# SOME THEOREMS CONCERNING THE STRONG LAW OF LARGE NUMBERS FOR NON-HOMOGENEOUS MARKOV CHAINS

BY M. ROSENBLATT-ROTH

*University of Bucharest*

## 1. Introduction.

1. *Preliminary*. This paper deals with the problem of finding (necessary or sufficient) conditions for the strong law of large numbers in the case of a Markov chain.

The results proved in this paper are of classical form, i.e. they come very close to those of Cantelli, Borel, Khintchine, Kolmogorov for mutually independent random variables; these classical results themselves remain true for a very large class of non-homogeneous Markov chains (for which $\alpha_i > \rho > 0$, $i \varepsilon I = (1, 2, \cdots)$). In the same way we obtain new results for homogeneous Markov chains ($\alpha_i = \rho > 0$, $i \varepsilon I$); these results contain as particular cases the analogous results for mutually independent random variables ($\alpha_i = 1$, $i \varepsilon I$).

A part of these results was announced in preliminary papers ([11]–[13]).

We express our results by means of *the ergodic coefficient of a stochastic transition function* ([2], [1]); in [9] can be found various of its definitions and properties that we shall use here.

2. *Notations and definitions*. Let $(\mathfrak{A}_i, \Sigma_i)$ be a measurable space, $x_i$ the elements of $\mathfrak{A}_i$, $A_i$ the measurable sets, elements in the $\sigma$-algebra $\Sigma_i (i \varepsilon I)$. If the sequence of random variables $\xi_i (i \varepsilon I)$ is a Markov chain, let us consider that it has the stochastic transition functions $P_i(x_i, A_{i+1})$ with domains of definition $(\mathfrak{A}_i, \Sigma_i, \mathfrak{A}_{i+1}, \Sigma_{i+1}) (i \varepsilon I)$. We denote by $\alpha_i = \alpha(P_i)$ the ergodic coefficient of $P_i$ and by $\alpha_{ij} = \alpha(P_{ij})$ that of the transition function $P_{ij}(x_i, A_j)$ for the time interval $i, j (i + 1 < j)$. We shall suppose that all the variances $D\xi_i (i \varepsilon I)$ are finite and we set

$$(1) \qquad \alpha^{(n)} = 1 - \eta_n = \min_{1 \leq i < n} \alpha_i, \qquad D_n = \sum_{i=1}^{n} D\xi_i .$$

We assume that $\alpha_i > 0 (i \varepsilon I)$, because in many important formulae (Basic Lemma, Lemma 1, Theorem 1) $\alpha^{(n)}$ appears in the denominator.

*The random variables $\sigma_i (i \varepsilon I)$ are called strongly stable, if there is some numerical sequence $d_i (i \varepsilon I)$ so that for $n \to \infty$, $\sigma_n - d_n$ converges to zero with probability* 1. In this case it is possible, [7], to take $d_i = m\sigma_i$ ($m$—the median); the $\sigma_i (i \varepsilon I)$ are called *normally strongly stable* if it is possible to take $d_i = M\sigma_i$ ($M$—the expectation). Let

$$(2) \qquad S_n = \sum_{i=1}^{n} \xi_i, \qquad \sigma_n = n^{-1} S_n, \qquad \mathfrak{U}_n = \max_{1 \leq s \leq n} |S_s - MS_s|, \quad (n \varepsilon I).$$

Received 8 January 1963; revised 27 September 1963.

*The sequence $\xi_i (i \varepsilon I)$ is said to satisfy the strong law of large numbers* (S.L.) *if the sequence $\sigma_i (i \varepsilon I)$ is strongly stable, and the normal strong law of large numbers* (N.S.L.) *if $\sigma_i (i \varepsilon I)$ is normally strongly stable.*

The sequences of random variables $\xi_i, \xi_i' (i \varepsilon I)$ are equivalent if

$$\sum_{i=1}^{\infty} P(\xi_i \neq \xi_i') < +\infty.$$

In this paper we shall consider everywhere that the sequence of random variables $\xi_i (i \varepsilon I)$ is a Markov chain. We shall use the

Basic Lemma [9], [10].

$$(3) \quad C' \alpha^{(n)} D_n \leqq DS_n \leqq C[\alpha^{(n)}]^{-1} D_n ; \qquad C = 16(1 + 6^{\frac{1}{2}}), \qquad C' = 10^{-2}$$

**2. Results.** If $1 < l \varepsilon I$, if $m, u = 0, 1, 2, \cdots$, and if $s$ and $\sigma$ are related to $m$ and $u$ by $s = l^m$, $\sigma = l^u$, set

$$(4) \qquad \omega_i^{-2} = \sum_{m=u}^{\infty} (s^2 \alpha^{(ls)})^{-1} \qquad \text{for all } i \text{ with } \sigma \leqq i < l\sigma.$$

Theorem 1. *If $n\alpha^{(n)} \to \infty (n \to \infty)$, then the condition*

$$(5) \qquad \sum_{n=1}^{\infty} \omega_n^{-2} D\xi_n < +\infty$$

*for some $l$ is sufficient for the* N.S.L. *If some number $A$ exists, so that for all $m \geqq m_0$*

$$(6) \qquad \alpha^{(ls)} \geqq A\alpha^{(s)} \text{ and } l^{-2} < A \leqq 1$$

*then in place of* (5) *we may take*

$$(7) \qquad \sum_{n=1}^{\infty} [n^2 \alpha^{(n)}]^{-1} D\xi_n < +\infty.$$

Theorem 2. *If $\alpha_i > \rho > 0 (i \varepsilon I)$, in place of* (7) *we may take*

$$(8) \qquad \sum_{n=1}^{\infty} n^{-2} D\xi_n < +\infty.$$

*This condition is the best in the sense that, if for some sequence of constants $b_n > 0$, the series*

$$(9) \qquad \sum_{n=1}^{\infty} n^{-2} b_n$$

*diverges, it is possible to construct a Markov chain $\xi_i (i \varepsilon I)$ (non degenerated into a sequence of mutually independent random variables) with $D\xi_i = b_i$, $\alpha_i > \rho > 0 (i \varepsilon I)$ and which does not verify the* N.S.L.

