On the Average Run Length to False Alarm in Surveillance Problems which Possess an Invariance Structure

Benjamin Yakir[1]
Department of Statistics
Hebrew University of Jerusalem
Jerusalem 91950
ISRAEL
msby@mscc.huji.ac.il

[1]This work was supported in part by grants of the Israel Science Foundation.

Abstract

Surveillance can be based, in some change-point detection problems, on a sequence of invariant statistics. Gordon and Pollak (1997) prove that, under certain conditions, the Average Run Length (ARL) to false alarm of invariance-based Shiryayev-Roberts detection schemes is asymptotically the same as that of the dual classical scheme that is based on the original sequence of observations. In this paper we give alternative conditions under which the two ARL coincide and demonstrate that these conditions are satisfied in cases where Gordon and Pollak's conditions are difficult to check.

Key words and phrases: Quality control, control charts, change-point detection, invariant statistics.

AMS (1991) subject classification: Primary 62L10, Secondary 62N10.

Short Title: ARL to False Alarm in Problems with Invariance Structure.

1 Introduction

The setting of the classical change-point problem has initial observations which are independent and identically distributed, until a change occurs at some unknown point in time ν. Subsequently, the distribution changes, though the observations are again independent and identically distributed. One obtains the observations sequentially, with the goal of raising an alarm as soon as it becomes clear that the distribution has changed. The pre-change distribution is assumed to be known. Classical surveillance schemes invariably make heavy use of this knowledge, and ignorance of the precise specification of the pre-change distribution typically renders them inoperable.

In practice, there are many situations in which the pre-change distribution is only partially specified, or not specified at all. A typical example is discussed in Wilson *et. al.* (1979). There, a scheme for monitoring the quality of laboratory tests is constructed. Samples are sent at regular intervals for assay. One is on the lookout for a change of variance. (The observations are assumed to be normally distributed.) The scenario is such that the initial variance is unknown.

In this setting, the pre-change distribution is known up to a nuisance parameter. A naive approach to this problem would call for estimation of the unknown parameter and subsequent use of classical procedures with the true value of the parameter replaced by its estimate. (This assumes the possibility of obtaining a learning sample from the pre-change distribution.) The difficulty with this approach is that the operating characteristics of classical schemes are very sensitive to misspecification of distributional parameters.

(See Section 2.4 in van Dobben de Bruyn, 1968. See also Table 1 in Gordon and Pollak, 1995.)

This difficulty can be both overcome and (sometimes) circumvented. Overcoming this difficulty requires an analysis of the operating characteristics which takes into account the fact that there are parameters being estimated. This approach was taken by Siegmund and Venkatraman (1992). See also Lai (1995). Circumvention can be done if the problem possesses invariance properties. The idea is to base surveillance on a sequence of invariant statistics instead of on the original observations. The invariance causes the pre-change distribution of the sequence to be devoid of unknown parameters, thereby making the pre-change distribution (of the invariant statistics) known. This approach was taken by Pollak and Siegmund (1991) in a parametric setting and by Macdonald (1990), Gordon and Pollak (1994, 1995, 1997), Bell, Gordon and Pollak (1992) in a nonparametric one.

While the invariance approach is appealing, it does entail difficulties. The dependence between the invariant statistics makes evaluation of operating characteristics more difficult. Gordon and Pollak (1997) prove a general theorem which states (under certain conditions) that the ARL to false alarm of invariance-based Shiryayev-Roberts detection schemes is asymptotically the same as that of the parallel classical scheme for the case where the pre-change parameters are known. Gordon and Pollak (1997) require that three conditions (A, B and C) be satisfied.

Problems which can be solved using Gordon and Pollak's theorem are detection of a change in the mean of a normal distribution with known variance where the initial value of the mean is unknown, detection of a change in the scale parameter of a gamma distribution with unknown initial scale (Gordon and Pollak, 1997), detection of a decrease in the variance of a normal distribution where the mean and the initial variance are unknown (Damian, 1994) and a variety of nonparametric detection schemes (Bell, Gordon and Pollak, 1992; Gordon and Pollak, 1994, 1995, 1997). Nonetheless, there are a number of problems (such as detection of an increase in the variance of a normal distribution where the mean and the initial variance are unknown, and detection of a change in the mean of a normal distribution where the variance and the initial mean are unknown) for which Gordon and Pollak's theorem seems to be very hard to apply. The main difficulty lies with showing fulfillment of Gordon and Pollak's condition C.

In this paper, an alternative to Gordon and Pollak's theorem is presented. Essentially, conditions A and B are (roughly) preserved, but condition C is relaxed, facilitating proofs in cases where Gordon and Pollak's theorem is hard to apply. This alternative theorem is shown to handle the two cases mentioned in the previous paragraph as being examples of situations where Gordon and Pollak's theorem is apparently hard to apply.

The difference between the approach studied in this paper and that of Gordon and Pollak is essentially the same as the difference between Pollak's (1987) and Yakir's (1995) approaches to proving the basic asymptotic properties of the ARL to false alarm of the simple Shiryayev-Roberts schemes.

2 General theory

We consider surveillance for a change in the case where the problem admits an invariance structure.

To fix ideas, consider the case of monitoring for an increase of a normal mean, where none of the parameters are known. Prior to a change, the distribution of the observations is $N(\mu_0, \sigma^2)$ and post-change it is $N(\mu_0 + \mu, \sigma^2)$, where μ_0, μ, σ^2 are unknown, $\mu > 0$, and the observations are independent. Since μ_0 and σ^2 are unknown, one cannot apply a standard Cusum or Shiryayev-Roberts control chart. One way of circumventing this problem is to exploit the invariance structure (invariance under increasing affine transformations) and base surveillance on a sequence of invariant statistics instead of the raw observations. More explicitly, suppose one can obtain a learning sample $X_{-m}, X_{-m+1}, \ldots, X_{-1}$, $m \geq 2$, of pre-change observations and monitoring commences with the observations $X_1, X_2, \ldots$. The statistics

$$T_n = \frac{X_n - \bar{X}_{-m}}{S_{-m}}, \qquad n = 1, 2, \ldots, \tag{1}$$

where

$$\bar{X}_{-m} = \frac{\sum_{i=1}^{m} X_{-i}}{m}, \quad S_{-m} = \sqrt{\frac{1}{m-1}\sum_{i=1}^{m}(X_{-i} - \bar{X}_{-m})^2}, \tag{2}$$

form a sequence of invariant statistics. The pre-change distribution of the sequence is fully known, so that likelihood-ratio based schemes (such as Cusum or Shiryayev-Roberts) can be applied.

