Теорія Ймовір. та Матем. Статист.
Вип. 70, 2004

Theor. Probability and Math. Statist.
No. 70, 2005, Pages 167–176
S 0094-9000(05)00640-X
Article electronically published on August 12, 2005

# ASYMPTOTIC BEHAVIOR OF MEDIAN ESTIMATORS OF MULTIPLE CHANGE POINTS

UDC 519.21

G. SHURENKOV

Abstract. We consider the problem of posterior estimation of multiple change points in the case of only two distributions. We find the asymptotic distribution of the difference between the median estimator of a single change point and the true change point and show that the distribution does not change if the unknown parameter is estimated by a median of the sample. We generalize the results to the case of multiple change points.

## 1. Introduction

There is an extensive literature devoted to the problem of the estimation of change points (see, for example, [5]). In particular, this problem appears when analyzing geological or telemetry data. We consider the posterior problem of the estimation of multiple change points for the model of only two distributions on the sample. The median estimator considered in [7] for the case of only one change point is also suitable for our problem. This estimator requires comparatively small amount of information about the distributions, and it can be used even in the case where the only information available is that the medians of the distributions are different. In the latter case the unknown parameter can be estimated by the median of the sample. This estimator is rough; nevertheless it can be used as first approximation in a more precise procedure of estimation of change points.

The asymptotic behavior of the difference between the median estimator and the true change point is found in the paper. The median estimator is an example of the so-called DP estimators, that is, those constructed by using the dynamic programming algorithms (see [6]). The limit distributions of estimators of change points are found in [2] in the case where distributions are known and there is only one change point. These results are generalized in [3] for DP estimators for the case where a restricted amount of information is available about the distributions and there are multiple change points.

In Section 3, we find the asymptotic distribution of the median estimator for the case where the unknown parameter is estimated by the median of the sample (the estimator is no more a standard DP estimator in this case). In Section 4, we apply the technique described in [3] to generalize these results to the case of multiple change points.

2000 *Mathematics Subject Classification.* Primary 62G20; Secondary 94A13.
*Key words and phrases.* Estimation of change points, limit distribution, dynamic programming algorithm, sampling median.

## 2. Setting of the problem

Consider a sequence of independent random variables $\{\zeta_1, \dots, \zeta_N\}$ and assume that the distribution of every random variable $\zeta_i$, $1 \le i \le N$, is either $F_1$ or $F_2$. Assume that $\operatorname{med} F_1 \ne \operatorname{med} F_2$. By $F_1$ we denote the distribution with the smaller median. Let $\mathsf{P}(\zeta_j \in A) = F_{h_j^0}(A)$, where $h^0 = \{h_j^0, j = 1, \dots, N\}$ is a nonrandom sequence such that $h_j^0 \in \{1, 2\}$ and $h_j^0 = \text{const}$ if $k_i = [\theta_i N] < j \le [\theta_{i+1} N]$ for some nonrandom numbers $0 = \theta_0 < \theta_1 < \dots < \theta_R < \theta_{R+1} = 1$ called the change moments; the $k_i$ are called the change points. Sequences $h$ of numbers of distributions are called the trajectories; $h^0$ is called the true trajectory of the sequence $\{\zeta_1, \dots, \zeta_N\}$.

Let $\hat{m} = \operatorname{med}\{\zeta_1, \dots, \zeta_N\}$. Consider the following functions:

$$\phi_m(x, 1) = \mathbb{1}_{\{x > m\}}, \qquad \phi_m(x, 2) = \mathbb{1}_{\{x < m\}}, \qquad \pi_N(g, l) = \pi_N \mathbb{1}_{g \ne l}, \quad \pi_N > 0,$$

and introduce the functional

$$J_m(h) = \sum_{i=1}^{N} (\pi_N(h_i, h_{i-1}) + \phi_m(\zeta_i, h_i)), \qquad m \in \mathbf{R}. \tag{1}$$

Consider an estimator $\tilde{h}$ of $h^0$ defined by

$$\tilde{h} = \operatorname*{argmin}_h J_m(h). \tag{2}$$

Estimators of the change points are constructed from the trajectory $\tilde{h}$ as follows:

$$k_1(\tilde{h}) = \min\left\{ l \colon \tilde{h}_l \ne \tilde{h}_j, 1 \le j < l \right\},$$
$$k_i(\tilde{h}) = \min\left\{ l \colon \tilde{h}_l \ne \tilde{h}_j, k_{i-1}(\tilde{h}) \le j < l \right\}$$

where $k_i(\tilde{h})$ is the change point $i$ of the trajectory $\tilde{h}$. By $R(\tilde{h})$ we denote the number of change points in the trajectory $\tilde{h}$.

