



HAL
open science

Non-asymptotic sequential change-point detection for Markov chains with applications in the epidemic statistical analysis

Evgenii Pchelintsev, Serguei M Pergamenchtchikov, Roman O Tenzin

► **To cite this version:**

Evgenii Pchelintsev, Serguei M Pergamenchtchikov, Roman O Tenzin. Non-asymptotic sequential change-point detection for Markov chains with applications in the epidemic statistical analysis. 2022. hal-03871920

HAL Id: hal-03871920

<https://hal.science/hal-03871920>

Preprint submitted on 25 Nov 2022

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.

Non-asymptotic sequential change-point detection for Markov chains with applications in the epidemic statistical analysis.

Evgenii A. Pchelintsev^{a,*}, Serguei M. Pergamenchchikov^b, Roman O. Tenzin^c

^a *International Laboratory of Statistics of Stochastic Processes and Quantitative Finance, National Research Tomsk State University, Tomsk, Russia*

^b *Laboratoire de Mathématiques Raphaël Salem, UMR 6085 CNRS-Université de Rouen Normandie, France and National Research Tomsk State University*

^c *Laboratoire de Mathématiques Raphaël Salem, UMR 6085 CNRS-Université de Rouen Normandie, France*

Abstract

We consider disruption detection problems for statistical models with dependent observations given by Markov chains over a finite time interval in the Bayesian setting for the uniform prior distribution. To do this, using the methods of optimal stopping, a new optimal sequential detection procedure is constructed that minimizes the average delay time in the class of sequential procedures with the probabilities of false alarms bounded by some fixed value. The main difference between the proposed detection procedure and the usual ones is that it is based not on the posterior probabilities, but on Roberts statistics. This makes it possible to provide optimal detection in a non-asymptotic sense over a finite time interval without using additional unknown parameters, in contrast to the well-known Bayesian procedures based on a priori geometric distribution containing an unknown parameter. Then we apply the constructed procedures to the problem of early detection of the beginning of the spread of the epidemic. To this end we use two epidemic models: the binomial models proposed in [2] and the models based on the Gaussian approximation introduced in [9].

Keywords: Disruption Markov models; Sequential detection; Optimal stopping; Markov epidemic models; Bayes procedure.

2010 MSC: Primary 62L10; 62L15; Secondary 60G40; 60J05; 60J20.

*Corresponding author. Email address: 4ubarok@gmail.com

1. Introduction

1.1. Motivations and contributions

The main motivation of this paper is to develop new optimal statistical methods for the early epidemic detection problems on the basis of the sequential analysis approach proposed by Baron in [1]. The main idea of this approach is that the onset of an epidemic is associated with exceeding a certain threshold of the probability of infection, and the problem is to detect the tipping point of this probability as early as possible and predict the epidemic before it spreads massively. To this end usually one uses the the quickest detection methods in times series based on the sequential analysis (see, e.g., [4, 7, 8, 13, 14] for details and further references). It should be noted, that for such change point problems one needs to use the statistical models with essentially dependent observations given by Markov processes (see, for example, [2]). In this paper we study sequential change-point detection problems for such models in Bayesian setting and our main goal is to develop non-asymptotic optimal detection methods for the Bayesian risks based on the uniform prior distribution of the change point moment. It should be emphasized, that usually the Bayesian detection procedure is based on the geometric prior distribution (see, for example, for the i.i.d case in [11] and for general non i.i.d case in [15] in the asymptotic setting). Unfortunately, the geometric distribution has an unknown parameter, which creates significant difficulties in the practical use of such Bayesian procedures. Moreover, note also, that usually for non i.i.d statistical models the optimality of the Bayesian sequential procedure is shown only in asymptotic setting, when the probability of the false alarm goes to zero. Therefore, the contribution of this paper consists in two parts. The first is the non asymptotic setting, i.e. we construct the optimal sequential detection procedure in the class of sequential procedures with any bounded false alarm probabilities on the finite time interval. The second one we develop the sequential detection procedures for the uniform prior distribution without additional unknown parameters.

1.2. Main tool

In this paper, we develop new non-asymptotic Bayesian optimal procedures for the quickest detection of the onset of epidemics in binomial epidemiological models on a finite time interval for the uniform prior distribution. This problem is studied using optimal stopping methods for homogeneous Markov processes. Based on the method of stochastic dynamic programming and modified Roberts statistics, we develop Bayesian detection procedures. Note that such methods provide ample opportunities for practical epidemiological analysis since a uniform (not informative) distribution over a given finite time interval does not contain any parameters and is the most adequate approach to the problem of early detection of epidemics in the absence of information about the distribution of the moments of the beginning of an epidemic. For comparison, we note, as established in [6], that in the case of non-asymptotic detection for i.i.d. observations the CUSUM procedures are optimal in the minimax sense. It should also be noted that in [9] it is established that the CUSUM procedures cannot be used for epidemic binomial models because the Kullback-Leibler information is zero.

1.3. Organization of the paper

The work is organized as follows. In Section 2, we describe the observation model and formulate the problem, then in Section 3 develop optimal stopping methods for homogeneous Markov processes. In Section 4 we construct optimal procedures for detection problems. In Section 5 we apply the developed methods to study the problems of early detection of epidemics. In Appendix, we provide all the auxiliary tools.

2. Markov Model

For the disruption detection problem in Markov processes we use the change - point Markov model proposed in [5, 7, 8]. To this end we assume that our observations are a Markov process $(X_n)_{n \geq 0}$ on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ with values in a measurable state $(\mathcal{X}, \mathcal{B}, \mu)$, where μ is a σ - finite measure on \mathcal{B} . As usual here, to describe the accessible information to the time moment $n \geq 0$ we will use the fields generated by the observations, i.e. $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_n = \sigma\{X_1, \dots, X_n\}$ for $n \geq 1$. Moreover, for any $n \geq 1$ we denote the corresponding n th powers of the field \mathcal{B} by $\mathcal{B}_n = \underbrace{\mathcal{B} \otimes \dots \otimes \mathcal{B}}_n$ and we also set $\mathcal{B}_\infty = \sigma\{\cup_{n \geq 0} \mathcal{B}_n\}$. In this model, we assume that the disruption occurs at some

integer $\nu \geq 0$, i.e. the random variables $(X_n)_{0 \leq n \leq \nu}$ and $(X_n)_{n \geq \nu+1}$ are homogeneous Markov chain with a transition densities $f^*(y|x)$ and $f(y|x)$ respectively, i.e. for any $A \in \mathcal{B}$

$$\mathbf{P}(X_{n+1} \in A | X_n = x) = \int_A f^*(y|x) \mu(dy) \quad \text{for } 0 \leq n \leq \nu$$

and

$$\mathbf{P}(X_{n+1} \in A | X_n = x) = \int_A f(y|x) \mu(dy) \quad \text{for } n \geq \nu + 1.$$

It should be noted here, that in the case when $\nu \geq 1$ the density of the random vector (X_1, \dots, X_n) in \mathbb{R}^n for any $1 \leq n \leq \nu$ is given as

$$\mathbf{q}^*(y_1, \dots, y_n) = \prod_{i=1}^n f^*(y_i | y_{i-1}), \quad (2.1)$$

i.e. for any $A \in \mathcal{B}_n$

$$\mathbf{P}((X_1, \dots, X_n) \in A) = \int_A \mathbf{q}^*(y_1, \dots, y_n) \mu(dy_1) \dots \mu(dy_n).$$

If $\nu = 0$, then for all $n \geq 1$ the density is $\mathbf{q}(y_1, \dots, y_n) = \prod_{i=1}^n f(y_i | y_{i-1})$, i.e. for any $A \in \mathcal{B}_n$

$$\mathbf{P}((X_1, \dots, X_n) \in A) = \int_A \mathbf{q}(y_1, \dots, y_n) \mu(dy_1) \dots \mu(dy_n).$$

Finally, if $\nu < n$ the density of the random vector (X_1, \dots, X_n) in \mathbb{R}^n is defined as

$$\mathbf{p}_\nu(y_1, \dots, y_n) = \prod_{i=1}^{\nu} f^*(y_i | y_{i-1}) \prod_{i=\nu+1}^n f(y_i | y_{i-1}), \quad (2.2)$$

i.e. for any $A \in \mathcal{B}_n$

$$\mathbf{P}((X_1, \dots, X_n) \in A) = \int_A \mathbf{p}_\nu(y_1, \dots, y_n) \mu(dy_1) \dots \mu(dy_n).$$

We denote by \mathbf{P}^* and \mathbf{P}_ν the probability measures on \mathcal{B}_∞ generated by the distribution families $(\mathbf{q}^*(y_1, \dots, y_n))_{n \geq 1}$ and $(\mathbf{p}_\nu(y_1, \dots, y_n))_{n \geq 1}$ respectively. It is clear, that in this case the corresponding Radon-Nikodym density for the restrictions $\mathbf{P}^{*,n}$ and \mathbf{P}_ν^n on \mathcal{B}_n for $\nu < n$ has the form

$$h_{\nu,n} = \frac{d\mathbf{P}_\nu^n}{d\mathbf{P}^{*,n}} = \frac{\mathbf{p}_\nu(y_1, \dots, y_n)}{\mathbf{q}^*(y_1, \dots, y_n)} = \prod_{i=\nu+1}^n \frac{f(y_i | y_{i-1})}{f^*(y_i | y_{i-1})} \quad (2.3)$$

and $h_{\nu,n} = 1$ for $\nu \geq n$.

