Theory of Stochastic Processes
Vol.13 (29), no.1-2, 2007, pp.251-266

GRIGORIJ SHURENKOV

THE LENGTH OF THE INTERVAL OF INDETERMINACY FOR THE ESTIMATE OF MULTIPLE CHANGE-POINTS

This article considers the problem of estimating the length of the interval of indeterminacy during construction of change-points' estimates using dynamical programming. It was proved that mathematical expectation of the length of the interval has asymptotically linear dependency on the penalty for a change of distribution when the number of estimations tends to infinity.

1. Introduction

The change-points' problem arises during analysis of geological, economical data and telemetrical information. There are several ways to the search for change-points. One of the ways is to consider a-posteriori estimates of change-points [1,2]. The other tries to find change-point in a minimal time while observations are still arriving [1,3-5]. In this article we consider the problem of the search for multiple change-points using dynamical programming. The algorithm of dynamical programming first was applied to a similar problem in [6]. In [7], there was suggested the fast algorithm of the search for multiple change-points, in which the problem was reduced to the search for the optimal sequence of numbers of distributions ('trajectories'). When using this algorithm we look for an argument of minimum of some functional from trajectory, that consists of a linear part and penalties for a change of distribution in a trajectory. The changes in the found trajectory will be the estimates for change-points. As that algorithm can be used in sequential search of that trajectory, the question arises if trajectory obtained on the current stage change when new data arrives, what is the length of the interval between the last observation and the observation till which the trajectory is finalized. In the work [8] was shown that for two distributions the length of the interval of indeterminacy has linear dependency on the penalty for the change when the number of observations tends to

2000 *Mathematics Subject Classifications.* Primary 62G20, Secondary 93E10
Key words and phrases. Change-points, dynamical programming

infinity. This article extends the results of [8] on the case of several distributions. It was proven that dependency between the length of the interval of indeterminacy and the penalty for the change is asymptotically linear when the penalty tends to infinity.

2. The problem

The family of random sequences $\Xi^N = \{\xi_1^N, \dots, \xi_N^N\}$, $N \geq 1$ is considered. $\{\xi_1^N, \dots, \xi_N^N\} = \{\xi_1, \dots, \xi_N\}$ is a sequence of independent random values. It is known, that every element of the sequence has a distribution from the set $\{F_1, \dots, F_K\}$. That distributions are unknown to us. $\mathsf{P}(\xi_j \in A) = F_{h_j^0}(A)$, where $h^0 = \{h_j^0, j = 1, \dots, N\}$ is a non-random sequence if numbers $h_j^0 \in \{1, \dots, K\}$, that has the following property $h_j^0 = const$ when $k_j = [\theta_i N] < j \leq [\theta_{i+1} N]$, where $0 = \theta_0 < \theta_1 < \dots < \theta_R < \theta_{R+1} = 1$ are fixed non-random numbers that are called change-moments, k_j are called change-points. Sequences of numbers of distributions h are called trajectories, h^0 is called the real trajectory of the sequence $\{\xi_1, \dots, \xi_N\}$. Numbers of distributions are called states.

Denote $\Gamma = \{1, \dots, K\}$. Let a set of functions be given

$$\phi(\cdot, j) \colon \mathbf{R} \mapsto \mathbf{R}, \quad j \in \Gamma.$$

Functions ϕ can be random. Denote

$$\mathcal{H}^N = \{(h_1, ..., h_N) \mid h_j \in \Gamma, j = 1, ..., N\}, \quad \mathcal{H} = \bigcup_N \mathcal{H}^N.$$

Let's introduce the functional $J \colon \mathcal{H} \mapsto \mathbf{R}$. The functional is specified for a trajectory $h \in \mathcal{H}^n$ by the following formula

$$J(h) = \sum_{i=1}^{n} (\phi(\xi_i, h_i) - \pi_N(h_i, h_{i-1})), \quad \pi_N(g, l) = \pi_N \, 1_{g \neq l},$$

where $\pi_N > 0$ is a non-random value. Consider

$$\hat{h} = \operatorname*{argmax}_{h \in \mathcal{H}^N} J(h). \tag{1}$$

as an estimate for h^0

Estimates for change-points are calculated using the trajectory $\hat{h}$ in the following way

$$\hat{k}_1 = k_1(\hat{h}) = \min \left\{ l \mid \hat{h}_l \neq \hat{h}_j, 1 \leq j < l \right\}$$
$$\hat{k}_i = k_i(\hat{h}) = \min \left\{ l \mid \hat{h}_l \neq \hat{h}_j, k_{i-1}(\hat{h}) \leq j < l \right\}$$

– i-th change-point in the trajectory $\hat{h}$. We can say that we consider the trajectory as series of transitions from one state to another. And i-th transition in $\hat{h}$ is the estimate for i-th change-point.
Definition 1. Trajectories g and h are called equivalent if $J(h) = J(g)$. So, $\hat{h}$ is defined accurate to equivalent trajectories. Denote

$$\begin{aligned} \bar{\theta} &= \min_{i \neq j} |\theta_i - \theta_j|, \\ \phi^{ki}(x) &= \phi(x,k) - \phi(x,i). \end{aligned}$$

3. Main results

Trajectory $\hat{h}$ (1) can be searched sequentially. Denote

$$g^j(l) = \operatorname*{argmax}_{h=\{h_0,\ldots,h_{j-1},h_j\},h_j=l} J(h),\ l \in \Gamma$$

Then

$$g_j^{j+1}(l) = \operatorname*{argmax}_{i\in\Gamma} \left(J(g^j(i)) + \phi(\xi_{j+1}, l) - \pi_N(i,l) \right),$$

$$\hat{h} = \operatorname*{argmax}_{h=h^N(l),l\in\Gamma} J(h).$$

Denote

$$\hat{h}^u = \operatorname*{argmax}_{h=g^u(l),l\in\Gamma} J(h)$$

Until some moment trajectories $g^j(l)$ can coincide, that means until that moment the optimal trajectory is known already. In the case of delays between observations' arrival, the question appears about the length of the interval of indeterminacy. Under that value we mean the number of available observations for which the optimal trajectory cannot be found at the current moment.

