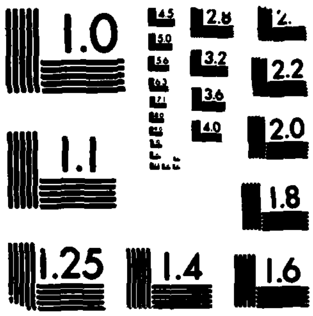
AD-A190 328 DETECTION OF THE NUMBER LOCATIONS AND MAGNITUDES OF JUMPS(U) PITTSBURGH UNIV PA CENTER FOR MULTIVARIATE ANALYSIS Y Q YIN AUG 87 TR-87-27 AFOSR-TR-87-1961 1/1

UNCLASSIFIED F49620-85-C-0008 F/G 12/3 NL

END
DATE
FILMED
4 8



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

AD-A190 328

(2)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 87-1961	2. GOVT ACCESSION NO	3. RECIPIENT'S CATALOG NUMBER DTIC FILE COPY
4. TITLE (and Subtitle) Detection of the number, locations and magnitudes of jumps		5. TYPE OF REPORT & PERIOD COVERED ~~technical~~ journal - August 1987
		6. PERFORMING ORG. REPORT NUMBER 87-27
7. AUTHOR(s) Y.Q. Yin		8. CONTRACT OR GRANT NUMBER(s) F49620-85-C-0008
9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Multivariate Analysis, University of Pittsburgh, Fifth Floor Thackeray, Pittsburgh, PA 15260		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304/A5
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research, Department of the Air Force, Bolling Air Force Base, DC 20332 NM		12. REPORT DATE August 1987
		13. NUMBER OF PAGES 12
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Same as 11		15. SECURITY CLASS. (of this report) unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

DTIC ELECTE JAN 1 5 1988 S D H

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Change points, consistency, random signals

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Consider a signal $x(t) = f(t) + w(t)$, $0 \le t \le 1$. Here the noise $w(t)$ is an independent process, and $f(t)$ is a function with only finitely many jumps, satisfies a Lipschitz' condition between any two consecutive jumps. This paper gives an algorithm to determine the number, locations and magnitudes of the jumps of $f(t)$. The consistency and speeds of convergence are obtained. (Keywords; stochastic processes; distributions; estimation; convergence).

DD FORM 1 JAN 73 1473

unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

87 12

AFOSR-TR- 87-1961

DETECTION OF THE NUMBER, LOCATIONS AND MAGNITUDES OF JUMPS*

Y.Q. Yin

Center for Multivariate Analysis
University of Pittsburgh

Center for Multivariate Analysis

University of Pittsburgh

DETECTION OF THE NUMBER, LOCATIONS
AND MAGNITUDES OF JUMPS*

Y.Q. Yin

Center for Multivariate Analysis
University of Pittsburgh

August 1987

Technical Report 87-27

Center for Multivariate Analysis
University of Pittsburgh
515 Thackeray Hall
Pittsburgh, PA 15260

*Research sponsored by the Air Force Office of Scientific Research (AFSC) under contract #F49620-85-C-0008. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

DETECTION OF THE NUMBER, LOCATIONS AND MAGNITUDES OF JUMPS

Y. Q. YIN

§1. Introduction. Jump detection or change-point detection is a very important problem in statistics and engineering. The problem can be stated in the following way:

Let $x(t) = f(t) + w(t)$ be a stochastic process, $0 \leq t \leq 1$. Here $w(t)$ is the noise process, $Ew(t) = 0$, and $f(t) = Ex(t)$ is a function with only finitely many discontinuities $t_1, \cdots t_q$. Suppose these discontinuities are all interior points of $[0,1]$ and for each $i = 1, \cdots q$, $f(t_i + 0)$, $f(t_i - 0)$ exist and $f(t_i + 0) \neq f(t_i - 0)$. For definiteness, we suppose f is left continuous everywhere. Our problem is to estimate

1. the number q of discontinuities,
2. the positions $t_1, \cdots, t_q$ of these discontinuities;
3. the magnitudes $f(t_i + 0) - f(t_i - 0)$ of the jumps, $i = 1, \cdots, q$;

based on a sample $x(\frac{k}{n})$, $k = 0, 1, \cdots n$.