In this paper we consider the detection problem in the Bayes setting, i.e. we assume that the change-point ν is a integer random variable with the values in $\mathcal{I}_N = \{0, \dots, N\}$ and independent on the observations $(X_k)_{k \geq 1}$. In this paper we use the uniform prior distribution, i.e.

$$\pi_j = \pi_* = \mathbf{P}(\nu = j) = \frac{1}{N+1} \quad \text{for } 0 \leq j \leq N. \quad (2.4)$$

We need the Bayesian probability measure on the σ -field $\mathcal{I}_N \otimes \mathcal{B}_\infty$ which for any $I \subseteq \mathcal{I}_N$ and $A \in \mathcal{B}_\infty$ defined as

$$\tilde{\mathbf{P}}(I \times A) = \sum_{i \in I} \pi_i \mathbf{P}_i(A). \quad (2.5)$$

It should be noted here, that, in view of Lemma 8 the posterior probability distribution used for the Bayesian procedures can be calculated for $n \geq 0$ as

$$\nu_n = \tilde{\mathbf{P}}(\nu \leq n | \mathcal{F}_n) = \frac{R_n}{R_n + \bar{\pi}_n}, \quad (2.6)$$

where $R_n = \sum_{i=0}^n \pi_i h_{i,n}$ and $\bar{\pi}_n = \sum_{i=n+1}^N \pi_i$. We remind, that R_n is called Roberts statistics (see, for example, in [10]) and for $n \geq 1$ it can be rewritten as

$$R_n = \eta_n \sum_{i=0}^{n-1} \pi_i h_{i,n-1} + \pi_n = \eta_n R_{n-1} + \pi_n, \quad (2.7)$$

where $\eta_j = \eta(X_j, X_{j-1})$, $\eta(y, x) = f(y, x)/f^*(y, x)$ and $R_0 = \pi_*$.

To formulate the detection problem we denote by \mathcal{M}_α the set of all $(\mathcal{F}_n)_{n \geq 0}$ stopping times with values in \mathcal{I}_N for which *Probability of a False Alarm (PFA)* is less than some fixed threshold $0 \leq \alpha \leq 1$, i.e.

$$\tilde{\mathbf{P}}(\tau < \nu) \leq \alpha. \quad (2.8)$$

For any $\tau \in \mathcal{M}_\alpha$ we set the average delay risk as

$$\mathcal{R}(\tau) = \tilde{\mathbf{E}}(\tau - \nu)_+, \quad (2.9)$$

where $\tilde{\mathbf{E}}$ is the expectation with respect to the Bayesian measure (2.5) and $(x)_+ = \max(x, 0)$. The detection problem is to minimise this risk, i.e.

$$\inf_{\tau \in \mathcal{M}_\alpha} \mathcal{R}(\tau). \quad (2.10)$$

To study this problem, we will use the Lagrange multiplier method, i.e. for some fixed $\lambda > 0$ we consider the following minimization problem

$$\min_{\tau \in \mathcal{M}} \left(\lambda \tilde{\mathbf{E}}(\tau - \nu)_+ + \tilde{\mathbf{P}}(\tau < \nu) \right), \quad (2.11)$$

where $\mathcal{M} = \mathcal{M}_1$ is the class of all possible $(\mathcal{F}_n)_{n \geq 0}$ stopping times with values in \mathcal{I}_N . As is shown in Appendix A.2

$$\min_{\tau \in \mathcal{M}} \left(\lambda \tilde{\mathbf{E}}(\tau - \nu)_+ + \tilde{\mathbf{P}}(\tau < \nu) \right) = 1 - \max_{\tau \in \mathcal{M}} \mathbf{E}^* G_\tau, \quad (2.12)$$

where $G_n = R_n - \lambda \sum_{i=0}^{n-1} R_i$ and \mathbf{E}^* is the expectation over the probability \mathbf{P}^* . Thus, the problem (2.11) can be represented as

$$\max_{\tau \in \mathcal{M}} \mathbf{E}^* G_\tau. \quad (2.13)$$

Our main condition is the following.

C_{*}) The sequence $(Z_n)_{0 \leq n \leq N}$ with $Z_n = (R_n, X_n)$ is homogeneous Markov chain with the values in $\mathcal{Z} = \mathbb{R} \times \mathcal{X}$.

Now, we can rewrite the random variables $G = (G_k)_{0 \leq k \leq N}$ defined in (2.12) as

$$G_k = \mathbf{U}_k(g) = g(Z_k) - \sum_{j=0}^{k-1} \mathbf{c}(Z_j), \quad (2.14)$$

where $g(r, x) = r$ and $\mathbf{c}(r, x) = \lambda r$. Thus, we represented the optimisation problem (2.13) as the optimal stopping problem for the Markov chain $(Z_n)_{0 \leq n \leq N}$.

Remark 1. Note that for the first time such problems were studied [11] for the independent observations and geometric prior distribution. In this paper we extend this results for the uniform prior distribution and for the Markov chains $(X_n)_{0 \leq n \leq N}$.

Remark 2. As to the condition **C_{*}**) it should be noted that in general the process $Z = (Z_n)_{0 \leq n \leq N}$ is not Markovian. In this paper, we study the problem (2.13) on the basis of the optimal stopping methods developed for the Markov processes (see, for example, in [11] and the references therein). As we will see below, for the epidemic models this condition holds true.

3. Optimal stopping method for Markov chains

In this section we describe the method developed in [11] to study the problem (2.13), (2.14) for arbitrary Markov homogeneous chain $(Z_n)_{0 \leq n \leq N}$ with the values in some space state $(\mathcal{Z}, \mathcal{B}_{\mathcal{Z}})$ and for arbitrary functions g and \mathbf{c} for which for all $z \in \mathcal{Z}$

$$\max_{0 \leq n \leq N} \mathbf{E}_z^* |g(Z_n)| < \infty \quad \text{and} \quad \max_{0 \leq n \leq N} \mathbf{E}_z^* |\mathbf{c}(Z_n)| < \infty, \quad (3.1)$$

where \mathbf{E}_z^* denotes the expectation under condition that $Z_0 = z$. Here we denote by $(\mathcal{F}_n)_{0 \leq n \leq N}$ the fields generated by the chain $(Z_k)_{0 \leq k \leq N}$, i.e. $\mathcal{F}_0 = \{\emptyset, \Omega\}$ and $\mathcal{F}_n = \sigma\{Z_1, \dots, Z_n\}$ for $1 \leq n \leq N$. In this section we study the optimization problem

$$\sup_{\tau \in \mathcal{M}} \mathbf{E}_z^* \mathbf{U}_{\tau}(g), \quad (3.2)$$

where the sequence $(\mathbf{U}_n(g))_{0 \leq n \leq N}$ is given in (2.14). To this end we use the stochastic dynamic programming method (see, for example, [11]). For this we need to study for any $0 \leq n \leq N$ the following problems

$$\mathbf{s}_n(g)(z) = \sup_{\tau \in \mathcal{T}_n} \mathbf{E}_z^* \mathbf{U}_{\tau}(g), \quad (3.3)$$

where $\mathcal{T}_n = \{\tau \in \mathcal{M} : \tau \leq n \text{ } \mathbf{P}^* \text{ - a.s.}\}$. It should be noted, that in view of (2.14) for any $z \in \mathcal{Z}$ we obtain

$$\mathbf{s}_0(g)(z) = g(z) \quad \text{and} \quad \mathbf{s}_N(g)(z) = \sup_{\tau \in \mathcal{M}} \mathbf{E}_z^* \mathbf{U}_{\tau}(g).$$

To study the problems (3.3), first of all, one needs to find the Bellman equations for this case. To this end, for any $\mathcal{Z} \rightarrow \mathbb{R}$ functions h for which $\mathbf{E}_z^* |h(Z_1)| < \infty$ we set the following mapping

$$\mathbf{Q}(h)(z) = \max(h(z), \mathbf{T}(h)(z) - \mathbf{c}(z)), \quad z \in \mathcal{Z}, \quad (3.4)$$

where the transition operator $\mathbf{T}(h)(z) = \mathbf{E}_z^* h(Z_1)$. Now we calculate the powers for this operator.