Here we study the ARL to false alarm of Shiryayev-Roberts procedures for a general setting of a surveillance problem having an invariance structure. For a formal definition of the general invariance structure, see Gordon and Pollak (1997).

In the general setting the sequence of raw observations will be denoted by $X_{-m}, X_{-m+1}, \ldots, X_{-1}, X_1, X_2, \ldots$, where $X_{-m}, X_{-m+1}, \ldots, X_{-1}$ are a learning sample of size m from the pre-change distribution. The sequence of invariant statistics, upon which the surveillance will be based, will be denoted by $T_1, T_2, \ldots$ where $T_n = T_n(X_{-m}, \ldots, X_n)$. For purposes of defining likelihood ratios, it is assumed that one can choose a "representative" of the set of possible post-change distributions without hurting the invariance of the problem and the sequence $\{T_i\}$. (In the aforementioned example, one can choose a value $\delta > 0$ and pretend that $\mu = \delta\sigma$ for purposes of defining a likelihood ratio. For practical purposes, δ would be a value such that there would be serious interest in detecting an increase in mean of at least δ standard deviations.) Note that the pre-change properties of any scheme based on the sequence $T_1, T_2, \ldots$ are the same for all possible values of the nuisance parameters, so that in order to study the properties one may choose a convenient set of nuisance parameters. (In the example, it would be natural to take $\mu_0 = 0$, $\sigma = 1$.) Henceforth, we assume that such a choice was made.

We denote by P_k the measure under which $X_{-m}, \ldots, X_{-1}, X_1, \ldots, X_{k-1}$ are pre-change observations and $X_k, X_{k+1}, \ldots$ are post-change, where all nuisance parameters have been set for convenience and a representative have been chosen. (In the example, this would mean that $\delta > 0$ is fixed and $\mu_0 = 0$, $\sigma = 1$.) P_∞ will denote the distribution when there is no change ($k = \infty$). dP_k and dP_∞ are the appropriate densities with respect to some σ-finite measure. Denote by $\mathcal{F}_n$ the σ-field generated by $T_1, T_2, \ldots, T_n$.

The main result of this article – Theorem 1 – states that the asymptotic first-order properties of the ARL to false alarm of the Shiryayev-Roberts procedure based on the sequence of invariant statistics are the same as those of the parallel procedure which would have been used had all nuisance parameters been known. It is necessary, therefore, to differentiate the notation of the two cases.

Define, for $1 \le k \le n$, the likelihood-ratio statistics

$$\begin{aligned}\Lambda_k(n) &= \frac{\mathrm{dP}_k(T_1,\ldots,T_n)}{\mathrm{dP}_\infty(T_1,\ldots,T_n)},\\ \Lambda_k^{\mathrm{fs}}(n) &= \frac{\mathrm{dP}_k(X_1,\ldots,X_n)}{\mathrm{dP}_\infty(X_1,\ldots,X_n)}.\end{aligned}$$

The (invariant) Shiryayev-Roberts statistics and stopping time are

$$\begin{aligned}R(n) &= \sum_{k=1}^{n} \Lambda_k(n), \qquad n = 1,2,\ldots,\\ N_A &= \inf\{n : R(n) \ge A\}.\end{aligned}$$

The Shiryayev-Roberts statistics and stopping time when the nuisance parameters are fully specified are

$$\begin{aligned}R^{\mathrm{fs}}(n) &= \sum_{k=1}^{n} \Lambda_k^{\mathrm{fs}}(n), \qquad n = 1,2,\ldots,\\ N_A^{\mathrm{fs}} &= \inf\{n : R^{\mathrm{fs}}(n) \ge A\}.\end{aligned}$$

In Theorem 1 the P_∞-asymptotic properties of N_A/A and N_A^{fs}/A are compared. In the process of proving the theorem auxiliary stopping times are used. Given r, $r = r(A)$ such that $\log A \ll r(A) \ll A$, the (invariant and truncated) Shiryayev-Roberts statistics and stopping time are

$$\begin{aligned}Q_A(n) &= \sum_{k=\lfloor n/r \rfloor r+1}^{n} \Lambda_k(n), \qquad n = 1,2,\ldots,\\ \tau_A &= \inf\{n : Q_A(n) \ge A\}.\end{aligned}$$

The (fully specified and truncated) Shiryayev-Roberts statistics and stopping time are

$$\begin{aligned}Q_A^{\mathrm{fs}}(n) &= \sum_{k=\lfloor n/r \rfloor r+1}^{n} \Lambda_k^{\mathrm{fs}}(n), \qquad n = 1,2,\ldots,\\ \tau_A^{\mathrm{fs}} &= \inf\{n : Q_A^{\mathrm{fs}}(n) \ge A\}.\end{aligned}$$

Consider the following conditions:

Condition 1 *There exists a function $r = r(A)$ such that for any given $\epsilon_1 > 0$ one can find constants $\theta_1 > 1$ and A_1. For this function and constants, and for all $A \geq A_1$ and $t \geq \epsilon_1 A$ one can find an event $B_1 = B_1(X_{t+1}, \ldots, X_{t+r})$ such that:*

$$\mathrm{P}_\infty \left((X_{t+1}, \ldots, X_{t+r}) \in B_1, \sup_{t<k\leq n\leq t+r} \left| \frac{\Lambda_k(n)}{\Lambda_k^{fs}(n)} - 1 \right| > \epsilon_1 \right) \leq \frac{\epsilon_1 r}{A} e^{-\frac{\theta_1 t}{A}}$$

and

$$\mathrm{P}_\infty \left((X_{t+1}, \ldots, X_{t+r}) \notin B_1 \right) \leq \epsilon_1 \frac{r}{A}.$$

The function $r = r(A)$ should be such that, as $A \to \infty$, $r(A)/A \to 0$ but $r(A)/\log A \to \infty$.

Condition 2 *Given the function $r = r(A)$ from Condition 1 and given any $\epsilon_2 > 0$ and $C_2 < \infty$ one can find A_2 such that the relation*

$$\sum_{k=1}^{n} \mathrm{P}_k(N_A > n) \leq \epsilon_2 r$$

holds for all $A \geq A_2$ and $C_2 A \geq n \geq \epsilon_2 A$.

