DISCRETE CAPACITY AND HIGHER-ORDER DIFFERENCES OF TWO-STATE MARKOV CHAINS

A. YU. SHAHVERDIAN

ABSTRACT. The paper studies the time-homogeneous two-state Markov chains; the states are assumed to be binary symbols 0 and 1. The higher-order absolute differences taken from progressive states of a given chain are considered. A discrete capacity of subsets of natural series is defined and a limiting theorem for these differences, formulated in terms of Wiener criterion type relation, is presented.

1. Introduction

In this paper an application of the suggested in [1]-[7] difference analysis to studying the two-state Markov chains is presented. The difference analysis is a method for studying irregular and random time series, based on consideration of higher-order absolute differences taken from the series’ progressive terms. This method allowed us to reveal some new aspects in dynamical systems: e.g., the higher-order-difference version for Lyapunov exponent [3] and bistability of higher-order differences, taken from periodic time series [6], have been established.

We study time-homogeneous Markov chains $\xi = (\xi_n)_{n=0}^{\infty}$, whose state space $X = \{x\}$ consists of two different items; more precisely, we suppose that $X = \{0, 1\}$, that is, each component $\xi_n$ of $\xi$ (which describes the chain at the moment $n$) is a random binary variable.

The main result of this paper, Theorem 1, is a limiting theorem for such chains: it assert the existence of the limit of $k$th order absolute differences taken from progressive terms of a given series $(\xi_n)_{n=0}^{\infty}$, when $k$ converges to $\infty$ remaining on ”large” subsets $E \subseteq \mathbb{N}$ of natural series $\mathbb{N}$. The ”size” of such sets $E$ is described in terms of some discrete capacity: such sets $E$ are thick sets, defined by means of Wiener criterion type relation from potential theory (see, e.g., [8] and [9]). The limiting process, whose existence asserts Theorem 1, is the equi-distributed random sequence.

The paper consists of three sections. The next Section 2 describes the statement of the considered problem, in Section 3 we present the definitions of discrete capacity, thin and thick sets, and formulate our Theorem 1.

Date: November 11, 2017.

Key words and phrases. Markov chain, Higher-order differences, Ergodic theorems, Wiener criterion.

2010 Mathematics Subject Classification: 31C40, 31C45, 31CD05, 60J10, 60J45.
2. Statement of the problem

Let us explain the statement of the problem which we study. Let

\[ \xi = (\xi_0, \xi_1, \ldots, \xi_n, \ldots) \]

be some random sequence whose components \( \xi_n \) take binary values \( x \) from \( X = \{0, 1\} \) with some positive probabilities \( p_n(x) \), \( P(\xi_n = x) = p_n(x) \) \( (p_n(0) + p_n(1) = 1) \). Then \( k \)th order \( (k \geq 0) \) absolute differences \( \xi_n^{(k)} \), defined recurrently as: \( \xi_n^{(0)} \equiv \xi_n \) and

\[ \xi_n^{(k)} = |\xi_n^{(k-1)} - \xi_n^{(k-1)}| \quad (n \geq 0), \]

also take binary values with some probabilities \( p_n^{(k)}(x) \),

\[ P(\xi_n^{(k)} = x) = p_n^{(k)}(x) \quad (p_n^{(k)}(0) + p_n^{(k)}(1) = 1); \]

hence, one can consider \( k \)th order difference random binary sequence

\[ \xi^{(k)} = (\xi_0^{(k)}, \xi_1^{(k)}, \ldots, \xi_n^{(k)}, \ldots). \]

We are interested in existence of the limit of \( \xi^{(k)} \) when \( k \) goes to infinity. Let some infinite \( \Lambda \subseteq \mathbb{N} \) be given; we say that \( \xi^{(k)} \) converge to a random binary sequence \( \xi^{(\infty)} \), if \( p_n^{(k)}(x) \) \( (n \in \mathbb{N}, x \in X) \) tend to some numbers \( p_n^{(\infty)}(x) \) \( (p_n^{(\infty)}(0) + p_n^{(\infty)}(1) = 1) \) as \( k \to \infty \) and \( k \in \Lambda \) (convergence by probability on \( \Lambda \)). Given \( \Lambda \) the limiting process

\[ \xi^{(\infty)} = \xi^{(\infty)}_{\Lambda} = (\xi_0^{(\infty)}, \xi_1^{(\infty)}, \ldots, \xi_n^{(\infty)}, \ldots) \]

(so-called partial limit) is defined as random sequence, whose components \( \xi_n^{(\infty)} \) take the values \( x \in X \) with the probabilities \( p_n^{(\infty)}(x) \).

