

Chapter 7

An Exact Formula for the Average Run Length to False Alarm of the Generalized Shiryaev–Roberts Procedure for Change-Point Detection under Exponential Observations

Wenyu Du, Grigory Sokolov, and Aleksey S. Polunchenko

Abstract We derive analytically an exact closed-form formula for the standard minimax Average Run Length (ARL) to false alarm delivered by the Generalized Shiryaev–Roberts (GSR) change-point detection procedure devised to detect a shift in the baseline mean of a sequence of independent exponentially distributed observations. Specifically, the formula is found through direct solution of the respective integral (renewal) equation, and is a general result in that the GSR procedure’s non-negativ headstart is not restricted to a bounded range, nor is there a “ceiling” value for the detection threshold. Apart from the theoretical significance (in change-point detection, exact closed-form performance formulae are typically either difficult or impossible altogether to get, especially for the GSR procedure), the obtained formula is also useful to a practitioner: in cases of practical interest, the formula is a function linear in both the detection threshold and the headstart, and, therefore, the ARL to false alarm of the GSR procedure can be easily computed.

7.1 Introduction

Quickest change-point detection is concerned with the design and analysis of reliable statistical machinery for rapid detection of changes that may spontaneously affect a “live” process, continuously monitored via sequentially made observations. See, e.g., [24] or [33, Part II]. A quickest change-point detection procedure is a stopping time adapted to the observed data, and is a rule whereby one is to stop and “sound an alarm” that the characteristics of the observed process may have (been) changed. A “good” (i.e., optimal or nearly optimal) detection procedure is one that

W. Du (✉) · G. Sokolov · A.S. Polunchenko
Department of Mathematical Sciences, State University of New York at Binghamton,
Binghamton, NY 13902-6000, USA
e-mail: wdu1@binghamton.edu

G. Sokolov
e-mail: gsokolov@binghamton.edu

A.S. Polunchenko
e-mail: aleksey@binghamton.edu

minimizes (or nearly minimizes) the desired detection delay penalty, subject to a constraint on the false alarm risk. For an overview of the major optimality criteria see, e.g., [18, 23, 32, 38] or [33, Part II].

A problem particularly persistent in applied change-point detection (e.g., in quality control) is evaluation of detection procedures' performance. To that end, the ideal would be to have the needed performance metrics expressed exactly and in a closed and simple form. However, this is generally quite difficult mathematically, if at all possible. Part of the reason is that the renewal equations that many popular performance metrics satisfy are Fredholm integral equations of the second kind (possibly written as equivalent differential equations), and such equations seldom allow for an analytical solution. As a result, the standard practice has been to evaluate the performance numerically (one particularly popular approach has been to devise an asymptotic approximation of some sort). Nevertheless, some exact performance formulae have been derived explicitly, although primarily for the "mainstream" detection methods. For instance, a number of characteristics of the celebrated CUSUM "inspection scheme" (due to [13]) have been expressed explicitly, e.g., in [1, 2, 6, 7, 25, 37],¹ although for only a handful of scenarios. Likewise, exact closed-form formulae for various performance metrics of the famous EWMA chart (due to [26]) in an exponential scenario have been established, e.g., in [3, 12, 21] (see footnote 1).

However, the corresponding progress made to date for the classical Shiryaev–Roberts (SR) procedure (due to [27–29]) is far more modest (except for the continuous-time case), and especially little has been done for the Generalized SR (GSR) procedure, which was introduced recently in [11] as a "headstarted" version of the classical SR procedure. Since the latter is a special case of the GSR procedure (when the headstart is zero), from now on we will follow [34] and use the term "GSR procedure" to refer to both procedures. As a matter of fact, to the best of our knowledge, exact and explicit formulae for a small subset of characteristics of the GSR procedure have been obtained only in [4, 9, 10, 14, 22, 23, 35, 40]. The purpose of this work is to add on to this list. Specifically, we obtain an exact, closed-form formula for the standard (minimax) Average Run Length (ARL) to false alarm delivered by the GSR procedure devised to detect a jump in the common baseline mean of a sequence of independent exponentially distributed observations. The formula is found analytically, through direct solution of the respective renewal (integral) equation, and is valid for an arbitrary (nonnegative) headstart, with the detection threshold not restricted from above. Furthermore, the formula is remarkably simple (it is a function linear in the detection threshold and in the headstart) and, unlike its complicated and cumbersome CUSUM and EWMA counterparts, *can* be used to compute the GSR procedure's ARL to false alarm (in the exponential scenario) essentially "*by hand*". This would clearly be of aid to a practitioner.

