Proceedings of the
44th IEEE Conference on Decision and Control, and
the European Control Conference 2005
Seville, Spain, December 12-15, 2005



# Asymptotically Optimal Sequential Change-Point Detection under Composite Hypotheses

Boris Brodsky and Boris Darkhovsky

***Abstract*— The problem of sequential detection of a change-point in the density function of observations from a sequence of independent random variables is considered when both before and after a change-point this density function belongs to a certain family of distributions, i.e. in the general situation of composite hypotheses. A new quality criterion for change-point detection is proposed. The asymptotic a priori lower bound for this criterion is established for any method of change-point detection. A method of change-point detection is proposed for which this lower bound is attained asymptotically so that the method can be called asymptotically optimal. In particular, for the case of a simple hypothesis before a change-point, this method coincides with the generalized cumulative sums (CUSUM) method.**

## I. INTRODUCTION

The problem of sequential change-point detection has first appeared in 1930s in connection with statistical quality control tests. The first diagnostic test was proposed by Shewhart (the so called "Shewhart chart", [1]). Shewhart introduced the fundamental concept of a "state of statistical control", in which the behavior of some suitable chosen quality characteristic at time $t$ has a given probability distribution. To detect significant departures from this state, he introduced a process inspection scheme that takes samples of fixed size at regular intervals of time and computes from the sample at time $t$ a suitably chosen statistic, which can be presented graphically in the form of a control chart.

In 1950s Girshik and Rubin ([2]), and later Page ([3]) have proposed some more effective methods of sequential change-point detection (called nowadays the GRSh (Girshick-Rubin-Shyriaev) and CUSUM (cumulative sums) test respectively). These tests were based upon the modified Wald theory of sequential hypotheses testing.

In works of Shyriaev (see [4]) the problem of sequential change-point detection was formulated as an extremal problem both for discrete and continuous time and the optimal solution of this problem was found. Lorden [5] has found an asymptotically optimal method of sequential change-point detection which minimizes the average delay time in detection given an upper boundary for the average time before a "false alarm" without a priori assumptions about the density function (d.f.) of a change-point. Pollak ([6]) proved that the method of Girshick and Rubin can be obtained as a certain limit from the method proposed by Shiryaev. He also demonstrated that this method is asymptotically optimal in the sense of Lorden's criterion. Moustakides ([7]) proved that Page's CUSUM procedure is strictly optimal (not only asymptotically) in Lorden's formulation of sequential change-point detection problem. For continuous time this result was obtained by Shiryaev ([8]).

Lai ([9],[10],[11]) considered the change-point problem in the more general situation of dependent random variables. In [9] it was shown that the Neyman–Pearson type procedure with the "moving window" of observations is asymptotically optimal. In [10] different boundaries for the "false alarm" probability were introduced instead of upper boundaries for the average time before the "false alarm" and the asymptotic optimality of Page's CUSUM method was proved. Lai ([11],[12]) developed information-theoretic bounds for sequential multihypothesis testing and fault detection in stochastic systems.

Different modifications and generalizations of the CUSUM method can be found in [13] and [14]. In [15] the sequential change-point problem in Bayesian statement was investigated for dependent and non-stationary observations.

The case of composite hypotheses for the problem of sequential change-point detection is the most interesting for applications, especially for the so called fault detection (FD) problem for dynamic systems. Real-time FD is a decision problem in which the healthy or faulty state of a system has to be inferred from the observation of the available data. The main difficulty in FD is that the observed data depend on nuisances (unknown or uncertain parameters, unknown inputs, unknown initial conditions, unobservable states). The FD problem for general non-linear dynamic systems can be reduced (under natural assumptions) to the change-point detection problem for two collections of probabilistic distributions: one collection corresponds to observable variables *before* the change-point and another collection corresponds to the variables *after* the change-point. The idea of such reduction (for the off-line problem) can be found in [16].

A special variant of the change-point problem for composite hypotheses (when the mathematical expectation of observed Gaussian random variables changes from zero to $\delta$ or $-\delta$) was considered in [17]. For this problem, the asymptotic optimality in Lorden's sense of the generalized likelihood ratio statistic was established. The general form of this statistic can be found in Lai ([18]). The detailed review of the literature on this topic can be also found there.