Theorem 1 *If Conditions 1 and 2 hold, then the limit (in P_∞-distribution) of N_A/A, as $A \to \infty$, is exponential with scale λ, where*

$$\lambda = \lim_{A\to\infty} A/\mathrm{E}_\infty N_A^{fs}.$$

Moreover, $\mathrm{E}_\infty N_A/A \to_{A\to\infty} 1/\lambda$.

Remark 1: The constant λ satisfies $0 < \lambda < 1$. Its exact value can be computed by standard renewal theory. (See Pollak, 1987).

□

Remark 2: Condition 1 is similar in nature to Condition A in Gordon and Pollak (1997). To see the connections between Condition B in Gordon and Pollak (1997), which deals with the P_∞-behavior of the likelihood ratios that define the statistic $R(n)$, and Condition 2 here notice that

$$\sum_{k=1}^{n} \mathrm{P}_k(N_A > n) = \mathrm{E}_\infty \left[R(n) \mathbb{1}(N_A > n) \right].$$

$\square$

Before proving Theorem 1 we will state and prove two lemmas that deal with the properties of the invariant and truncated Shiryayev-Roberts stopping time τ_A.

Lemma 1 *Given $r = r(A)$, let t be an integer multiple of r. Then for any A, for which $r/A < 1$,*

$$\mathrm{P}_\infty(\tau_A > t) \geq e^{-\frac{1}{1-(r/A)} \frac{t}{A}}.$$

Proof : It is easy to see that

$$\begin{aligned} &\mathrm{P}_\infty(\tau_A \leq t \,|\, \tau_A > t-r) = \\ &\qquad \mathrm{P}_\infty \left(\sup_{t-r<n\leq t} Q_A(n) \geq A \middle| \tau_A > t-r \right) \\ &= \mathrm{E}_\infty \left[\mathrm{P}_\infty \left(\sup_{t-r<n\leq t} Q_A(n) \geq A \middle| \mathcal{F}_{t-r} \right) \middle| \tau_A > t-r \right]. \end{aligned}$$

However, $\{Q_A(n) : t-r < n \leq t\}$ is a sub-martingale with respect to the measure $\mathrm{P}_\infty(\cdot \,|\, \mathcal{F}_{t-r})$ and the filter $\{\mathcal{F}_n : t-r < n \leq t\}$. Moreover, $\mathrm{E}_\infty[Q_A(t) \,|\, \mathcal{F}_{t-r}] = r$. Hence, by Doob's Inequality,

$$\mathrm{P}_\infty \left(\sup_{t-r<n\leq t} Q_A(n) \geq A \middle| \mathcal{F}_{t-r} \right) \leq \frac{r}{A}.$$

Thus

$$\mathrm{P}_\infty(\tau_A \le t \,|\, \tau_A > t - r) \le \frac{r}{A}.$$

By induction one gets that

$$\mathrm{P}_\infty(\tau_A > t) \ge \left(1 - \frac{r}{A}\right)^{\frac{t}{r}}.$$

Finally, the relation $\log(1+x) \ge x/(1+x)$, valid for all $x > -1$, can be used to show that

$$\left(1 - \frac{r}{A}\right)^{\frac{t}{r}} \ge e^{-\frac{1}{1-(r/A)}\frac{t}{A}},$$

and the result follows.

□

Let $r = r(A)$ be such that $r(A)/A \to 0$ but $r(A)/\log A \to \infty$, as $A \to \infty$. For the fully specified case it can be shown that for any $\epsilon_1 > 0$ and for any large A:

$$\left| \frac{A}{r}\mathrm{P}_\infty\left(\sup_{t<n\le t+r} Q_A^{\mathrm{fs}}(n) \ge A \right) - \lambda \right| < \epsilon_1,$$

where λ was defined in Theorem 1. (See Yakir, 1995.)

Lemma 2 *If Condition 1 holds then the limit (in* P_∞*-distribution) of* τ_A/A*, as* $A \to \infty$*, is exponential with scale* λ*. Moreover, the* P_∞*-expectation of* τ_A/A *converges to* $1/\lambda$*.*

Proof : Assume that Condition 1 holds and that A is large. Let

$$B_2 = \left\{ \sup_{t<k\le n\le t+r} \left| \frac{\Lambda_k(n)}{\Lambda_k^{\mathrm{fs}}(n)} - 1 \right| \le \epsilon_1 \right\}.$$

On the event B_2 it is true that

$$(1-\epsilon_1) \sup_{t<n\le t+r} Q_A^{\mathrm{fs}}(n) \le \sup_{t<n\le t+r} Q_A(n) \le (1+\epsilon_1) \sup_{t<n\le t+r} Q_A^{\mathrm{fs}}(n).$$

For any $t \geq \epsilon_1 A$, t an integer multiple of r, one can write

$$\begin{aligned}
\mathrm{P}_\infty(t < \tau_A \leq t + r) &= \mathrm{P}_\infty\left(t < \tau_A, \sup_{t<n\leq t+r} Q_A(n) \geq A\right) \\
&\leq \mathrm{P}_\infty\left(t < \tau_A, (1+\epsilon_1) \sup_{t<n\leq t+r} Q_A^{\mathrm{fs}}(n) \geq A\right) \\
&\quad + \mathrm{P}_\infty((X_{t+1}, \ldots, X_{t+r}) \in B_1, B_2^c) \\
&\quad + \mathrm{P}_\infty((X_{t+1}, \ldots, X_{t+r}) \notin B_1, t < \tau_A) \\
&\leq \mathrm{P}_\infty(t < \tau_A)\frac{r}{A} \\
&\quad \times \left[(1+\epsilon_1)(\lambda+\epsilon_1) + \frac{\epsilon_1 e^{-\theta_1(t/A)}}{\mathrm{P}_\infty(t < \tau_A)} + \epsilon_1\right],
\end{aligned}$$

since $\{t < \tau_A\}$ is independent of $(X_{t+1}, \ldots, X_{t+r})$.

In a similar fashion

$$\begin{aligned}
\mathrm{P}_\infty(t < \tau_A \leq t + r) &\geq \mathrm{P}_\infty\left(t < \tau_A, (1-\epsilon_1) \sup_{t<n\leq t+r} Q_A^{\mathrm{fs}}(n) \geq A, B_1, B_2\right) \\
&\geq \mathrm{P}_\infty(t < \tau_A)\frac{r}{A} \\
&\quad \times \left[(1-\epsilon_1)(\lambda-\epsilon_1) - \frac{\epsilon_1 e^{-\theta_1(t/A)}}{\mathrm{P}_\infty(t < \tau_A)} - \epsilon_1\right].
\end{aligned}$$

The result follows Lemma 1 and induction (see the proof of Theorem 1 below).