In this paper, under mild conditions, we will give an algorithm to estimate q, i. e. , define an estimator $\hat{q}$ of q. For each $i = 1, \cdots q$, we give an algorithm to estimate the position of the ith discontinuity point t_i, i. e. , define an estimator $\hat{t}_i$ of t_i; we also define an estimator $D_{n\hat{t}_i}$ for the jump at the ith discontinuity point.

We will prove that all these estimators are strongly consistent, or in other words, these estimators converge to the corresponding parameters as n, the number of sample points, tends to infinity, with probability one.

We also get the speed of convergence, for example for $\hat{t}_i$ we get

$$|\hat{t}_i - t_i| = O\left(\frac{(\ln n)^{1+\alpha}}{n}\right)$$

for any $\alpha > 0$.

The complexity of computation of our algorithms is $O(n \log n)$ approximately.

The basic hypotheses are two:

1. $w(t)$ is a gaussian white noise.
2. There exists a positive constant $K > 0$, such that

$$|f(t) - f(s)| \leq K|t-s|, \text{ if no } t_i \text{ are in } [s,t].$$

The condition 1. can be relaxed to nongaussian white noises, but it would be difficult to relax 2.

In Section 2, we define the algorithms. In Section 3 we give the proofs.

Works on this topic mostly concentrate on the single jump problem, see the references listed at the rear of this paper. Especially no work has been done on the case when the number of jumps is unknown.

Acknowledgment. The author sincerely appreciates the encouragement and financial support of Dr. Krishnaiah, director of the Center for Multivariate Analysis, University of Pittsburgh.

§2. The Method for Detecting the Change Points. Let $x(t) = f(t) + w(t)$ be a stochastic process, $0 \leq t \leq 1$. Here $w(t)$ is a Gaussian white noise. $f(t) = Ex(t)$ has only finitely many discontinuities $t_1, \cdots, t_q$. All these discontinuities are in $(0,1)$ and of the first kind, i. e., $f(t_i + 0)$, $f(t_i - 0)$ exist for $i = 1, \ldots, q$.

Besides, suppose that there is a constant $K > 0$ such that $|f(s) - f(t)| \leq K|s-t$ for any interval $[s,t] \subset [0,1]$ not containing any discontinuities. Let

$$d_i = f(t_i + 0) - f(t_i - 0) \neq 0, \qquad i = 1, \cdots, q,$$

and suppose

$$|d_1| \geq |d_2| \geq \cdots \geq |d_q|, \text{ and if } |d_i| = |d_j|, \text{ and } i < j, \text{ then } t_i < t_j.$$

Let $m = m(n) \uparrow \infty$, $m/n \to 0$. Define

$$D_{nk} = \frac{x(\frac{k+1}{n}) + \cdots + x(\frac{k+m}{n})}{m} - \frac{x(\frac{k-1}{n}) + \cdots + x(\frac{k-m}{n})}{m},$$

for $m \leq k \leq n - m$.

Let h_n be a sequence of positive numbers, $h_n \to 0$, $\dfrac{m_n}{nh_n} \to 0$, for definiteness, let $h_n = (\ln_3 n / \ln_2 n)^{1/3}$, $m_n = \ln n (\ln_2 n)^{2/3} (\ln_3 n)^{1/3}$ where $\ln_1 n = 1 \wedge \ln n$, $\ln_k n = \ln_1(\ln_{k-1} n)$.

Define

$$I_1 = \arg\max_k \left\{ |D_{nk}| - \frac{k}{n} h_n \right\},$$

$$I_2 = \arg\max_k \left\{ |D_{nk}| - \frac{k}{n} h_n : |k - I_1| > 4m \right\},$$

$$I_3 = \arg\max_k \left\{ |D_{nk}| - \frac{k}{n} h_n : |k - I_1| > 4m \text{ and } |k - I_2| > 4m \right\}, \cdots.$$

If the definition is not unique, we choose the smallest one. At first, we state the main theorem, which will be proved in Section 3.

Theorem 3.1. (1) $\left|\frac{I_i}{n} - t_i\right| \leq \frac{2m}{n}$ for all large n, a. s, $i = 1, \cdots, q$.

(2) $D_{nI_i} \to d_i$, a. s., $i = 1, \cdots, q$.