Lemma 1. *For any $1 \leq n \leq N$ and any $\mathcal{Z} \rightarrow \mathbb{R}$ function h for which $\max_{1 \leq k \leq n} \mathbf{E}_z^* |h(Z_k)| < \infty$ for any $z \in \mathcal{Z}$, the n th power $\mathbf{Q}^n(h)(z)$ is well defined and, moreover, there exists a constant $\mathbf{U}_n^* > 0$ for which*

$$|\mathbf{Q}^n(h)(z)| \leq \mathbf{U}_n^* \left(\sum_{j=0}^n \mathbf{E}_z^* |h(Z_j)| + \sum_{j=0}^N \mathbf{E}_z^* |\mathbf{c}(Z_j)| \right), \quad \forall z \in \mathcal{Z}. \quad (3.5)$$

Proof. We show this lemma by the induction. Indeed, for $n = 1$ this inequality directly follows the definition (3.4). Assume now, that this lemma holds true for some fixed $1 \leq n \leq N - 1$. Now, one needs to show it for $n + 1$. Let h be a $\mathcal{Z} \rightarrow \mathbb{R}$ function for which $\max_{1 \leq k \leq n+1} \mathbf{E}_z^* |h(Z_k)| < \infty$. In this case note, that for the function $\check{h} = \mathbf{Q}(h)$ we can show, that

$$\max_{1 \leq k \leq n} \mathbf{E}_z^* |\check{h}(Z_k)| < \infty. \quad (3.6)$$

Indeed, from the definition (3.4) it follows, that for any $z \in \mathcal{Z}$

$$|\check{h}(z)| \leq |h(z)| + \mathbf{T}(|h|)(z) + |\mathbf{c}(z)|. \quad (3.7)$$

Taking into account here, that the expectation $\mathbf{E}_z^* \mathbf{T}(|h|)(Z_k) = \mathbf{E}_z^* |h|(Z_{k+1})$ for any $1 \leq k \leq n$ we obtain the property (3.6). Therefore, we can use the upper bound (3.5) for the function \check{h} , i.e.

$$|\mathbf{Q}^{n+1}(h)(z)| = |\mathbf{Q}^n(\check{h})(z)| \leq \mathbf{U}_n^* \left(\sum_{j=0}^n \mathbf{E}_z^* |\check{h}(Z_j)| + \sum_{j=0}^N \mathbf{E}_z^* |\mathbf{c}(Z_j)| \right).$$

Therefore, using here the upper bound (3.7), we obtain the inequality (3.5) for $n + 1$ which completes the proof. \square

Lemma 2. Let h be a $\mathcal{Z} \rightarrow \mathbb{R}$ function for which $\max_{0 \leq n \leq N} \mathbf{E}_z^* |h(Z_n)| < \infty$ for all $z \in \mathcal{Z}$. Then, for all $1 \leq n \leq N$ and $z \in \mathcal{Z}$

$$\mathbf{Q}^n(h)(z) = \max \left(h(z), \mathbf{T}(\mathbf{Q}^{n-1}(h))(z) - \mathbf{c}(z) \right). \quad (3.8)$$

Proof. First of all, we note, that in view of Lemma 1 all powers of the mapping \mathbf{Q} for the function h are well defined and from the inequality (3.5) it follows, that $|\mathbf{T}(\mathbf{Q}^k(h))(z)| < \infty$ for any $1 \leq k \leq N-1$ and $z \in \mathcal{Z}$. Moreover, using the definition (3.4) it is easy to deduce that $\mathbf{Q}^k(h)(z) \geq h(z)$ and, therefore, $\mathbf{T}(\mathbf{Q}^k(h))(z) \geq \mathbf{T}(h)(z)$, for any $1 \leq k \leq N-1$ and $z \in \mathcal{Z}$. To prove the representation (3.8) we use the induction method. Indeed, for $n=1$, taking into account, that $\mathbf{Q}^0(h) = h$, the representation (3.8) is the definition of the operator \mathbf{Q} . According to the induction assume now, that the property (3.8) holds true for some fixed $2 \leq n \leq N-1$ and for any $z \in \mathcal{Z}$. Therefore, to show this lemma it is sufficient to establish the property (3.8) for $n+1$. To do this, setting $\check{h}(z) = \mathbf{Q}(h)(z)$ and using the induction assumption, we get

$$\begin{aligned} \mathbf{Q}^{n+1}(h)(z) &= \mathbf{Q}^n(\check{h})(z) = \max \left(\check{h}(z), \mathbf{T}(\mathbf{Q}^{n-1}(\check{h}))(z) - \mathbf{c}(z) \right) \\ &= \max \left(h(z), \mathbf{T}(h)(z) - \mathbf{c}(z), \mathbf{T}(\mathbf{Q}^n(h))(z) - \mathbf{c}(z) \right) \\ &= \max \left(h(z), \mathbf{T}(\mathbf{Q}^n(h))(z) - \mathbf{c}(z) \right), \end{aligned}$$

i.e. we get the equality (3.8) for $n+1$. Hence, Lemma 2. \square

Lemma 3. For any $n \geq 0$, for any $\mathcal{Z} \rightarrow \mathbb{R}$ function h with $\max_{0 \leq k \leq n} \mathbf{E}_z^* |h(Z_k)| < \infty$ and for any $z \in \mathcal{Z}$

$$\mathbf{s}_n(h)(z) \leq \mathbf{Q}^n(h)(z), \quad (3.9)$$

where the functions $\mathbf{s}_n(h)(z)$ are defined in (3.3).

Proof. We will proceed this proof by induction. For $n=0$ this is true by the definition, i.e. $\mathbf{s}_0(h)(z) = h(z)$ for any $\mathcal{Z} \rightarrow \mathbb{R}$ function h . Assume now, that this lemma holds true for some $n \geq 1$. We show it for $n+1$. So, let h be some $\mathcal{Z} \rightarrow \mathbb{R}$ function for which $\max_{0 \leq k \leq n+1} \mathbf{E}_z^* |h(Z_k)| < \infty$. Then, using the upper bound (3.7) for the function $\check{h}(z) = \mathbf{Q}(h)(z)$, we get, that $\max_{0 \leq k \leq n} \mathbf{E}_z^* |\check{h}(Z_k)| < \infty$. Therefore, according to the induction assumption

$$\mathbf{s}_n(\check{h})(z) \leq \mathbf{Q}^n(\check{h})(z) = \mathbf{Q}^{n+1}(h)(z). \quad (3.10)$$

Let now $\tau \in \mathcal{T}_{n+1}$. Then, for any $z \in \mathcal{Z}$

$$\begin{aligned} \mathbf{E}_z^* \mathbf{U}_\tau(h) &= \mathbf{E}_z^* \left(\mathbf{U}_\tau(h) \mathbf{1}_{\{\tau \leq n\}} + \mathbf{U}_{n+1}(h) \mathbf{1}_{\{\tau > n\}} \right) \\ &= \mathbf{E}_z^* \left(\mathbf{U}_{\hat{\tau}}(h) \mathbf{1}_{\{\tau \leq n\}} + \mathbf{1}_{\{\tau > n\}} \bar{\mathbf{U}}_n(h) \right) = \mathbf{E}_z^* \left(\mathbf{U}_{\hat{\tau}}(h) \mathbf{1}_{\{\tau \leq n\}} + \mathbf{1}_{\{\tau > n\}} \bar{\mathbf{U}}_{\hat{\tau}}(h) \right), \end{aligned}$$

where the stopping time $\hat{\tau} = \min(\tau, n)$ belongs to \mathcal{T}_n and

$$\bar{\mathbf{U}}_n(h) = \mathbf{E}^* (\mathbf{U}_{n+1}(h) | \mathcal{F}_n) = \mathbf{E}^* (h(Z_{n+1}) | \mathcal{F}_n) - \sum_{i=0}^n \mathbf{c}(Z_i) = \mathbf{T}(h)(Z_n) - \sum_{i=0}^n \mathbf{c}(Z_i).$$

Therefore, taking into account, that

$$\max(\mathbf{U}_n(h), \bar{\mathbf{U}}_n(h)) = \max \left(h(Z_n), \mathbf{T}(h)(Z_n) - \mathbf{c}(Z_n) \right) - \sum_{i=0}^{n-1} \mathbf{c}(Z_i) = \mathbf{Q}(h)(Z_n) - \sum_{i=0}^{n-1} \mathbf{c}(Z_i),$$

we get, that

$$\mathbf{E}_z^* \mathbf{U}_\tau(h) \leq \mathbf{E}_z^* \max(\mathbf{U}_{\hat{\tau}}(h), \bar{\mathbf{U}}_{\hat{\tau}}(h)) = \mathbf{E}_z^* \left(\check{h}(Z_{\hat{\tau}}) - \sum_{i=0}^{\hat{\tau}-1} \mathbf{c}(Z_i) \right) \leq \mathbf{s}_n(\check{h})(z)$$

and through the inequality (3.10) we obtain the upper bound (3.9) for $n+1$. Hence Lemma 3. \square

Now we set

$$Y_k = \mathbf{Q}^{n-k}(g)(Z_k) - \sum_{i=0}^{k-1} \mathbf{c}(Z_i). \quad (3.11)$$

We study the properties of this process.