To estimate that number for every moment s, consider a random value

$$\chi(s) = \min \left\{ t > s \mid g_s^t(i) = g_s^t(1), i \in \Gamma \right\}.$$

$\chi(s)$ is the first moment at which the optimal trajectory is known till the moment s. We are interested in estimation of the difference $\Delta\chi(s) = \chi(s) - s$.

Denote

$$\chi(s,k) = \min\{t > s \mid g_s^t(i) = k, i \in \Gamma\}, \quad k \in \Gamma,$$

тоді

$$\chi(s) = \min_{k\in\Gamma} \chi(s,k).$$

Definition 2. The state i is called *non-favorable* at the moment t if the trajectory $g^{t+1}(i)$ ends with a transition and *favorable* in the complement case.

Definition 3. The trajectory h is called *better* that the trajectory g if $J(h) > J(g)$.
Definition 4. The state i is called *optimal* at the moment t, if $J(g^t(i)) > J(g^t(j)), \quad j \in \Gamma, \quad j \neq i$, the state i is called *almost optimal* at the moment t , if $J(g^t(i)) \geq J(g^t(j)), \quad j \in \Gamma$.
It is obvious that a state is almost optimal at the moment t if there's a trajectory equivalent to $\hat{h}^t$ that ends in that state. If all such trajectories are ending with the same state then that state is optimal.
Definition 5. Let $\alpha \in \mathbf{R}$. The state i is called α - *optimal* at the moment t if

$$J(g^t(i)) > J(g^t(j)) - \alpha, \quad j \in \Gamma, \quad j \neq i.$$

The state i is called *almost* α - *optimal* at the moment t if

$$J(g^t(i)) \geq J(g^t(j)) - \alpha, \quad j \in \Gamma, \quad j \neq i.$$

We denote the set of all α-optimal states at the moment t as $\mathcal{O}^t_\alpha$.

Introduce the probability measure $\mathsf{P}_k(\cdot)$, in which $\xi_1, \ldots \xi_N$ have the distribution F_k. Denote as E_k mathematical expectation by that measure.

Formulate the conditions

A. $\kappa = \min_{i \neq j} \mathsf{E}\phi^{ji}(\zeta_j) > 0$, де ζ_j is a random value that has the distribution F_j

B. $\phi^{ji}(\zeta_j) \leq C < \infty, \quad i, j \in \Gamma$

C. $\mathsf{D}\phi^{ji}(\zeta_j) < \sigma^2 < \infty, \quad i, j \in \Gamma$

D. functions $\phi(x, j)$ are independent from $\xi_1, \ldots \xi_N$

Consider the sequence of stopping epochs

$$w_0^{ki} = w_0^{ki}(t_0) = t_0, w_1^{ki}(t_0) = \min\left\{t > t_0 \mid \max_{t_0 < u \leq t} \sum_{j=t_0+1}^{u} \phi^{ki}(\xi_j) > \pi_N\right\}$$
$$w_0^k = w_0^k(t_0) = t_0, \quad w_l^{ki}(t_0) = w_l^{ki}(t_0) = w_1^{ki}(w_{l-1}^k(t_0))$$
$$w_l^k = w_l^k(t_0) = w_l^k(t_0) = \max_{i \neq k} w_l^{ki}(t_0)$$
$$v_l^k = v_l^k(t_0) = v_l^k(t_0) = w_{3l}^k(t_0)$$

and the set of events

$$\mathcal{E} = \{E(t_1, t_2) \mid t_1 < t_2\},$$

$$E(t_1, t_2) = \left\{\min_{i \neq k} \min_{t_1 < l_1 \leq l_2 \leq t_2} \sum_{j=l_1}^{l_2} \phi^{ki}(\xi_j) > -\frac{\pi_N}{2}\right\}, \quad t_1 < t_2$$

Denote

$$v_{\mathcal{E}}^k = v_{\mathcal{E}}^k(t_0) = v_{\mathcal{E}}^k(t_0,\phi) = \min\{v_1^k(t) \mid E(s, v_1^k(s)) \textit{ occurs }, \quad t \geq t_0\}$$
$$\tilde{v}_{\mathcal{E}}^k = \tilde{v}_{\mathcal{E}}^k(t_0) = \tilde{v}_{\mathcal{E}}^k(t_0,\phi) = \min\{v_l^k(t_0) | E(v_{l-1}^k, v_l^k) \textit{ occurs }\}$$

It is evident that $v_{\mathcal{E}}^k(t_0) \leq \tilde{v}_{\mathcal{E}}^k(t_0)$.
Theorem 1. *Let t_0 be a random value.*

$$\chi(s) \leq \min_{k\in\Gamma} v_{\mathcal{E}}^k(t_0), \quad \textit{якщо} \quad t_0 \geq s.$$