To the best of our knowledge, the sequential change-

This work was not supported by any organization

B.Brodsky is with Central Institute for Mathematics and Economics RAS, Moscow, Russia bbrodsky@yandex.ru

B. Darkhovsky is with Institute for Systems Analysis RAS, Moscow, Russia darbor@isa.ru

point detection problem in the general context of composite hypotheses, when the d.f. of observations not only *after* the change-point but also *before* it is unknown and belongs to a certain family of distributions, was not considered in the literature.

This paper has two major objectives:

1) to propose a new criterion of quality for sequential change-point detection methods for composite hypotheses. This criterion is based upon the a priori inequality proved by Brodsky and Darkhovsky, [19] (see also [20] and corresponds very well to the intuitive requirements of effectiveness of sequential change-point detection (here it is possible to derive an analogy with the classic Rao-Cramer inequality for an arbitrary estimate of an unknown parameter);

2) to propose an asymptotically optimal detection method (based on the lower bound for the criterion of quality) for the general change-point detection problem with a change of one composite hypothesis to another composite hypothesis. The detection method was earlier announced in our paper [21].

## II. Problem Statement. Assumptions

### A. Problem Statement

On the probability space $(\Omega, \mathcal{F}, \mathbf{P})$ a sequence of independent random vectors $\{\xi_k\}_{k=1}^{\infty}$ is considered. Let $\theta = (\theta_0, \theta_1)$, $\theta \in \Theta$, $\theta_0 \in \Theta_0, \theta_1 \in \Theta_1$, where $\Theta$ is a certain parametric set which belongs to some open set $U$ in the finite dimensional space, $\Theta = \Theta_0 \bigcup \Theta_1$, $\Theta_0 \bigcap \Theta_1 = \emptyset$. Suppose that the density function (d.f.) $\xi$ with respect to some $\sigma$-finite measure $\mu$ is equal to $f(\theta_0, x), \theta_0 \in \Theta_0$ before an unknown change-point, and is equal to $f(\theta_1, x), \theta_1 \in \Theta_1$ after this change-point. The d.f. $f$ is known and defined for all parameter values from $U$.

In what follows we denote by $\mathbf{P}_{m,\theta}(\mathbf{E}_{m,\theta})$ the measure (mathematical expectation) corresponding to a sequence $\{\xi_k\}_{k=1}^{\infty}$ with the change-point at the instant $m$ and the fixed value of the parameter $\theta = (\theta_0, \theta_1)$ (so the d. f. of observations $\xi_n$ is equal to $f(\theta_0, x)$ if $n < m$ and $f(\theta_1, x)$ if $n \geq m$). Symbols $\mathbf{P}_{\infty,\theta}(\mathbf{E}_{\infty,\theta})$ correspond to an observed sequence without change-points. Therefore, in our notations for an arbitrary point $\theta = (\theta_0, \theta_1)$ the measure $\mathbf{P}_{\infty,\theta}$ corresponds to an observed sequence with d.f. $f(\theta_0, x)$ and the measure $\mathbf{P}_{1,\theta}$ corresponds to an observed sequence with d.f. $f(\theta_1, x)$.

The problem consists in sequential detection of a change-point $m$ in the sequence of independent random variables for the case of composite hypotheses $\mathbf{H}_0$ and $\mathbf{H}_1$ about the density function of observations before and after the change-point respectively.

For all known methods of sequential change-point detection, a certain "large parameter" $N$ can be defined (see [20]). In what follows we consider the asymptotics of sequential methods as $N \to \infty$. Below for the asymptotically optimal method, this "large parameter" $N$ will be explicitly defined.

Suppose $a$ is a certain change-point detection method (depending on $N$) and $d_N^a(n)$ is its decision function such that $d_N^a(n) = 1\,(d_N^a(n) = 0)$ corresponds to the decision about the presence (absence) of a change at the instant $n, \quad \tau_N^a = \min\{n : d_N^a(n) = 1\}$ is the stopping time w.r.t. the natural flow of $\sigma$-algebras generated by observations.