If

$$\operatorname{med} F_1 < m < \operatorname{med} F_2,$$

then the above estimators are consistent; see [1, 4]. If the medians of the distributions are unknown, then $m$ can be estimated by the median of the sample $\hat{m} = \operatorname{med} \zeta_j$. It can be proved that if $\hat{m}$ is substituted for $m$, then the estimators still are consistent. As an estimator of $h$ we take the statistic

$$\hat{h} = \operatorname*{argmin}_h J_{\hat{m}}(h). \tag{3}$$

The estimators of the change points are $\hat{k}_{j,N} = k_j(\hat{h})$, and the estimator of the number of change points is $\hat{R} = R(\hat{h})$. The problem is to find the limit of distributions of the differences $\hat{k}_{j,N} - k_j$ as $N \to \infty$. First we consider this problem for the case where there is only one change point.

## 3. The case of only one change point

We solve the problem under the assumption that $\operatorname{med} \xi_1 < \operatorname{med} \eta_1$ where the distribution of $\xi_1$ is $F_1$, while the distribution of $\eta_1$ is $F_2$. The sequence $\zeta_j$ can be divided in this case into two consecutive parts, namely

$$\zeta_j = \begin{cases} \xi_{k-j+1}, & j \le k, \\ \eta_{j-k}, & j > k, \end{cases}$$

where the distribution of $\xi_j$ is $F_1$, while the distribution of $\eta_j$ is $F_2$. Thus we deal with the sequence $\{\xi_k, \xi_{k-1}, \dots, \xi_1, \eta_1, \dots, \eta_{n-k}\}$. It is easy to check that in this case the analog of estimator (2) is given by

$$\hat{k}_N = \operatorname*{argmax}_{l=1,\dots,N} \sum_{j=1}^{l} r(\zeta_j, m) \tag{4}$$

where

$$r(x) = r(x,m) = \mathbb{1}_{x<m} - \mathbb{1}_{x>m}.$$

The symbol argmax stands for the least $l$ for which the maximum is attained. The estimator for the change moment $\theta$ is $\hat{\theta}_N = \hat{k}_N/N$. If $m$ is estimated by the median of the sample, then we get the estimator for the change point

$$\hat{k}_N = \operatorname*{argmax}_{l=1,\dots,N} \sum_{j=1}^{l} r(\zeta_j, \hat{m}). \tag{5}$$

Let the symbol $C_m^n$ stand for the binomial coefficient $\binom{m}{n}$. Put

$$H_{u_1v_2}(p) = \begin{cases} \sum_{p/2<j\le p} \left(C_{p-1}^{j-1} - C_{p-1}^{j}\right) u_2^j v_2^{p-j} \left(1 - \frac{u_2}{v_2}\right)\left(1 - \left(\frac{u_1}{v_1}\right)^{2j-p}\right), & p>0,\\ \left(1-\frac{u_2}{v_2}\right)\left(1-\frac{u_1}{v_1}\right)v_1, & p=0,\\ \sum_{|p|/2\le j\le p} \left(C_{|p|}^{j} - C_{|p|}^{j+1}\right) u_1^j v_1^{|p|-j+1}\left(1-\frac{u_1}{v_1}\right) & \\ \qquad\qquad \times\left(1-\left(\frac{u_2}{v_2}\right)^{2j-|p|+1}\right), & p<0, \end{cases}$$

where $u_1 \in (0,1)$, $v_1 = 1-u_1$, $v_2 \in (0,1)$, and $u_2 = 1 - v_2$.

**Theorem 1.** *Let the distributions $F_1$ and $F_2$ be continuous in the interval*

$$[\operatorname{med}\xi_1, \operatorname{med}\eta_1].$$

*If $\hat{k}_N$ is defined by* (5), *then*

$$H^*(p) := \lim_{N\to\infty} \mathsf{P}\left(\hat{k}_N - k = p\right) = H_{u_1v_2}(p)$$

*where $u_1 = 1 - F_1(\bar{m})$ and $v_2 = 1 - F_2(\bar{m})$. Here the point $\bar{m}$ is the median of the distribution $\bar{F} = \theta F_1 + (1-\theta)F_2$.*

*Remark.* A certain number $\lambda \in (0,1)$ plays the role of $\theta$ in the case of multiple change points.

