



**HAL**  
open science

## On-line detection of a part of a sequence with unspecified distribution

Wojciech Sarnowski, Krzysztof Szajowski

► **To cite this version:**

Wojciech Sarnowski, Krzysztof Szajowski. On-line detection of a part of a sequence with unspecified distribution. *Statistics and Probability Letters*, 2009, 78 (15), pp.2511. <10.1016/j.spl.2008.02.040>. <hal-00564090>

**HAL Id: hal-00564090**

**<https://hal.science/hal-00564090v1>**

Submitted on 8 Feb 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



HAL Authorization

## Accepted Manuscript

On-line detection of a part of a sequence with unspecified distribution

Wojciech Sarnowski, Krzysztof Szajowski

PII: S0167-7152(08)00157-0  
DOI: 10.1016/j.spl.2008.02.040  
Reference: STAPRO 5007

To appear in: *Statistics and Probability Letters*

Received date: 9 January 2006  
Revised date: 15 March 2007  
Accepted date: 14 February 2008

Please cite this article as: Sarnowski, W., Szajowski, K., On-line detection of a part of a sequence with unspecified distribution. *Statistics and Probability Letters* (2008), doi:10.1016/j.spl.2008.02.040

This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.



# On-line detection of a part of a sequence with unspecified distribution<sup>\*</sup>

Wojciech Sarnowski<sup>a,\*</sup> Krzysztof Szajowski<sup>a</sup>

<sup>a</sup>Wrocław University of Technology, Institute of Mathematics and Computer Science, Wybrzeże Wyspiańskiego 27, 50-370 Wrocław, Poland

---

## Abstract

We register a Markov process. At two random moments  $\theta_1, \theta_2$ , where  $\theta_1 < \theta_2$ , the distribution of observed sequence changes. It is known before  $\theta_1$  and after  $\theta_2$ . Between these instants is unknown, chosen randomly from a set of distributions. The optimal stopping rule which stops observation of the sequence between disorders  $\theta_1$  and  $\theta_2$  is identified.

*Key words:* Disorder problem, sequential detection, optimal stopping, Markov process, change point.

*1991 MSC:* Primar 60G40, 60K99; Secondary 90D60

---

## 1 Introduction

In this paper we present a generalization of the double disorder problem considered by Yoshida [10] and Szajowski [7]. The considerations are inspired by the problem regarding how can we protect ourselves against a second fault in a technological system after the occurrence of an initial fault. The proposed procedure assumes that the distribution of observations between disorders is unknown in advance and it is chosen randomly by "nature" (see Bojdecki & Hosza [2]). In order to find the optimal stopping rule we use the approach of

---

<sup>\*</sup> 21 March 2008

<sup>\*</sup> Corresponding author

*Email addresses:* Wojciech.Sarnowski@pwr.wroc.pl (Wojciech Sarnowski), Krzysztof.Szajowski@pwr.wroc.pl (Krzysztof Szajowski).

*URLs:* <http://www.im.pwr.wroc.pl/~sarnowski> (Wojciech Sarnowski), <http://neyman.im.pwr.wroc.pl/~szajow> (Krzysztof Szajowski).

maximizing the probability of stopping between these two faults. The formulation of the problem can be found in section 2. The main result is presented in section 3.

## 2 Formulation of the problem

Let  $(X_n, n \in \mathbf{N})$  be an observable sequence of random variables defined on the space  $(\Omega, \mathcal{F}, P)$  with values in  $(\mathbf{E}, \mathcal{B})$ , where  $\mathbf{E}$  is a subset of  $\mathbf{R}$ . On  $(\mathbf{E}, \mathcal{B})$  there are  $\sigma$ -additive measures  $\{\mu_x\}_{x \in \mathbf{E}}$ . Space  $(\Omega, \mathcal{F}, P)$  supports variables  $\theta_1, \theta_2$ . They are  $\mathcal{F}$ -measurable variables with values in  $\mathbf{N}$ . We assume the following distributions:

$$P(\theta_1 = j) = p_1^{j-1} q_1, P(\theta_2 = k \mid \theta_1 = j) = p_2^{k-j-1} q_2; k > j, j = 1, 2, \dots (1)$$

Moreover we introduce  $\varepsilon$ , which is  $\mathcal{F}$ -measurable, independent of  $\theta_1, \theta_2$ .  $\varepsilon$  has distribution:  $P(\varepsilon = i) = e_i; i = 1, 2, \dots, d; d < \infty; \sum_{i=1}^d e_i = 1$ . Additionally we consider  $d+2$  Markov processes on  $(\Omega, \mathcal{F}, P)$ :  $(X_n^1, \mathcal{G}_n^1, \mu_x^1), (X_n^{2,i}, \mathcal{G}_n^{2,i}, \mu_x^{2,i}), i = 1, \dots, d, (X_n^3, \mathcal{G}_n^3, \mu_x^3)$  defined in the following way:

$$X_n = X_n^1 \cdot \mathbb{I}_{\{\theta_1 > n\}} + X_n^{2,i} \cdot \mathbb{I}_{\{\theta_1 \leq n < \theta_2, \varepsilon = i\}} + X_n^3 \cdot \mathbb{I}_{\{\theta_2 \leq n\}}. (2)$$

$\sigma$ -fields  $\mathcal{G}_n^1, \mathcal{G}_n^{2,i}, \mathcal{G}_n^3$  are the smallest  $\sigma$ -fields for which  $(X^1), (X^{2,i}), (X^3)$  are adapted (respectively).