¹By no means is this an exhaustive list of available papers on the subject.

7.2 Preliminaries

The centerpiece of this work is the (minimax) Average Run Length (ARL) to false alarm of the Generalized Shiryaev–Roberts (GSR) detection procedure (due to [11]) considered in the context of the basic minimax quickest change-point detection problem (see, e.g., [8, 14]). As a performance metric, the ARL to false alarm was apparently introduced in [13]; see also, e.g., [8].

Let $f_\infty(x)$ and $f_0(x)$ denote, respectively, the observations' pdf in the pre- and post-change regime. Let $\Lambda_n \triangleq f_0(X_n)/f_\infty(X_n)$ be the “instantaneous” likelihood ratio (LR) for the n -th data point, X_n . The GSR procedure (due to [11]) is then formally defined as the stopping time

$$\mathcal{S}_A^r \triangleq \inf\{n \geq 1: R_n^r \geq A\}, \quad \text{such that } \inf\{\emptyset\} = \infty, \quad (7.1)$$

where $A > 0$ is a detection threshold used to control the false alarm risk, and

$$R_{n+1}^r = (1 + R_n^r)\Lambda_{n+1} \quad \text{for } n = 0, 1, \dots \text{ with } R_0^r = r \geq 0, \quad (7.2)$$

is the GSR detection statistic. We remark that $R_0^r = r \geq 0$ is a design parameter referred to as the headstart and, in particular, when $R_0^r = r = 0$, the GSR procedure is equivalent to the classical Shiryaev–Roberts (SR) procedure (due to [27–29]); a brief account of the SR procedure's history may be found, e.g., in [16]. Albeit “young” (the GSR procedure was proposed in 2011), it has already been shown (see, e.g., [17, 22, 30, 34, 35]) to possess very strong optimality properties, not exhibited by the CUSUM scheme or the EWMA chart; in fact, in certain scenarios, the latter two charts have been found experimentally to be inferior to the GSR procedure.

Let \mathbb{P}_∞ (\mathbb{E}_∞) be the probability measure (expectation) induced by the observations in the pre-change regime, i.e., when $X_n \propto f_\infty(x)$ for all $n \geq 1$. The ARL to false alarm of the GSR procedure is defined as $\text{ARL}(\mathcal{S}_A^r) \triangleq \mathbb{E}_\infty[\mathcal{S}_A^r]$. A key property of the GSR statistic (7.2) is that the sequence $\{R_n^r - n - r\}_{n \geq 0}$ is a zero-mean \mathbb{P}_∞ -martingale, i.e., $\mathbb{E}_\infty[R_n^r - n - r] = 0$ for all $n \geq 0$ and all r . This and Doob's Optional stopping (sampling) theorem (see, e.g., [33, Theorem 2.3.1, p. 31]) imply that $\mathbb{E}_\infty[R_{\mathcal{S}_A^r}^r - \mathcal{S}_A^r - r] = 0$, so that $\text{ARL}(\mathcal{S}_A^r) = \mathbb{E}_\infty[R_{\mathcal{S}_A^r}^r] - r \geq A - r$. As a result, to ensure that $\text{ARL}(\mathcal{S}_A^r) \geq \gamma$ for a desired $\gamma > 1$, it suffices to pick A and r from the solution set of the inequality $A - r \geq \gamma$ and such that $A > 0$ and $r \geq 0$.

A more accurate result is the approximation $\text{ARL}(\mathcal{S}_A^r) \approx (A/\xi) - r$ valid for sufficiently large $A > 0$; see, e.g., [15, Theorem 1] or [34]. To define ξ , let $S_n \triangleq \sum_{i=1}^n \log \Lambda_i$ for $n \geq 1$, and let $\tau_a \triangleq \inf\{n \geq 1: S_n \geq a\}$ for $a > 0$ (again, with the understanding that $\inf\{\emptyset\} = \infty$). Then $\kappa_a \triangleq S_{\tau_a} - a$ is the so-called “overshoot” (excess over the level $a > 0$ at stopping), and $\xi \triangleq \lim_{a \rightarrow \infty} \mathbb{E}_0[e^{-\kappa_a}]$, and is referred to as the “limiting average exponential overshoot”; here \mathbb{E}_0 denotes the expectation under the probability measure induced by the observations in the post-change regime, i.e., when $X_n \propto f_0(x)$ for all $n \geq 1$. In general, ξ is clearly between 0 and 1, and is a model-dependent constant, which falls within the scope of nonlinear renewal theory; see, e.g., [39], [38, Section II.C] or [33, Section 2.6].