If conditions A-D hold then

$$\mathsf{E}\Delta\chi(t) \leq \frac{6(\pi_N + C)(K-1)(1+\theta^{-1})}{C_1\kappa},$$

where

$$C_1 = 1 - \frac{48\sigma^2(K-1)^2}{\pi_N}\frac{\pi_N + C}{\pi_N\kappa}.$$

Lemma 1.

$$\exists u \quad 0 \leq u \leq t \quad g^t(k) = \left\{\hat{h}_1^u, \ldots \hat{h}_u^u, k, \ldots, k\right\}$$

Proof. Suppose that the lemma does not hold. Trajectories with the following shape $\{k, \ldots, k\}$ satisfy the condition given in the lemma's condition. So let $u+1$ be the last transition in the trajectory $g^t(k)$. Denote $h' = \{g_1^t(k), \ldots, g_u^t(k)\}$. Trajectory

$$h'' = \left\{\hat{h}_1^u, \ldots \hat{h}_u^u, k, \ldots, k\right\}$$

will be better than $g^t(k)$:

$$J(g^t(k)) = J(h') - \pi_N + \sum_{j=u+1}^{t} \phi(\xi_j, k) < J(\hat{h}^u) - \pi_N + \sum_{j=u+1}^{t} \phi(\xi_j, k) = J(h'').$$

We came to contradiction. The lemma is proved.
So, the last transition to state k in the trajectory $g^t(k)$ is performed from the almost optimal state.
Lemma 2. *A state that is almost optimal at the moment t also is favorable at the moment t.*
Proof. Let k be an almost optimal state at the moment t. $J(\hat{h}_t^t) = J(g^t(k))$. Suppose that it is non-favorable. Then trajectory $g^{t+1}(k)$ ends with a transition. Consider trajectories

$$\tilde{g}^t(k) = \left\{g_1^{t+1}(k), ..., g_t^{t+1}(k)\right\}, \quad \tilde{g}^{t+1}(k) = \left\{g_1^t(k), ..., g_t^t(k), k\right\}.$$

Compare trajectories $g^{t+1}(k)$ and $\tilde{g}^{t+1}(k)$. $g^{t+1}(k)$ is not worse than $\tilde{g}^{t+1}(k)$ by the definition.

$$J(g^{t+1}(k)) \geq J(\tilde{g}^{t+1}(k)) = J(g^t(k)) + \phi(\xi_{t+1}, k).$$

The following equality holds

$$J(g^{t+1}(k)) = J(\tilde{g}^t(k)) - \pi_N + \phi(\xi_{t+1}, k).$$

because trajectory $g^{t+1}(k)$ ends with a transition. But k is almost optimal at the moment t

$$J(g^t(k)) \geq J(\tilde{g}^t(k)).$$

So,

$$0 \geq J(\tilde{g}^t(k)) - J(g^t(k)) = J(g^{t+1}(k)) - J(\tilde{g}^{t+1}(k)) + \pi_N \geq \pi_N.$$

We came to contradiction. The lemma is proved.

Lemma 3. *State i is non-favorable at moment t if and only if the following condition holds*

$$J\left(\hat{h}^t\right) - J\left(g^t(i)\right) > \pi_N \tag{2}$$

Proof. Necessity. Let the state i at the moment t to be non-favorable. We get the following from the previous lemma $g^t(i) \neq \hat{h}^t$. Then $J(g^t(i)) < J(\hat{h}^t)$. For $g^{t+1}(i)$ to be better than $\{g_1^t(i), \ldots, g_t^t(i), i\}$ the following condition is needed

$$J(g^{t+1}(i)) = J(\hat{h}^t) - \pi_N + \phi(\xi_{t+1}, i) > J(g^t(i)) + \phi(\xi_{t+1}, i)$$

So,

$$J(\hat{h}^t) - J(g^t(i)) > \pi_N.$$

Sufficency. Let inequality (2) to hold. Suppose that trajectory $g^{t+1}(i)$ does not end with a transition. Consider the following trajectory

$$\tilde{g}^{t+1}(i) = \left\{\hat{h}_1^t, ..., \hat{h}_t^t, i\right\}.$$

That trajectory is better than $g^{t+1}(i)$

$$J(\tilde{g}^{t+1}(i)) - J(g^{t+1}(i)) = J(\hat{h}^t) - \pi_N + \phi(\xi_{t+1}, i) - J(g^t(i)) - \phi(\xi_{t+1}, i) > 0.$$

As $g^{t+1}(i)$ is the best from all the trajectories that end at the moment t in the state i, we came to contradiction. So the state i at the moment t is non-favorable. The lemma is proved.

Remark. A state is non-favorable at the moment t if and only if it is not almost π_N-optimal at the moment.

Lemma 4. *Let* $i, k \in \Gamma$, $i \neq k$, $t > t_0$, $t_0 \leq u < t$, $x \in \mathbf{R}$. $\alpha, \alpha_u \in [0, \pi_N]$. *Denote*

$$C_{ki}^{tu}(x, \alpha) = \left\{ \sum\nolimits_{j=u+1}^{t} \phi^{ki}(\xi_j) > \alpha - \pi_N + x, \quad k \in \mathcal{O}_\alpha^u \right\},$$

$$\mathsf{D}_{ki}^{tu}(x) = \left\{ \sum\nolimits_{j=u+1}^{t} \phi^{ki}(\xi_j) > x \right\}.$$

The following random event

$$\bigcap_{t_0 \leq u < t} \left(C_{ki}^{tu}(x, \alpha_u) \cup D_{ki}^{tu}(x) \right) \cap D_{ki}^{tt_0}(\pi_N + x)$$

implies

$$J\left(g^t(k)\right) - J\left(g^t(i)\right) > x$$

Proof. Let trajectory $g^t(i)$ contain change-point at the interval from t_0 to t. Let u be the last change-point in $g^t(i)$.