We make inference only based on observable sequence  $(X_n, n \in \mathbf{N})^1$ . Thus for further consideration we define filtration  $\{\mathcal{F}_n\}_{n \in \mathbf{N}}$  where  $\mathcal{F}_n = \sigma(X_0, X_1, \dots, X_n)$ . Variables  $\theta_1, \theta_2$  and  $\varepsilon$  are not stopping time with respect to  $\mathcal{F}_n$  and  $\sigma$ -fields  $\mathcal{G}_n^i$ . Moreover we assume that  $\theta_1, \theta_2$  are independent of  $(X_n^1, n \in \mathbf{N}), (X_n^{2,i}, n \in \mathbf{N}), i = 1, \dots, d$  and  $(X_n^3, n \in \mathbf{N})$ . Measures  $\mu$  satisfy the relations:  $\mu_x^1(dy) = 1 \cdot \mu_x(dy), \mu_x^{2,i}(dy) = f_x^i(y) \mu_x(dy), i = 1, \dots, d, \mu_x^3(dy) = g_x(y) \mu_x(dy)$ . Where  $f_x^i(\cdot) \neq 1, g_x(\cdot) \neq 1$ . Furthermore the functions  $f_x^1(\cdot), \dots, f_x^d(\cdot), g_x(\cdot)$  are different and  $g_x(y)/f_x^i(y) < \infty$  for all  $i = 1, \dots, d$  and all  $x, y \in \mathbf{E}$ . We assume that the measures  $\mu_x^1, \mu_x^{2,1}, \dots, \mu_x^{2,d}, \mu_x^3, x \in \mathbf{E}$  are known in advance and we have that  $P(X_1^1 \in A \mid X_0 = x) = \mu_x^1(A)$  for every  $A \in \mathcal{B}$ .

The model presented has the following heuristic justification: two disorders take place in the observed sequence  $(X_n)$ . They affect distributions by changing their parameters. Disorders occur at two random moments of time  $\theta_1$  and

<sup>1</sup> It should be emphasized that the sequence  $(X_n, n \in \mathbf{N})$  is not markovian one under admitted assumption as it has been written in [7], [9] and [4]. However the sequence has Markov property given  $\theta_1$  and  $\theta_2$  (see [8] and [5]).

$\theta_2$ . They split the sequence of observations into three segments. The first segment is described by  $(X_n^1)$ , the second one - for  $\theta_1 \leq n < \theta_2$  - by  $(X_n^2)$ . The third is given by  $(X_n^3)$  and is observed when  $n \geq \theta_2$ . The second epoch of observation is the important one. We do not know the distribution of observations in advance. After the first disorder takes the place there is a "switch" from the initial distribution to one of  $d$  possible distributions (given by densities  $f_x^i$ ). The choice of this distribution is random. We model this using an unobservable variable  $\varepsilon$  that indicates which of them is drawn (if  $\varepsilon = j$ , then the  $\mu_x^j$  distribution is chosen with density  $f_x^j$  with respect to measure  $\mu_x$ ). Next, at the random time  $\theta_2$  the distribution of observations becomes  $\mu_x^3$ . We assume that the variables  $\varepsilon, \theta_1, \theta_2$  are unobservable.

Our aim is to stop the observed sequence between the two disorders. This can be interpreted as a strategy for protecting against a second failure when the first has already happened. In the mathematical approach - putting  $\mathcal{T}$  as the set of all stopping times with respect to the filtration  $\{\mathcal{F}_n\}_{n \in \mathbf{N}}$  - we want to control the probability  $V(\tau) = P(\tau < \infty, \theta_1 \leq \tau < \theta_2)$  by choosing the stopping time  $\tau^* \in \mathcal{T}$  for which

$$V(\tau^*) = \sup_{\tau \in \mathcal{T}} V(\tau) \quad (3)$$

### 3 Solution of the problem

For  $X_0 = x$  let us define:  $Z_n = P_x(\theta_1 \leq n < \theta_2 \mid \mathcal{F}_n)$  for  $n = 0, 1, 2, \dots$ ,  $Y_n = \text{esssup}_{\{\tau \in \mathcal{T}, \tau \geq n\}} P_x(\theta_1 \leq \tau < \theta_2 \mid \mathcal{F}_n)$  for  $n = 0, 1, 2, \dots$  and

$$\tau_0 = \inf\{n : Z_n = Y_n\} \quad (4)$$

Notice that, if  $Z_\infty = 0$ , then  $Z_\tau = P_x(\theta_1 \leq \tau < \theta_2 \mid \mathcal{F}_\tau)$  for  $\tau \in \mathcal{T}$ . Since  $\mathcal{F}_n \subseteq \mathcal{F}_\tau$  (when  $n \leq \tau$ ) we have

$$\begin{aligned} Y_n &= \text{ess sup}_{\tau \geq n} P_x(\theta_1 \leq \tau < \theta_2 \mid \mathcal{F}_n) = \text{ess sup}_{\tau \geq n} E_x(\mathbb{I}_{\{\theta_1 \leq \tau < \theta_2\}} \mid \mathcal{F}_n) \\ &= \text{ess sup}_{\tau \geq n} E_x(Z_\tau \mid \mathcal{F}_n) \end{aligned}$$