We now state the main equation that we shall deal with (and, in fact, solve analytically) in the next section in a certain exponential scenario. Let $P_\infty^A(t) \triangleq$

$\mathbb{P}_\infty(A_1 \leq t)$, $t \geq 0$, be the cdf of the LR under probability measure \mathbb{P}_∞ . Let $R_0^{r=x} = r = x \geq 0$ be fixed and define

$$\mathcal{K}_\infty(x, y) \triangleq \frac{\partial}{\partial y} \mathbb{P}_\infty(R_{n+1}^r \leq y | R_n^r = x) = \frac{\partial}{\partial y} P_\infty^A\left(\frac{y}{1+x}\right), \quad \text{for } x, y \geq 0, \quad (7.3)$$

i.e., the transition probability density kernel for the homogeneous Markov process $\{R_n^r\}_{n \geq 0}$ under probability measure \mathbb{P}_∞ .

From now on, let $\ell(x, A) \triangleq \text{ARL}(\mathcal{S}_A^{r=x})$. It is shown, e.g., in [11], that $\ell(x, A)$ is governed by the renewal equation

$$\ell(x, A) = 1 + \int_0^A \mathcal{K}_\infty(x, y) \ell(y, A) dy, \quad (7.4)$$

where $x \geq 0$ and $A > 0$. The question of existence and uniqueness of solution for this equation has been answered in the affirmative, e.g., in [11]. It is this equation, viz. the exact solution thereof in a specific exponential scenario, that is the centerpiece of this work.

Equation (7.4) is a Fredholm (linear) integral equation of the second kind. Since for such equations an analytical solution is rarely a possibility, they are usually solved numerically. Numerical schemes specifically for Eq. (7.4) have been developed and applied, e.g., in [11, 20, 36]. However, it turns out that in a certain exponential scenario it is possible to solve (7.4) analytically, and, more importantly, the solution is a simple linear function of x and A , just as one would expect from the approximation $\text{ARL}(\mathcal{S}_A^r) \approx (A/\xi) - r$ mentioned earlier. This is the main result of this paper, it generalizes [5, Proposition 1], and the details are given in the next section.

7.3 The Main Result

We are now in a position to establish the main result of this work, i.e., derive analytically an exact closed-form formula for the ARL to false alarm exhibited by the GSR procedure (7.1)–(7.2) “tasked” to detect a change in the baseline (common) mean of a series of independent exponentially distributed observations. More concretely, suppose the observations’ pre- and post-change pdf’s are

$$f_\infty(x) = e^{-x} \mathbb{1}_{\{x \geq 0\}} \quad \text{and} \quad f_0(x) = \frac{1}{1+\theta} e^{-x/(1+\theta)} \mathbb{1}_{\{x \geq 0\}}, \quad (7.5)$$

respectively, where $\theta > 0$, a known parameter with an obvious interpretation: it is the magnitude of the shift in the mean of the exponential distribution, so that the higher (lower) the value of θ , the more (less) contrast the mean shift is, and the easier (harder) it is to detect. We shall from now on refer to this scenario as the $\mathcal{E}(1)$ -to- $\mathcal{E}(1+\theta)$ model, to reflect not only the throughout “exponentiality” of the data, but also that their mean is 1 pre-change and $1+\theta > 1$ post-change. For a motivation to consider this model, see, e.g., [4, 31], or [33, Section 3.1.6].

To “tailor” the general equation (7.4) on the ARL to false alarm to the $\mathcal{E}(1)$ -to- $\mathcal{E}(1 + \theta)$ model, the first step is to find $\Lambda_n \triangleq f_0(X_n)/f_\infty(X_n)$. To that end, it is easy to see from (7.5) that

$$\Lambda_n = \frac{1}{1 + \theta} \exp\left\{\frac{\theta}{1 + \theta} X_n\right\}, \quad n \geq 1, \quad (7.6)$$

and we note that since $X_n \geq 0$ w.p. 1 for all $n \geq 1$ under any probability measure, it can be deduced that $\Lambda_n \geq 1/(1 + \theta)$ w.p. 1 for all $n \geq 1$, also under any probability measure. The latter inequality is a circumstance with consequences, which are illustrated in the following two results.