□

Proof of Theorem 1: Let $\epsilon_2 > 0$ be a given small number. Define, for any A, $t_0 = t_0(A) = \lfloor \epsilon_2 A/r \rfloor r$, where $r = r(A)$ is an integer, and $\lfloor x \rfloor$ is the integer part of x. It can be shown, using a measure transformation, that

$$\begin{aligned}
\mathrm{P}_\infty(N_A \leq t_0) &= \sum_{n=1}^{t_0} \mathrm{P}_\infty(N_A = n) \\
&= \sum_{n=1}^{t_0} \sum_{k=1}^{n} \mathrm{E}_k\left[\frac{\mathbb{1}(N_A = n)}{R(n)}\right]
\end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=1}^{t_0}\sum_{n=k}^{t_0} \mathrm{E}_k\left[\frac{\mathbb{I}(N_A = n)}{R(N_A)}\right] \\
&= \sum_{k=1}^{t_0} \mathrm{E}_k\left[\frac{\mathbb{I}(k \le N_A \le t_0)}{R(N_A)}\right].
\end{aligned}
$$

It can be concluded, since $R(N_A) \ge A$, that $\mathrm{P}_\infty(N_A \le t_0) \le t_0/A \le \epsilon_2$. Hence, $\mathrm{P}_\infty(N_A > t_0) \ge 1 - \epsilon_2$.

Let C_2 be a given large number. Consider any t, an integer multiple of r, such that $t_0 \le t$ but $t + r \le C_2 A$. It is easy to see that

$$
\begin{aligned}
\mathrm{P}_\infty(t < N_A \le t + r) &\ge \mathrm{P}_\infty\left(t < N_A, \sup_{t < n \le t+r} Q_A(n) \ge A\right) \\
&\ge \mathrm{P}_\infty(t < N_A) \times \frac{r}{A} \\
&\quad \times \left[(1-\epsilon_1)(\lambda - \epsilon_1) - \frac{\epsilon_1 e^{-\theta_1 (t/A)}}{\mathrm{P}_\infty(t < N_A)} - \epsilon_1\right].
\end{aligned}
$$

Hence,

$$
\begin{aligned}
&\mathrm{P}_\infty(N_A \le t + r \,|\, N_A > t) \ge \\
&\quad \frac{r}{A} \times \left[(1-\epsilon_1)(\lambda - \epsilon_1) - \frac{\epsilon_1 e^{-\theta_1 (t/A)}}{\mathrm{P}_\infty(t < N_A)} - \epsilon_1\right]. \qquad (3)
\end{aligned}
$$

Likewise,

$$
\begin{aligned}
\mathrm{P}_\infty(t < N_A \le t + r) &\le \mathrm{P}_\infty\left(t < N_A, \sup_{t < n \le t+r} Q_A(n) \ge (1-\epsilon)A\right) \\
&\quad + \mathrm{P}_\infty\left(t < N_A, \sup_{t < n \le t+r} \{R(n) - Q_A(n)\} \ge \epsilon A\right) \\
&\le \frac{\mathrm{P}_\infty(t < N_A)}{(1-\epsilon)} \times \frac{r}{A} \\
&\quad \times \left[(1+\epsilon_1)(\lambda + \epsilon_1) + \frac{\epsilon_1 e^{-\theta_1 (t/A)}}{\mathrm{P}_\infty(t < N_A)} + \epsilon_1\right] \\
&\quad + \mathrm{P}_\infty\left(t < N_A, \sup_{t < n \le t+r} \{R(n) - Q_A(n)\} \ge \epsilon A\right).
\end{aligned}
$$

Consider the stopping time $T_{\epsilon,A,t}$, where

$$T_{\epsilon,A,t} = \inf\{t < n : R(n) - Q_A(n) \geq \epsilon A\}.$$

It follows that,

$$\mathrm{P}_\infty\left(t < N_A, \sup_{t<n\leq t+r}\{R(n) - Q_A(n)\} \geq \epsilon A\right) = \mathrm{P}_\infty(t < N_A, T_{\epsilon,A,t} \leq t+r)$$

and

$$\begin{aligned}
\mathrm{P}_\infty(t < N_A, T_{\epsilon,A,t} \leq t+r) &= \sum_{n=t+1}^{t+r} \mathrm{P}_\infty(t < N_A, T_{\epsilon,A,t} = n) \\
&= \sum_{n=t+1}^{t+r}\sum_{k=1}^{t} \mathrm{E}_k\left[\frac{\mathbb{1}(t < N_A, T_{\epsilon,A,t} = n)}{R(n) - Q_A(n)}\right] \\
&= \sum_{k=1}^{t} \mathrm{E}_k\left[\frac{\mathbb{1}(t < N_A, T_{\epsilon,A,t} \leq t+r)}{R(T_{\epsilon,A,t}) - Q_A(T_{\epsilon,A,t})}\right] \\
&\leq \frac{1}{\epsilon A}\sum_{k=1}^{t} \mathrm{P}_k(t < N_A).
\end{aligned}$$

It follows, applying Condition 2, that

$$\begin{aligned}
&\mathrm{P}_\infty(N_A \leq t+r \,|\, N_A > t) \leq \\
&\quad \frac{r}{A} \times \left[\frac{1+\epsilon_1}{1-\epsilon}(\lambda+\epsilon_1) + \frac{\epsilon_1}{1-\epsilon}\left(\frac{e^{-\theta_1(t/A)}}{\mathrm{P}_\infty(t < N_A)} + 1\right) + \frac{\epsilon_2}{\epsilon \mathrm{P}_\infty(t < N_A)}\right] \quad (4)
\end{aligned}$$

Given any δ, $1-\lambda > \delta > 0$, choose ϵ, ϵ_1, and then ϵ_2, all small enough to ensure that

$$\begin{aligned}
1 - e^{-\frac{r}{A}(\lambda+\delta)} &\geq \frac{r}{A}\left[\frac{1+\epsilon_1}{1-\epsilon}(\lambda+\epsilon_1) + \frac{\epsilon_1(2-\epsilon_2)}{(1-\epsilon)(1-\epsilon_2)} + \frac{\epsilon_2}{\epsilon e^{-C_2(\lambda+\delta)}}\right] \\
1 - e^{-\frac{r}{A}(\lambda-\delta)} &\leq \frac{r}{A}\left[(1-\epsilon_1)(\lambda-\epsilon_1) - \frac{\epsilon_1}{1-\epsilon_2} - \epsilon_1\right],
\end{aligned}$$

for all A such that r/A is small. It follows from (3), (4) and induction that

$$e^{-(\lambda-\delta)\frac{t-t_0}{A}} \geq \mathrm{P}_\infty(N_A > t) \geq (1-\epsilon_2)e^{-(\lambda+\delta)\frac{t}{A}},$$

for all t, $\epsilon_2 A \le t \le C_2 A$, t an integer times r. The limit in distribution of N_A/A is thus obtained. The limit $\mathrm{E}_\infty N_A/A \to 1/\lambda$ follows from the fact that N_A/A is dominated by τ_A/A.