**Lemma 1** *The stopping time  $\tau_0$  defined by formula (4) is the solution of problem (3).*

**Proof.** From the theorems presented in [1] it is enough to show that  $\lim_{n \rightarrow \infty} Z_n = 0$ . For all natural numbers  $n, k$ , where  $n \geq k$  for each  $x \in \mathbf{E}$  we have:

$$Z_n = E_x(\mathbb{I}_{\{\theta_1 \leq n < \theta_2\}} \mid \mathcal{F}_n) \leq E_x(\sup_{j \geq k} \mathbb{I}_{\{\theta_1 \leq j < \theta_2\}} \mid \mathcal{F}_n)$$

From Levy's theorem  $\limsup_{n \rightarrow \infty} Z_n \leq E_x(\sup_{j \geq k} \mathbb{I}_{\{\theta_1 \leq j < \theta_2\}} \mid \mathcal{F}_\infty)$  where  $\mathcal{F}_\infty = \sigma(\bigcup_{n=1}^\infty \mathcal{F}_n)$ . It is true that:  $\limsup_{j \geq k, k \rightarrow \infty} \mathbb{I}_{\{\theta_1 \leq j < \theta_2\}} = 0$  a.s. and by the dominated convergence theorem we get

$$\lim_{k \rightarrow \infty} E_x(\sup_{j \geq k} \mathbb{I}_{\{\theta_1 \leq j < \theta_2\}} \mid \mathcal{F}_\infty) = 0 \text{ a.s.}$$

what ends the proof of the Lemma.

For further considerations let us introduce the following notation: let  $\langle \underline{u}, \underline{v} \rangle = \sum_{i=1}^d u_i v_i$ . Thus  $\langle \cdot, \cdot \rangle$  stands for inner product. Let us now define some necessary processes and functions:  $\Pi_n^1 = P_x(\theta_1 \leq n \mid \mathcal{F}_n)$ ,  $\Pi_n^2 = P_x(\theta_2 \leq n \mid \mathcal{F}_n)$ , for  $n = 0, 1, 2, \dots$ ,

$$H(y, z, \alpha, \beta) = (1 - \alpha)p_1 + [(1 - \alpha)q_1 + (\alpha - \beta)p_2] \langle \underline{e}, \underline{f}_y(z) \rangle + [(\alpha - \beta)q_2 + \beta] g_y(z) \quad (5)$$

$$\Pi^1(y, z, \alpha, \beta) = 1 - (1 - \alpha)p_1 H^{-1}(y, z, \alpha, \beta) \quad (6)$$

$$\Pi^2(y, z, \alpha, \beta) = [(\alpha - \beta)q_2 + \beta] g_y(z) H^{-1}(y, z, \alpha, \beta) \quad (7)$$

Using these definitions we formulate and prove a technical Lemma which will be useful when proving the main theorem.

**Lemma 2** For each  $x \in \mathbf{E}$  and each Borel function  $u : \mathbf{R} \rightarrow \mathbf{R}$  the equations given below are true

$$\Pi_{n+1}^1 = \Pi^1(X_n, X_{n+1}, \Pi_n^1, \Pi_n^2) \quad (8)$$

$$\Pi_{n+1}^2 = \Pi^2(X_n, X_{n+1}, \Pi_n^1, \Pi_n^2), \quad (9)$$

with boundary condition:  $\Pi_0^1 = \Pi_0^2 = 0$

$$E_x(u(X_{n+1})(1 - \Pi_{n+1}^1) \mid \mathcal{F}_n) = (1 - \Pi_n^1)p_1 \int_{\mathbf{E}} u(y) \mu_{X_n}(dy), \quad (10)$$

$$E_x(u(X_{n+1})(\Pi_{n+1}^1 - \Pi_{n+1}^2) \mid \mathcal{F}_n) = [(1 - \Pi_n^1)q_1 + (\Pi_n^1 - \Pi_n^2)p_2] \int_{\mathbf{E}} u(y) \langle \underline{e}, \underline{f}_{X_n}(y) \rangle \mu_{X_n}(dy), \quad (11)$$

$$E_x(u(X_{n+1})\Pi_{n+1}^2 \mid \mathcal{F}_n) = [q_2(\Pi_n^1 - \Pi_n^2) + \Pi_n^2] \int_{\mathbf{E}} u(y) g_{X_n}(y) \mu_{X_n}(dy) \quad (12)$$