Lemma 7.1 *For the $\mathcal{E}(1)$ -to- $\mathcal{E}(1 + \theta)$ model (7.5), the pre-change transition probability density kernel, $\mathcal{K}_\infty(x, y)$, defined by (7.3), is given by the formula:*

$$\mathcal{K}_\infty(x, y) = \theta^{-1} (1 + \theta)^{-1/\theta} y^{-2-1/\theta} (1 + x)^{1+1/\theta} \mathbb{1}_{\{y \geq (1+x)/(1+\theta)\}}, \quad (7.7)$$

where it is understood that $x \geq 0$.

Proof The desired result can be established directly from (7.3), i.e., the definition of the pre-change transition probability density kernel, $\mathcal{K}_\infty(x, y)$, combined with (7.6), i.e., the formula for the LR specific to the $\mathcal{E}(1)$ -to- $\mathcal{E}(1 + \theta)$ model (7.5). The presence of the indicator function in the right-hand side of (7.7) is an implication of the aforementioned inequality $\Lambda_n \geq 1/(1 + \theta)$ valid w.p. 1 for all $n \geq 1$ and under any probability measure. \square

Now, with (7.7) put in place of $\mathcal{K}_\infty(x, y)$ in the general equation (7.4) the latter takes on the form

$$\ell(x, A) = 1 + \theta^{-1} (1 + \theta)^{-1/\theta} (1 + x)^{1+1/\theta} \int_{(1+x)/(1+\theta)}^A y^{-2-1/\theta} \ell(y, A) dy, \quad (7.8)$$

where $x \geq 0$ and $A > 0$, and we recall that $\ell(x, A) \triangleq \mathbb{E}_\infty[\mathcal{S}_A^{r=x}]$. It is this equation that we shall now attempt solve explicitly. To that end, a natural point of departure here would be the aforementioned approximation $\text{ARL}(\mathcal{S}_A^r) \approx (A/\xi) - r$, where ξ is the limiting average exponential overshoot formally defined in the preceding section. It is known (see, e.g., [31]) that $\xi = 1/(1 + \theta) \in (0, 1)$ for the $\mathcal{E}(1)$ -to- $\mathcal{E}(1 + \theta)$ model (7.5). Hence, at least for large enough A 's, the solution to (7.8) should behave roughly as $\ell(x, A) \approx A(1 + \theta) - x$. As will be shown shortly, this is, in fact, precisely the behavior of the solution, without A having to be large. However, the aforementioned fact that $\Lambda_n \geq 1/(1 + \theta)$ w.p. 1 under any measure makes things a bit complicated.

Lemma 7.2 *For the $\mathcal{E}(1)$ -to- $\mathcal{E}(1 + \theta)$ model (7.5), at each epoch $n \geq 0$ and under any probability measure, the GSR statistic R_n^r has a deterministic lower bound, i.e.,*

$R_n^r \geq B_n^r$ w.p. 1, for each $n \geq 0$ and under any probability measure, where

$$B_n^r \triangleq \frac{1}{\theta} \left[1 - \frac{1}{(1+\theta)^n} \right] + \frac{r}{(1+\theta)^n}, \quad n \geq 0, \quad (7.9)$$

and r is the GSR statistic's headstart, i.e., $R_0^r = r \geq 0$.

Proof It is merely a matter of “unfolding” the recursion $R_n^r = (1 + R_{n-1}^r)A_n$, $n \geq 1$, one term at a time, and applying, at each step, the inequality $A_n \geq 1/(1 + \theta)$ valid w.p. 1 under any probability measure. \square