Let the state k to be α-optimal at the moment $u-1$ and

$$\sum_{j=u}^{t} \phi^{ki}\left(\xi_j\right) > x - \pi_N + \alpha.$$

For $g^u(i)$ the following statement holds

$$J(g^u(i)) \leq J(g^{u-1}(k)) - \pi_N + \alpha + \phi(\xi_u, i)$$

because at the moment $u-1$ state i is non-favorable, and k is α-optimal with $\alpha \geq 0$.

$$J(g^{u-1}(k)) + \sum_{j=u}^{t} \phi(\xi_j, k) - J(g^u(i)) - \sum_{j=u+1}^{t} \phi\left(\xi_j, i\right) > x.$$

From $J(g^t(k)) \geq J(g^{u-1}(k)) + \sum_{j=u}^{t} \phi(\xi_j, k)$ we get

$$J\left(g^t(k)\right) - J\left(g^u(i)\right) - \sum_{j=u+1}^{t} \phi\left(\xi_j, i\right) > x.$$

Let the event

$$\sum_{j=u}^{t} \phi^{ki}\left(\xi_j\right) > x.$$

hold. In any case

$$J(g^u(k)) \geq J(\hat{h}_{u-1}) - \pi_N + \phi(\xi_u, k), \quad J(g^u(i)) = J(\hat{h}_{u-1}) - \pi_N + \phi(\xi_u, i).$$

Then

$$J\left(g^u(k)\right) + \sum_{j=u+1}^{t} \phi\left(\xi_j, k\right) - J(g^u(i)) - \sum_{j=u+1}^{t} \phi(\xi_j, i) > x$$

From the previous inequality we get

$$J(g^t(k)) - J(g^u(i)) - \sum_{j=u+1}^{t} \phi(\xi_j, i) > x$$

Let $g^t(i)$ not to have change-point at the interval $[t_0, t]$ and

$$\sum_{j=t_0+1}^{t} \phi^{ki}\left(\xi_j\right) > \pi_N + x$$

As $J(g^t(k)) \geq J(g^{t_0}(i)) + \sum_{j=t_0+1}^{t} \phi(\xi_j, k) - \pi_N$, then

$$J(g^t(k)) - J(g^u(i)) - \sum_{j=u+1}^{t} \phi(\xi_j, i) > x.$$

The lemma is proved.

Lemma 5. *Let the state k be optimal on $[t_0, w_1^k(t_0)]$, then trajectories $g^{w_1^k(t_0)}(i)$ have transition from the state k to the state i on that interval.*

Proof. As the state k is optimal at the moment t_0, we have $J(g^{t_0}(k)) - J(g^{t_0}(i)) > 0$ for $i \neq k$. Remark that

$$\sum_{j=t_0+1}^{w_1^{ki}(t_0)} \phi^{ki}(\xi_j) > \pi_N.$$

Show that there is a trajectory with a transition on the interval $(t_0, w_1^k(t_0)]$ that is better than any trajectory without transitions. Consider the trajectory

$$\tilde{g}(i) = \left\{g_1^{t_0}(k), \ldots g_{t_0}^{t_0}(k), k, \ldots, k, i, \ldots i\right\},$$

where transition from the state k to i happens at the moment $w_1^k(t_0)$. Show that trajectory is better than

$$\bar{g}(i) = \left\{g_1^{t_0}(i), \ldots g_{t_0}^{t_0}(i), i, \ldots i\right\}.$$

But

$$J(\tilde{g}(i)) - J(\bar{g}(i)) = J(g^{t_0}(k)) - J(g^{t_0}(i)) + \sum_{j=t_0+1}^{w_1^{ki}(t_0)} \phi^{ki}(\zeta_j) - \pi_N > 0.$$

So, $g^{w_1^k(t_0)}(i)$ is optimal among the trajectories that end at the state i and it cannot have no transitions on the interval $(t_0, w_1^k(t_0)]$. As a transition is always made from an almost optimal state, trajectories $g^{w_1^k(t_0)}(i)$ have a transition from the state k to the state i. Lema is proved.

Lemma 6. *Let t_0 be a random value. The following inequality holds*

$$\chi(s,k) \le v_{\mathcal{E}}^k(t_0), \quad \textit{if} \quad t_0 \ge s.$$

Proof. Let the random event $E(v_{l-1}^k(t_0), v_l^k(t_0))$ take place. Denote $w_l = w_l^k(t_0)$. Show that on the interval $[w_{3l-2}, w_{3l})$ the state k is $\frac{\pi_N}{2}$-optimal. Prove that the following event takes place

$$\bigcap_{i\neq k} \bigcap_{w_{3l-3}<u\le t} D_{ki}^{tu}\left(-\frac{\pi_N}{2}\right) \cap D_{ki}^{tw_{3l-3}}\left(\frac{\pi_N}{2}\right)$$

for $w_{3l-2} \le t < w_{3l}$. Then Lemma 4 implies the state k is $\frac{\pi_N}{2}$-optimal on the interval mentioned above. The random event

$$D_{ki}^{tu}\left(-\frac{\pi_N}{2}\right) = \left\{\sum_{j=u+1}^{t} \phi^{ki}(\xi_j) > -\frac{\pi_N}{2}\right\}$$