**Proof.** It is obvious that  $\Pi_0^1 = \Pi_0^2 = 0$ . Notice that since  $\theta_1 < \theta_2$  we obtain the inclusion  $\{\omega : \theta_1 \leq n\} \supseteq \{\omega : \theta_2 \leq n\}$  and:

$$Z_n = P_x(\theta_1 \leq n < \theta_2 \mid \mathcal{F}_n) = \Pi_n^1 - \Pi_n^2 \quad (13)$$

Next suppose that  $A_i \in \mathcal{F}_i$ ,  $i \leq n + 1$  and let us consider the probability

$$P_x(\theta_1 > n + 1 \mid X_i \in A_i, i \leq n + 1) = \frac{P_x(\theta_1 > n + 1, X_{n+1} \in A_{n+1} \mid X_i \in A_i, i \leq n)}{P_x(X_{n+1} \in A_{n+1} \mid X_i \in A_i, i \leq n)}$$

This follows from Bayes' formula. Let us transform the probability appearing in the numerator:

$$\begin{aligned} P_x(\theta_1 > n + 1, X_{n+1} \in A_{n+1} \mid X_i \in A_i, i \leq n) \\ = P_x(\theta_1 > n \mid X_i \in A_i, i \leq n) \cdot p_1 \cdot \mu_{X_n}^1(A_{n+1}) \end{aligned}$$

Now we split the probability in the denominator into three parts

$$\begin{aligned} P_x(X_{n+1} \in A_{n+1} \mid X_i \in A_i, i \leq n) &= P_x(\theta_1 > n, X_{n+1} \in A_{n+1} \mid X_i \in A_i, i \leq n) \\ &+ P_x(\theta_1 \leq n < \theta_2, X_{n+1} \in A_{n+1} \mid X_i \in A_i, i \leq n) \\ &+ P_x(\theta_2 \leq n, X_{n+1} \in A_{n+1} \mid X_i \in A_i, i \leq n) \end{aligned}$$

The first part is:

$$\begin{aligned} P_x(\theta_1 > n, X_{n+1} \in A_{n+1} \mid X_i \in A_i, i \leq n) \\ = P_x(\theta_1 > n, \theta_1 = n + 1, X_{n+1} \in A_{n+1} \mid X_i \in A_i, i \leq n) \\ + P_x(\theta_1 > n, \theta_1 \neq n + 1, X_{n+1} \in A_{n+1} \mid X_i \in A_i, i \leq n) \\ = P_x(\theta_1 > n \mid X_i \in A_i, i \leq n) \left[ q_1 \sum_{i=1}^d \mu_{X_n}^{2,i}(A_{n+1}) e_i + \mu_{X_n}^1(A_{n+1}) p_1 \right] \end{aligned}$$

The second one:

$$\begin{aligned} P_x(\theta_1 \leq n < \theta_2, X_{n+1} \in A_{n+1} \mid X_i \in A_i, i \leq n) \\ = P_x(\theta_1 \leq n < \theta_2, \theta_2 = n + 1, X_{n+1} \in A_{n+1} \mid X_i \in A_i, i \leq n) \\ + P_x(\theta_1 \leq n < \theta_2, \theta_2 \neq n + 1, X_{n+1} \in A_{n+1} \mid X_i \in A_i, i \leq n) \\ = (P_x(\theta_1 \leq n \mid X_i \in A_i, i \leq n) - P_x(\theta_2 \leq n \mid X_i \in A_i, i \leq n)) \\ \times \left[ q_2 \mu_{X_n}^3(A_{n+1}) + p_2 \sum_{i=1}^d \mu_{X_n}^{2,i}(A_{n+1}) e_i \right] \end{aligned}$$

And the third:

$$P_x(\theta_2 \leq n, X_{n+1} \in A_{n+1} \mid X_i \in A_i, i \leq n) = P_x(\theta_2 \leq n \mid X_i \in A_i, i \leq n) \mu_{X_n}^3(A_{n+1})$$

Thus, taking into account (5) we obtain for  $\Pi_{n+1}^1 = 1 - P_x(\theta_1 > n + 1 \mid \mathcal{F}_{n+1})$

$$\Pi_{n+1}^1 = 1 - \left[ (1 - \Pi_n^1) p_1 \right] H^{-1}(X_n, X_{n+1}, \Pi_n^1, \Pi_n^2)$$