*Proof.* According to the definition of the estimator $\hat{k}$ the difference $\hat{k}_N - k$ can be rewritten as follows:

$$\hat{k}_N - k = \operatorname*{argmax}_{-k\le l\le N-k} \left\{ \sum_{j=1}^{l} r(\eta_j),\ l>0;\ 0,\ l=0, -\sum_{j=1}^{|l|} r(\xi_j),\ l<0 \right\}.$$

Then

$$\hat{k}_N - k = p > 0 \Leftrightarrow \begin{cases} \sum_{j=l+1}^{p} r(\eta_j) > 0, & 0 \le l \le p-1,\\ \sum_{j=p+1}^{l} r(\eta_j) \le 0, & p \le l \le N-k,\\ \sum_{j=1}^{p} r(\eta_j) > -\sum_{j=1}^{l} r(\xi_j), & 0 < l \le k. \end{cases}$$

Put

$$S_l = S_l(m) = \sum_{j=1}^{l} r(\eta_{p+1-j}), \qquad T_l = T_l(m) = \sum_{j=1}^{l} r(\eta_{p+j}),$$

$$U_l = U_l(m) = -\sum_{j=1}^{l} r(\xi_j).$$

The distribution of $\hat{k}_N - k$ can be represented as follows:

$$\begin{aligned} &\mathsf{P}\left(\hat{k}_N - k = p\right) \\ &\qquad = \mathsf{P}\big(S_l > 0,\, 1 \le l \le p;\; T_l \le 0,\, 1 \le l \le N-k-p; S_p > U_l,\, 1 \le l \le k\big). \end{aligned}$$

To evaluate the latter expression we first consider a simpler case where $m$ is fixed (that is, we treat the estimator $\hat{k}_N$ defined by (4)).

**Lemma 1.** *Let* $\operatorname{med}\xi_j < m < \operatorname{med}\eta_j$. *Then* $H(p) := \lim_{N\to\infty} \mathsf{P}(\hat{k}_N - k = p)$*, and the asymptotic distribution of* $\hat{k}_N - k$ *is* $H_{u_1 v_2}(p)$ *where* $u_1 = 1 - F_1(m)$ *and* $v_2 = 1 - F_2(m)$.

*Proof.* If $m$ is nonrandom, then the sums defined above are independent random variables, thus

$$\begin{aligned} &\mathsf{P}\left(\hat{k}_N - k = p\right) \\ &\qquad = \mathsf{P}\left(S_l > 0,\, 1 \le l \le p;\; T_l \le 0,\, 1 \le l \le N-k-p; S_p > U_l,\, 1 \le l \le k\right) \\ &\qquad = \sum_{p/2 < j \le p} \mathsf{P}\left(S_l > 0,\, 1 \le l \le p-1,\, S_p = 2j - p\right) \\ &\qquad\qquad \times \mathsf{P}(T_l \le 0, 1 \le l \le N-k-p)\mathsf{P}(U_l \le 2j-p-1,\, 1 \le l \le k), \end{aligned}$$

that is, the probability $\mathsf{P}\big(\hat{k}_N - k = p\big)$ is represented as the sum of products of three factors.

I) Consider the first of the factors. The sequence $S_l$, $1 \le l \le p$, is a random walk for which $u_2 = \mathsf{P}(\eta_j < m)$ is the probability to move to the right and $v_2 = \mathsf{P}(\eta_j > m)$ is the probability to move to the left. Thus the first factor is the probability that the random walk $S_l$ starts at 0 and walks above the zero level up to the moment $p$ when its state becomes $x = 2j - p$. Applying the reflection principle we get the desired probability:

$$\mathsf{P}\big(S_l > 0,\, 1 \le l \le p-1,\, S_p = 2j-p\big) = \left(C_{p-1}^{j-1} - C_{p-1}^{j}\right) u_2^j v_2^{p-j}. \tag{6}$$

Note that the result remains true for the case of $x = p$, too. Indeed, $j = p$ in this case and there is a unique trajectory that reaches $x$ at the moment $p$. The above result holds, since $C_{p-1}^{p-1} = 1$ and $C_{p-1}^{p} = 0$.

II) Now we evaluate the limit of $\mathsf{P}(T_l(m) \le 0,\, 1 \le l \le N-k-p)$ as $N - k \to \infty$.