occurs as the event

$$E(v_{l-1}^k, v_l^k) = \left\{\min_{i\neq k} \min_{v_{l-1}^k<l_1\le l_2\le v_l^k} \sum_{j=l_1}^{l_2} \phi^{ki}\left(\xi_j\right) > -\frac{\pi_N}{2}\right\}$$

take place. From that event and the definition of w_l^{ki} we have

$$\sum_{j=w_{3l-3}+1}^{t} \phi^{ki}\left(\xi_j\right) = \sum_{j=w_{3l-3}+1}^{w_{3l-2}^{ki}} \phi^{ki}\left(\xi_j\right) + \sum_{j=w_{3l-2}^{ki}+1}^{t} \phi^{ki}\left(\xi_j\right) > \pi_N - \frac{\pi_N}{2} = \frac{\pi_N}{2}.$$

So, the event $D_{ki}^{tw_{3l-3}}\left(\frac{\pi_N}{2}\right)$ takes place. And the state k is $\frac{\pi_N}{2}$-optimal.

Prove that on the interval $[w_{3l-1}, w_{3l})$ the state k is optimal. Show that the event

$$\bigcap_{i\neq k} \bigcap_{w_{3l-3}<u\le w_{3l-2}} D_{ki}^{tu}\left(0\right) \cap \bigcap_{w_{3l-2}<u\le t} C_{ki}^{tu}\left(0, -\frac{\pi_N}{2}\right) \cap D_{ki}^{tw_{3l-3}}\left(\pi_N\right).$$

takes place for $w_{3l-1} \le t < w_{3l}$. Then optimality will follow from the Lemma 4. Let $u \in [w_{3l-3}, w_{3l-2})$. The event $D_{ki}^{tu}(0)$ occurs because

$$\sum_{j=w_{3l-2}+1}^{t} \phi^{ki}(\xi_j) > \frac{\pi_N}{2} \quad \text{та} \quad \sum_{j=u}^{w_{3l-2}} \phi^{ki}(\xi_j) > -\frac{\pi_N}{2}.$$

Let $u \in [w_{3l-2}, w_{3l-1})$. The event $C_{ki}^{tu}\left(0, -\frac{\pi_N}{2}\right)$ takes place as the event $E(v_{l-1}^k, v_l^k)$ occurs and the state k is $\frac{\pi_N}{2}$-optimal on the interval $[w_{3l-2}, w_{3l-1})$. The event $D_{ki}^{tw_{3l-3}}(\pi_N)$ takes place as

$$\sum_{j=w_{3l-3}+1}^{w_{3l-2}} \phi^{ki}(\xi_j) + \sum_{j=w_{3l-2}+1}^{w_{3l-1}^{ki}} \phi^{ki}(\xi_j) + \sum_{j=w_{3l-1}^{ki}+1}^{t} \phi^{ki}(\xi_j) > \frac{\pi_N}{2} + \pi_N - \frac{\pi_N}{2} =$$
$$= \pi_N, \quad t \in [w_{3l-1}, w_{3l}).$$

So, the state k is optimal on the interval $[w_{3l-1}, w_{3l})$, and condition of the Lema 5 holds. Trajectories $g^{v_l}(i)$ coincide till the moment w_{3l-1} and

$$\chi(s,k) \le v_{\mathcal{E}}^k(t_0).$$

The lemma is proved.
Denote $E_j = E(v_{j-1}^k, v_j^k)$.
Lema 7. *Let the condition D hold and ξ_j have the distribution F_k then*

$$\mathsf{E}_k\, \tilde{v}_{\mathcal{E}}^k \le \frac{\mathsf{E}_k\, v_1^k}{\mathsf{P}_k(E(0, v_1^k))}$$

Proof. Let us introduce conditional mathematical expectation by event

$$\mathsf{E}(\xi \mid A) = \int \xi(\omega) \mathsf{P}(d\omega \mid A)$$

Expand sequence $\xi_1, ...\xi_N$ with independent random values $\xi_{N+1}, \ldots$ that have distribution F_k. $\tilde{v}_{\mathcal{E}}^k$ will only become greater after that procedure. Then

$$\mathsf{E}_k\, \tilde{v}_{\mathcal{E}}^k \le \sum_{l=1}^{\infty} \mathsf{E}_k(v_l^k \mid \cap_{j=1}^{l-1}\overline{E_j} \cap E_l) \mathsf{P}_k\left(\cap_{j=1}^{l-1}\overline{E_j} \cap E_l\right)$$
$$= \sum_{l=1}^{\infty}\sum_{j=1}^{l} \mathsf{E}_k(\Delta v_j^k \mid \cap_{j=1}^{l-1}\overline{E_j} \cap E_l) \mathsf{P}_k\left(\cap_{j=1}^{l-1}\overline{E_j} \cap E_l\right)$$
$$= \mathsf{E}_k(v_1^k \mid \overline{E}) \mathsf{P}_k(E) \sum_{l=0}^{\infty} l(1 - \mathsf{P}_k(E)^l + \mathsf{E}_k(v_1^k \mid E) =$$
$$\mathsf{E}_k(v_1^k \mid \overline{E}) \mathsf{P}_k(\overline{E}) \mathsf{P}_k(E)^{-1} + \mathsf{E}_k(v_1^k \mid E) = \frac{\mathsf{E}_k(v_1^k)}{\mathsf{P}_k(E)}$$

The lemma is proved.