(3) $D_{nI_i} = O(h_n)$, a. s. for $i > q$.

We see from this theorem that the quantities defined above are strongly consistent estimators of the change points $t_1, \cdots, t_q$ and the magnitudes d_i of changements. Furthermore we get the convergence rate

$$\left|\frac{I_i}{n} - t_i\right| \leq \frac{2m}{n}, \quad \text{for large } n, \text{ a. s.}$$

from (1).

Theorem 3.1 does not supply a method to estimate the integer q, explicitly. But based on Theorem 3.1, we can construct a strongly consistent estimator of q in the following manner.

Let

$$(4) \qquad G_{nk} = \frac{1}{2^{k+1}}|D_{n,I_{k+1}}| + \frac{1}{2^{k+2}}|D_{n,I_{k+2}}| + \cdots + kc_n.$$

Here $c_n > 0$ with the properties $c_n \to 0$ and $h_n/c_n \to 0$. Let

$$\hat{q}_n = \arg\min_k G_{nk}.$$

Suppose we have proved Theorem 3.1. We are going to prove

Theorem 2.1. $\hat{q}_n \to q$, a. s.

Proof. If $k < q$, then, almost surely, as $n \to \infty$

$$\begin{aligned} G_{nk} - G_{nq} \geq \frac{1}{2^{k+1}}|D_{nI_{k+1}}| + \cdots + \frac{1}{2^q}|D_{nI_q}| - \frac{1}{2^k}h_n \\ + (k-q)c_n \to \frac{1}{2^{k+1}}|d_{k+1}| + \cdots + \frac{1}{2^q}|d_q| > 0, \end{aligned}$$

by (2) of Theorem 3.1. That means $\hat{q}_n \neq k$ for large n.

If $k > q$, then, noticing $|D_{nI_i}| \downarrow$, as $i \uparrow$, by (3) of Theorem 3.1, we have for a constant $C > 0$,

$$\begin{aligned} G_{nk} - G_{nq} &= -\frac{1}{2^{q+1}}|D_{nI_{q+1}}| - \cdots - \frac{1}{2^k}|D_{nI_k}| + (k-q)c_n \\ &\geq -\frac{1}{2^q}|D_{nI_{q+1}}| + (k-q)c_n \\ &\geq -\frac{C}{2^q}h_n + (k-q)c_n = c_n\left((k-q) - \frac{Ch_n}{2^{q+1}c_n}\right) \\ &\geq c_n\frac{1}{2}(k-q) > 0 \end{aligned}$$

for n sufficiently large, almost surely. In this case $\hat{q}_n \neq k$ either. So, almost surely, for n sufficiently large, $\hat{q}_n = q$.

§3. Proof of the Main Theorem. At first we prove an elementary lemma.

Lemma. Suppose $f(t)$ is defined in the interval $(t_0 - a, t_0 + a)$ for a positive number a. Suppose $f(t_0 \pm 0)$ exist and are finite. Let I_n and m_n be two sequences of positive integers such that $I_n/n \to t_0$, $m_n/n \to 0$, $m_n \to \infty$. Let

$$A_n = -\frac{1}{m}\left(f(\frac{I-m}{n}) + \cdots + f(\frac{I-1}{n}) - f(\frac{I+1}{n}) - \cdots - f(\frac{I+m}{n})\right),$$

where $I = I_n$, $m = m_n$. Then, from $\liminf_{n\to\infty} |A_n| \geq |f(t_0 + 0) - f(t_0 - 0)|$, we can deduce that $A_n \to f(t_0 + 0) - f(t_0 - 0)$.

Proof. Without loss of generality, suppose $\frac{I}{n} < t_0$ for all n. Let

$$k_n = max\left\{k : \frac{I+k}{n} < t_0, 0 \leq k \leq m\right\}.$$

Fix $\epsilon > 0$, $\exists \delta > 0$ such that $t_0 - \delta < t < t_0 \implies |f(t) - f(t_0 - 0)| < \epsilon$, and $t_0 < t < t_0 + \delta \implies |f(t) - f(t_0 + 0)| < \epsilon$.