□

3 Examples

Consider a setting in which the observations are independent and the in-control distribution is normal with mean μ_0 and variance σ^2. One can envision a number of surveillance problems:

(i) a change in mean,

(ii) a change in variance,

(iii) a change in both mean and variance.

Nuisance parameters in all three cases can be either:

(a) an unknown initial mean,

(b) an unknown initial variance or

(c) both initial mean and initial variance unknown.

Consider problems (i) and (ii) in the above list: The case of no nuisance parameters (i.e. the in-control distribution is fully specified) was handled by Pollak (1987).

Case (i.a) – detecting a change in mean when the initial mean is unknown but the initial variance is known – was studied by Pollak and Siegmund (1991). (See also Gordon and Pollak, 1997.)

Case (ii.a) – detecting a change in variance when the initial variance is known but the initial mean is unknown – was handled by Gordon and Pollak (1997).

Case (ii.c) – detecting a change in variance when both the initial mean and initial variance are unknown – was tackle by Damian (1994), who solved the problem of detecting a decrease in variance using Gordon and Pollak's (1997) Theorem 1.

The asymptotics of the ARL to false alarm in the other cases has not been worked out. The difficulty of applying Gordon and Pollak's Theorem 1 lies in showing that their Condition C is satisfied.

To show that our Theorem 1 can handle such cases, we fully work out two examples:

Example 1: Case (i.c) – detecting a change in mean when the initial mean and variance are both unknown.

Example 2: Case (ii.c) – detecting an increase in variance when the initial mean and variance are both unknown.

Again, we assume that $X_{-m}, X_{-m+1}, \ldots, X_{-1}$ is a learning sample of m independent observations from the in-control distribution. Ensuing observations are $X_1, X_2, \ldots$. Surveillance is based on the statistics $\{T_n\}$, defined in (1). This sequence is a sequence of invariant statistics both for Example 1 and for Example 2. We use $\sum_{i=-m}^{n}$ to denote sums of the form $\sum_{i=-m}^{-1} + \sum_{i=1}^{n}$.

In Example 2, an explicit form of the likelihood ratio is available. Therefore it will be developed first. We consider the case where the representative post-change distribution is $N(\mu, c^2\sigma^2)$, where $c^2 > 1$ has a fixed (known) value. We apply Theorem 1, assuming for convenience that the pre-change distribution is $N(0,1)$. (As mentioned above, this entails no loss of generality.) Straightforward calculations yield

$$\Lambda_k^{\text{fs}}(n) = c^{-(n-k+1)} \exp\left\{\frac{1}{2}\left(1-\frac{1}{c^2}\right)\sum_{i=k}^{n} X_i^2\right\}$$

and

$$\Lambda_k(n) = \frac{c^{-(n-k+1)}\sqrt{m+n}}{\sqrt{k-1+\frac{m+n-k+1}{c^2}}} \times \left(\frac{m-1+\sum_{i=1}^{n} T_i^2 - \frac{\left(\sum_{i=1}^{n} T_i\right)^2}{m+n}}{m-1+\sum_{i=1}^{k-1} T_i^2 + \frac{\sum_{i=k}^{n} T_i^2}{c^2} - \frac{\left(\sum_{i=1}^{k-1} T_i + \frac{\sum_{i=k}^{n} T_i}{c^2}\right)^2}{m+k-1+\frac{n-k+1}{c^2}}}\right)^{\frac{m+n-1}{2}}.$$

It is easy to see that the value of the statistic $\Lambda_k(n)$ does not change if we add a constant to all of the X_i's or multiply by a positive constant. It follows, by subtracting $\bar{X}_{-m}$ from each of the X_i's in the expression bellow, and dividing them by S_{-m}, that

$$\Lambda_k(n) = \frac{c^{-(n-k+1)}\sqrt{m+n}}{\sqrt{m+k-1+\frac{n-k+1}{c^2}}} \times \left(\frac{\sum_{i=-m}^{n} X_i^2 - \frac{\left(\sum_{i=-m}^{n} X_i\right)^2}{m+n}}{\sum_{i=-m}^{k-1} X_i^2 + \frac{1}{c^2}\sum_{i=k}^{n} X_i^2 - \frac{\left(\sum_{i=-m}^{k-1} X_i + \frac{1}{c^2}\sum_{i=k}^{n} X_i\right)^2}{m+k-1+(n-k+1)/c^2}}\right)^{\frac{m+n-1}{2}}.$$

Hence,

$$\frac{\Lambda_k(n)}{\Lambda_k^{\text{fs}}(n)} = \frac{\exp\left\{-\frac{1}{2}\left(1-\frac{1}{c^2}\right)\sum_{i=k}^{n} X_i^2\right\}}{\sqrt{\frac{m+k-1}{m+n}+\frac{1}{c^2}\left(1-\frac{m+k-1}{m+n}\right)}} \times$$

$$\left(1+\frac{u_k(n)-\left(1-\frac{1}{c^2}\right)\sum_{i=k}^n X_i^2}{(m+n-1)\hat{\sigma}_n^2}\right)^{-\frac{m+n-1}{2}},$$

were

$$\hat{\sigma}_n^2=\frac{1}{m+n-1}\left[\sum_{i=-m}^{n} X_i^2-\frac{(\sum_{i=-m}^n X_i)^2}{m+n}\right]$$

and

$$u_k(n)=\frac{\left(\sum_{i=-m}^n X_i\right)^2}{m+n}-\frac{\left(\sum_{i=-m}^{k-1} X_i+\frac{1}{c^2}\sum_{i=k}^n X_i\right)^2}{m+k-1+(n-k+1)/c^2}.$$