We cannot directly apply the continuity of probability, since we consider a scheme of series, namely $\zeta_j = \zeta_j^N$. Let $T_l^\infty$ be an unbounded random walk with parameters $u_2$ and $v_2$. Then

$$\begin{aligned} \mathsf{P}\left(T_l^N \le 0,\, 1 \le l \le N-k-p\right) &= \mathsf{P}\left(T_l^\infty \le 0, 1 \le l \le N-k-p\right) \\ &\to \mathsf{P}\left(T_l^\infty \le 0,\, l \ge 1\right). \end{aligned}$$

Thus the desired limit is the probability that an unbounded random walk with parameters $u_2$ and $v_2$ does not cross the zero level from below. To evaluate this probability we find the distribution of the maximum of $T_l^\infty$. Put

$$p_j = \mathsf{P}\left(\max_N T_N^\infty = j\right).$$

It is easy to check that $p_j$ satisfies the following recurrence relation:

$$p_j = u_2 p_{j-1} + v_2 p_{j+1}, \qquad j \geq 1. \tag{7}$$

Solving this equation we get

$$p_j = \left(1 - \frac{u_2}{v_2}\right)\left(\frac{u_2}{v_2}\right)^j,$$

since $p_j$ is a distribution. The probability we want to evaluate is the probability that $\max T_n^\infty = 0$. Thus

$$\mathsf{P}\big(T_l^\infty \leq 0,\ l \geq 1\big) = p_0 = 1 - u_2/v_2. \tag{8}$$

III) Consider the third factor. The sequence

$$U_l = \sum_{j=1}^{l}(-r(\xi_j)) = \sum_{j=1}^{l}\big(\mathbb{1}_{\{\xi_j > m\}} - \mathbb{1}_{\{\xi_j < m\}}\big)$$

is a random walk with parameters $u_1 = \mathsf{P}(\xi_j < m)$ and $v_1 = \mathsf{P}(\xi_j > m)$.

Similarly to the preceding case, the limit of the probability

$$\mathsf{P}(U_l \leq 2j - p - 1, 1 \leq l \leq k)$$

as $k \to \infty$ is given by

$$\mathsf{P}(U_l^\infty \leq 2j - p - 1, l \geq 1) = \mathsf{P}\left(\max_{l \geq 1} U_l^\infty \leq 2j - p - 1\right) = 1 - \left(\frac{u_1}{v_1}\right)^{2j-p}.$$

Therefore if $p > 0$, then

$$\lim_{n\to\infty} \mathsf{P}\left(\hat{k}_n - k = p\right) = \sum_{p/2 < j \leq p}\left(C_{p-1}^{j-1} - C_{p-1}^{j}\right) u_2^j v_2^{p-j}\left(1 - \frac{u_2}{v_2}\right)\left(1 - \left(\frac{u_1}{v_1}\right)^{2j-p}\right).$$

The case of $p = 0$ is simpler:

$$\begin{aligned} \mathsf{P}\left(\hat{k}_n = k\right) &= \mathsf{P}\left(\sum_{j=1}^{l} r(\eta_{p-j+1}) \leq 0,\ 1 \leq l \leq n-k;\ -\sum_{j=1}^{l} r(\xi_j) < 0,\ 1 \leq l \leq k\right) \\ &\to \mathsf{P}\left(T_l^\infty \leq 0,\ U_l^\infty < 0\right) = \left(1 - \frac{u_2}{v_2}\right)\left(1 - \frac{u_1}{v_1}\right) v_1. \end{aligned} \tag{9}$$

The case of $p < 0$ is similar to the case of $p > 0$. Thus $H(p) = H_{u_1 v_2}(p)$. □

We turn back to the proof of Theorem 1 and study the limit distribution $H^*(p)$ of the random variable $\hat{k}_N - k$ for the case where the median of the sample $\hat{m} = \operatorname{med}\{\zeta_1, \ldots, \zeta_N\}$ is substituted for $m$.

It is easy to check that the distribution function of the distribution $\hat{F}_N$ converges uniformly to

$$\bar{F} = \lambda F_1 + (1 - \lambda) F_2$$

where $\lambda$ is some number of the interval $(0, 1)$, and $\hat{m}$ converges in probability to $\bar{m} = \operatorname{med} \bar{F}$ if $F_1$ and $F_2$ are continuous in a neighborhood of $\bar{m}$. Recall that $\operatorname{med} F_1 < \bar{m} < \operatorname{med} F_2$. The latter inequalities do not depend on the number of changes.