Using (6), it can be seen that (8) is satisfied. Applying similar reasoning and transformations to the process  $\Pi_{n+1}^2$  we get:

$$\begin{aligned} \Pi_{n+1}^2 &= P_x(\theta_2 \leq n + 1 \mid \mathcal{F}_{n+1}) = \frac{P_x(\theta_2 \leq n + 1, X_{n+1} \in A \mid \mathcal{F}_n)}{P_x(X_{n+1} \in A \mid \mathcal{F}_n)} \\ &= \left[ (\Pi_n^1 - \Pi_n^2) q_2 + \Pi_n^2 \right] g_{X_n}(X_{n+1}) H^{-1}(X_n, X_{n+1}, \Pi_n^1, \Pi_n^2) \end{aligned}$$

which leads to formula (9). In the case of the conditional expectations we show that equation (12) holds. To do this we need to first define  $\sigma$ -field  $\tilde{\mathcal{F}}_n = \sigma(\varepsilon, \theta_1, \theta_2, X_0, \dots, X_n)$ . Notice that  $\mathcal{F}_n \subset \tilde{\mathcal{F}}_n$ . We have:

$$\begin{aligned}
E_x(u(X_{n+1})\Pi_{n+1}^2 \mid \mathcal{F}_n) &= E_x(E_x(u(X_{n+1})\mathbf{1}_{\{\theta_2 \leq n+1\}} \mid \tilde{\mathcal{F}}_n) \mid \mathcal{F}_n) \\
&= E_x\left(\int_{\mathbf{E}} u(y)P_x(dy \mid \tilde{\mathcal{F}}_n, \theta_2 \leq n+1)P_x(\theta_2 \leq n+1 \mid \tilde{\mathcal{F}}_n) \mid \mathcal{F}_n\right) \\
&= \int_{\mathbf{E}} u(y)g_{X_n}(y)\mu_{X_n}(dy)P_x(\theta_2 \leq n+1 \mid \mathcal{F}_n) \\
&= \int_{\mathbf{E}} u(y)g_{X_n}(y)\mu_{X_n}(dy)(P_x(\theta_2 = n+1, \theta_1 \leq n < \theta_2 \mid \mathcal{F}_n) + P_x(\theta_2 \leq n \mid \mathcal{F}_n)) \\
&= \left((\Pi_n^1 - \Pi_n^2)q_2 + \Pi_n^2\right) \int_{\mathbf{E}} u(y)g_{X_n}(y)\mu_{X_n}(dy)
\end{aligned}$$

We used the properties of conditional expectation here. Similar transformations give us equations (10) and (11). This proves Lemma 2.

Notice that Lemma 2 enables us to calculate  $E_x(u(X_{n+1}) \mid \mathcal{F}_n)$  in easy way. Combining (10)-(12) we have

$$E_x(u(X_{n+1}) \mid \mathcal{F}_n) = \int_{\mathbf{E}} u(y)H(X_n, y, \Pi_n^1, \Pi_n^2)\mu_{X_n}(dy) \quad (14)$$

**Lemma 3** *System  $X^x = \{X_n^x\}$ , where  $X_n^x = (X_{n-1}, X_n, \Pi_n^1, \Pi_n^2)$  forms a family of random Markov functions.*

**Proof.** Define a function:

$$\varphi(x_1, x_2, \alpha_1, \alpha_2; z) = (x_2, z, \Pi^1(x_2, z, \alpha_1, \alpha_2), \Pi^1(x_2, z, \alpha_1, \alpha_2)) \quad (15)$$

Observe that

$$X_n^x = \varphi(X_{n-2}, X_{n-1}, \Pi_{n-1}^1, \Pi_{n-1}^2; X_n) = \varphi(X_{n-1}^x; X_n)$$

Hence  $X_n^x$  can be interpreted as function of previous state  $X_{n-1}^x$  and random variable  $X_n$ . Moreover, applying (14), we get that conditional distribution of  $X_n$  given  $\sigma$ -field  $\mathcal{F}_{n-1}$  depends only on  $X_{n-1}^x$ . According to [6] (pp. 102-103) system  $X^x$  is a family of random Markov functions. This fact implies that we can reduce initial problem (3) to the problem of optimal stopping four-dimensional process  $(X_{n-1}, X_n, \Pi_n^1, \Pi_n^2)$  with reward

$$h(x_1, x_2, \alpha_1, \alpha_2) = \alpha_1 - \alpha_2 \quad (16)$$

The reward function results from equation (13). Thanks to Lemma 3 we construct the solution using standard tools of optimal stopping theory (c.f [6]).

For any Borel function  $v : \mathbf{E}^2 \times [0, 1]^{d+1} \longrightarrow [0, 1]$  let us define two operators:

$$\begin{aligned}
T_x v(y, z, \alpha, \beta) &= E_x(v(X_n, X_{n+1}, \Pi_{n+1}^1, \Pi_{n+1}^2) \mid X_{n-1} = y, X_n = z, \Pi_n^1 = \alpha, \Pi_n^2 = \beta) \\
Q_x v(y, z, \alpha, \beta) &= \max\{v(y, z, \alpha, \beta), T_x v(y, z, \alpha, \beta)\}
\end{aligned}$$

From well known theorems of optimal stopping theory ([6]), we infer that the solution of the problem posed is the Markov time  $\tau_0$ :

$$\tau_0 = \inf\{h(X_n, X_{n+1}, \Pi_{n+1}^1, \Pi_{n+1}^2) \geq h^*(X_n, X_{n+1}, \Pi_{n+1}^1, \Pi_{n+1}^2)\}$$

where:

$$h^*(y, z, \alpha, \beta) = \lim_{k \rightarrow \infty} Q_x^k h(y, z, \alpha, \beta)$$