Lemma 4. *The process (3.11) is such, that $Y_n = G_n$ and*

$$Y_k = \max(G_k, \mathbf{E}^*(Y_{k+1} | \mathcal{F}_k)), \quad \text{for } 0 \leq k \leq n-1. \quad (3.12)$$

Remark 3. Note now that the equations (3.12) are the Bellman equations for the problem (2.13) and Lemma 4 means, that the process (3.11) is the Snell envelop (Minimal Excessive Functions) (see, for example, [11]).

Note that according to the general optimal stopping theory (see, for example, Theorem 3.2 from [3]) the solutions for the problems (3.3) are given as

$$\tau_n^* = \min \{0 \leq k \leq n : Y_k = G_k\} = \min \{0 \leq k \leq n : \mathbf{Q}^{n-k}(g)(Z_k) = g(Z_k)\}. \quad (3.13)$$

Now, using Lemma 4 one can show the following result.

Theorem 1. *For all $n \geq 1$ and $z \in \mathcal{Z}$ the stopping time (3.13) is the solution for the problem (2.13), (2.14), i.e.*

$$\sup_{\tau \in \mathcal{T}_n} \mathbf{E}_z^* G_\tau = \mathbf{E}_z^* G_{\tau_n^*}. \quad (3.14)$$

Proof. First, we show, that the stopping Snell envelop $M_k = Y_{k \wedge \tau_n^*}$ is a martingale. Indeed, note that the conditions (3.1) provide, that $\max_{0 \leq k \leq n} \mathbf{E}^* |Y_k| < \infty$. Moreover, taking into account, that $M_{k+1} = Y_{\tau_n^*} \mathbf{1}_{\{\tau_n^* \leq k\}} + Y_{(k+1)} \mathbf{1}_{\{\tau_n^* > k\}}$ and, that in view of the definitions (3.11) and (3.13) the conditional expectation $\mathbf{E}^*(Y_{k+1} | \mathcal{F}_k) = Y_k$ on the set $\{k < \tau_n^*\}$, we obtain

$$\mathbf{E}^*(M_{k+1} | \mathcal{F}_k) = Y_{\tau_n^*} \mathbf{1}_{\{\tau_n^* \leq k\}} + \mathbf{1}_{\{\tau_n^* > k\}} \mathbf{E}^*(Y_{k+1} | \mathcal{F}_k) = Y_{\tau_n^*} \mathbf{1}_{\{\tau_n^* \leq k\}} + Y_k \mathbf{1}_{\{\tau_n^* > k\}} = Y_{k \wedge \tau_n^*} = M_k.$$

Therefore, $M_0 = \mathbf{E}_z^* M_n$, i.e. $Y_0 = \mathbf{E}_z^* Y_{\tau_n^*}$. Note that, from (3.11) and (3.13) we obtain, that $Y_0 = \mathbf{Q}^n(g)(z)$ and $Y_{\tau_n^*} = G_{\tau_n^*}$, i.e. $\mathbf{Q}^{(n)}(g)(z) = \mathbf{E}_z^* G_{\tau_n^*} \leq \sup_{\tau \in \mathcal{T}_n} \mathbf{E}_z^* G_\tau = \mathbf{s}_n(z)$. Thus, Lemma 3 implies

$$\mathbf{s}_n(z) = \sup_{\tau \in \mathcal{T}_n} \mathbf{E}_z^* G_\tau = \mathbf{Q}^n(g)(z) = \mathbf{E}_z^* G_{\tau_n^*}$$

and we get Theorem 1. \square

Remark 4. It should be noted, that Theorem 1 can be shown through the induction method (see, for example, the proof of Theorem 3.2 from [3]) or Theorem 1 in Chapter 2 from [11]).

4. Optimal sequential detection procedure

Now we apply the sequential procedure (3.13) to the problem (2.11). In this case for any $\lambda > 0$ this moment has the following form

$$\mathbf{t}_\lambda^* = \min \{k \geq 0 : \mathbf{Q}_\lambda^{N-k}(g)(R_k, X_k) = R_k\}, \quad (4.1)$$

where $g(r, x) = r$, the process $(R_k)_{0 \leq k \leq N}$ is defined in (2.7) with $\pi_j = 1/(N+1)$ for $0 \leq j \leq N$ and, moreover, the mapping (3.4) for any $r \in \mathbb{R}_+$ and $x \in \mathcal{X}$ is defined as

$$\mathbf{Q}_\lambda(h)(r, x) = \max(h(r, x), \mathbf{T}(h)(r, x) - \lambda r) \quad \text{and} \quad \mathbf{T}(h)(r, x) = \mathbf{E}_{r,x}^* h(R_1, X_1). \quad (4.2)$$

Here $\mathbf{E}_{r,x}^*(\cdot) = \mathbf{E}^*(\cdot | R_0 = r, X_0 = x)$. First, we have to study the properties of the stopping times (4.1).

Lemma 5. *The family $(\mathbf{t}_\lambda^*)_{\lambda \geq 0}$ is a.s. decreasing, i.e. $\mathbf{t}_{\lambda_1}^* \leq \mathbf{t}_{\lambda_0}^*$ a.s. for $\lambda_1 > \lambda_0 \geq 0$.*

Theorem 2. *Assume, that there exist $0 < \lambda_\alpha < \infty$ such that for $0 < \alpha < 1$*

$$\widetilde{\mathbf{P}}(\mathbf{t}_{\lambda_\alpha}^* < \nu) = \alpha. \quad (4.3)$$

Then the stopping time (4.1) with $\lambda = \lambda_\alpha$ is a solution of the problem (2.10), i.e.

$$\widetilde{\mathbf{E}}(\mathbf{t}_{\lambda_\alpha}^* - \nu)_+ = \inf_{\tau \in \mathcal{M}_\alpha} \widetilde{\mathbf{E}}(\tau - \nu)_+. \quad (4.4)$$

Proof. First of all, note that the stopping time (4.1) is equal to N for $\lambda = 0$, i.e. $\mathbf{t}_0^* = N$. Indeed, in this case

$$\mathbf{T}(g)(r, x) = \mathbf{E}_{r,x}^* R_1 = r \mathbf{E}^*(\eta_1 | X_0 = x) + \pi_* = r \int_{\mathcal{X}} \frac{f(y, x)}{f^*(y, x)} f^*(y, x) \mu(dy) + \pi_* = r \int_{\mathcal{X}} f(y, x) \mu(dy) + \pi_* = r + \pi_*$$

and $\pi_* = 1/(N + 1)$. Therefore, $\mathbf{Q}_0(g) = g + \pi_*$. Note now here, that from (3.8) we obtain, that the n th power of the mapping \mathbf{Q}_0 can be represented as

$$\mathbf{Q}_0^n(g) = \max(g, \mathbf{T}(\mathbf{Q}_0^{n-1}(g))).$$

Using this representation and the induction method, we can show that $\mathbf{Q}_0^n(g) = g + n\pi_*$ for any $n \geq 1$, i.e. $\mathbf{t}_0^* = N$ a.s. and, therefore, in view of Lemma 5, $\lambda_\alpha > 0$ for $\alpha > 0$. Moreover, for any $\tau \in \mathcal{M}_\alpha$ we obtain, that

$$\widetilde{\mathbf{E}}(\tau - \nu)_+ \geq \frac{1}{\lambda_\alpha} (\lambda_\alpha \widetilde{\mathbf{E}}(\tau - \nu)_+ + \widetilde{\mathbf{P}}(\tau < \nu) - \alpha).$$