We prove that $H^*(p)$ coincides with $H(p)$ in the case of $p > 0$ (the proof for other cases is the same).

Note that $r(x, m)$ is a nondecreasing function of $m$. Thus the sums $S_l$ and $T_l$ also are nondecreasing, while $U_l$ is nonincreasing with respect to $m$. This allows one to get lower and upper estimates of $\mathsf{P}(\hat{k}_N(\hat{m}) - k = p)$ for sufficiently large $N$ (under the condition that $\mathsf{P}(|\hat{m}_N - \bar{m}| > \delta) < \varepsilon$):

$$\begin{aligned}
&\mathsf{P}\big(S_l(\bar{m}-\delta) > 0; T_l(\bar{m}+\delta) \le 0; S_p(\bar{m}-\delta) > U_l(\bar{m}-\delta)\big) - \varepsilon \\
&\quad \le \mathsf{P}\left(\hat{k}_N - k = p\right) \\
&\quad \le \mathsf{P}\big(S_l(\bar{m}+\delta) > 0;\ T_l(\bar{m}-\delta) \le 0;\ S_p(\bar{m}+\delta) > U_l(\bar{m}+\delta)\big) + \varepsilon.
\end{aligned}$$

Put $v_1(x) = \mathsf{P}(\xi_1 < x)$, $u_1(x) = \mathsf{P}(\xi_1 > x)$, $v_2(x) = \mathsf{P}(\eta_1 > x)$, $u_2(x) = \mathsf{P}(\eta_1 < x)$, $m_- = \bar{m} - \delta$, and $m_+ = \bar{m} + \delta$.

Now we find the asymptotic distributions of the sums:

$$\begin{aligned}
&\mathsf{P}\big(S_l(m_-) > 0;\ T_l(m_+) \le 0;\ S_p(m_-) > U_l(m_-)\big) \\
&\quad = \sum_{p/2 < j \le p} \mathsf{P}(S_l(m_-) > 0;\ S_p(m_-) = 2j - p)\mathsf{P}(T_l(m_+) \le 0) \\
&\qquad\qquad \times \mathsf{P}\left(U_l(m_-) \le 2j - p - 1\right) \\
&\quad \to \sum_{p/2 < j \le p} \left(C_{p-1}^{j-1} - C_{p-1}^{j}\right) v_2^{p-j}(m_-) u_2^j(m_-) \\
&\qquad\qquad \times \left(1 - \frac{u_2(m_+)}{v_2(m_+)}\right)\left(1 - \left(\frac{u_1(m_-)}{v_1(m_-)}\right)^{2j-p}\right) \\
&\quad := H_{-\delta}(p),
\end{aligned} \tag{10}$$

$$\begin{aligned}
&\mathsf{P}(S_l(m_+) > 0;\ T_l(m_-) \le 0;\ S_p(m_+) > U_l(m_+)) \\
&\quad \to \sum_{p/2 < j \le p} \left(C_{p-1}^{j-1} - C_{p-1}^{j-1}\right) v_2^{p-j}(m_+) u_2^j(m_+) \left(1 - \frac{u_2(m_-)}{v_2(m_-)}\right) \\
&\qquad\qquad \times \left(1 - \left(\frac{u_1(m_+)}{v_1(m_+)}\right)^{2j-p}\right) \\
&\quad := H_{+\delta}(p).
\end{aligned} \tag{11}$$

Note that $v_i(x) \to v_i$, $u_i(x) \to u_i$, and $x \to \bar{m}$, whence

$$H_{-\delta}(p) - \varepsilon \to H(p), \qquad H_{+\delta}(p) + \varepsilon \to H(p)$$

as $\delta \to 0$ and $\varepsilon \to 0$. Hence relations (10)–(11) imply

$$H^*(p) = H(p).$$

Therefore the theorem is proved. □

## 4. The case of multiple change points

**Theorem 2.** *Let $F_1$ and $F_2$ be continuous in the interval* $[\operatorname{med} F_1, \operatorname{med} F_2]$ *and*

$$\pi_N / \ln N \to \infty, \qquad N \to \infty.$$

*Then*

$$\lim_{N \to \infty} \mathsf{P}\left(\hat{k}_{j,N} - k_j = p\right) = \bar{H}(p)$$

*for $1 \le j \le R$ where*

$$\bar{H}(p) := \begin{cases} H_{u_1 v_2}(p), & h_{k_j}^0 = 1, \\ H_{u_2 v_1}(-p), & h_{k_j}^0 = 2. \end{cases}$$

*Remark.* We do not assume that the number of changes in the sequence is known. When evaluating the moments of change, we estimate the number of changes. Note that the number of changes is equal to $R$ after a certain random moment $n(\omega) < \infty$.