Lemma 8. If the conditions A-B and D hold and ξ_j have the distribution F_k then

$$\mathsf{E}_k\, w_1^k(t_0) - t_0 \le \frac{(K-1)(\pi_N + C)}{\kappa}.$$

Proof. In the case of homogeneously distributed $\xi_1, \dots \xi_N$

$$\mathsf{E} w_1^k(t_0) - t_0 = \mathsf{E} w_1^k(0).$$

From $w_1^k(0) = \max_{i\neq k} w_1^{ki}(0) \le \sum_{i\neq k} w_1^{ki}(0)$ follows

$$\mathsf{E} w_1^k(0) \ge \sum_{i\neq k} \mathsf{E} w^{ki}.$$

Estimate $\mathsf{E} w^{ki}$. Using Vald's lemma for the case of non-homogeneous random values we get

$$\mathsf{E} \sum_{j=1}^{w^{ki}} \phi^{ki}(\xi_j) \le \frac{\pi_N + C}{\kappa},$$

and

$$\mathsf{E} w_1^k(0) \le \frac{(K-1)(\pi_N + C)}{\kappa}.$$

The lemma is proved.

Remark. If $\xi_1, \dots \xi_N$ are homogeneous then $\Delta w_j(t_0)$ are too and

$$\mathsf{E}_k\, v_1^k(t_0) - t_0 \le \frac{3(K-1)(\pi_N + C)}{\kappa}.$$

Lemma 9.*If $\xi_1, \dots \xi_N$ have the distribution F_k and conditions A-D hold then*

$$\mathsf{P}\left(\min_{i\neq k} \min_{1\le l_1 \le l_2 \le v_1^k} \sum_{j=l_1}^{l_2} \phi^{ki}(\xi_j) > -\frac{\pi_N}{2} \right) \ge 1 - \frac{48\sigma^2 (K-1)^2}{\pi_N} \frac{\pi_N + C}{\kappa \pi_N}$$

Proof. Estimate the probability of the complementary event.

$$\mathsf{P}\left(\min_{i\neq k} \min_{1\le l_1 \le l_2 \le v_1^k} \sum_{j=l_1}^{l_2} \phi^{ki}(\xi_j) \le -\frac{\pi_N}{2} \right) \le \sum_{i\neq k} \mathsf{P}\left(\min_{1\le l_1 \le l_2 \le v_1^k} \sum_{j=l_1}^{l_2} \phi^{ki}(\xi_j) \le -\frac{\pi_N}{2} \right)$$

Estimate the probability under the sign of sum

$$\mathsf{P}\left(\min_{1\le l_1\le l_2\le v_1^k}\sum_{j=l_1}^{l_2}\phi^{ki}(\xi_j)\le -\frac{\pi_N}{2}\right)\le$$

$$\mathsf{P}\left(\min_{1\le l_1\le l_2\le v_1^k}\sum_{j=l_1}^{l_2}\left(\phi^{ki}(\xi_j)-\mathsf{E}\phi^{ki}(\xi_j)\right)\le -\frac{\pi_N}{2}\right)\le$$

$$\mathsf{P}\left(\max_{1\le l_1\le l_2\le v_1^k}\left|\sum_{j=l_1}^{l_2}\left(\phi^{ki}(\xi_j)-\mathsf{E}\phi^{ki}(\xi_j)\right)\right|\ge \frac{\pi_N}{2}\right)\le$$

$$\mathsf{P}\left(\max_{1\le l_1\le v_1^k}\left|\sum_{j=1}^{l_1}\left(\phi^{ki}(\xi_j)-\mathsf{E}\phi^{ki}(\xi_j)\right)\right|\ge \frac{\pi_N}{4}\right)$$

From Kolmogorov's inequality we get

$$\mathsf{P}\left(\max_{1\le l_1\le v_1^k}\left|\sum_{j=1}^{l_1}\left(\phi^{ki}(\xi_j)-\mathsf{E}\phi^{ki}(\xi_j)\right)\right|\ge \frac{\pi_N}{4}\right)\le\frac{16}{\pi_N^2}\,\mathsf{D}\sum_{j=1}^{v_1^k}\phi^{ki}(\xi_j)\le$$

$$\frac{16\sigma^2\mathsf{E}v_1^k}{\pi_N^2}\le\frac{48\sigma^2(K-1)(\pi_N+C)}{\kappa\pi_N^2}$$

So

$$\mathsf{P}\left(\min_{i\ne k}\min_{1\le l_1\le l_2\le v_1^k}\sum_{j=l_1}^{l_2}\phi^{ki}(\xi_j)>-\frac{\pi_N}{2}\right)\ge 1-\frac{48\sigma^2(K-1)^2(\pi_N+C)}{\kappa\pi_N^2}.$$

The lemma is proved.

Lemma 10._If the following condition holds for_ $t_0\ge t$ _:_

$$\chi(t)\le\min_{k\in\Gamma}v_{\mathcal{E}}^k(t_0),$$

then

$$\mathsf{E}\Delta\chi(t)\le 2(1+\bar\theta^{-1})\max_{k\in\Gamma}\mathsf{E}_k\,v_{\mathcal{E}}^k$$

Proof. Denote

$$v(s)=v_{\mathcal{E}}^k(s),\quad \partial e\quad k=h_s^0$$

Denote as $\rho(s)$ the number of change-points between moments s and $v(s)$. Denote as k^s the first change-point after s, k_n^s - n-th change-point after s.

$$\mathsf{E}\Delta\chi(t)=\mathsf{E}\Delta\chi(t)\,1_{\rho(t)=0}+\mathsf{E}\Delta\chi(t)\,1_{\rho(t)=1}+\mathsf{E}\Delta\chi(t)\,1_{\rho(t)>1}=a_1+a_2+a_3$$