Let us introduce the following classes of such stopping times (and corresponding methods of detection)

$$\mathcal{M} = \{\tau_N : \sup_{\theta_1 \in \Theta_1} \mathbf{E}_{1,\theta} \tau_N < \infty\} \tag{1}$$

Consider tests generated by stopping times from the set $\mathcal{M}$. Suppose $\theta_1^*$ is an arbitrary point from the set $\Theta_1$. Starting from the point $\theta_1^*$ we define the point $\theta_0^*(\theta_1^*) = \theta_0^*(\cdot)$ as the argument of the maximum of the following criterion

$$\max_{\theta_0 \in \Theta_0} \left( \int \ln \frac{f(\theta_1^*, x)}{f(\theta_0, x)} f(\theta_1^*, x) d\mu(x) \right)^{-1} = \mathbf{I}_{01}^{-1}\left(\theta_0^*(\cdot), \theta_1^*\right) \tag{2}$$

(below we make assumptions which guarantee that this maximum is attained). If the set of maximum points in this criterion consists of more than one point, then we choose an arbitrary point from this set as $\theta_0^*(\cdot)$ (we will see below that this choice does not influence the optimality rule).

In the sequel we use the notation $\theta^* = (\theta_0^*(\cdot), \theta_1^*)$.

If the point $\theta_1^*$ is the true parameter of the 1-d.f. *after* the change-point, then the point $\theta_0^*(\cdot)$ corresponds to the least favourable alternative *before* the change-point: the Kullback-Leibler distance $\mathrm{dist}\,(\mathbf{P}_{1,\theta^*}, \mathbf{P}_{\infty,\theta^*})$ between distributions $\mathbf{P}_{1,\theta^*}$ and $\mathbf{P}_{\infty,\theta^*}$ is minimal in virtue of the definition of $\theta_0^*(\cdot)$.

Define for any method of detection $a$ and an arbitrary point $\theta^* = (\theta_0^*(\cdot), \theta_1^*)$

$$\alpha_N^a(\theta^*) = \sup_n \mathbf{P}_{\infty,\theta^*}\{d_N^a(n) = 1\}. \tag{3}$$

The value $\alpha_N^a$ is correctly defined for an *arbitrary* "true" parameter $\theta_1^*$ of the d.f. *after* the change-point and can be interpreted as the maximal (for the given point $\theta_1^* \in \Theta_1$ and for the given method $a$) probability of the false alarm under the least favourable alternative *before* the change-point.

Now consider the following *new criterion* of quality in sequential change-point detection:

$$\mathcal{K}_N^a(\theta^*, m) = \frac{\mathbf{E}_{m,\theta^*}(\tau_N^a - m)^+}{|\ln \alpha_N^a(\theta^*)|} \tag{4}$$

(here and below $x^+ = \max(x, 0)$).

Let us explain the sense of this criterion. For all methods of sequential change-point detection, the "large parameter" $N$ can be chosen in such a way that the "false alarm" probability diminishes exponentially as $N \to \infty$ (see [20]). On the other hand, for an arbitrary $\theta = (\theta_0, \theta_1)$ the average time before the false alarm $\mathbf{E}_{\infty,\theta}\tau_N^a$ is of the asymptotic order $(\alpha_N^a(\theta))^{-1}$ as $N \to \infty$ (see details below).

Therefore this criterion $\mathcal{K}_N^a(\theta^*, m)$ represents the ratio (in a certain scale) of the average delay time in change-point detection to the average time before the false alarm for the change-point $m$ and the parameter $\theta^*$. Remind that the average delay time in change-point detection and the average time before the false alarm are the main qualitative characteristics of any method of sequential change-point detection. Usually the problem of sequential change-point detection is formulated as follows: to minimize the average delay time given the lower constraint on the average time before the false alarm (or the upper constraint on the false alarm probability). However, in our opinion, the proposed criterion $\mathcal{K}_N^a(\theta^*, m)$ characterizes the quality of a change-point detection method no worse than these conventional criteria and corresponds very well to the pragmatic sense of the change-point detection problem.

Moreover, the advantage of the proposed criterion consists in the fact that it allows for the asymptotic lower bound (as $N \to \infty$) that depends on the parameter $\theta^*$ only (here it is possible to derive an analogy with the classic Rao-Cramer inequality for an arbitrary estimate of an unknown parameter). This fact in its turn enables us to formulate the problem of the asymptotically optimal change-point detection: to find a method for which the a priori lower bound on $\mathcal{K}_N^a(\theta^*, m)$ is attained as $N \to \infty$.