Of course

$$Q_x^k v(y, z, \alpha, \beta) = \max\{Q_x^{k-1} v, T_x Q_x^{k-1} v\} = \max\{v, T_x Q_x^{k-1} v\}$$

To obtain a clearer formula for  $\tau_0$ , we formulate the basic result of this paper:

**Theorem 1 (a)** *The solution of problem (3) is given by:*

$$\tau^* = \inf\{n : (X_n, X_{n+1}, \Pi_{n+1}^1, \Pi_{n+1}^2) \in B^*\} \quad (17)$$

Set  $B^*$  is of the form:

$$\begin{aligned} B^* = & \{(y, z, \alpha, \beta) : (\alpha - \beta) \geq (1 - \alpha) \\ & \times \left[ p_1 \int_{\mathbf{E}} R^*(y, u, \Pi^1(y, u, \alpha, \beta), \Pi^2(y, u, \alpha, \beta)) \mu_y(du) \right. \\ & \left. + q_1 \int_{\mathbf{E}} S^*(y, u, \Pi^1(y, u, \alpha, \beta), \Pi^2(y, u, \alpha, \beta)) \langle \underline{e}, \underline{f}_y(u) \rangle \mu_y(du) \right] \\ & \left. + (\alpha - \beta) p_2 \int_{\mathbf{E}} S^*(y, u, \Pi^1(y, u, \alpha, \beta), \Pi^2(y, u, \alpha, \beta)) \langle \underline{e}, \underline{f}_y(u) \rangle \mu_y(du) \right\} \end{aligned}$$

Where:

$$R^*(y, z, \alpha, \beta) = \lim_{k \rightarrow \infty} R^k(y, z, \alpha, \beta), \quad S^*(y, z, \alpha, \beta) = \lim_{k \rightarrow \infty} S^k(y, z, \alpha, \beta)$$

Functions  $R^k$  and  $S^k$  are defined recursively:

$$R^1(y, z, \alpha, \beta) = 0, \quad S^1(y, z, \alpha, \beta) = 1$$

$$\begin{aligned} R^{k+1}(y, z, \alpha, \beta) = & (1 - \mathbb{I}_{\mathcal{R}_k}(y, z, \alpha, \beta)) \quad (18) \\ & \times \left( p_1 \int_{\mathbf{E}} R^k(y, u, \Pi^1(y, u, \alpha, \beta), \Pi^2(y, u, \alpha, \beta)) \mu_y(du) \right. \\ & \left. + q_1 \int_{\mathbf{E}} S^k(y, u, \Pi^1(y, u, \alpha, \beta), \Pi^2(y, u, \alpha, \beta)) \langle \underline{e}, \underline{f}_y(u) \rangle \mu_y(du) \right), \end{aligned}$$

$$\begin{aligned} S^{k+1}(y, z, \alpha, \beta) = & \mathbb{I}_{\mathcal{R}_k}(y, z, \alpha, \beta) + (1 - \mathbb{I}_{\mathcal{R}_k}(y, z, \alpha, \beta)) \quad (19) \\ & \times p_2 \int_{\mathbf{E}} S^k(y, u, \Pi^1(y, u, \alpha, \beta), \Pi^2(y, u, \alpha, \beta)) \langle \underline{e}, \underline{f}_y(u) \rangle \mu_y(du) \end{aligned}$$

Where the set  $\mathcal{R}_k$  is:

$$\begin{aligned}
 \mathcal{R}_k &= \{(y, z, \alpha, \beta) : h(y, z, \alpha, \beta) \geq T_x Q_x^{k-1} h(y, z, \alpha, \beta)\} \\
 &= \{(y, z, \alpha, \beta) : (\alpha - \beta) \geq (1 - \alpha) \\
 &\quad \times \left[ p_1 \int_{\mathbf{E}} R^k(y, u, \Pi^1(y, u, \alpha, \beta), \Pi^2(y, u, \alpha, \beta)) \mu_y(du) \right. \\
 &\quad \left. + q_1 \int_{\mathbf{E}} S^k(y, u, \Pi^1(y, u, \alpha, \beta), \Pi^2(y, u, \alpha, \beta)) \langle \underline{e}, \underline{f}_y(u) \rangle \mu_y(du) \right] \\
 &\quad \left. + (\alpha - \beta) p_2 \int_{\mathbf{E}} S^k(y, u, \Pi^1(y, u, \alpha, \beta), \Pi^2(y, u, \alpha, \beta)) \langle \underline{e}, \underline{f}_y(u) \rangle \mu_y(du) \right\}
 \end{aligned}$$

(b) *The value problem. The optimal value for (3) is given by the formula*

$$\begin{aligned}
 V(\tau^*) &= p_1 \int_{\mathbf{E}} R^*(x, u, \varphi_x(u), 0) \mu_x(du) \\
 &\quad + q_1 \int_{\mathbf{E}} S^*(x, u, \varphi_x(u), 0) \langle \underline{e}, \underline{f}_x(u) \rangle \mu_x(du)
 \end{aligned}$$

$$\text{where: } \varphi_x(u) = 1 - \frac{p_1}{p_1 + q_1 \langle \underline{e}, \underline{f}_x(u) \rangle}$$