Now, applying here Theorem 1 with $n = N$ to the sequence $(G_k)_{0 \leq k \leq N}$ defined in (2.14) and using the definition of λ_α , we obtain, that

$$\widetilde{\mathbf{E}}(\tau - \nu)_+ \geq \frac{1}{\lambda_\alpha} (\lambda_\alpha \widetilde{\mathbf{E}}(\mathbf{t}_{\lambda_\alpha}^* - \nu)_+ + \widetilde{\mathbf{P}}(\mathbf{t}_{\lambda_\alpha}^* < \nu) - \alpha) = \widetilde{\mathbf{E}}(\mathbf{t}_{\lambda_\alpha}^* - \nu)_+,$$

i.e. this implies (4.4) and, hence, Theorem 2. \square

Now we study the condition (4.3). To this end note, that the function

$$\mathbf{F}(\lambda) = \widetilde{\mathbf{P}}(\mathbf{t}_\lambda^* < \nu) = \frac{1}{N+1} \sum_{m=0}^{N-1} \mathbf{P}^*(\mathbf{t}_\lambda^* \leq m). \quad (4.5)$$

Now we set

$$\Lambda_0 = \{\lambda \geq 0 : \mathbf{Q}_\lambda^N(g)(\pi_*, X_0) > \pi_*\} \quad \text{and} \quad \lambda_{max} = \sup\{\lambda > 0 : \mathbf{Q}_\lambda^N(g)(\pi_*, X_0) > \pi_*\}. \quad (4.6)$$

Lemma 6. *The set Λ_0 is an interval, i.e. $\Lambda_0 = [0, \lambda_{max}[$ and $0 < \lambda_{max} < 1$.*

Note here, that from Lemma 5 it follows, that the function \mathbf{F} is increasing, i.e. $\mathbf{F}(\lambda_1) \geq \mathbf{F}(\lambda_0)$ for $\lambda_1 > \lambda_0 > 0$ and, that $\mathbf{F}(0) = 0$. Moreover, Lemma 6 implies, that $\mathbf{F}(\lambda) = N/(N+1)$ for all $\lambda \geq \lambda_{max}$. So, to avoid the trivial solutions for the problem (2.10) we assume, that $0 < \alpha < N/(N+1)$. Therefore, the equation (4.3) for $0 < \lambda < \lambda_{max}$ can be represented as

$$\mathbf{F}(\lambda) = \alpha. \quad (4.7)$$

Now we study this function.

Lemma 7. *The function $F(\cdot)$ is continuous on Λ_0 .*

Remark 5. It should be noted, that if the function \mathbf{F} is continuous, then this equation has a solution λ_α for any $0 < \alpha < N/(N+1)$. If there are many roots we choose any.

5. Epidemiy detection problem

In this paper we study the epidemiological statistical models proposed in [2]. Denoting the number of susceptible people at the time n by X_n and the last time moment before the epidemics beginning by ν assume, that $(X_n)_{1 \leq n \leq \nu}$ and $(X_n)_{n > \nu}$ are homogeneous Markov processes with the values in the finite space (\mathcal{X}, μ) , $\mathcal{X} = \{0, \dots, D\}$, where $D \in \mathbb{N}$ is the number of susceptible people at the initial time $n = 0$. Moreover, in this case we set $\mu\{0\} = \dots = \mu\{D\} = 1$. In this model, the conditional $X_n | X_{n-1}$ densities for $n \leq \nu$ and for $n > \nu$ are defined respectively as

$$f^*(y|x) = \binom{x}{y} (\theta_*)^{x-y} (1 - \theta_*)^y \mathbf{1}_{\{x \geq y\}} \quad \text{and} \quad f(y|x) = \binom{x}{y} \theta^{x-y} (1 - \theta)^y \mathbf{1}_{\{x \geq y\}}, \quad (5.1)$$

where $0 < \theta_* < \theta < 1$. It should be noted that, for any $\mathcal{X}^m \rightarrow \mathbb{R}$ function U we get

$$\mathbf{E}^* U(X_1, \dots, X_m) = \sum_{(k_1, \dots, k_m) \in \mathcal{X}^m} U(k_1, \dots, k_m) \mathbf{q}_m^*(k_1, \dots, k_m), \quad (5.2)$$

where

$$\mathbf{q}_m^*(k_1, \dots, k_m) = \prod_{l=1}^m f^*(k_l | k_{l-1}) = \prod_{l=1}^m \binom{k_{l-1}}{k_l} (\theta_*)^{k_{l-1} - k_l} (1 - \theta_*)^{k_l} \mathbf{1}_{\{k_{l-1} \geq k_l\}} \quad \text{and} \quad k_0 = D.$$

Moreover, note, that the process $(X_n)_{n \geq 1}$ can be represented as the Galton-Watson process (see, for example, in [12]), i.e. setting $X_0 = D$ for $n \geq 1$

$$X_n = S_{n, X_{n-1}} \quad \text{and} \quad S_{n, m} = \sum_{i=1}^m \xi_{n, i}, \quad (5.3)$$

where $(\xi_{n, i})_{1 \leq n \leq N, i \geq 1}$ is i.i.d. sequence Bernoulli's random variables with $\mathbf{P}(\xi_{n, i} = 1) = 1 - \vartheta_n$ and $\vartheta_n = \theta_* \mathbf{1}_{\{n \leq \nu\}} + \theta \mathbf{1}_{\{\nu > n\}}$. Note, that in this case the function η defined in (2.7) has the form

$$\eta(y, x) = (\kappa_1)^{x-y} (\kappa_2)^y \mathbf{1}_{\{x \geq y\}}, \quad (5.4)$$

where $\kappa_1 = \theta/\theta_*$ and $\kappa_2 = (1 - \theta)/(1 - \theta_*)$.

Proposition 1. *For the process (5.3) and the uniform prior distribution (2.4) the condition \mathbf{C}_* holds true.*

We have to find now the transition mapping \mathbf{T} defined in (3.4). To this end, using the difference equation (2.7) and the definition (5.3) we can obtain directly, that for any bounded $\mathbb{R}_+ \times \mathcal{X} \rightarrow \mathbb{R}$ function h and for any $z = (r, x) \in \mathbb{R}_+ \times \mathcal{X} \rightarrow \mathbb{R}$

$$\mathbf{T}(h)(z) = \mathbf{E}^*(h(R_1, X_1) | R_0 = r, X_0 = x) = \mathbf{E}_x^* \widetilde{h}(S_{1, x}, z), \quad (5.5)$$

where $\widetilde{h}(y, z) = h((\kappa_1)^{x-y} (\kappa_2)^y r + \pi_*, y)$ and $\pi_* = 1/(N + 1)$. Therefore, for any $z = (r, x) \in \mathbb{R}_+ \times \mathcal{X}$ this mapping can be calculated as

$$\mathbf{T}(h)(z) = \sum_{j=0}^x \widetilde{h}(j, z) \binom{x}{j} (\theta_*)^{(x-j)} (1 - \theta_*)^j = \sum_{j=0}^x h((\kappa_1)^{x-j} (\kappa_2)^j r + \pi_*, j) \binom{x}{j} (\theta_*)^{(x-j)} (1 - \theta_*)^j. \quad (5.6)$$

Thus, to calculate the optimal stopping time \mathbf{t}_λ^* defined in (4.1) one needs to calculate all values $(\mathbf{Q}_\lambda^{N-n}(g)(Z_n))_{0 \leq n \leq N}$ for the process $(Z_n = (R_n, X_n))_{0 \leq n \leq N}$ and the $\mathbb{R}_+ \times \mathcal{X} \rightarrow \mathbb{R}$ function $g(z) = g(r, x) = r$. Note, that in this case in view of the property (3.8) we obtain for any $1 \leq m \leq N$

$$\mathbf{Q}_\lambda^m(g)(z) = \max(r, \mathbf{T}(\mathbf{Q}_\lambda^{m-1}(g))(z) - \lambda r). \quad (5.7)$$