### B. Assumptions

For any $\theta = (\theta_0, \theta_1)$ consider the following random sequence

$$\eta_n(\theta) = \ln \frac{f(\theta_1, \xi_n)}{f(\theta_0, \xi_n)}$$

Everywhere below we assume that the following conditions are satisfied:

1) $\Theta$ is a compact set[1];

2) $\mu\{x : f(\theta_1, x) \neq f(\theta_2, x)\} > 0$ if $\theta_1 \neq \theta_2$;

3) For $\mu$-a.e. $x$ the functions $f(\theta_1, x)$, $f(\theta_0, x)$ are continuous with respect to $\theta \in \Theta$ and are not equal to zero;

4) The function

$$(\mathbf{E}_{1,\theta}\, \eta_n(\theta))^{-1} = \left( \int \ln \frac{f(\theta_1, x)}{f(\theta_0, x)} f(\theta_1, x) d\mu(x) \right)^{-1}$$

is continuous with respect to $\theta_0 \in \Theta_0$ for any $\theta_1 \in \Theta_1$;

5) For $\tilde{\theta} = (\tilde{\theta}_0, \tilde{\theta}_1)$

$$\infty > \sup_{\tilde{\theta}_1 \in \Theta_1} \sup_{\theta \in \Theta} \mathbf{E}_{1,\tilde{\theta}}\, \eta_n(\theta) \geq \inf_{\tilde{\theta}_1 \in \Theta_1} \inf_{\theta \in \Theta} \mathbf{E}_{1,\tilde{\theta}}\, \eta_n(\theta) > 0;$$

6) For any $\theta \in \Theta$, $\tilde{\theta} \in \Theta$ the uniform (w.r.t. $\theta \in \Theta$) Cramer condition is satisfied

$$\sup_{\theta \in \Theta} \mathbf{E}_{\infty,\tilde{\theta}} \exp\{t\eta(\theta)\} < \infty \quad \text{for} \quad |t| < H_0(\tilde{\theta}_0)$$

where $\inf_{\tilde{\theta}_0 \in \Theta_0} H_0(\tilde{\theta}_0) > 0$.

[1]this assumption is used only to simplify the description

7) For any $\theta_0^* \in \Theta_0$ the function

$$\kappa(t, \theta, \theta_0^*) = \ln \int \left( \frac{f(\theta_1, x)}{f(\theta_0, x)} \right)^t f(\theta_0^*, x) \mu(dx)$$

has only two zeros: 0 and $t^*(\theta, \theta_0^*) > 0$, the function $t^*(\cdot, \theta_0^*)$ is continuous for any $\theta_0^* \in \Theta_0$ and $\min_{\theta \in \Theta} t^*(\theta, \theta_0^*) > 0$.

## III. Main results

### A. Basic inequality

Let $a$ be a certain change-point detection method with the "large parameter" $N$, the decision rule $d_N^a$ and the stopping time $\tau_N^a$ generated by this decision rule. Let $\theta^* = (\theta_0^*(\cdot), \theta_1^*)$ be an arbitrary (defined above) point from $\Theta$.

*Theorem 1:* Let for any point $\theta^* \in \Theta$ and any fixed $m$ there exist the limits

$$\lim_{N \to \infty} \frac{\mathbf{E}_{m,\theta^*} (\tau_N^a - m)^+}{N} < \infty, \quad \lim_{N \to \infty} \frac{|\ln \alpha_N^a(\theta^*)|}{N} > 0$$

Then

$$\mathcal{K}^a(\theta^*, m) = \lim_{N \to \infty} \frac{\mathbf{E}_{m,\theta^*} (\tau_N^a - m)^+}{|\ln \alpha_N^a(\theta^*)|} \geq I_{01}^{-1}(\theta^*) \qquad (5)$$

*Remark.* In [20] it was shown that for all known methods of sequential change-point detection, given the above-made assumptions, there exist all limits mentioned in the formulation of this Theorem. Therefore inequality (5) is valid for all known methods of sequential change-point detection. Besides, all these limits exist for mixing random sequences and therefore the results can be generalized to this case.