**Proof.** Part (a). Part (a) results from Lemma 2 - the problem reduces to the problem of optimal stopping of the Markov process  $(X_{n-1}, X_n, \Pi_n^1, \Pi_n^2)$  with payoff function  $h(y, z, \alpha, \beta) = \alpha - \beta$ . Given (11) with the function  $u$  equal to unity we get:

$$\begin{aligned}
 T_x h(y, z, \alpha, \beta) &= E \left( \Pi_{n+1}^1 - \Pi_{n+1}^2 \mid \mathcal{F}_n \right) \Big|_{X_{n-1}=y, X_n=z, \Pi_n^1=\alpha, \Pi_n^2=\beta} \\
 &= (\Pi_n^1 - \Pi_n^2) p_2 \int_{\mathbf{E}} \langle \underline{e}, \underline{f}_{X_n}(u) \rangle \mu_{X_n}(du) + (1 - \Pi_n^1) q_1 \Big|_{X_{n-1}=y, X_n=z, \Pi_n^1=\alpha, \Pi_n^2=\beta} \\
 &= (1 - \alpha) q_1 + (\alpha - \beta) p_2
 \end{aligned}$$

From the definition of  $R^1$  and  $S^1$  it is clear that

$$h(y, z, \alpha, \beta) = \alpha - \beta = (1 - \alpha) R^1(y, z, \alpha, \beta) + (\alpha - \beta) S^1(y, z, \alpha, \beta)$$

Also  $\mathcal{R}_1 = \{(y, z, \alpha, \beta) : h(y, z, \alpha, \beta) \geq T_x h(y, z, \alpha, \beta)\}$ . From the definition of  $Q_x$  and the facts above we obtain:

$$Q_x h(y, z, \alpha, \beta) = (1 - \alpha) R^2(y, z, \alpha, \beta) + (\alpha - \beta) S^2(y, z, \alpha, \beta)$$

where:  $R^2(y, z, \alpha, \beta) = q_1(1 - \mathbb{I}_{\mathcal{R}_1}(y, z, \alpha, \beta))$  and  $S^2(y, z, \alpha, \beta) = p_2 + (1 - p_2)\mathbb{I}_{\mathcal{R}_1}(y, z, \alpha, \beta)$ . Suppose the following induction hypothesis holds:

$$Q_x^{k-1} h(y, z, \alpha, \beta) = (1 - \alpha) R^k(y, z, \alpha, \beta) + (\alpha - \beta) S^k(y, z, \alpha, \beta)$$

where  $R^k$  and  $S^k$  are given by equations (18), (19) respectively. We will show that:

$$Q_x^k h(y, z, \alpha, \beta) = (1 - \alpha) R^{k+1}(y, z, \alpha, \beta) + (\alpha - \beta) S^{k+1}(y, z, \alpha, \beta)$$

From the induction assumption and equations (10), (11) we obtain:

$$\begin{aligned}
T_x Q_x^{k-1} h(y, z, \alpha, \beta) &= T_x (1 - \alpha) R^k(y, z, \alpha, \beta) \\
&+ T_x (\alpha - \beta) S^k(y, z, \alpha, \beta) = (1 - \alpha) p_1 \int_{\mathbf{E}} R^k(y, u, \alpha, \beta) \mu_y(du) \\
&+ [(1 - \alpha) q_1 + (\alpha - \beta) p_2] \int_{\mathbf{E}} S^k(y, u, \alpha, \beta) \langle \underline{e}, \underline{f}_y(u) \rangle \mu_y(du) \\
&= (1 - \alpha) \left[ p_1 \int_{\mathbf{E}} R^k(y, u, \alpha, \beta) \mu_y(du) + q_1 \int_{\mathbf{E}} S^k(y, u, \alpha, \beta) \right. \\
&\times \left. \langle \underline{e}, \underline{f}_y(u) \rangle \mu_y(du) \right] + (\alpha - \beta) p_2 \int_{\mathbf{E}} S^k(y, u, \alpha, \beta) \langle \underline{e}, \underline{f}_y(u) \rangle \mu_y(du)
\end{aligned} \tag{20}$$

Notice that

$$(1 - \alpha) R^{k+1}(y, z, \alpha, \beta) + (\alpha - \beta) S^{k+1}(y, z, \alpha, \beta)$$

is equal  $\alpha - \beta = h(y, z, \alpha, \beta) = Q_x^k h(y, z, \alpha, \beta)$  for  $(y, z, \alpha, \beta) \in \mathcal{R}_k$  and, taking into account (20), it is equal  $T_x Q_x^{k-1} h(y, z, \alpha, \beta) = Q_x^k h(y, z, \alpha, \beta)$  for  $(y, z, \alpha, \beta) \notin \mathcal{R}_k$ , where  $\mathcal{R}_k$  is given by (20). Finally we get:

$$Q_x^k h(y, z, \alpha, \beta) = (1 - \alpha) R^{k+1}(y, z, \alpha, \beta) + (\alpha - \beta) S^{k+1}(y, z, \alpha, \beta)$$

This proves (18) and (19). Using the monotone convergence theorem and theorems of optimal stopping theory ([6]) we conclude that the optimal stopping time  $\tau^*$  is given by (17).