*Proof.* Put

$$\mathcal{H}(d_N) = \{h = (h_1, \dots, h_N)\colon R(h) = R, k_j(h) \in (k_j - d_N, k_j + d_N], j \in [1, R]\}.$$

Using the trajectory

$$\check{h} = \operatorname*{argmin}_{h \in \mathcal{H}(d_N)} J(h)$$

we construct auxiliary estimators $\check{k}_j = k_j(\check{h})$. Consider the events

$$C_N(d_N) = \left\{\hat{R} = R, \hat{h}_{\hat{k}_j} = h^0_{k_j}, \hat{k}_j \in [k_j - d_N + 1, k_j + d_N]\right\}.$$

It can be proved that if $F_1$ and $F_2$ are continuous at every point between their medians and

$$\frac{\pi_N}{\ln N} \to \infty, \qquad \frac{d_n}{\pi_N} \to \infty, \quad \text{and} \quad \frac{d_N}{N} \to 0 \tag{12}$$

as $N \to \infty$, then the event $C_N$ occurs almost surely starting with some random $N < \infty$. A similar assertion is proved in [1] for estimators generated by functions $\phi(\zeta_j, h)$ of a general form. The difference between the case of this paper and the case of the paper [1] is that in [1] the random variables $\phi(\zeta_j, h)$ are assumed to be independent for different $j$. Below we prove the analog of the lemma in [1] that does not use the assumption on the independence (other parts of the proof of the lemma in [1] can be adopted to our case with minor changes).

**Lemma 2.** *Let $a_n > 0$ and $a_n / \ln n \to \infty$ as $n \to \infty$. Then there exists $N(\omega)$ such that the events*

$$A_n = \left\{\max_{1 \le l_1 \le l_2 \le n,\ g=1,2} \sum_{j=l_1}^{l_2} \left(\phi(\xi_j, h^0_j) - \phi(\xi_j, g)\right) \le a_n\right\}$$

*occur for all $n > N$.*

*Proof.* Since

$$\phi\left(\xi_j, h^0_j\right) - \phi(\xi_j, g) = 0$$

for $h^0_j = g$ and $\sum_{j=l_1}^{l_2} \left(\phi(\xi_j, h^0_j) - \phi(\xi_j, g)\right)$ can be represented as a sum of no more than $R$ terms whose indices belong to the intervals of homogeneity of $h^0_j$, we restrict the proof of the inequality to the case of $h^0_j = h \ne g$.

In what follows we need the *Vapnik–Chervonenkis inequality* (see [8])

$$\mathsf{P}\left(\sup_y \left|\frac{1}{n} \sum_{j=1}^{n} \mathbb{1}_{\{\zeta_j < y\}} - G(y)\right| > \varepsilon\right) \le 6(2n+1) \exp\left\{-\frac{\varepsilon^2 (n-1)}{2}\right\}$$

where the $\zeta_j$ are independent identically distributed random variables with the distribution $G$. A generalization of the Vapnik–Chervonenkis inequality holds for sequences of random variables whose distributions belong to a family of two distributions

$$\mathsf{P}\left(\sup_y \left|\sum_{j=l_1}^{l_2} \left(\mathbb{1}_{\{\zeta_j < y\}} - \mathsf{P}(\zeta_j < y)\right)\right| > x\right) \le 12(2n+1) \exp\left\{-\frac{x^2}{16(l_2 - l_1 + 1)}\right\}, \quad n \ge 2,$$

where $[l_1, l_2] \subset [1, N]$.