Let us return to the question about the sense of the criterion $\mathcal{K}^a(\theta^*, m)$. In [22] it was demonstrated that if the false alarm probability does not exceed $\epsilon$ then with $\epsilon \to 0$ the average time before the false alarm is estimated from below by the value $(2\epsilon)^{-1}(1 + o(1))$ and this estimate is sharp. On the other hand, in [20] it was shown that the false alarm probability tends to zero *exponentially* as $N \to \infty$ for all known methods of change-point detection and, due to above inequality, the average time before the false alarm is asymptotically equivalent to $\exp N$. Besides, in [20] it was shown also that for all known methods of change-point detection $\mathbf{E}_{m,\theta^*} (\tau_N^a - m)^+ \sim N$ as $N \to \infty$. It follows from here that the criterion $\mathcal{K}^a(\theta^*, m)$ represents — in a certain scale — the upper estimate for the limit ratio of the average delay time for an arbitrary hypothesis $\theta_1^* \in \Theta_1$ *after* the change-point to the average time *before* the false alarm under the least favourable alternative for the point $\theta_1^*$ (i.e. for the nearest in sense of Kullback-Leibler distance hypothesis $\theta_0^*(\cdot) \in \Theta_0$) *before* the change-point)

Clearly, this ratio should be kept as small as possible (i.e. the smallest delay time in detection and the largest time before the false alarm). However, from Theorem 1 it follows that for the proposed criterion there exists the a priori lower bound that depends only on the pair $\theta^* = (\theta_0^*(\cdot), \theta_1^*)$ (the situation here is similar to the classic Rao-Cramer inequality with respect to arbitrary estimates of a parameter). It is natural to call a method of sequential change-point detection *adaptive asymptotically optimal* if inequality (5) turns into a strict equality for it. The term "adaptive" here means that the proposed quality criterion attains asymptotically its a priori lower boundary for any (*a priori unknown*) hypothesis after the change-point.

### B. Asymptotically optimal detection method

Suppose $b > 0$ is an arbitrary number. Denote by $\Theta(b) \subset \Theta$, $\Theta(b) = (\Theta_0(b), \Theta_1(b))$ a *finite* $1/b$-network in $\Theta$, $\#\Theta(b) = R(b)$.

Put

$$\mathcal{L}(n,\theta) = \max_{1\le k\le n} \sum_{i=k}^{n} \ln \frac{f(\theta_1, x_i)}{f(\theta_0, x_i)} \tag{6}$$

Define the stopping time

$$T_N(b) = \inf\left\{ n : L(n,b) = \min_{\theta_0\in\Theta_0(b)} \max_{\theta_1\in\Theta_1(b)} \mathcal{L}(n,\theta) > N \right\}, \tag{7}$$

and the corresponding decision rule

$$d_N(n,b) = \begin{cases} 1 & \text{if } L(n,b) > N \\ 0 & \text{if } L(n,b) \le N \end{cases} \tag{8}$$

Evidently, if the sets $\Theta_0, \Theta_1$ consist of one point, then the just described method of change-point detection turns into the classic CUSUM method. Therefore it is natural to call it the *minimax method of cumulative sums* (MMCS). Below we add the index MMCS to all objects relating to this method (in particular, stopping time (7), decision rule (8) and the corresponding false alarm probability). The term MMCS reflects as well the fact that the corresponding method minimizes the maximal (for any given parameter $\theta_1^*$) asymptotic loss in quality (in the sense of above new criterion) of change-point detection among all methods.

Note that in (7) we can interchange operations max and min.

Denote

$$\mathcal{K}_b^{MMCS}(\theta^*, m) = \lim_{N\to\infty} \frac{\mathbf{E}_{m,\theta^*}\left(T_N^{MMCS}(b) - m\right)^+}{|\ln \alpha_N^{MMCS}(\theta^*)|}$$

*Theorem 2:* Let $\theta^*$ be any above-defined point in $\Theta$. Then for any $\epsilon > 0$ there exist $b(\epsilon)$, $\rho(\epsilon)$ ($b(\epsilon) \uparrow \infty$, $\rho(\epsilon) \downarrow 0$ as $\epsilon \to 0$) such that method $d_N^{MMCS}(n,\epsilon)$, corresponding to the stopping time $T_N^{MMCS}(b(\epsilon)) = T_N^{MMCS}(\epsilon)$, satisfies the relationship

$$\mathbf{I}_{01}^{-1}(\theta^*) + \rho(\epsilon) \ge \mathcal{K}_\epsilon^{MMCS}(\theta^*, m) \ge \mathbf{I}_{01}^{-1}(\theta^*) \tag{9}$$