Let N be such that as $n > N$, $\left|\frac{I}{n} - t_0\right| < \frac{\delta}{2}$, $\frac{m}{n} < \frac{\delta}{2}$. So, as $n > N$

$$\begin{aligned} A_n &< -f(t_0 - 0) + \epsilon + (f(t_0 - 0) + \epsilon)\frac{k_n}{m} + \left(1 - \frac{k_n}{m}\right)(f(t_0 + 0) + \epsilon) \\ &= \left(1 - \frac{k_n}{m}\right)(f(t_0 + 0) - f(t_0 - 0)) + 2\epsilon. \end{aligned}$$

In the same way

$$A_n > \left(1 - \frac{k_n}{m}\right)(f(t_0 + 0) - f(t_0 - 0)) - 2\epsilon.$$

Thus,

$$|A_n| \leq \left|1 - \frac{k_m}{m}\right| |f(t_0 + 0) - f(t_0 - 0)| + 2\epsilon.$$

Therefore we must have $\frac{k_n}{m} \to 0$, thus $A_n \to f(t_0+0) - f(t_0-0)$.

Theorem 3.1. (1) $|I_k/n - t_k| \le \frac{2m}{n}$, for all large n, a. s., for $k = 1, \cdots, q$;

(2) $D_{nI_k} \to d_k$, a. s., for $k = 1, \cdots, q$;

(3) $D_{nI_k} = O(h_n)$, a. s., for $k > q$.

Proof. 1. In this part, we prove $\left|\frac{I_1}{n} - t_1\right| \le \frac{2m}{n}$ for all large n, a. s. Introduce the following notations

$\Delta_k = \frac{1}{m}\left\{f\left(\frac{k+1}{n}\right) + \cdots + f\left(\frac{k+m}{n}\right)\right\} - \frac{1}{m}\left\{f\left(\frac{k-1}{n}\right) + \cdots + f\left(\frac{k-m}{n}\right)\right\}$,

$W_k^+ = \frac{1}{m}\left\{w\left(\frac{k+1}{n}\right) + \cdots + w\left(\frac{k+m}{n}\right)\right\}$,

$W_k^- = \frac{1}{m}\left\{w\left(\frac{k-1}{n}\right) + \cdots + w\left(\frac{k-m}{n}\right)\right\}$.

Let $\hat{k}$ be such that $|\hat{k}/n - t_1| = \min_k |k/n - t_1|$, $\hat{k}$ depends only on n.

At first, we note that for $|k/n - t_1| > 2m/n$,

$$\begin{aligned}(4)\quad & P\left(|D_{n\hat{k}}| - \tfrac{\hat{k}}{n}h_n \le |D_{nk}| - \tfrac{k}{n}h_n\right) \\ & \le P\left(|\Delta_{\hat{k}}| - |\Delta_k| + \left(\tfrac{k-\hat{k}}{n}\right)h_n \le |W_k^+| + |W_k^-| + |W_{\hat{k}}^+| + |W_{\hat{k}}^-|\right).\end{aligned}$$

Because $\left|\frac{k}{n} - t_1\right| > \frac{2m}{n}$, the points $\frac{k-m}{n}, \cdots, \frac{k-1}{n}, \frac{k+1}{n}, \cdots, \frac{k+m}{n}$ are all on the same side of t_1. If n is larger than some nonrandom number $N_1 > 0$, for any k with $\left|t_1 - \frac{k}{n}\right| > \frac{2m}{n}$, the interval $\left[\frac{k-m}{n}, \frac{k+m}{n}\right]$ can contain at most one discontinuous point t_i. There are three possibilities:

(a) No t_i in $\left[\frac{k-m}{n}, \frac{k+m}{n}\right]$,

(b) $t_i \in \left[\frac{k-m}{n}, \frac{k+m}{n}\right]$, $|d_i| = |d_1|$, then $t_i > t_1$, and $k > \hat{k}$,

(c) $t_i \in \left[\frac{k-m}{n}, \frac{k+m}{n}\right]$, $|d_i| < |d_1|$.

For case (a) above,

$$|\Delta_{\hat{k}}| \ge |d_1| - 2K\frac{m}{n}, \quad |\Delta_k| \le K\frac{m+1}{n}$$

so,

$$|\Delta_{\hat{k}}| - |\Delta_k| + \frac{k-\hat{k}}{n}h_n \ge |d_1| - 3K\frac{m+1}{n} + \frac{k-\hat{k}}{n}h_n > \frac{1}{2}|d_1|,$$

for $n \ge N_2 > 0$, N_2 is a nonrandom constant.