Expand a_2.

$$\mathsf{E}\Delta\chi(t)\,1_{\rho(t)=1} = \mathsf{E}\left(\chi(t) - k^t\right) 1_{\rho(t)=1} + \mathsf{E}\left(k^t - t\right) 1_{\rho(t)=1} \leq \\ \mathsf{E}\max(\chi(t) - k^t, 0)\,1_{\rho(t)=1} + \mathsf{E}(k^t - t)\,1_{\rho(t)=1} = a_{21} + a_{22}$$

1) Estimate $a_1 + a_{22}$, using inequalities $k^t > t$ and $\chi(t) \leq v(t)$

$$\mathsf{E}\Delta\chi(t)\,1_{\rho(t)=0} + \mathsf{E}(k^t - t)\,1_{\rho(t)=1} \leq \mathsf{E}\Delta\chi(t)\,1_{\rho(t)=0} + \mathsf{E}(k^t - t)\,1_{\rho(t)\geq 1} \leq \\ \mathsf{E}\left(\Delta\chi(t)\,1_{v(t)<k^t} + (k^t - t)\,1_{v(t)\geq k^t}\right) \leq \mathsf{E}\min\left(\Delta v(t), k^t - t\right)$$

Let's estimate the right side of inequality. Obtained expression is determined by homogeneous random values ξ_j from the interval $(t, k^t]$.

$$\mathsf{E}\min(\Delta v(t), k^t - t) = \mathsf{E}_{h_t^0}\min(\Delta v(t), k^t - t) \leq \mathsf{E}_{h_t^0}\, v(t) \leq \max_{k\in\Gamma} \mathsf{E}_k\, v_{\mathcal{E}}^k(0),$$

2) Estimate a_{21}

$$\mathsf{E}\max\left(\chi(t) - k^t, 0\right) 1_{\rho(t)=1} \leq \mathsf{E}\min\left(\max(\chi(t) - k^t, 0), k_2^t - k^t\right) \\ \leq \mathsf{E}\min\left(\max(v(k^t) - k^t, 0), k_2^t - k^t\right) = \mathsf{E}\min\left(\Delta v(k^t), k_2^t - k^t\right)$$

From the previous case we get

$$\mathsf{E}\min(\Delta v(k^t), k_2^t - k^t) \leq \max_{k\in\Gamma} \mathsf{E}_k\, \Delta v_{\mathcal{E}}^k(0)$$

3) Estimate a_3: $\mathsf{E}(\Delta\chi(t)\,1_{\rho(t)>1}) \leq (N+1)\mathsf{P}\left(\Delta v(k^t) > \theta N\right)$. We get the following formula from Markov's inequality

$$\mathsf{P}\left(\Delta v(k^t) > \bar{\theta} N\right) \leq \frac{\mathsf{E}\,\Delta v(k^t)}{\bar{\theta} N}$$

So

$$\mathsf{E}(\Delta\chi(t)\,1_{\rho(t)>1}) \leq \frac{2\max_k \mathsf{E}_k\, v(0)}{\bar{\theta}}$$

Lema is proved.

Proof of theorem 1. From lemma 6 follows

$$\chi(s) \leq \min_{k\in\Gamma} v_{\mathcal{E}}^k(t_0),$$

where $t_0 \geq s$. When ξ_j have the distribution F_k we can use lemma 7, from which we get

$$\mathsf{E}_k\, v_{\mathcal{E}}^k(t_0) \leq \mathsf{E}_k\, \tilde{v}_{\mathcal{E}}^k(t_0) \leq \frac{\mathsf{E}_k\, v_1^k(0)}{\mathsf{P}_k(E(0, v_1^k))}.$$

From lemma 9 follows

$$\mathsf{P}_k(E(0, v_1^k)) \geq 1 - \frac{48\sigma^2(K-1)^2}{\pi_N}\frac{\pi_N + C}{\pi_N \kappa}.$$

Then

$$\mathsf{E}_k\, v^k_{\mathcal{E}}(t_0) \le \frac{\mathsf{E}_k\, v^k_1(0)}{C_1}.$$

From the remark to the lemma 8 follows

$$\mathsf{E}_k\, v^k_1 \le \frac{3(K-1)(\pi_N + C)}{\kappa}.$$

Then

$$\mathsf{E}_k\, v^k_{\mathcal{E}}(t_0) \le \frac{3(K-1)(\pi_N + C)}{C_1 \kappa}.$$

Using lemma 10 we get

$$\mathsf{E}\,\chi(s) \le 2(1+\bar{\theta}^{-1}) \max_{k\in\Gamma} \mathsf{E}_k\, v^k_{\mathcal{E}} \le \frac{6(1+\bar{\theta}^{-1})(K-1)(\pi_N + C)}{C_1 \kappa}.$$

The theorem is proved.

3. Generalization for the case $\mathsf{E}\,\phi^{ki}(\zeta_j) = \infty$

Generalize obtained results for the case of unconstrained difference $\phi(\zeta_j, j) - \phi(\zeta_j, i)$. We assume that difference can have infinite mathematical expectation.