From here it follows, that for any $0 \leq m \leq N - 1$

$$\{\mathbf{t}_\lambda^* \leq m\} = \left\{ \min_{0 \leq j \leq m} (\mathbf{Q}_\lambda^{N-j}(g)(R_j, X_j) - R_j) = 0 \right\} = \left\{ \min_{0 \leq j \leq m} (\mathbf{T}(\mathbf{Q}_\lambda^{N-1-j}(g))(R_j, X_j) - (1 + \lambda)R_j) \leq 0 \right\},$$

where $R_0 = \pi_*$ and $X_0 = D$. Therefore, using the distribution (5.2), we obtain, that for $1 \leq m \leq N - 1$

$$\mathbf{P}^*(\mathbf{t}_\lambda^* \leq m) = \sum_{(k_1, \dots, k_m) \in \mathcal{X}^m} U_{m, \lambda}(k_1, \dots, k_m) \mathbf{q}_m^*(k_1, \dots, k_m), \quad (5.8)$$

where $k_0 = D$, $U_{m, \lambda}(k_1, \dots, k_m) = \mathbf{1}_{\{\min_{0 \leq j \leq m} (\mathbf{Q}_\lambda^{N-j}(g)(\mathbf{r}_j, k_j) - \mathbf{r}_j) = 0\}}$ and

$$\mathbf{r}_j = \frac{1}{N + 1} \sum_{l=0}^j \prod_{l=l+1}^j (\kappa_1)^{k_{l-1} - k_l} (\kappa_2)^{k_l} \mathbf{1}_{\{k_{l-1} \geq k_l\}}.$$

We remind, that by the definition $\prod_{l=k}^j = 1$ for $k > j$. Therefore, the function (4.5) can be represented as

$$\mathbf{F}(\lambda) = \frac{1}{N+1} \mathbf{1}_{\{\mathbf{Q}_\lambda^N(g)(\pi_*, D) = \pi_*\}} + \sum_{m=1}^{N-1} \sum_{(k_1, \dots, k_m) \in \mathcal{X}^m} U_{m,\lambda}(k_1, \dots, k_m) \mathbf{q}_m^*(k_1, \dots, k_m).$$

Note, that $\mathbf{P}^*(\mathbf{t}_\lambda^* > 0) = 1$ for $\lambda \in \Lambda_0$, i.e. for such λ

$$\mathbf{F}(\lambda) = \sum_{m=1}^{N-1} \sum_{(k_1, \dots, k_m) \in \mathcal{X}^m} \mathbf{1}_{\{\min_{1 \leq j \leq m} (\mathbf{Q}_\lambda^{N-j}(g)(\mathbf{r}_j, k_j) - \mathbf{r}_j) = 0\}} \mathbf{q}_m^*(k_1, \dots, k_m). \quad (5.9)$$

It should be noted, that in view of Lemma 7 this function is right continuous, therefore, to study the equation (4.7) for $0 < \alpha < N/(N+1)$ we set

$$\bar{\lambda}_\alpha = \inf\{0 \leq \lambda \leq \lambda_{max} : \mathbf{F}(\lambda) \geq \alpha\}. \quad (5.10)$$

It is clear, that if there exist a solution of the equation (4.7), then $\bar{\lambda}_\alpha$ is the solution also, i.e. $\mathbf{F}(\bar{\lambda}_\alpha) = 0$ and we can take $\lambda_\alpha^* = \bar{\lambda}_\alpha$. Note here, that for any $1 \leq l \leq N$

$$\mathbf{r}_{min} := \min_{0 \leq j \leq N} \mathbf{r}_j \leq R_l \leq \max_{0 \leq j \leq N} \mathbf{r}_j := \mathbf{r}_{max}, \quad (5.11)$$

where \mathbf{r}_j is defined in (5.8). Note now also, that

$$\mathbf{Q}_\lambda(g)(r, x) = \max(r, (1-\lambda)r + \pi_*) = (1-\lambda)r + \pi_* > r$$

for $0 \leq \lambda < \pi_*/\mathbf{r}_{max}$ and $\mathbf{r}_{min} \leq r \leq \mathbf{r}_{max}$. Therefore, through the induction method, using the equality (5.7), we can show that for any $2 \leq m \leq N$, $\mathbf{r}_{min} \leq r \leq \mathbf{r}_{max}$ and $x \in \mathcal{X}$

$$\mathbf{Q}_\lambda^m(g)(r, x) > r. \quad (5.12)$$

From here we obtain, that $\mathbf{F}(\lambda) = 0$ for any $0 \leq \lambda < \pi_*/\mathbf{r}_{max}$, i.e. $\bar{\lambda}_\alpha > 0$. In the same way, we can show $\mathbf{Q}_\lambda^m(g)(r, x) = r$ for $\lambda \geq \pi_*/\mathbf{r}_{min}$ and $1 \leq m \leq N$, i.e. $\lambda_{max} \leq \pi_*/\mathbf{r}_{min} < \infty$ and, therefore, $\bar{\lambda}_\alpha < \infty$. Indeed, in practice the values of the observations X_n are sufficiently large, i.e., $D \rightarrow \infty$. In this case the calculation of this function (5.9) will take a lot too long to complete. So, to overcome this difficulty we propose to pass to epidemic model introduced in [9] which is based on the Gaussian approximation of the model (5.3), i.e. we can represent the (5.3) as

$$X_n = (1 - \vartheta)X_{n-1} + \sum_{j=1}^{X_{n-1}} (\xi_{n,j} - 1 + \vartheta),$$

where $\vartheta = \theta$ and θ^* in the post-change and pre-change modes, respectively. Using the Gaussian approximation for the last sum for sufficiently large values of X_{n-1}

$$\frac{1}{\sqrt{X_{n-1}}} \sum_{j=1}^{X_{n-1}} (\xi_{n,j} - 1 + \vartheta) \sim \mathcal{N}(0, \sigma_\vartheta^2), \quad \sigma_\vartheta^2 = \vartheta(1 - \vartheta),$$

we obtain the following model

$$X_n = (1 - \vartheta_n)X_{n-1} + \sigma_n \sqrt{|X_{n-1}|} \zeta_n, \quad X_0 = D, \quad (5.13)$$

where $\vartheta_n = \theta_* \mathbf{1}_{\{n \leq \nu\}} + \theta \mathbf{1}_{\{\nu > n\}}$, $\sigma_n = \sigma_{\vartheta_n}$ and $(\zeta_n)_{n \geq 1}$ are i.i.d. $\mathcal{N}(0, 1)$ random variables. We assume here, that $\theta_* + \theta < 1$. In this case, the space $(\mathcal{X}, \mathcal{B}, \mu)$ is $\mathcal{X}^n = \mathbb{R}_* = \mathbb{R} \setminus \{0\}$, $\mathcal{B} = \mathcal{B}(\mathbb{R}_*)$ is the Borel field and μ is the Lebesgue measure on $\mathcal{B}(\mathbb{R}_*)$. It should be noted, that to avoid large values for the process we normalise it over the initial value, i.e. we

pass from the original observations $(X_n)_{n \geq 0}$ to the process $(\tilde{X}_n = X_n/D)_{n \geq 0}$ which obeys the same equation with $\tilde{X}_0 = 1$. Obviously,

$$f^*(y|x) = \frac{1}{\sigma_{\theta_*} \sqrt{2\pi|x|}} \exp\left\{-\frac{\mathbf{a}_*^2(y, x)}{2}\right\} \quad \text{and} \quad f(y|x) = \frac{1}{\sigma_\theta \sqrt{2\pi|x|}} \exp\left\{-\frac{\mathbf{a}^2(y, x)}{2}\right\}, \quad (5.14)$$

where

$$\mathbf{a}_*(y, x) = \frac{y - (1 - \theta_*)x}{\sigma_{\theta_*} \sqrt{|x|}} \quad \text{and} \quad \mathbf{a}(y, x) = \frac{y - (1 - \theta)x}{\sigma_\theta \sqrt{|x|}}.$$

From here we obtain, that the function (2.7) can be represented as

$$\eta(y, x) = \frac{\sigma_{\theta_*}}{\sigma_\theta} \exp\left\{-\frac{\mathbf{a}_*^2(y, x)}{2} - \frac{\mathbf{a}^2(y, x)}{2}\right\}. \quad (5.15)$$

It should be noted, that the operator (5.5) can be represented as

$$\mathbf{T}(h)(z) = \mathbf{E}^*(h(R_1, X_1) | R_0 = r, X_0 = x) = \frac{1}{\sigma_{\theta_*} \sqrt{2\pi|x|}} \int_{\mathbb{R}_*} \tilde{h}(y, z) \exp\left\{-\frac{\mathbf{a}_*^2(y, x)}{2}\right\} dy, \quad (5.16)$$

where $\tilde{h}(y, z) = h(\eta(y, x)r + \pi_*, y)$ and $\pi_* = 1/(N + 1)$. Note here, that in this case for any $\mathcal{X}^m \rightarrow \mathbb{R}$ function U

$$\mathbf{E}^* U(X_1, \dots, X_m) = \int_{\mathcal{X}^m} U(x_1, \dots, x_m) \mathbf{q}_m^*(x_1, \dots, x_m) dx_1, \dots, dx_m, \quad (5.17)$$

where $x_0 = D$ and

$$\mathbf{q}_m^*(x_1, \dots, x_m) = \prod_{i=1}^m f^*(x_i | x_{i-1}) = \frac{1}{(\sigma_{\theta_*} \sqrt{2\pi})^m} \prod_{i=1}^m \frac{1}{|x_{i-1}|^{m/2}} \exp\left\{-\frac{\mathbf{a}_*^2(x_i, x_{i-1})}{2}\right\}.$$