Put $\Phi_j(1) = \mathsf{P}(\zeta_j > x)\big|_{x=\hat{m}}$, $\Phi_j(2) = \mathsf{P}(\zeta_j < x)\big|_{x=\hat{m}}$, and $\zeta_{jg} = \phi(\zeta_j, g) - \Phi_j(g)$. If $h \neq g$, then

$$\zeta_{jh} = -\zeta_{jg},$$

whence

$$\mathsf{P}\left(\sum_{j=l_1}^{l_2} (\zeta_{jh} - \zeta_{jg}) > x\right) = \mathsf{P}\left(\sum_{j=l_1}^{l_2} \zeta_{jh} > \frac{x}{2}\right) \leq 12(2n+1)\exp\left\{-\frac{x^2}{64(l_2-l_1+1)}\right\}.$$

Assume that $|\hat{m} - \bar{m}| < \delta$. Since $F_1$ and $F_2$ are continuous at each point between their medians, there exists $\delta > 0$ such that if $|\hat{m} - \bar{m}| < \delta$, then $\Phi_j(g) - \Phi_j(h)$ is greater than some $\varkappa > 0$ and

$$\phi(\xi_j, h) - \phi(\xi_j, g) = \zeta_{jh} - \zeta_{jg} + \Phi_j(h) - \Phi_j(g) \leq \zeta_{jh} - \zeta_{jg} - \varkappa.$$

Then we set $x = a_n + \varkappa(l_2 - l_1 + 1)$ and get

$$\begin{aligned} p(l_1, l_2, n) &= \mathsf{P}\left(\sum_{j=l_1}^{l_2} (\phi(\xi_j, g) - \Phi_j(g)) \geq a_n, |\hat{m} - m| < \delta\right) \\ &\leq 12(2n+1)\exp\left\{-\frac{\left(a_n + \varkappa(l_2 - l_1+1)^2\right)}{64(l_2-l_1+1)}\right\}, \end{aligned}$$

since

$$\exp\left\{-(a_n + \varkappa y)^2/64y\right\} \leq \exp\left\{-a_n\varkappa/16\right\}$$

for all $y > 0$. Now we estimate $p(l_1, l_2, N)$ and proceed in the same way as in the proof in [1], namely we apply the Borel–Cantelli lemma, the estimate

$$\mathsf{P}\left(\overline{A_N}, |\hat{m} - \bar{m}| < \delta\right) \leq \sum_{1 \leq l_1 \leq l_2 \leq N} p(l_1, l_2, N),$$

and the convergence in probability of $\hat{m}$ to $\bar{m}$. The lemma is proved. □

Thus $\mathsf{P}(C_N) \to 1$ and $\mathsf{P}(\overline{C_N}) \to 0$. Note that $\{\check{k}_i \neq \hat{k}_i\} \subset \overline{C_N}$, that is, we need to determine the asymptotic distribution of $\check{k}_{j,N} - k_j$ (it is the desired distribution).

Put $\check{k}_N = (\check{k}_{1,N}, \dots, \check{k}_{R,N})$, $\bar{k} = (k_1, \dots, k_R)$, and $\bar{p} = (p_1, \dots, p_R)$. Note that

$$\check{h} = \operatorname*{argmin}_{h \in \mathcal{H}(d_N)} J(h) = \operatorname*{argmin}_{h \in \mathcal{H}(d_N)} \left(\sum_{j=1}^N \phi(\zeta_j, h_j) + \pi_N R\right) = \operatorname*{argmin}_{h \in \mathcal{H}(d_N)} \sum_{j=1}^N \phi(\zeta_j, h_j).$$

If $N$ is sufficiently large, then the intervals $(k_i - d_n, k_i + d_N]$ are disjoint and the sum in the latter relation splits into $k+1$ terms:

$$\sum_{j=1}^N \phi(\zeta_j, h_j) = \sum_{i=1}^R \sum_{j=k_i-d_N+1}^{k_i+d_N} \phi(\zeta_j, h_j) + \sum_{j \notin \bigcup(k_i - d_N, k_i + d_N]} \phi(\zeta_j, h_j).$$

The last term does not depend on $h$. Thus it does not change the argument of the minimum and we omit it:

$$\check{h} = \operatorname*{argmin}_{h \in \mathcal{H}(d_N)} \sum_{i=1}^R \sum_{j=k_i-d_N+1}^{k_i+d_N} \phi(\zeta_j, h_j). \tag{13}$$

Other terms depend on the trajectory on their own intervals only, so that the minimal trajectory can be determined step by step. The parts of the trajectory corresponding to the intervals $[k_i - d_N + 1, k_i + d_N]$ are denoted by $h^i = (h_{k_i - d_N + 1}, \dots, h_{k_i + d_N})$. Note that every part $h^i$ of the trajectory contains only one change point. The part of the trajectory for which the change occurs at the position $l$ is denoted by $h^i(l)$. The symbol $\mathcal{H}^i$ denotes the set of all such parts. Thus

$$\mathsf{P}(\check{k}_N - \bar{k} = \bar{p}) = \mathsf{P}\left( h^i(d_N + p_i) = \operatorname*{argmin}_{h^i \in \mathcal{H}^i} \sum_{j=k_i-d_N+1}^{k_i+d_N} \phi(\zeta_j, h_j),\ i = 1, \dots, R \right). \tag{14}$$