Therefore, it follows from the Theorem that MMCS method corresponding to the stopping time

$$T_N^{MMCS} = \inf\left\{ n : \min_{\theta_0\in\Theta_0} \max_{\theta_1\in\Theta_1} \mathcal{L}(n,\theta) > N \right\}, \tag{10}$$

is *adaptive asymptotically optimal*.

Note that if the hypothesis *before the change-point* is simple (it means that the set $\Theta_0$ consists of one point), then the proposed adaptive asymptotically optimal test turns into the generalized likelihood ratio test (GLRT) (see, for example, [12]). However, in the general case of two composite hypotheses the GLRT test *is not asymptotically optimal*.

## IV. EXPERIMENTAL RESULTS

In this section some results of a small simulation study is given in order to assess the efficiency of the proposed minimax CUSUM method of sequential change-point detection in the general situation of composite hypotheses both before and after a change-point.

The following data were analyzed. The Gaussian sequence was simulated with the d.f. $\mathcal{N}(\theta, 1)$. Under the null hypothesis $\mathbf{H}_0$: $\theta = \theta_0 \in [0,1]$, under the alternative hypothesis $\mathbf{H}_1$: $\theta = \theta_1 \in [2,5]$. The minimax method of cumulative sums (MMCS) was analyzed.

The objective was to estimate characteristics of the proposed method in two main regimes of change-point detection:

– the "false alarm" regime: here some value of $\theta_0$ from the interval $[0,1]$ was fixed and the average time before the "false alarm": $\mathbf{E}T$ was estimated;

– the change-point detection regime: here the average delay time in change-point detection for different values of $\theta_1$ from the interval $[2,5]$ was estimated.

Each value in cells of these tables is obtained as the average of 5000 Monte Carlo trials. Here $\mathbf{E}T$ is the average time between "false alarms" (which is asymptotically equal to $1/p$, where $p$ is the probability of the error decision); $\ln \mathbf{E}T$ is the logarithm of $\mathbf{E}T$; $\mathbf{E}\tau$ is the average delay time in change-point detection; $\mathbf{E}\tau/\ln \mathbf{E}T$ is the ratio criterion of efficiency.

Table 1.

| | $N$ | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| $\theta_0$ | $ET$ | 138.7 | 577.2 | 2203.2 | 8091.1 |
| 0.5 | $\ln ET$ | 4.93 | 6.36 | 7.70 | 9.00 |
| $\theta_1$ | $E\tau$ | 3.89 | 5.53 | 7.34 | 9.30 |
| 2 | $E\tau/\ln ET$ | 0.79 | 0.86 | 0.95 | 1.03 |
| $\theta_1$ | $E\tau$ | 1.68 | 2.15 | 2.49 | 3.04 |
| 3 | $E\tau/\ln ET$ | 0.34 | 0.33 | 0.32 | 0.33 |
| $\theta_1$ | $E\tau$ | 1.15 | 1.33 | 1.49 | 1.63 |
| 4 | $E\tau/\ln ET$ | 0.23 | 0.20 | 0.19 | 0.18 |

Results presented in Table 1 allow us to conclude that the proposed minimax method of cumulative sums is practically efficient in the general situation of sequential change-point detection under composite hypotheses both before and after a change-point. The ratio $E\tau/\ln ET$ of the average delay time in change-point detection to the logarithm of the average time before a "false alarm" tends to a certain limit as the threshold $N$ increases.

## V. CONCLUSION

The problem of sequential change-point detection under composite hypotheses is considered. In the general situation it is assumed that the density function of observations not only *after* a change-point but also *before* it is unknown and belongs to a certain family of distributions. This problem is very important for applications, in particular, in the case when it is necessary to detect faults in dynamical systems. The problem of fault detection for dynamical systems can be reduced to the change-point problem under composite hypotheses which are characterized by a finite-dimensional vector with such components as uncertain parameters of a system, unknown inputs, unknown initial conditions, and unknown states of this system. Therefore, the true hypothesis before a change-point is unknown as well as the true hypothesis after the change-point. The advantage of the proposed criterion consists in the fact that it is possible to prove the asymptotic lower bound for this criterion and therefore to find the asymptotically optimal method of change-point detection.