Therefore, the function (5) is defined as

$$\mathbf{F}(\lambda) = \pi_* \mathbf{1}_{\{\mathbf{Q}_\lambda^N(g)(\pi_*, D) = \pi_*\}} + \pi_* \sum_{m=1}^{N-1} \int_{\mathcal{X}^m} U_{m,\lambda}(x_1, \dots, x_m) \mathbf{q}_m^*(x_1, \dots, x_m) dx_1, \dots, dx_m, \quad (5.18)$$

where the functions $U_{m,\lambda}$ are defined in (5.8) with $\mathbf{r}_j = \pi_* \sum_{i=0}^j \prod_{l=i+1}^j \eta(x_l, x_{l-1})$ and $x_0 = D$.

Acknowledgment

This work is supported by RSF Grant No. 22-21-00302 (National Research Tomsk State University, Russia).

A. Appendix

A.1. Posterior distributions

Lemma 8. *The posterior probabilities $v_n = \tilde{\mathbf{P}}(v \leq n | \mathcal{F}_n)$ are given in (2.6).*

Proof. Let ξ_n be some random bounded variable measurable with respect to \mathcal{F}_n . Then using the definition (2.5) and (2.3), we obtain, that

$$\tilde{\mathbf{E}} \xi_n \mathbf{1}_{\{v \leq n\}} = \sum_{j=0}^n \pi_j \mathbf{E}_j \xi_n = \mathbf{E}^* \sum_{j=0}^n \pi_j h_{j,n} \xi_n = \mathbf{E}^* R_n \xi_n,$$

where R_n is defined in (2.6). Therefore, taking into account that $\sum_{j=0}^N \pi_j h_{j,n} = R_n + \bar{\pi}_n$ we obtain that the last expectation equals to

$$\mathbf{E}^* R_n \xi_n = \tilde{\mathbf{E}} \frac{R_n}{R_n + \bar{\pi}_n} \xi_n$$

and we obtain the equality (2.6). Hence Lemma 8. \square

A.2. Proof of the equality (2.12)

First we show that the optimization problem (2.11) can be represented as

$$\min_{\tau \in \mathcal{M}} \left(\lambda \widetilde{\mathbf{E}}(\tau - \nu)_+ + \widetilde{\mathbf{P}}(\tau < \nu) \right) = 1 - \max_{\tau \in \mathcal{M}} \widetilde{\mathbf{E}} \mathbf{V}_\tau, \quad (\text{A.1})$$

where $\mathbf{V}_n = \nu_n - \lambda \sum_{i=0}^{n-1} \nu_i$. To this end note, that $(\tau - \nu)_+ = \sum_{j=0}^{\tau-1} \mathbf{1}_{\{\nu \leq j\}}$. Therefore,

$$\widetilde{\mathbf{E}}(\tau - \nu)_+ = \sum_{j=0}^{N-1} \widetilde{\mathbf{E}} \mathbf{1}_{\{j < \tau\}} \mathbf{1}_{\{\nu \leq j\}} = \sum_{j=0}^{N-1} \widetilde{\mathbf{E}} \mathbf{1}_{\{j < \tau\}} \widetilde{\mathbf{E}}(\mathbf{1}_{\{\nu \leq j\}} | \mathcal{F}_j) = \widetilde{\mathbf{E}} \sum_{j=0}^{\tau-1} \nu_j.$$

Therefore, taking into account, that $\widetilde{\mathbf{P}}(\nu \leq \tau) = \widetilde{\mathbf{E}} \nu_\tau$ we obtain the equality (A.1). Moreover, for any bounded \mathcal{F}_n measurable random variable ξ_n from (2.6) we obtain

$$\widetilde{\mathbf{E}} \nu_n \xi_n = \sum_{i=0}^N \pi_i \mathbf{E}_i \nu_n \xi_n = \mathbf{E}^* \xi_n \nu_n \sum_{i=0}^N \pi_i h_{i,n} = \mathbf{E}^* \xi_n \nu_n (R_n + \bar{\pi}_n) = \mathbf{E}^* \xi_n R_n.$$

Therefore,

$$\begin{aligned} \widetilde{\mathbf{E}} \mathbf{V}_\tau &= \widetilde{\mathbf{E}} \nu_\tau - \lambda \widetilde{\mathbf{E}} \sum_{i=0}^{\tau-1} \nu_i = \sum_{n=0}^N \widetilde{\mathbf{E}} \nu_n \mathbf{1}_{\{\tau=n\}} - \lambda \sum_{i=0}^{N-1} \widetilde{\mathbf{E}} \mathbf{1}_{\{j < \tau\}} \nu_j \\ &= \sum_{n=0}^N \mathbf{E}^* R_n \mathbf{1}_{\{\tau=n\}} - \lambda \sum_{i=0}^{N-1} \mathbf{E}^* \mathbf{1}_{\{j < \tau\}} R_j = \mathbf{E}^* \left(R_\tau - \lambda \sum_{i=0}^{\tau-1} R_i \right) = \mathbf{E}^* G_\tau. \end{aligned}$$

Hence the equality (2.12). \square

A.3. Proof of Proposition 1

To show this proposition, it suffices to check for any bounded $\mathbb{R} \times \mathcal{X} \rightarrow \mathbb{R}$ function h

$$\mathbf{E}^*(h(Z_n) | Z_1, \dots, Z_{n-1}) = \mathbf{E}^*(h(Z_n) | Z_{n-1}),$$

i.e. for any $\mathbf{a}_i \in \mathbb{R}$, $\mathbf{b}_i \in \mathcal{X}$, $1 \leq i \leq n-1$, one needs to show, that

$$\mathbf{E}^*(h(Z_n) | Z_1 = (\mathbf{a}_1, \mathbf{b}_1), \dots, Z_{n-1} = (\mathbf{a}_{n-1}, \mathbf{b}_{n-1})) = \mathbf{E}^*(h(Z_n) | Z_{n-1} = (\mathbf{a}_{n-1}, \mathbf{b}_{n-1})). \quad (\text{A.2})$$

Indeed, using the definition (2.7), we obtain, that the first expectation in (A.2) equals to

$$\mathbf{E}^*(H_n(X_n) | Z_1 = (\mathbf{a}_1, \mathbf{b}_1), \dots, Z_{n-1} = (\mathbf{a}_{n-1}, \mathbf{b}_{n-1})),$$

where $H_n(x) = h(\eta(x, \mathbf{b}_{n-1}) \mathbf{a}_{n-1} + \pi_*)$ and the function η is defined in (5.4). Moreover, taking into account, the definition (5.3), we obtain that $\mathbf{E}^*(H_n(X_n) | Z_1 = (\mathbf{a}_1, \mathbf{b}_1), \dots, Z_{n-1} = (\mathbf{a}_{n-1}, \mathbf{b}_{n-1})) = \mathbf{E}^* H_n(S_{n, \mathbf{b}_{n-1}})$. Similarly, we can deduce that $\mathbf{E}^*(h(Z_n) | Z_{n-1} = (\mathbf{a}_{n-1}, \mathbf{b}_{n-1})) = \mathbf{E}^* H_n(S_{n, \mathbf{b}_{n-1}})$, which implies the property (A.2). Hence Proposition 1. \square

A.4. Proof of Lemma 4

We show this lemma by the back induction method. First note, that from the definition (3.11) and taking into account that by the convention $\sum_{j=0}^{-1} = 0$ we get $Y_n = G_n$. Let now $1 \leq m \leq n-1$. Note, that from (3.11) and (3.8) one can deduce, that

$$Y_{m-1} = \mathbf{Q}^{(n-m+1)}(g)(Z_{m-1}) - \sum_{i=0}^{m-2} c(Z_i)$$

$$\begin{aligned}
&= \max \left(g(Z_{m-1}), \mathbf{T}(\mathbf{Q}^{(n-m)}(g))(Z_{m-1}) - c(Z_{m-1}) - \sum_{i=0}^{m-2} c(Z_i) \right) \\
&= \max \left(G_{m-1}, \mathbf{T}(\mathbf{Q}^{(n-m)}(g))(Z_{m-1}) - \sum_{i=0}^{m-1} c(Z_i) \right).
\end{aligned}$$