The latter result means that the distribution of every change point can be found separately, that is,

$$\mathsf{P}\left(\check{k}_{j,N} - k_i = p_j\right) = \mathsf{P}\left( h^j(d_N + p_j) = \operatorname*{argmin}_{h^j \in \mathcal{H}^j} \sum_{l=k_j-d_N+1}^{k_j+d_N} \phi(\zeta_l, h_l) \right). \tag{15}$$

There are two possible cases:

1) If $h^0_{k_j} = 1$, then

$$\operatorname*{argmin}_{h^j \in \mathcal{H}^j} \sum_{l=k_j-d_N+1}^{k_j+d_N} \phi(\zeta_l, h_l) = \operatorname*{argmax}_{h^j \in \mathcal{H}^j} \sum_{l=k_j-d_N+1}^{k_j+p_j} r(\zeta_j, \hat{m}).$$

The latter relation coincides with (5); thus we apply Theorem 1 and conclude that the distribution of $\hat{k}_{j,N} - k_j$ equals $H_{u_1 v_2}(p)$.

2) If $h^0_{k_j} = 2$, then

$$\begin{aligned} \operatorname*{argmin}_{h^j \in \mathcal{H}^j} \sum_{l=k_j-d_N+1}^{k_j+d_N} \phi(\zeta_l, h_l) &= \operatorname*{argmin}_{h^j \in \mathcal{H}^j} \sum_{l=k_j-d_N+1}^{k_j+p_j} r(\zeta_j, \hat{m}) \\ &= \operatorname*{argmax}_{h^j \in \mathcal{H}^j} \sum_{l=k_j-d_N+1}^{k_j+p_j} r(-\zeta_j, -\hat{m}). \end{aligned}$$

Applying Theorem 1 to $-\zeta_j$ we prove that the distribution of $\hat{k}_{j,N} - k_j$ equals $H_{u_2 v_1}(-p)$. Therefore the theorem is proved. □

## Bibliography


1. R. E. Maĭboroda and O. V. Sugakova, *A fast algorithm for detecting multiple change points*, Teor. Ĭmovīr. Mat. Stat. **57** (1997), 103–108; English transl. in Theory Prob. Math. Statist. **57** (1998), 109–114. MR1806888 (2003b:62100)
2. A. A. Borovkov, *Asymptotically optimal solutions in the change-point problem*, Teor. Veroyatnost. Primenen. **43** (1998), no. 4, 625–654; English transl. in Theory Probab. Appl. **43** (1999), no. 4, 539–561. MR1692429 (2001g:62044)
3. R. E. Maĭboroda and O. V. Sugakova, *The limit distribution of DP estimators of multiple change points*. (to appear)
4. O. V. Sugakova, *A search for change points in a flow of independent observations*, Teor. Ĭmovīr. Mat. Stat. **55** (1996), 167–172; English transl. in Theory Probab. Math. Statist. **55** (1997), 181–186. MR1641589 (99g:62066)
5. Marc Raimondo, *Minimax estimation of sharp change points*, Ann. Stat. **26** (1998), no. 4, 1379–1397. MR1647673 (99i:62076)
6. V. V. Mottl', I. B. Muchnik, and V. G. Yakovlev, *Optimal segmenting of experimental curves*, Avtomat. i Telemekh. **8** (1983), 84–95. (Russian)

7. R. E. Maĭboroda, *The median estimator of disorder in the case of weakly dependent observations*, Teor. Veroyatnost. i Mat. Statist. **43** (1990), 82–87; English transl. in Theory Probab. Math. Statist. **43** (1991), 87–91. MR1069318 (91m:62079)
8. V. N. Vapnik and A. Ya. Chervonenkis *Theory of Pattern Recognition*, "Nauka", Moscow, 1974. (Russian) MR0474638 (57:14274)

Department of Probability Theory and Mathematical Statistics, Mechanics and Mathematics Faculty, National Taras Shevchenko University, Academician Glushkov Avenue 6, Kyiv 03127, Ukraine

*E-mail address*: skorohod@i.com.ua

Received 14/MAR/2003

Translated by V. SEMENOV