For case (b),

$$|\Delta_{\hat{k}}| \geq |d_1| - 2K\frac{m}{n}, \quad |\Delta_k| \leq |d_1| + 2K\frac{m}{n},$$

so, for a constant $c > 0$,

$$|\Delta_{\hat{k}}| - |\Delta_k| + \frac{(k-\hat{k})}{n}h_n \geq -4K\frac{m}{n}$$
$$+ (t_i - t_1)h_n + o(h_n) \geq ch_n$$

when $n \geq N_3 > 0$, N_3 is nonrandom, c is a positive constant.

For case (c),

$$|\Delta_{\hat{k}}| - |\Delta_k| + \frac{(k-\hat{k})}{n}h_n \geq |d_1| - |d_i| + o(1)$$
$$\geq \frac{1}{2}(|d_1| - |d_i|) > 0$$

when $n \geq N_4 > 0$, N_4 is nonrandom.

Thus, for $n \geq max(N_1, N_2, N_3, N_4)$, and for some constants $c_1 > 0$, $c_2 > 0$,

$$P\left(|D_{n\hat{k}}| - \frac{\hat{k}}{n}h_n \leq |D_{nk}| - \frac{k}{n}h_n\right)$$
$$\leq 4P\left(b_1 h_n \leq \left|\frac{v_1 + \cdots + v_m}{m}\right|\right)$$
$$\leq 4P\left(\frac{b_1}{\sigma}h_n\sqrt{m} \leq |z|\right)$$
$$\leq 8\frac{1}{\sqrt{2\pi}\frac{b_1}{\sigma}h_n\sqrt{m}}e^{-\frac{1}{2}\frac{b_1^2}{\sigma^2}h_n^2 m}$$
$$\leq b_2 e^{-b_3 \ln n \ln_3 n} = b_2 n^{-b_3 \ln_3 n}$$

if we choose $h_n = \left(\frac{\ln_3 n}{\ln_2 n}\right)^{1/3}$, $m = \ln n(\ln_2 n)^{2/3}(\ln_3 n)^{1/3}$. Here b_1, b_2, b_3 are positive constants, $\ln_2 x = \ln\ln x$, $\ln_3 x = \ln\ln x$. Thus, the series

$$\sum_n P\left(|D_{n\hat{k}}| - \frac{\hat{k}}{n}h_n \leq |D_n k| - \frac{k}{n}h_n, \text{for some } k \text{ with } \left|\frac{k}{n} - t_1\right| > \frac{2m}{n}\right)$$
$$\leq \sum_n nP\left(|D_{n\hat{k}}| - \frac{\hat{k}}{n}h_n \leq |D_{nk}| - \frac{k}{n}h_n\right) < \infty.$$

By Borel-Cantelli Lemma,

$$P\Big(\exists N, s.t. n \geq N, |\frac{k}{n} - t_1| > \frac{2m}{n}$$
$$\Longrightarrow |D_{n\hat{k}}| - \frac{\hat{k}}{n}h_n > |D_{nk}| - \frac{k}{n}h_n\Big) = 1$$

or,

$$P\Big(\exists, N, s.t. n \geq N, |D_{n\hat{k}}| - \frac{\hat{k}}{n}h_n \leq |D_{nk}| - \frac{k}{n}h_n$$
$$\Longrightarrow |\frac{k}{n} - t_1| \leq \frac{2m}{n}\Big) = 1.$$

Since $|D_{nI_1}| - \frac{I_1}{n}h_n \geq |D_{n\hat{k}}| - \frac{\hat{k}}{n}h_n$,

$$P\left(\exists N, s.t. n \geq N \Longrightarrow \left|\frac{I_1}{n} - t_1\right| \leq \frac{2m}{n}\right) = 1.$$

This proves case $k = 1$ for (1).

By

$$|D_{nI_1}| - \frac{I_1}{n}h_n \geq |D_{n\hat{k}}| - \frac{\hat{k}}{n}h_n$$

we have

$$\liminf |D_{nI_1}| \geq |d_1|.$$

So, by the elementary lemma, we get $D_{nI_1} \to d_1$ a. s.