At this point note that since $1 + \theta > 1$, the lower bound sequence $\{B_n^r\}_{n \geq 0}$ given by (7.9) is such that (a) for $r \leq 1/\theta$, it increases monotonically with n , i.e., $r \equiv B_0^r \leq B_1^r \leq B_2^r \leq \dots$, when $r \leq 1/\theta$, and (b) $\lim_{n \rightarrow \infty} B_n^r = 1/\theta$, irrespective of $R_0^r = r \geq 0$. Hence, when $A < 1/\theta$, the GSR statistic, $\{R_n^r\}_{n \geq 0}$, is guaranteed to either hit or exceed the level $A > 0$ within at most m steps, where $m \equiv m(r, A, \theta)$ is found from the inequality $B_m^r \geq A$, i.e.,

$$m \equiv m(r, A, \theta) \triangleq \begin{cases} \lceil (\log \frac{1-\theta r}{1-\theta A}) / \log(1+\theta) \rceil, & \text{for } r < A (< 1/\theta); \\ 1, & \text{for } r \geq A, \end{cases}$$

with $\lceil x \rceil$ denoting the usual “ceiling” function. Therefore, the general solution to (7.8) is dependent upon whether $A < 1/\theta$ or $A \geq 1/\theta$. In the latter case, the (exact) solution is given by the following theorem, which is the main result of this paper.

Theorem 7.1 For the $\mathcal{E}(1)$ -to- $\mathcal{E}(1 + \theta)$ model (7.5), if the detection threshold, $A > 0$, is set so that $A \geq 1/\theta$, then the ARL to false alarm of the GSR procedure is given by the formula:

$$\ell(x, A) = 1 + (1 + \theta) \left(A - \frac{1+x}{1+\theta} \right) \mathbb{1}_{\{(1+x)/(1+\theta) \leq A\}}, \quad (7.10)$$

and it is understood that $x \geq 0$.

Proof It is sufficient to insert (7.10) into Eq. (7.8) and directly verify that the latter does, in fact, “check out”. The condition that $A \geq 1/\theta$ “protects” against the situation described in Lemma 7.2 and in the discussion following it. \square

The special case of Theorem 7.1 when $R_0^{r=x} = r = x \geq 0$ (i.e., when there is no headstart) was previously established in [4, Proposition 1] using the memorylessness of the exponential distribution. It is also noteworthy that formula (7.10) as well as Eq. (7.8) are actually valid for $x \geq -1$; the same can also be said about the general equation (7.4).

We conclude this section with a brief analysis of the case when $A < 1/\theta$. Recall that the integral in the right-hand side of (7.8) plays no role, unless $(1+x)/(1+\theta) < A$. For this condition to hold when $A < 1/\theta$, it must be the case that $(1+x)/(1+\theta) < 1/\theta$, i.e., that $x < 1/\theta$. Hence, if $A < 1/\theta$, then $\ell(x, A) \equiv 1$ for all

$x \geq 1/\theta$. To obtain $\ell(x, A)$ explicitly for $x < 1/\theta$, note that if $x < 1/\theta$, the function $h(x) \triangleq (1+x)/(1+\theta)$, i.e., the lower limit of integration in the integral in the right-hand side of (7.8), is such that $h(x) \geq x$. As a result, the nature of the integral equation becomes such that the unknown function, $\ell(x, A)$, is dependent *solely* upon the values it assumes for higher x 's, and since $\ell(x, A) \equiv 1$ for $x \geq 1/\theta$, one can iteratively work out backwards the solution for any $x \geq 0$. However, this process involves formidable integrals, and only the first few steps seem to be feasible to actually carry out.

While an explicit formula for the ARL to false alarm of the GSR procedure when $A < 1/\theta$ turned out to be problematic to get, from a practical standpoint it might not be worthwhile altogether, for the formula for $A \geq 1/\theta$ alone, i.e., Theorem 7.1, is sufficient. Specifically, since $\text{ARL}(\mathcal{S}_A^r) \geq A - r$, the formula for the ARL to false alarm when $A > 1/\theta$, i.e., formula (7.10), will never yield ARL's lower than $(1/\theta) - r$. However, the size of this “blind spot” is not necessarily large, unless θ is very small, which is to say that the change in the mean in the $\mathcal{E}(1)$ -to- $\mathcal{E}(1+\theta)$ model (7.5) is faint and not worthy of detection to begin with. As an illustration of this point, consider the original SR procedure ($r = 0$) and suppose that θ is 0.01, which, from a practical standpoint, can hardly be considered a “change” in the first place. Yet, since $1/\theta$ in this case is 100, the linear formula for the ARL to false alarm will never yield a value of 100 or less. However, this is unlikely to be of inconvenience to a practitioner, as in most applications the ARL to false alarm is set to be at least in the hundreds, and, when $\theta = 0.01$, these levels of the ARL to false alarms *would be* obtainable through formula (7.10).

