generated by Q_1 after observing $\{Y_1^{s_1}, \cdots, Y_k^{s_k}\}$. After making each observation, we can update the posterior probability using Bayesian rule:

$$\pi_1 = \frac{\pi_0 q_1(Y_1^1)}{\pi_0 q_1(Y_1^1) + (1 - \pi_0) q_0(Y_1^1)} \tag{2}$$

$$\pi_{k+1} = \frac{\pi_k q_1(Y_{k+1}^{s_{k+1}})}{\pi_k q_1(Y_{k+1}^{s_{k+1}}) + (1 - \pi_k) q_0(Y_{k+1}^{s_{k+1}})} \mathbf{1}(\phi_k = 0) + \frac{\pi_0 q_1(Y_{k+1}^{s_{k+1}})}{\pi_0 q_1(Y_{k+1}^{s_{k+1}}) + (1 - \pi_0) q_0(Y_{k+1}^{s_{k+1}})} \mathbf{1}(\phi_k = 1),$$

in which $\mathbf{1}(\cdot)$ is the indicator function.

Theorem 1. *The optimal stopping time for* (1) *is specified by a parameter* π_U^**, whose value depends on the cost of sampling* c*, and is given by* $\tau_{opt} = \inf\{k : \pi_k > \pi_U^*\}$*. And at time* k*, we switch to another sequence if, and only if,* $\pi_k < \pi_0$*.*

It is now easy to see the equivalence between the optimal test in Theorem 1 and the CUSUM test proposed by Page (1954) and shown to be optimal for a non-Bayesian statistical change-point detection problem by Moustakides (1986). More specifically, let $L_k = q_1(Y_k^{s_k})/q_0(Y_k^{s_k})$, then under the condition that $\phi_k = 1$ if $\pi_k < \pi_0$ and $\phi_k = 0$ if $\pi_k \geq \pi_0$, the recursive formula in (2) is the equivalent to the following recursive formula:

$$R_1 = \log(L_1), \tag{3}$$

$$R_{k+1} = (R_k + \log(L_{k+1}))\, \mathbf{1}(R_k \geq 0) + \log(L_{k+1}) \mathbf{1}(R_k < 0) = \max\{R_k, 0\} + \log(L_{k+1}).$$

In terms of R_k, the optimal solution is to switch to the next sequence if $R_k < 0$ (this corresponds to a reset event in the CUSUM test, which is to reset R_k to zero, if $R_k < 0$), and to stop when $R_k \geq (1 - \pi_0)\pi_U^*/(\pi_0(1 - \pi_U^*))$. Hence the test in Theorem 1 is equivalent to a CUSUM test with parameter $\exp\left((1 - \pi_0)\pi_U^*/(\pi_0(1 - \pi_U^*)\right)$, in which we switch to another sequence if a reset event occurs in the CUSUM test, and we stop and claim that the sequence under examination is generated by Q_1 when the CUSUM test stops.

3 Acknowledgment

This research was supported in part by the Qatar National Research Fund under Grant NPRP 08-522-2-211.

References

Moustakides, G. V. (1986). Optimal stopping times for detecting changes in distributions. *Ann. Statist. 14*, 1379-1387.

Page, E. S. (1954). Continuous inspection schemes. *Biometrika. 41*, 100-115.

Poor, H. V. and Hadjiliadis, O. (2009). *Quickest Detection*. Cambridge University Press, Cambridge, UK.

Wald, A. (1945). Sequential tests of statistical hypotheses. *Ann. Math. Statist. 16*, 117-186.

Wald, A. and Wolfowitz, J. (1948). Optimum character of the sequential probability ratio test. *Ann. Math. Statist. 19*, 326-339.