2. In this part we prove $|\frac{I_i}{n} - t_i| \leq \frac{2m}{n}$ for all large n, a. s.. and $D_{nI_i} \to d_i$ a. s., if $2 \leq i \leq q$. But we carry out the proof only for the case $i = 2$.

Let $\hat{k}$ be such that $|\frac{\hat{k}}{n} - t_2| = \min_k |\frac{k}{n} - t_2|$, and suppose k is such that $|\frac{k}{n} - t_1| > \frac{2m}{n}$ and $|\frac{k}{n} - t_2| > \frac{2m}{n}$. Suppose for these $\hat{k}$ and k, $|D_{n\hat{k}}| - \frac{\hat{k}}{n}h_n \leq |D_{nk}| - \frac{k}{n}h_n$, so that

$$|\Delta_{\hat{k}}| - |\Delta_k| + \frac{k - \hat{k}}{n}h_n \leq |W_k^+| + |W_k^-| + |W_{\hat{k}}^+| + |W_{\hat{k}}^-|.$$

Since $|\frac{k}{n} - t_1| > \frac{2m}{n}$, $|\frac{k}{n} - t_2| > \frac{2m}{n}$, t_1 and t_2 are not in the interval $[\frac{k-m}{n}, \frac{k+m}{n}]$. $\exists N_1 > 0$ nonrandom, such that as $n \geq N_1$, $[\frac{k-m}{n}, \frac{k+m}{n}]$ contains at most one discontinuity point. There are three possibilities:

(a) No t_i in $[\frac{k-m}{n}, \frac{k+m}{n}]$,

(b) $t_i \in [\frac{k-m}{n}, \frac{k+m}{n}]$, $|d_i| = |d_2|$, of course $i > 2$, $t_i > t_2$.

(c) $t_i \in [\frac{k-m}{n}, \frac{k+m}{n}]$, $|d_i| < |d_2|$.

For (a), we have

$$|\Delta_{\hat{k}}| \geq |d_2| - K\frac{2m}{n}, \quad |\Delta_k| \leq K\frac{2m}{n},$$

and then as $n \geq N_2$, N_2 nonrandom,

$$|\Delta_{\hat{k}}| - |\Delta_k| + \frac{k-\hat{k}}{n}h_n \geq |d_2| - \frac{4Km}{n} - h_n \geq \frac{1}{2}|d_2|.$$

For (b),

$$|\Delta_{\hat{k}}| \geq |d_2| - K\frac{2m}{n}, \quad |\Delta_k| \leq |d_2| + K\frac{2m}{n},$$

and as $n \geq N_3$, N_3 nonrandom, we have

$$\begin{aligned} |\Delta_{\hat{k}}| - |\Delta_k| + \frac{k-\hat{k}}{n}h_n &\geq -K\frac{4m}{n} + \frac{k-\hat{k}}{n}h_n \\ &\geq \left(-K\frac{4m}{nh_n} + t_i - t_2 + o(1)\right)h_n \geq bh_n, \end{aligned}$$

for constant $b > 0$.

For case (c), just in the same way, we can show that as $n \geq N_4$, N_4 nonrandom,

$$|\Delta_{\hat{k}}| - |\Delta_k| + \frac{k-\hat{k}}{n}h_n \geq \frac{1}{2}(|d_2| - |d_i|) > 0.$$

therefore as $n \geq max(N_1, N_2, N_3, N_4)$,

$$\begin{aligned} P\left(|D_{n\hat{k}}| - \frac{\hat{k}}{n}h_n \leq |D_{nk}| - \frac{k}{n}h_n\right) & \\ &\leq 4P\left(b_1 h_n \leq \left|\frac{v_1 + \cdots + v_m}{m}\right|\right) \\ &\leq 4P(b_2 h_n\sqrt{m} \leq |z|) \leq b_2 n^{-b_3 \ln_2 n} \end{aligned}$$

for positive constants b_1 b_2, b_3. Here $v_1, \cdots, v_m$, z have the same meanings as before.