7.4 Concluding Remarks

This contribution is part of the authors' ongoing effort (manifested, e.g., in [19, 20], and, with other collaborators, e.g., in [11, 22, 34, 35]) to “pave the way” for further research on the theory and application of the GSR procedure. To that end, case studies involving “stress-testing” the GSR procedure on real data are still an “uncharted territory” and would be of particular interest. Hopefully, the result obtained in this work, the data-analytic advantages pointed out in [5], and the strong optimality properties established, e.g., in [17, 22, 30, 34, 35], will help the GSR procedure rightly stand out as the top tool for change-point detection.

Acknowledgements The authors would like to thank Prof. Sven Knoth of the Helmut Schmidt University, Hamburg, Germany, and Prof. Ansgar Steland of the RWTH Aachen University, Aachen, Germany, for the invitation to contribute this work to the 12-th German–Polish Workshop on Stochastic Models, Statistics and Their Applications. Constructive feedback provided by Dr. Ron Kenett of Israel-based KPA Ltd. (www.kpa-group.com), by Prof. William H. Woodall of Virginia Polytechnic Institute, Blacksburg, Virginia, USA, and by the two anonymous referees is greatly appreciated as well.

The effort of A.S. Polunchenko was supported, in part, by the Simons Foundation (www.simonsfoundation.org) via a Collaboration Grant in Mathematics (Award #304574) and by the Research Foundation for the State University of New York at Binghamton via an Interdisciplinary Collaboration Grant (Award #66761).

Last but not least, A.S. Polunchenko is also indebted to the Office of the Dean of the Harpur College of Arts and Sciences at the State University of New York at Binghamton for the support provided through the Dean's Research Semester Award for Junior Faculty granted for the Fall semester of 2014.

References

1. DeLucia J, Poor HV (1997) Performance analysis of sequential tests between Poisson processes. *IEEE Trans Inf Theory* 43(1):221–238
2. Gan FF (1992) Exact run length distributions for one-sided exponential CUSUM schemes. *Stat Sin* 2:297–312
3. Gan FF (1998) Designs of one- and two-sided exponential EWMA charts. *J Qual Technol* 30(1):55–69
4. Kenett R, Pollak M (1986) A semi-parametric approach to testing for reliability growth, with application to software systems. *IEEE Trans Reliab* 35(3):304–311
5. Kenett R, Pollak M (1996) Data-analytic aspects of the Shiryayev–Roberts control chart: Surveillance of a non-homogeneous Poisson process. *J Appl Stat* 23(1):125–138
6. Knoth S (1995) Quasistationäre CUSUM-Verfahren bei Erlangverteilung. Ph.D. thesis, TU. Chemnitz–Zwickau, Germany (in German)
7. Knoth S (1998) Exact average run lengths of CUSUM schemes for Erlang distributions. *Seq Anal* 17(2):173–184
8. Lorden G (1971) Procedures for reacting to a change in distribution. *Ann Math Stat* 42(6):1897–1908
9. Mei Y (2006) Comments on “A note on optimal detection of a change in distribution,” by Benjamin Yakir. *Ann Stat* 34(3):1570–1576
10. Mevorach Y, Pollak M (1991) A small sample size comparison of the CUSUM and Shiryayev–Roberts approaches to changepoint detection. *Am J Math Manag Sci* 11(3&4):277–298
11. Moustakides GV, Polunchenko AS, Tartakovsky AG (2011) A numerical approach to performance analysis of quickest change-point detection procedures. *Stat Sin* 21(2):571–596
12. Novikov A (1990) On the first exit time of an autoregressive process beyond a level and an application to the “disorder” problem. *Theory Probab Appl* 35(2):269–279
13. Page ES (1954) Continuous inspection schemes. *Biometrika* 41(1&2):100–115
14. Pollak M (1985) Optimal detection of a change in distribution. *Ann Stat* 13(1):206–222
15. Pollak M (1987) Average run lengths of an optimal method of detecting a change in distribution. *Ann Stat* 15(2):749–779
16. Pollak M (2009) The Shiryayev–Roberts changepoint detection procedure in retrospect—Theory and practice. In: *Proceedings of the 2nd International Workshop on Sequential Methodologies*, University of Technology of Troyes, Troyes, France
17. Pollak M, Tartakovsky AG (2009) Optimality properties of the Shiryayev–Roberts procedure. *Stat Sin* 19:1729–1739
18. Polunchenko AS, Sokolov G, Du W (2013) Quickest change-point detection: A bird's eye view. In: *Proceedings of the 2013 Joint Statistical Meetings*, Montréal, Québec, Canada
19. Polunchenko AS, Sokolov G, Du W (2014) Efficient performance evaluation of the Generalized Shiryayev–Roberts detection procedure in a multi-cyclic setup. *Appl Stoch Models Bus Ind* 30(6):723–739
20. Polunchenko AS, Sokolov G, Du W (2014) An accurate method for determining the pre-change run-length distribution of the Generalized Shiryayev–Roberts detection procedure. *Seq Anal* 33(1):112–134
21. Polunchenko AS, Sokolov G, Tartakovsky AG (2014) Optimal design and analysis of the Exponentially Weighted Moving Average chart for exponential data. *Sri Lankan J Appl Statist*. doi:10.4038/sljastats.v5i4.7784