It should be emphasized that the case of composite hypotheses before and after a change-point was not considered in the literature.

A new asymptotically optimal method of change-point detection is proposed for which the a priori information-theoretical lower bound for the main functional of quality of change-point detection is attained. In a particular case when the hypothesis *before the change-point* is simple the proposed method turns into the generalized likelihood ratio test but it is not true in the general case of two composite hypotheses (before and after the change-point).

## REFERENCES


[1] Shewhart, W.A. (1931). *Economic control of quality of manufactured produst*. Van Nostrand Reinhold, New York.

[2] Girshik, M.A. and H. Rubin (1952). A Bayes approach to a quality control model. *Ann.Math.Stat.*, **23**, no.1, 114-125.

[3] Page, E.S. (1954). Continuous inspection schemes. *Biometrika*, **41**, 100-114.

[4] Shiryaev, A.N. (1976).*Optimal Stopping Rules*. Nauka, Moscow.

[5] Lorden, G. (1971). Procedures for reacting to a change in distribution. *Ann.Math.Stat.*, **42**, 1897-1908.

[6] Pollak, M. (1985). Optimal detection of a change in distribution. *Ann.Stat.*, **13**, 206-227.

[7] Moustakides, G.V. (1986). Optimal stopping for detecting changes in distribution. *Ann.Stat.*, **14**, 379-1387.

[8] Shiryaev, A.N. (1996). Minimax optimality of the CUSUM method for the continuous time. *Adv. Math. Sci.*, **310**, no. 4, 173-174.

[9] Lai, T.L. (1995). Sequential change-point detection in quality control and dynamical systems (with discussion). *J.Roy.Statist.Ser., B*, **57**, 613-658.

[10] Lai, T.L. (1998). Information bounds and quick detection of parameter changes in stochastic systems. *IEEE Trans.Inform. Theory*, **44**, 2917-2929.

[11] Lai, T.L. and J. Z. Shan. (1999). Efficient recursive algorithms for detection of abrupt changes in signals and systems. *IEEE Trans. Automat.Contr.*, **44**, 952-966.

[12] Lai, T.L. (2000). Sequential multiple hypothesis testing and efficient fault detection-isolation in stochastic systems. *IEEE Trans.Inform.Theory*, **46**, 595-608.

[13] Basseville, M and I.V. Nikiforov. (1993). *Detection of Abrupt Changes: Theory and Applications*. Prentice Hall, Englewood Cliffs, NJ.

[14] Nikiforov, I.V. (1996). A generalized change detection problem, *IEEE Trans.Inform. Theory*, **41**, 171-181.

[15] Tartakovsky, A.G. and V.V. Veeravalli (2004). General Asymptotic Bayesian Theory of Quickest Change detection. *Theory of Probability and Applications*, **49**, no. 3, 538-582.

[16] Darkhovsky, B.S. and M.Staroswiecki. (2003). The fault Detection problem for Dynamic Systems: General Results. In: *Proc. of IFAC Safeprocess'03*, 193-198.

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

[18] Lai, T.L. (2001). Sequential analysis: some classical problems and new challenges. *Statistica Sinica*, **11**, no. 2, 303-408.

[19] Brodsky, B.E. and B.S. Darkhovsky. (1990). Comparative study of some nonparametric methods of sequential change-point detection. *Theory of Probability and Applications*, **35**, no. 4, 655-668.

[20] Brodsky, B.E. and B.S. Darkhovsky (2000). *Non-Parametric Statistical Diagnosis. Problems and Methods*. Kluwer, Dordrecht.

[21] Brodsky, B.E. and B.S. Darkhovsky (2005). Asymptotically Optimal Methods for Sequential Testing of Composite Hypotheses. *Journal of Statistical Planning and Inference*, vol. 133, no. 1, 2005, pp 123-138.

[22] Darkhovsky, B.S. and B.E. Brodsky (1987). Nonparametric method of the quickiest detection of a change in mean of a random sequence. *Theory of Probability and Applications*, **32**, no. 4, 703-711.