We study time-homogeneous Markov chains \( \xi \), that is, when for \( x, x_1, y \in X \)

\[ P(\xi_n = y | \xi_{n-1} = x, \xi_{n-2} = x_1, \ldots, \xi_0 = x_{n-1}) = P(\xi_n = y | \xi_{n-1} = x) \quad (1) \]

(Markov property) and there is some function \( \pi(x, y) \) on \( X \times X \) such that

\[ P(\xi_n = y | \xi_{n-1} = x) = \pi(x, y) \quad \text{for } n \geq 1 \text{ and } x, y \in X \quad (2) \]

(homogeneity). Some computations testify, that if for such \( \xi \) an infinite \( \Lambda \subseteq \mathbb{N} \) is chosen arbitrarily, the limiting process \( \xi^{(\infty)}_{\Lambda} \) may not exist; on the other hand, a theorem announced in [7] asserts that if \( \Lambda = \{2^m - 1 : m \geq 0\} \), then \( \xi^{(\infty)}_{\Lambda} \) exists. The problem which studies the present paper is the following (descriptively): how "large" can be the sets \( \Lambda \subseteq \mathbb{N} \) which permit the existence of \( \xi^{(\infty)}_{\Lambda} \), and how their "size" can be described? This paper considers the chains for which

\[ \pi(x, y) \neq 0, \quad \pi(0, 0) \neq \pi(1, 1), \quad \text{and} \quad \pi(0, 0) + \pi(1, 1) \neq 1. \quad (3) \]

We claim that for time-homogeneous binary Markov chains the problem stated is resolved in terms of some discrete capacity defined on \( 2^\mathbb{N} \) and corresponding thin (fine) and thick sets. The capacity \( C \), considered here, is a modification of the discrete capacity used in [4]. The solution to our problem
is given by Theorem 1, which is formulated in terms of thick sets, defined by means of well-known in potential theory Wiener criterion type relation.

3. Some definitions and main theorem

We consider binary Markov chains \( \xi = (\xi_n)_{n=0}^\infty \) whose state space \( X \) consists of two binary symbols, \( X = \{0, 1\} \), and for which Eq. (1) holds. We assume that the chains \( \xi \) are time-homogeneous, which means that one-step transition probabilities \( P(\xi_n = y|\xi_{n-1} = x) \) do not depend on time \( n \), i.e., for some \( \pi(x, y) \) Eq. (2) holds; it is also assumed that some initial distribution of probabilities \( P(\xi_0 = x) \) on \( X \) is given.

To proceed to formulation of our Theorem 1, we first present the notions of discrete capacity \( C \) and associated with this capacity thin and thick sets.

The capacity \( C \) is assigned on \( 2^\mathbb{N} \); to define it, we consider binary codes of natural numbers. Let \( k \in \mathbb{N}, (k \geq 1) \) and \((\varepsilon_0, \ldots, \varepsilon_p)\) be the binary code of \( k \): \( k = \sum_{i=0}^{p} \varepsilon_i 2^i \) where \( p \geq 0, \varepsilon_i \in \{0, 1\} \) and \( \varepsilon_p = 1 \) (binary expansion of \( k \)). Let \( \nu(k) \) denotes the maximal of such \( m (0 \leq m \leq p) \), for which all the coefficients \( \varepsilon_i, 0 \leq i \leq m \) of binary expansion of \( k \) are equal to 1.

**Definition 1.** For \( e \subseteq \mathbb{N} \) we define

\[
C(e) = \sum_{k \in e} \nu(k). \tag{4}
\]

A set \( e \subseteq \mathbb{N} \) is called thin (or, fine) set (\( F \)-set) if the relation

\[
\sum_{p=1}^{\infty} 2^{-p} C(e \cap K_p) < \infty, \tag{5}
\]

where \( K_p = \{k \in \mathbb{N} : 2^p \leq k < 2^{p+1}\} \), holds. If the set \( e \subseteq \mathbb{N} \) is not thin (i.e., Eq. (5) is failed), \( e \) is called thick set (\( T \)-set).