Denote $b_k(n)=(1-1/c^2)\sum_{i=k}^n X_i^2$. It follows that

$$\begin{aligned} 2\log\frac{\Lambda_k(n)}{\Lambda_k^{\text{fs}}(n)} &\geq -\frac{1}{\hat{\sigma}_n^2}[b_k(n)|\hat{\sigma}_n^2-1|+|u_k(n)|]-\mathrm{O}\left(\frac{n-k}{n+m}\right), \\ 2\log\frac{\Lambda_k(n)}{\Lambda_k^{\text{fs}}(n)} &\leq \frac{1}{\hat{\sigma}_n^2}\frac{b_k(n)|\hat{\sigma}_n^2-1|+|u_k(n)|}{1+\frac{u_k(n)-b_k(n)}{(m+n-1)\hat{\sigma}_n^2}}+\mathrm{O}\left(\frac{n-k}{n+m}\right), \end{aligned}$$

where $\mathrm{O}(x)/x$ is a bounded function in x.

Let $B_1=\{\sum_{i=t+1}^{t+r} X_i^2\leq 4r\}$. It follows that over the event B_1 the random variables $b_k(n)$, $|\sum_{i=t+1}^{k-1} X_i|$ and $|\sum_{i=k}^n X_i|$ are all bounded by a constant times r. The bound is uniform in k and n, where $t<k\leq n\leq t+r$. Hence, over the event B_1, $|u_k(n)|\leq Z^2\times\mathrm{O}(r/t)+|Z|\times\mathrm{O}\left(r/\sqrt{t}\right)+\mathrm{O}(r^2/t)$, with $Z=\sum_{i=-m}^t X_i/\sqrt{t+m}$. Therefore, given any $\epsilon'>0$, $\mathrm{P}_\infty\left(B_1,\sup_{t<k\leq n\leq t+r}|u_k(n)|>\epsilon'\right)<e^{-dt/r^2}$, for some positive $d=d(\epsilon')$. The distribution of both $(m+n-1)\hat{\sigma}_n$ and $\sum_{i=t+1}^{t+r} X_i^2$ is χ^2. The moment generating function of which and large deviation theory can be used to show that for any $\epsilon'>0$

$$\begin{aligned} \mathrm{P}_\infty\left(\sup_{t<n\leq t+r}|\hat{\sigma}_n^2-1|>\epsilon'/r\right) &\leq re^{-gt/r^2}, \quad \text{and} \\ \mathrm{P}_\infty\left(\sum_{i=t+1}^{t+r} X_i^2>4r\right) &\leq e^{-r/2}. \end{aligned}$$

with $g=g(\epsilon')>0$. Condition 1 is accounted for by taking $r=r(A)$ such that $r^2\log A/A\to 0$, as $A\to\infty$, but $r/\log A\to\infty$.

In order to verify Condition 2 notice that

$$\mathrm{P}_k(N_A > n) = \mathrm{P}_k(\max_{j\le n} \textstyle\sum_{l=1}^{j} \Lambda_l(j) < A) \le \mathrm{P}_k\left(\Lambda_k(j) < A\right),$$

for any j, $k \le j \le n$. In particular, let $j = j_k = k + \lfloor \epsilon'' r \rfloor$, for some $\epsilon_2/3 > \epsilon'' > 0$ and for all k such that $\epsilon_2 r/3 < k < n - \epsilon_2 r/3$. Consider the events

$$\begin{aligned} B_2 &= \{|\textstyle\sum_{i=k}^{j} X_i^2 - (j-k+1)c^2| \le \epsilon''(j-k+1)\}, \\ B_3 &= \{\hat\sigma_j^2 \le 1 + \epsilon''\}, \\ B_4 &= \{u_k(j) \ge -\epsilon''(j-k+1)\}, \end{aligned}$$

Note that

$$2\log \Lambda_k(j) \ge (1-c^{-2})\hat\sigma_j^{-2} \sum_{i=k}^{j} X_i^2 - \log(c^2)(j-k+1) + u_k(j)/\hat\sigma_j^2 - \mathrm{o}\,(1)\,,$$

where $\mathrm{o}\,(1) \to 0$ as $\epsilon'' \to o$. It follows that if ϵ'' is small enough but $\epsilon'' r/\log A$ is large then

$$\mathrm{P}_k(\Lambda_k(j) < A) \le \mathrm{P}_k(B_2^c) + \mathrm{P}_k(B_3^c, B_2) + \mathrm{P}_k(B_4^c, B_2).$$

Large deviation arguments can be used to show that $\mathrm{P}_k(B_2^c) \le \exp\{-d_2 r\}$, for some positive d_2 that depends on ϵ''. On the event B_2 the relation $\hat\sigma_j^2 \le \sum_{i=-m}^{k-1} X_i^2/(j+m-1) + (c^2+\epsilon'')(j-k+1)/(j+m-1)$ holds. This can be used to show that $\mathrm{P}_k(B_3^c, B_2) \le \exp\{-d_3 k\}$, for some positive d_3 (that depends on ϵ''). Finally, on the event B_2,

$$u_k(j) = Z^2 \times \mathrm{O}\left((j-k)/k\right) + Z \times \mathrm{O}\left((j-k)/\sqrt{k}\right) + \mathrm{o}\,(1)\,,$$

where $Z = \sum_{i=-m}^{k-1} X_i/\sqrt{k+m-1}$, $\mathrm{O}\,(\cdot)$ is a bounded function, and $\mathrm{o}\,(1) \to 0$ as $\epsilon'' \to 0$. It follows that $\mathrm{P}_k(B_4^c, B_2) \le \exp\{-d_4 k\}$, for yet another $d_4 > 0$.

The above claims can be summed up in order to concluded that for some $d > 0$

$$\sum_{k=1}^{n} \mathrm{P}_k(N_A > n) \le (2/3)\epsilon_2 r + 3AC_2 e^{-dr}.$$

Condition 2 thus follows, provided that $r = r(A)$ is such that $r/\log A \to \infty$, as $A \to \infty$.

Consider next Example 1 – detecting a change in mean when the initial mean and variance are both unknown. We consider the case where the representative post-change distribution is $N(\mu + \delta\sigma, \sigma^2)$, where δ has a fixed (known) value. We apply Theorem 1, again assuming for convenience (without loss of generality) that the pre-change distribution is $N(0, 1)$.