Therefore the series

$$\sum_n P\Big(|D_{n\hat{k}}| - \frac{\hat{k}}{n}h_n \le |D_{nk}| - \frac{k}{n}h_n, \quad \text{for some } k \text{ with } |\frac{k}{n} - t_1| > \frac{2m}{n} \text{ and } |\frac{k}{n} - t_2| > \frac{2m}{n}\Big) < \infty.$$

Thus

$$P\Big(\exists N : n \ge N \implies |D_{n\hat{k}}| - \frac{\hat{k}}{n}h_n > |D_{nk}| - \frac{k}{n}h_n \quad \text{for all } k \text{ with } |\frac{k}{n} - t_1| > \frac{2m}{n} \text{ and } |\frac{k}{n} - t_2| > \frac{2m}{n}\Big) = 1.$$

Suppose the event in the last $P(\quad)$ is true, and $|\frac{I_1}{n} - t_1| \le \frac{2m}{n}$ is true for large n. Then, $|\hat{k} - I_1| = n|\frac{\hat{k}}{n} - \frac{I_1}{n}| \ge n(|t_2 - t_1| - \frac{1}{2n} - \frac{2m}{n}) > 4m$, if n is large. So, $|D_{nI_2}| - \frac{I_2}{n}h_n \ge |D_{n\hat{k}}| - \frac{\hat{k}}{n}h_n$, and $|\frac{I_2}{n} - t_1| \le \frac{2m}{n}$ or $|\frac{I_2}{n} - t_2| \le \frac{2m}{n}$, for large n a. s. But $|\frac{I_2}{n} - t_1| \ge |\frac{I_2 - I_1}{n}| - |\frac{I_1}{n} - t_1| > \frac{4m}{n} - \frac{2m}{n} = \frac{2m}{n}$, so $|\frac{I_2}{n} - t_2| \le \frac{2m}{n}$ for large n a. s.

From $|D_{nI_2}| - \frac{I_2}{n}h_n \ge |D_{n\hat{k}}| - \frac{\hat{k}}{n}h_n$, we have $\liminf_{n\to\infty} |D_{nI_2}| \ge |d_2|$. By the lemma, we must have $D_{nI_2} \to d_2$, a. s.

3. In this part we prove that $D_{nI_{q+1}} = O(h_n)$, a. s. By definition, $|I_{q+1} - I_i| > 4m$, $i = 1, \cdots, q$ for sufficiently large n, a. s. Thus, in this case,

$$\left|\frac{I_{q+1}}{n} - t_i\right| > \frac{4m}{n} - \frac{2m}{n} = \frac{2m}{n}.$$

Therefore in $\left[\frac{I_{q+1}-m}{n}, \frac{I_{q+1}+m}{n}\right]$ there are no discontinuity points, so

$$\begin{aligned}
|D_{nI_{q+1}}| &\leq |\Delta_{I_{q+1}}| + |W^{+}_{I_{q+1}}| + |W^{-}_{I_{q-1}}| \\
&\leq K\frac{2m}{n} + |W^{+}_{I_{q+1}}| + |W^{-}_{I_{q+1}}|,
\end{aligned}$$

and

$$\begin{aligned}
&\sum P\left(|D_{nI_{q+1}}| \geq h_n\right) \\
&\qquad \leq \sum P\left(h_n \leq K\frac{2m}{n} + |W^{+}_{I_{q+1}}| + |W^{-}_{I_{q+1}}|\right) \\
&\qquad \leq b\sum_n nP\left(b_1 h_n \leq \left|\frac{v_1 + \cdots + v_m}{m}\right|\right) \\
&\qquad \leq b\sum_n nP(b_2\sqrt{m}h_n \leq |z|) < \infty.
\end{aligned}$$

Therefore

$$|D_{nI_{q+1}}| = O(h_n), \quad a.\ s.$$

References

1. E. Y. Chow, Analytical Studies of the Generalized Likelihood Ratio Technique for Failure Detection, M. S. Thesis, Dept. Elec. Eng. Comput. Sci., M. I .T., 1976.
2. D. V. Hinkley, Inference about the Change-point from Cumulative Sum-tests, Biometrika, vol. 58, pp. 509-523(1971).
3. R. L. Kirlim, A. Mghaddamjoo, A Robust Running Window Detector and Estimator for Step Signals in Contaminated Gaussian Noise IEE Trans., vol. SP 34 No.4, pp 816-823.
4. E. S. Page, Continuous Inspection Schemes, Biometrika, vol. 41, pp 100-114(1954).

FILMED

4 8