The \( C(e) \) from Eq. (1) can be expressed in terms of binomial coefficients as follows. Let (for given \( k \geq 1 \)) \( \mu(k) \) denotes the maximal of such \( m (0 \leq m \leq k) \), for which all the binomial coefficients \( \binom{k}{i}, 0 \leq i \leq m \) (first \( m \) entries of \( k \)th line \((\binom{k}{0}, \binom{k}{1}, \ldots, \binom{k}{k})\) of the Pascale triangle), are odd numbers; one can prove that

\[
\mu(k) = 2^{\nu(k)}
\]

and, therefore,

\[
C(e) = \sum_{k \in e} \log_2 \mu(k).
\]

Since for infinite collection of bounded sets \( e \subset \mathbb{N} \) and some positive constant we have \( C(e) \leq \text{const}. C(\partial e) \) (cp. [4]; such inequality is mentioned also in [10] when defining a capacity of clusters from \( \mathbb{N} \times \mathbb{N} \), used in some models [11]-[12] of self-organized criticality), which is a characteristic property of classical capacities (e.g., [13]), we call \( C \) a capacity. We note that \( C \) is differed from
discrete capacity, considered in denumerable Markov chains and random walk (see, e.g., [13]).

The next Proposition 1 contains some formal properties of capacity $C$ and fine and thick sets (which we abbreviate as $\mathcal{F}$-sets and $\mathcal{T}$-sets, respectively); we note that $C(e) \geq 0$ for arbitrary $e \subseteq \mathbb{N}$.

**Proposition 1.** The next statements (a)-(f) are true: (a) $C(\emptyset) = 0$ and $C(\mathbb{N}) = \infty$. (b) If $e_1 \subseteq e_2$ then $C(e_1) \leq C(e_2)$. (c) $C(\{2^p \leq k < 2^{p+1}\}) = (1 + o(1))2^p (p \to \infty)$. (d) The $\mathbb{N}$ is $\mathcal{T}$-set. (e) Every finite subset of $\mathbb{N}$ is $\mathcal{F}$-set and finite union of $\mathcal{F}$-sets is $\mathcal{F}$-set. (f) If $e$ is $\mathcal{T}$-set and $e'$ is $\mathcal{F}$-set, then $e \cup e'$ and $e \setminus e'$ are $\mathcal{T}$-sets.

By using the next Proposition 2 one can construct more complicated examples of thin and thick subsets of $\mathbb{N}$.

**Proposition 2.** Let for $p \geq 1$ the natural numbers $0 \leq s_p \leq p, s_p \to \infty (p \to \infty)$ be given and $E \subseteq \mathbb{N}$ be defined as

$$E = \bigcup_{p=1}^{\infty} \{2^p \leq k < 2^{p+1} : \nu(k) \geq s_p\}. \quad (6)$$

Then $E$ is $\mathcal{F}$-set if and only if for $s_p$ the condition

$$\sum_{p=1}^{\infty} \frac{s_p}{2^p} < \infty$$

holds.

**Definition 2.** A number $a$ is called thick limit point ($\mathcal{T}$-limit point or $\mathcal{T}$-cluster point) of a given infinite numerical sequence $a_k, k \geq 0$ if there is a $\mathcal{T}$-set $E \subseteq \mathbb{N}$ such that

$$\lim_{k \to \infty; k \in E} a_k = a.$$

A random binary sequence $\xi = (\xi_n)_{n=0}^{\infty}$ is called $\mathcal{T}$-limit process for a given infinite series of random binary sequences $\xi_k = (\xi_{n,k})_{n=0}^{\infty}, k \geq 0$ if for $x \in X$ and $n \geq 0$ the probability $P(\xi_n = x)$ is $\mathcal{T}$-limit point for the sequence of probabilities $P(\xi_{n,k} = x), k \geq 0$.

The following Theorem 1 is the main result of this paper.