Recall that

$$\hat{\sigma}_n^2 = \frac{1}{n+m-1}\left[\sum_{i=-m}^{n} X_i^2 - \frac{\left(\sum_{i=-m}^{n} X_i\right)^2}{m+n}\right].$$

and define $\bar{X}_n = \sum_{i=-m}^{n} X_i/(m+n)$. It is shown in an appendix that the likelihood ratio of the invariant statistics for the case of detecting a change in the mean is given by

$$\Lambda_k(n) = \mathrm{E}\exp\left\{\sqrt{W}\delta\frac{\sum_{i=k}^{n}(X_i - \bar{X}_n)}{\hat{\sigma}_n} - (n-k+1)\frac{\delta^2}{2} + \frac{(n-k+1)^2\delta^2}{2(n+m)}\right\},$$

where expectation is with respect to W. The random variable W is independent of the observations X_i, $-m \le i \le n$, and has Gamma distribution with both shape and scale equal to $(m+n-1)/2$.

The fully-specified likelihood ratio for this case is given by

$$\Lambda_k^{\mathrm{fs}}(n) = \exp\left\{\delta\sum_{i=k}^{n} X_i - (n-k+1)\delta^2/2\right\}.$$

Hence,

$$\frac{\Lambda_k(n)}{\Lambda_k^{\mathrm{fs}}(n)} = \exp\left\{\frac{(n-k+1)^2\delta^2}{2(n+m)} - \delta(n-k+1)\bar{X}_n\right\}$$

$$\times \mathrm{E} \exp\left\{ \delta \sum_{i=k}^{n} (X_i - \bar{X}_n)(\sqrt{W}/\hat{\sigma}_n - 1) \right\}.$$

The crucial part in the estimation of the ratio between the invariant and the fully-specified likelihood ratios — thus showing Conditions 1 and 2 — depends on bounding the term $\Delta_k(n) = \Delta_k(X_{-m}, \ldots, X_n, n)$, where

$$\Delta_k(n) = \mathrm{E}\left[\exp\left\{\left(\sqrt{W}/\hat{\sigma}_n - 1\right) v_k(n)\right\}\right] - 1,$$

and $v_k(n) = \delta\left(\sum_{i=k}^{n} X_i - (n-k+1)\bar{X}_n\right)$.

It follows that

$$\begin{aligned} |\Delta_k(n)| &= \mathrm{E}\left[\exp\left\{\left(\sqrt{W}/\hat{\sigma}_n - 1\right) v_k(n)\right\} \mathbb{I}(\sqrt{W}/\hat{\sigma}_n > 1)\right] \\ &\quad -\mathrm{E}\left[\exp\left\{\left(\sqrt{W}/\hat{\sigma}_n - 1\right) v_k(n)\right\} \mathbb{I}(\sqrt{W}/\hat{\sigma}_n < 1)\right] \\ &\quad +\mathrm{P}(\sqrt{W}/\hat{\sigma}_n < 1) - \mathrm{P}(\sqrt{W}/\hat{\sigma}_n > 1), \end{aligned}$$

where, again, the computation of the probability and of the expectation is with respect to W. However, on the event $\{\sqrt{W}/\hat{\sigma}_n > 1\}$,

$$\sqrt{W}/\hat{\sigma}_n - 1 < \frac{W/\hat{\sigma}_n^2 - 1}{2}$$

and $\mathrm{E}\left[\exp\left\{(1/2)(W/\hat{\sigma}_n^2 - 1)v_k(n)\right\} \mathbb{I}(W/\hat{\sigma}_n^2 > 1)\right]$ is equal to

$$\exp\left\{-\frac{v_k(n)}{2}\right\}\left(1 - \frac{v_k(n)}{\hat{\sigma}_n^2(m+n-1)}\right)^{-\frac{m+n-1}{2}}$$
$$\times \mathrm{P}\left(W/\hat{\sigma}_n^2 > 1 - \frac{v_k(n)}{\hat{\sigma}_n^2(m+n-1)}\right)$$

Likewise, on the event $\{\sqrt{W}/\hat{\sigma}_n < 1\}$,

$$\sqrt{W}/\hat{\sigma}_n - 1 > W/\hat{\sigma}_n^2 - 1$$

and $\mathrm{E}\left[\exp\left\{(W/\hat{\sigma}_n^2 - 1)v_k(n)\right\} \mathbb{I}(W/\hat{\sigma}_n^2 < 1)\right]$ is equal to

$$\exp\{-v_k(n)\}\left(1 - \frac{2v_k(n)}{\hat{\sigma}_n^2(m+n-1)}\right)^{-\frac{m+n-1}{2}}$$

$$\times \mathrm{P}\left(W/\hat{\sigma}_n^2 < 1 - \frac{2v_k(n)}{\hat{\sigma}_n^2(m+n-1)}\right)$$

The above discussion leads to the following approximation:

$$\begin{aligned}
|\Delta_k(n)| &\leq \left|\exp\left\{-\frac{v_k(n)}{2}\right\}\left(1 - \frac{v_k(n)}{\hat{\sigma}_n^2(m+n-1)}\right)^{-\frac{m+n-1}{2}} - 1\right| \\
&+ \left|\exp\left\{-v_k(n)\right\}\left(1 - \frac{2v_k(n)}{\hat{\sigma}_n^2(m+n-1)}\right)^{-\frac{m+n-1}{2}} - 1\right| \\
&+ \mathrm{P}\left(\left|W/\hat{\sigma}_n^2 - 1\right| \leq \frac{2\,|v_k(n)|}{\hat{\sigma}_n^2(m+n-1)}\right).
\end{aligned}$$

However, the mode of a $\Gamma(s,s)$ distribution is attained at $(s-1)/s$. It follows that the last expression in the above approximation is bounded by a constant times $|v_k(n)|/\sqrt{m+n+1}$.

This approximation, together with arguments parallel to those used for Example 2, can be applied in order to show that conditions 1 and 2 hold for Example 1.

Acknowledgment. This work has benefited greatly from discussions with Moshe Pollak, to which the author is indebted.