Moreover, in view of the definition of the transition mapping \mathbf{T} in (3.4) and taking into account, that $(Z_k)_{0 \leq k \leq n}$ is homogeneous Markov chain and denoting $\tilde{z} = Z_{m-1}$, we obtain that for any $\mathbb{R}_+ \times \mathcal{X} \rightarrow \mathbb{R}$ function h

$$\mathbf{T}(h)(Z_{m-1}) = \mathbf{E}_{\tilde{z}} h(Z_1) = \mathbf{E}(h(Z_m) | Z_{m-1}) = \mathbf{E}(h(Z_m) | \mathcal{F}_{m-1}).$$

Therefore, this implies, that

$$\mathbf{T}(\mathbf{Q}^{(n-m)}(g))(Z_{m-1}) - \sum_{i=0}^{m-1} c(Z_i) = \mathbf{E} \left(\mathbf{Q}^{(n-m)}(g)(Z_m) - \sum_{i=0}^{m-1} c(Z_i) \mid \mathcal{F}_{m-1} \right) = \mathbf{E}(Y_m | \mathcal{F}_{m-1})$$

and we obtain that $Y_{m-1} = \max(G_{m-1}, \mathbf{E}(Y_m | \mathcal{F}_{m-1}))$. Hence Lemma 4. \square

A.5. Proof of Lemma 5

First, we show, that for any nonnegative $\mathbb{R}_+ \times \mathbb{R} \rightarrow \mathbb{R}_+$ function h , for any integer $n \geq 1$, any $r \in \mathbb{R}_+$ and $x \in \mathcal{X}$ the family $(\mathbf{Q}_\lambda^n(h)(r, x))_{\lambda \geq 0}$ defined in (4.2) is decreasing, i.e.

$$\mathbf{Q}_{\lambda_1}^n(h)(r, x) \leq \mathbf{Q}_{\lambda_0}^n(h)(r, x) \quad \text{for } \lambda_1 > \lambda_0 \geq 0. \quad (\text{A.3})$$

Indeed, the definition (4.2) implies the inequality (A.3) for $n = 1$. Then, for $n \geq 2$ using the induction method this inequality can be obtained from the definition (3.8). \square

A.6. Proof Lemma 6

First we note, that the function $\mathbf{Q}_\lambda^N(g)(\pi_*, X_0)$ is decreasing and continuous in λ . Therefore, if λ_1 belongs to Λ_0 , then $\lambda \in \Lambda_0$ for any $0 \leq \lambda \leq \lambda_1$. Taking into account, that $\mathbf{Q}_0^N(g)(\pi_*, X_0) = g(\pi_*, X_0) + N\pi_* = (1 + N)\pi_*$ and that

$$\lim_{\lambda \rightarrow 0} \mathbf{Q}_\lambda^N(g)(\pi_*, X_0) = \mathbf{Q}_0^N(g)(\pi_*, X_0),$$

we can conclude, that there exists $\lambda_1 > 0$ for which $\mathbf{Q}_{\lambda_1}^N(g)(\pi_*, X_0) > \pi_*$, i.e. $\lambda_{max} \geq \lambda_1 > 0$. Moreover, note that for $\lambda \geq 1$ we obtain that $\mathbf{Q}_\lambda(g)(r, x) = r$ for any $r \geq \pi_*$ and $x \in \mathcal{X}$. Furthermore, by induction through the property (4.6) we can conclude that $\mathbf{Q}_\lambda^m(g)(r, x) = r$ for any $r \geq \pi_*$ and $x \in \mathcal{X}$ and $m \geq 2$. \square

A.7. Proof of Lemma 7

First, note, that similarly to (5.9) for $\lambda \in \Lambda_0$ the function (5.18) can be represented as

$$\mathbf{F}(\lambda) = \frac{1}{N+1} \sum_{m=1}^{N-1} \mathbf{F}_m(\lambda) \quad \text{and} \quad \mathbf{F}_m(\lambda) = \mathbf{P}^* \left(\min_{1 \leq j \leq m} (\mathbf{Q}_\lambda^{N-j}(g)(R_j, X_j) - R_j) = 0 \right). \quad (\text{A.4})$$

Taking into account here the representation (5.7), we obtain, that for any $2 \leq m \leq N-1$ and $\lambda \in \Lambda_0$

$$\mathbf{F}_m(\lambda) = \mathbf{P}^* \left(\min_{1 \leq j \leq m} (\mathbf{Q}_\lambda^{N-j}(g)(R_j, X_j) - R_j) = 0 \right) = \mathbf{P}^* \left(\min_{1 \leq j \leq m} \zeta_{\lambda, N-1-j}(R_j, X_j) \leq 0 \right),$$

where $\zeta_{\lambda, k}(r, x) = T(\mathbf{Q}_\lambda^k(g))(r, x) - (1 + \lambda)r$. Note, that the function $\zeta_{\lambda, k}(r, x)$ is continuous and is decreasing in view of Lemma 6 in $\lambda > 0$, $r > 0$ and $x \in \mathcal{X}$. Therefore, it is right continuous. \square

References

- [1] Baron, M. (2002) Bayes and asymptotically pointwise optimal stopping rules for the detection of influenza epidemics. - In *in Case Studies in Bayesian Statistics*, vol. 6, C. Gatsonis, R. E. Kass, A. Carriquiry, A. Gelman, D. Higdon, D. K. Pauler, and I. Verdinelli, eds., pp. 153–163, New York: Springer.
- [2] Baron, M., Choudhary K. and Yu, X. (2013) Change-Point Detection in Binomial Thinning Processes, with Applications in Epidemiology. - *Sequential Analysis: Design Methods and Applications*, **32**, 350- 367.
- [3] Chow, Y.S., Robbins, H. and Siegmund, D. *Great Expectations: The Theory of Optimal Stopping*. HOUGHTON MIFFLIN COMPANY BOSTON, New York, Atlanta, Geneva, Illinois, Dallas, Palo Alto, 1971.
- [4] Lai, T.L. (1998) Informations bounds and quick detection of parameters changes in stochastic systems. *IEEE Transactions on Information Theory*, **44**(7), 2917-2929
- [5] Lai, T.L. (2000) Sequential multiple hypothesis testing and efficient fault detection-isolation in stochastic systems. *IEEE Trans. Inform. Theory*, **46**, 595 – 608.
- [6] Moustakides G.V. (1986) Optimal stopping times for detecting changes in distributions. *Ann. Stat.*, **14**, 1379–1387
- [7] Pergamenchtchikov, S.M. and Tartakovsky, A.G. (2018) Asymptotically optimal point wise and minimax quickest change-point detection for dependent data. *Stat Inference Stoch Process*, **21**, 217 – 259
- [8] Pergamenchtchikov, S.M. and Tartakovsky, A. G. (2019) Asymptotically optimal pointwise and minimax change-point detection for general stochastic models with a composite post-change hypothesis. - *Journal of Multivariate Analysis*, **174**, November, 2019.
- [9] Pergamenchtchikov, S. M., Tartakovsky, A. G. and Spivak, V. (2022) Minimax and pointwise sequential changepoint detection and identification for general stochastic models. - *Journal of Multivariate Analysis*, **190**, July 2022, 104977. <https://doi.org/10.1016/j.jmva.2022.104977>
- [10] Pollak M, Tartakovsky A.G. (2009) Optimality properties of the Shiryaev–Roberts procedure. - *Stat. Sin.*, **19**, 1729 - 1739
- [11] Shiryaev, A. N. (2008). *Optimal Stopping Rules*, New York: Springer.
- [12] Spitzer, F., Kesten, H. and Ney, P. (1966) The Galton-Watson process with mean one and finite variance. *Probability theory and its applications*, **11**, 579 – 611.
- [13] Tartakovsky, A., Nikiforov, I. and Basseville, M. *Sequential Analysis: Hypothesis Testing and Changepoint Detection*. Chapman & Hall book, Taylor & Francis Group, 2015.
- [14] Tartakovsky, A. *Sequential Change Detection and Hypothesis Testing: General Non-i.i.d. Stochastic Models and Asymptotically Optimal Rules*, Chapman & Hall/CRC Monographs on Statistics and Applied Probability) 1st Edition., 2019.
- [15] Tartakovsky, A.G., Veeravalli, V.V. (2005) General asymptotic Bayesian theory of quickest change detection. - *Theory Probab. Appl.*, **49**, 458 – 497.