22. Polunchenko AS, Tartakovsky AG (2010) On optimality of the Shiryaev–Roberts procedure for detecting a change in distribution. *Ann Stat* 38(6):3445–3457
23. Polunchenko AS, Tartakovsky AG (2012) State-of-the-art in sequential change-point detection. *Methodol Comput Appl Probab* 44(3):649–684
24. Poor HV, Hadjilias O (2009) *Quickest detection*. Cambridge University Press, New York
25. Regula G (1975) Optimal CUSUM procedure to detect a change in distribution for the Gamma family. Ph.D. thesis, Case Western University, Cleveland, OH
26. Roberts SW (1959) Control chart tests based on geometric moving averages. *Technometrics* 1(3):239–250
27. Roberts SW (1966) A comparison of some control chart procedures. *Technometrics* 8(3):411–430
28. Shiryaev AN (1961) The problem of the most rapid detection of a disturbance in a stationary process. *Sov Math Dokl* 2:795–799
29. Shiryaev AN (1963) On optimum methods in quickest detection problems. *Theory Probab Appl* 8(1):22–46
30. Shiryaev AN, Zryumov PY (2009) On the linear and nonlinear generalized Bayesian disorder problem (discrete time case). In: Delbaen F, Rásonyi M, Stricker Ch (eds) *Optimality and risk—Modern trends in mathematical finance*. The Kabanov Festschrift. Springer, Berlin, pp 227–235
31. Tartakovsky AG, Ivanova IV (1992) Comparison of some sequential rules for detecting changes in distributions. *Probl Inf Transm* 28(2):117–124
32. Tartakovsky AG, Moustakides GV (2010) State-of-the-art in Bayesian changepoint detection. *Seq Anal* 29(2):125–145
33. Tartakovsky A, Nikiforov I, Basseville M (2014) *Sequential analysis: Hypothesis testing and changepoint detection*. CRC Press, Boca Raton
34. Tartakovsky AG, Pollak M, Polunchenko AS (2012) Third-order asymptotic optimality of the Generalized Shiryaev–Roberts changepoint detection procedure. *Theory Probab Appl* 56(3):457–484
35. Tartakovsky AG, Polunchenko AS (2010) Minimax optimality the Shiryaev–Roberts procedure. In: *Proceedings of the 5th International Workshop in Applied Probability*, Universidad Carlos III de Madrid, Colmenarejo Campus, Spain
36. Tartakovsky AG, Polunchenko AS, Moustakides GV (2009) Design and comparison of Shiryaev–Roberts- and CUSUM-type change-point detection procedures. In: *Proceedings of the 2nd International Workshop on Sequential Methodologies*, University of Technology of Troyes, Troyes, France
37. Vardeman S, Ray D (1985) Average run lengths for CUSUM schemes when observations are exponentially distributed. *Technometrics* 27(2):145–150
38. Veeravalli VV, Banerjee T (2013) Quickest change detection. In: Chellappa R, Theodoridis S (eds) *Academic press library in signal processing: Array and statistical signal processing*, vol 3. Academic Press, Oxford, pp 209–256
39. Woodroffe M (1982) *Nonlinear renewal theory in sequential analysis*. SIAM, Philadelphia
40. Yakir B (1997) A note on optimal detection of a change in distribution. *Ann Stat* 25(5):2117–2126