Appendix

We seek a convenient representation of the likelihood ratio of the invariant statistics for the case of a change in the mean. Let there be given the random couple (T,S). Consider two distinct (joint) distributions for this element. For ease of notations, assume that the distributions are represented by two continuous densities with a common support: $f_k(t,s)$ and $f_\infty(t,s)$. Assume, furthermore, that the marginal densities of S, both under P_k and under P_∞, are identical. Hence,

$$\frac{f_k(t)}{f_\infty(t)} = \frac{\int f_k(t,s)\mathrm{d}s}{\int f_\infty(t,s)\mathrm{d}s}$$

$$= \int \frac{f_k(t,s)}{f_\infty(t,s)} \frac{f_\infty(t,s)}{\int f_\infty(t,u)\mathrm{d}u} \mathrm{d}s$$

$$= \int \frac{f_k(t|s)}{f_\infty(t|s)} f_\infty(s|t)\mathrm{d}s.$$

If, in particular, $T = (T_1, \ldots, T_n)$ and $S = (U, V)$, where $U = \bar{X}_{-m}$ and $V = (m-1)S_{-m}^2$, then the conditional likelihood ratio becomes

$$\frac{f_k(t|s)}{f_\infty(t|s)} = \exp\left\{ \frac{\sqrt{v}}{\sqrt{m-1}} \delta \sum_{i=k}^{n} t_i - (n-k+1)\frac{\delta^2}{2} + (n-k+1)\delta u \right\} \quad (5)$$

The random variables U and V are independent. The marginal distribution of V is $\chi^2_{(m-1)}$ and the marginal distribution of U is $N(0, 1/m)$. Standard Bayesian argumentation can be used to show that the conditional distribution of U, given V and $\mathcal{F}_n$, is Gaussian. The conditional mean and variance of that distribution are given by

$$\mathrm{E}(U \mid v, \mathcal{F}_n) = -\frac{\sqrt{v}}{\sqrt{m-1}} \frac{\sum_{i=1}^n t_i}{n+m}, \quad \mathrm{var}(U \mid v, \mathcal{F}_n) = \frac{1}{n+m}.$$

The conditional distribution of V, given $\mathcal{F}_n$, is a Gamma distribution with shape parameter given by $(n+m-1)/2$ and scale parameter given by

$$\frac{1}{2(m-1)} \left[\sum_{i=1}^{n} t_i^2 + m - 1 - \frac{\left(\sum_{i=1}^n t_i\right)^2}{n+m} \right].$$

Integrating the conditional likelihood ratio (5) with respect to the conditional distribution of U yields

$$\exp\left\{ \frac{\sqrt{v}}{\sqrt{m-1}} \delta \sum_{i=k}^{n} \left(t_i - \frac{\sum_{i=1}^n t_i}{n+m} \right) - (n-k+1)\frac{\delta^2}{2} + \frac{(n-k+1)^2\delta^2}{2(n+m)} \right\}$$

Define the random variable W by

$$W = \frac{V}{m-1} \cdot \frac{1}{m+n-1} \cdot \left[\sum_{i=1}^{n} T_i^2 + m - 1 - \frac{\left(\sum_{i=1}^n T_i\right)^2}{m+n} \right].$$

Note that W is independent of $\mathcal{F}_n$ and has a Gamma distribution with both shape and scale equal to $(m+n-1)/2$. Moreover, it can be shown that

$$(n+m-1)\sum_{i=k}^{n}\left(T_i - \frac{\sum_{i=1}^{n} T_i}{n+m}\right) \Big/ \left[\sum_{i=1}^{n} T_i^2 + m - 1 - \frac{\left(\sum_{i=1}^{n} T_i\right)^2}{m+n}\right]$$

is equal to $\sum_{i=k}^{n}\left(X_i - \bar{X}_n\right)/\hat{\sigma}_n^2$, where $\bar{X}_n = \sum_{i=-m}^{n} X_i/(m+n)$ and

$$\hat{\sigma}_n^2 = \frac{1}{n+m-1}\left[\sum_{i=-m}^{n} X_i^2 - \frac{\left(\sum_{i=-m}^{n} X_i\right)^2}{m+n}\right].$$

Therefore,

$$\Lambda_k(n) = \mathrm{E}\exp\left\{\sqrt{W}\delta\frac{\sum_{i=k}^{n}\left(X_i - \bar{X}_n\right)}{\hat{\sigma}_n} - (n-k+1)\frac{\delta^2}{2} + \frac{(n-k+1)^2\delta^2}{2(n+m)}\right\},$$

where expectation is with respect to W.

Remark : The computation of the likelihood ratio $\Lambda_k(n)$ involves integration.

□

References

[1] Bell, C., Gordon, L. and Pollak M. (1992). An efficient nonparametric detection scheme with applications to surveillance of a Bernoulli process with unknown baseline. In Carlestin, E., Möller, H.G. and Siegmund, D., editors, *IMS Lecture Notes Monograph Series* **23** 7–27.

[2] Damian, D. (1994). Detecting a decrease in a normal variance with unknown initial mean and variance. Master's thesis, The Hebrew University.

[3] Gordon, L. and Pollak, M. (1994). An efficient sequential nonparametric scheme for detecting a change in distribution. *Ann. Statist.* **22** 763–804.

[4] Gordon, L. and Pollak, M. (1995). A robust surveillance scheme for stochastically ordered alternatives. *Ann. Statist.* **23** 1350–1375.

[5] Gordon, L. and Pollak, M. (1997). Average run length to false alarm for surveillance scheme designed with partially specified pre-change distribution. To appear in *Ann. Statist.*

[6] Lai, T. L. (1995). Sequential change-point detection in quality control and dynamical systems. *JRSSB* **57**(4) 613–658.

[7] Macdonald, D. (1990). A CUSUM procedure based on sequential ranks. *Navel Research Logistics* **37** 627–646.

[8] Pollak, M. (1987). Average run length of an optimal method of detecting a change in distribution. *Ann. Statist.* **13** 206–227.

[9] Pollak, M. and Siegmund, D. (1991). Sequential detecting of a change in a normal mean when the initial value is unknown. *Ann. Statist.* **19** 394–416.

[10] Siegmund, D. and Venkatraman, E. S. (1995). Using the generalized likelihood ratio statistics for sequential detection of a change-point. *Ann. Statist.* **23** 255–271.

[11] van Dobben de Bruyn, C. S. *Cumulative Sum Tests.* Griffin, London, 1968.

[12] Wilson, D. W., Griffiths, K., Kemp, K. W., Nix, A. B. J. and Rowlands, R. J. (1979). Internal quality control of radioimmunoassays: Monitoring of error. *J. Endocr.* **80** 365–372.

[13] Yakir, B. (1995). A note on the run length to false alarm of a change-point detection policy. *Ann. Statist.* **23** 272–281.