**Theorem 1.** Let $\xi = (\xi_n)_{n=0}^{\infty}$ be time-homogeneous binary Markov chain for which Eq. (3) holds. Then the equi-distributed random binary sequence is the $\mathcal{T}$-limit process for the sequence of higher-order differences $\xi^{(k)} = (\xi^{(k)}_{n})_{n=0}^{\infty}, k \geq 0$. More precisely, for $x \in X$ and $n \geq 0$ there is a $\mathcal{T}$-set $E \subseteq \mathbb{N}$ of the form (6) with $\sum_{p=1}^{\infty} s_p 2^{-s_p} = \infty$, for which

$$\lim_{k \to \infty; k \in E} P(\xi^{(k)}_n = x) = \frac{1}{2}.$$
In certain sense, Theorem 1 can be treated as the higher-order-difference version of the classical ergodic theorem for finite (two-state) Markov chains, where some notions from potential theory are now involved.

To the end, we present some characteristics of the sets $E$ from Theorem 1 formulated in terms of their density in natural series. For $m \geq 1$ we denote $E_m = \{k \in E : 1 \leq k \leq m\}$ and consider the ratio $\rho_m(E) = \frac{|E_m|}{m}$ where $|E_m|$ denotes the cardinality of $E_m$.

**Remark 1.** The sets $E \subseteq \mathbb{N}$ defined by Eq. (6) in Proposition 2 and presented in formulation of Theorem 4 are of zero density in natural series: $\rho_m(E) \to 0$ as $m \to \infty$. The sets $E$ defined by Eq. (6) can be such that the ratio $\rho_m(E)$ converges to 0 as slow as we please: given $0 < \delta_m \leq 1$, $\delta_m \downarrow 0$ the $\mathcal{T}$-set $E$ from Theorem 4 can be constructed in such a way that $\rho_m(E) \geq \delta_m$ for all $m \geq 1$.

**REFERENCES**

[1] Shahverdian, A. Yu., Apkarian, A. V. (1999). On irregular behavior of neural spike trains. *Fractals*, 7(1), 93-103.

[2] Shahverdian, A. Yu. (2000). The finite-difference method for analyzing one-dimensional nonlinear systems. *Fractals*, 8(1), 49-65.

[3] Shahverdian, A. Yu., Apkarian, A. V. (2007). A difference characteristic for one-dimensional nonlinear systems. *Comm. Nonlin. Sci. & Comput. Simul.*, 12, 233-242.

[4] Shahverdian, A. Yu. (2012). Minimal Lie algebra, fine limits, and dynamical systems. *Reports Armenian Natl. Acad. Sci.*, 112(2), 160-169.

[5] Shahverdian, A. Yu., Kilicman, A., Benosman, R. B. (2012). Higher difference structure of some discrete processes. *Adv. Difference Equations*, 202, 1-10.

[6] Shahverdian, A. Yu., Agarwal, R. P., Benosman, R. B. (2014). The bistability of higher-order differences of periodic signals. *Adv. Difference Equations*, 60, 1-9.

[7] Shahverdian, A. Yu. (2015). A theorem on higher-order differences of two-state Markov chains. *Proc. Intern. Conf. CSIT-2015*. Yerevan, 251-252 (reprinted in: *IEEE Conference Ser.*, *CSIT-2015*, 137-138).

[8] Brelot, M. (1971). Topologies and Boundaries in Potential Theory. *Springer*, Berlin.

[9] Shahverdian, A. Yu. (2011). Fine topology and estimates for potentials and subharmonic functions. *Computational Methods and Function Theory*, 11(1), 71-121.

[10] Shahverdian, A. Yu. (2011). Avalanches and memory in rotator networks. *Reports Armenian Natl. Acad. Sci.*, 111(3), 240-249.

[11] Shahverdian, A. Yu. (1997). Lattice animals and self-organized criticality. *Fractals*, 5(2), 199-213.

[12] Shahverdian, A. Yu., Apkarian, A. V. (2008) Avalanches in networks of weakly coupled phase shifting rotators. *Comm. Math. Sci.*, 6(1), 217-234.

[13] Dynkin, E. B., Yushkevich, A. A. (1969). Markov processes: Theorems and Problems. *Plenum Press*, New York.

**INSTITUTE FOR INFORMATICS AND AUTOMATION PROBLEMS OF NAS RA**

*E-mail address: svrdn@yerphi.am*