Formulate conditions

E. $\mathsf{E}\,\phi^{ji}(\zeta_j) > 0, \quad i \ne j$, where ζ_j is a random value that has the distribution F_j (including the case $\mathsf{E}\,\phi^{ji}(\zeta_j) = +\infty$)

F. $\mathsf{E}\,\phi^{ji}(\zeta_j)\, 1_{\phi^{ji}(\zeta_j)<0} > -\infty, \quad i, j \in \Gamma$

G. $\mathsf{D}\,\phi^{ji}(\zeta_j)\, 1_{\phi^{ji}(\zeta_j)<0} < \sigma^2 < \infty, \quad i, j \in \Gamma$

Denote

$$\phi^{ki}_y(x) = \begin{cases} \phi^{ki}(x), & y < x, \\ y, & y \ge x. \end{cases}$$

Theorem 2. *Let conditions D-G hold then exists* $\kappa > 0,\ C \in \mathbf{R}$

$$\mathsf{E}\,\Delta\chi(t) \le \frac{6(\pi_N + C)(K-1)(1+\bar{\theta}^{-1})}{C_2 \kappa},$$

where

$$C_2 = 1 - \frac{48(\sigma^2 + C^2)(K-1)^2}{\pi_N} \frac{\pi_N + C}{\pi_N \kappa}.$$

Proof. From theorem 1, lemma 10 and lemma 7 follows

$$\mathsf{E}\,\chi(s) - s \le 2(1+\bar{\theta}^{-1}) \frac{\mathsf{E}_k\, v^k_1(0)}{\mathsf{P}_k(E(0, v^k_1))}.$$

Estimate $\mathsf{E}_k\, v_1^k(0) = 3\,\mathsf{E}_k\, w_1^k(0)$. As was shown in lemma 8 $\mathsf{E}_k\, w_1^k(0) \le \sum_{i\neq k} \mathsf{E}_k\, w_1^{ki}(0)$.

Оскільки $\sum_{j=1}^t \phi_y^{ki}(\xi_j) \le \sum_{j=1}^t \phi^{ki}(\xi_j)$, то $w_1^{ki}(0,\phi) \le w_1^{ki}(0,\phi_y)$.

From Lebegue's theorem about monotonous convergence

$$\mathsf{E}\,\phi_y^{ki}(\zeta_j) \to \mathsf{E}\,\phi^{ki}(\zeta_j), \quad \text{when} \quad y \to \infty,$$

то

$$\exists \kappa > 0 \quad \exists C \in \mathbf{R}: \quad \mathsf{E}\,\phi_C^{ki}(\zeta_j) > \kappa.$$

Then from Vald's lemma follows

$$\mathsf{E}_k\, w_1^{ki}(0,\phi_C) \le \frac{\pi_N + C}{\kappa}.$$

So

$$\mathsf{E}_k\, v_1^k(0,\phi_C) = 3\,\mathsf{E}_k\, w_1^k(0,\phi_C) \le \frac{3(K-1)(\pi_N + C)}{\kappa}.$$

Estimate $\mathsf{P}_k(E(0,v_1^k))$

$$\mathsf{P}_k(E(0,v_1^k)) \ge \mathsf{P}_k\left(\min_{i\neq k}\ \min_{0<l_1\le l_2\le v_1^k} \sum_{j=l_1}^{l_2} \phi_C^{ki}(\xi_j) > -\frac{\pi_N}{2}\right)$$

As

$$\mathsf{D}\,\phi_C^{ki}(\zeta_k) \le \mathsf{D}\,\phi^{ki}(\zeta_k)_{\phi^{ki}(\zeta_k)<0} + \mathsf{E}\left(\phi_C^{ki}(\zeta_k)_{\phi^{ki}(\zeta_k)\ge 0}\right)^2 \le \sigma^2 + C^2$$

then from lemma 9 follows

$$\mathsf{P}_k(E(0,v_1^k)) \le C_2.$$

Then

$$\mathsf{E}\,\Delta\chi(s) \le \frac{6(\pi_N + C)(K-1)(1+\bar{\theta}^{-1})}{C_2\kappa}$$

The theorem is proved.

3. Conclusion.

It was proved that mathematical expectation of the length of interval indeterminacy that is retrieved during estimation of multiple change-points has asymptotically linear dependency on the penalty for change of the distributions.

Bibliography.

1. Borovkov, A.A., *Asymtotically optimal solutions of the problem of change-points.* Theory of probability and its applications, vol.43, No. 4, 1998, c.625-654
2. Mayboroda, R.E., Sugakova, O.V., *Limit distribution of DP-estimates of multiple change-moments.* Theory of probability and mathematical statistics. vol.69, 2003, c.96-105.
3. Yakir, B., Krieger, A.M., Pollak, M., *Detecting change in regression: first order optimality.* The Annals of Statistics, 1999, vol.27, No.6, p.1896-1913.
4. Poor, H.V., *Quickest detection with exponential penalty for delay.* The Annals of Statistics, 1998, vol.26, No.6, p.2179-2205.
5. Loader, C.R., *Change point estimation using nonparametric regression.* The Annals of Statistics, 1996, vol.24,No.4,p.1667-1678
6. Mottl, V.V., Muchnik, I.B., Yakovlev, V.G., *Optimal segmentation of experimental curves.* Automatics and telemechanics, No. 8, 1983, c.84-95.
7. Sugakova, O.V., *The search for change-points in the flow of independent observations.* Theory of probability and mathematical statistics. vol.55, 1997, c.181-186.
8. Shurenkov, G.V., *Length of interval of indeterminacy in change-points estimation using dynamic programming algorithm* in print.

Department of Probability Theory and Mathematical Statistics, Kyiv National Taras Shevchenko University, Kyiv, Ukraine
E-mail address: tskorohod@voliacable.com