**Proof.** Part (b). First, notice that  $\Pi_1^1 = \varphi_x(u) = 1 - \frac{p_1}{p_1 + q_1 \langle \underline{e}, \underline{f}_x(u) \rangle}$ .

Under the assumption  $\tau^* < \infty$  a.s. we get:

$$\begin{aligned}
P_x(\tau^* < \infty, \theta_1 \leq \tau^* < \theta_2) &= \sup_{\tau} E Z_{\tau} \\
&= E \max\{h(x, X_1, \Pi_1^1, \Pi_1^2), T_x h^*(x, X_1, \Pi_1^1, \Pi_1^2)\} = E \lim_{k \rightarrow \infty} Q_x^k h(x, X_1, \Pi_1^1, \Pi_1^2) \\
&= E \left[ (1 - \Pi_1^1) R^*(x, X_1, \Pi_1^1, \Pi_1^2) + (\Pi_1^1 - \Pi_1^2) S^*(x, X_1, \Pi_1^1, \Pi_1^2) \right] \\
&= p_1 \int_{\mathbf{E}} R^*(x, u, \varphi_x(u), 0) \mu_x(du) + q_1 \int_{\mathbf{E}} S^*(x, u, \varphi_x(u), 0) \langle \underline{e}, \underline{f}_x(u) \rangle \mu_x(du)
\end{aligned}$$

We used Lemma 2 here and simple calculations for  $\Pi_1^1, \Pi_1^2$ . This ends the proof.

#### 4 Remarks

It is notable that the solution of formulated problem depends only on two-dimensional vector of posterior processes. None additional process containing knowledge about  $\varepsilon$  is necessary. This is a consequence of considered payoff function (3). The presence of variable  $\varepsilon$  and random choice of the distribution on the second epoch are reflected in weighted average  $\langle \underline{e}, \underline{f}_y(u) \rangle$ . What implies that the problem can be reduced to the case when we just fix the distribution on the middle segment of observations.

The formulas obtained are very general and for this reason - quite complicated. We simplify the model by assuming that  $P(\theta_1 > 0) = 1$  and  $P(\theta_2 > \theta_1) = 1$ . However, it seems that some further simplifications can be made in special cases. Further research should be carried out in this direction. From a practical point of view, computer algorithms are necessary to construct  $B^*$  – the set in which we stop our observable sequence. In the paper an extension of the sequential approach proposed by [6] is applied. There variety of models related to disorder in random sequences are presented in the monography by [3].

## References

- [1] T. Bojdecki. Probability maximizing approach to optimal stopping and its application to a disorder problem. *Stochastics*, 3:61–71, 1979.
- [2] T. Bojdecki and J. Hosza. On a generalized disorder problem. *Stochastic Processes Appl.*, 18:349–359, 1984.
- [3] B.E. Brodsky and B.S. Darkhovsky. *Nonparametric Methods in Change-Point Problems*. Mathematics and its Applications (Dordrecht). 243. Dordrecht: Kluwer Academic Publishers. 224 p., Dordrecht, 1993.
- [4] P. Dube and R. Mazumdar. A framework for quickest detection of traffic anomalies in networks. Technical report, Electrical and Computer Engineering, Purdue University, November 2001. [citeseer.ist.psu.edu/506551.html](http://citeseer.ist.psu.edu/506551.html).
- [5] G.V. Moustakides. Quickest detection of abrupt changes for a class of random processes. *IEEE Trans. Inf. Theory*, 44(5):1965–1968, 1998.
- [6] A.N. Shiryaev. *Optimal Stopping Rules*. Springer-Verlag, New York, Heidelberg, Berlin, 1978.
- [7] K. Szajowski. Optimal on-line detection of outside observation. *J.Stat. Planning and Inference*, 30:413–426, 1992.
- [8] K. Szajowski. A two-disorder detection problem. *Appl. Math.*, 24(2):231–241, 1996.
- [9] B. Yakir. Optimal detection of a change in distribution when the observations form a Markov chain with a finite state space. In D. Siegmund E. Carlstein, H.-G. Müller, editor, *Change-point Problems. Papers from the AMS-IMS-SIAM Summer Research Conference held at Mt. Holyoke College, South Hadley, MA, USA, July 11–16, 1992*, volume 23 of *IMS Lecture Notes-Monograph Series*, pages 346–358, Hayward, California, 1994. Institute of Mathematical Statistics.
- [10] M. Yoshida. Probability maximizing approach for a quickest detection problem with complicated Markov chain. *J. Inform. Optimization Sci.*, 4:127–145, 1983.