Let us denote by $E_i$ the random event $\{|\xi_i - m\xi_i| > \epsilon\}$, by $E_i'$ its complement, and by $P(E_i \mid E_{i-1}')$ the conditional probability of $E_i$ given $E_{i-1}'$.

Theorem 3. *In the case of strong stability, for any $\epsilon > 0$*

$$(10) \qquad \sum_{i=1}^{\infty} P(E_i \mid E_{i-1}') < +\infty.$$

THEOREM 4. *If the* S.L. *is verified then* (10) *is true for the events* $E_i = \{|\xi_i - m\xi_i| > \epsilon i\}$. *If* $\alpha_i > \rho > 0$ $(i \varepsilon I)$ *and the* N.S.L. *is verified, then this condition is not necessary if in the definition of* $E_i$ *we take* $M\xi_i$ *and* $\varphi(i) = o(i)$ *instead of* $m\xi_i$ *and* $\epsilon i$.

Let $R$ be a real line and $\Omega = \{u_k\}$ some finite or denumerable system of non-overlapping Borel sets on it, so that the union of all $u_k$ is $R$. The totality of all real-valued random variables $\xi$ can be divided into disjunct classes $\Lambda(\Omega)$ so that the probability $P(|\xi| \varepsilon u_k)$ depends only on the class which contains $\xi$, but not on the random variable $\xi$. All the random variables $\xi$, contained in the same class $\Lambda(\Omega)$ will be called $\Omega$-*identically distributed.* Obviously, if all the random variables of some family are identically distributed, they are also $\Omega$-identically distributed in respect to any system $\Omega$. We will use here $u_k = \{k \leqq x < k + 1\}$ for all integers $k$.

THEOREM 5. *If the* $\xi_i (i \varepsilon I)$ *are* $\Omega$-*identically distributed, and*

$$\alpha^{(n)} \geqq cn^{-\beta} (0 \leqq \beta < 1),$$

*for the* N.S.L. *it is sufficient that in the same class with the* $\xi_i (i \varepsilon I)$ *there exists some random variable* $\xi$ *possessing a finite moment of order* $1 + \beta$.

LEMMA 1.

$$(11) \quad P(\mathfrak{U}_n > \epsilon) < \epsilon_1^{-2}[\alpha^{(n)}]^{-1} D_n ; \qquad \epsilon_1 = (20K)^{-1}\epsilon, \qquad K = 1 + 6^{\frac{1}{2}}.$$

*If* $\alpha_i > \rho > 0$ $(i \varepsilon I)$, *then*

$$(12) \qquad P(\mathfrak{U}_n > \epsilon) < \epsilon_1^{-2} D_n ; \qquad \epsilon_1 = (20K)^{-1}\epsilon\rho^{-1}.$$

LEMMA 2. *If for a sequence* $E_i(i \varepsilon I)$ *of random events connected in a Markov chain, the series* (10) *diverges, then with probability* 1 *an infinite set of the* $E_i$ *occurs.*

1. *Proof of Lemma* 1. Without any loss of generality we may suppose $M\xi_i = 0$ $(i \varepsilon I)$. Let us consider the random events

$$E_k = \{|S_i| < \epsilon; \ 1 \leqq i < k; \ |S_k| \geqq \epsilon\}; \qquad (1 \leqq k \leqq n);$$

$$E_0 = \{|S_i| < \epsilon, \ 1 \leqq i \leqq n\}.$$

Obviously $\{\mathfrak{U}_n \geqq \epsilon\} = \bigcup_{i=1}^n E_i$ and because the $E_i$ are disjoint,

$$P\{\mathfrak{U}_n \geqq \epsilon\} = \sum_{i=1}^n P(E_i).$$

For $1 \leqq k \leqq n$, from

$$S_n = S_k + \sum_{i=k+1}^n \xi_i , \qquad S_n^2 \geqq S_k^2 + 2\sum_{j=1}^k \sum_{i=k+1}^n \xi_j\xi_i$$

follows

$$M(S_n^2 \mid E_k) \geqq \epsilon^2 + 2 \sum_{j=1}^k \sum_{i=k+1}^n M(\xi_j\xi_i \mid E_k)$$

and because $MS_n = 0$, we have

$$DS_n = \sum_{k=0}^{n} M(S_n^2 \mid E_k)P(E_k) \geqq \sum_{k=1}^{n} P(E_k)\left[\epsilon^2 + 2\sum_{j=1}^{k}\sum_{i=k+1}^{n} M(\xi_j \xi_i \mid E_k)\right]$$

and

$$\epsilon^2 \sum_{i=1}^{n} P(E_i) \leqq DS_n + 2\sum_{k=1}^{n} P(E_k) I_k \tag{13}$$

where

$$I_k = \sum_{j=1}^{k}\sum_{i=k+1}^{n} |M(\xi_j\xi_i \mid E_k)|.$$

Using the formulae 7 and 8 from [9], we deduce

$$1 - \alpha_{ji} \leqq \eta_n^{i-j}. \tag{14}$$

Now let us consider two stochastic transition functions $P(\omega_1, A_2)$, $P^*(\omega_1, A_2)$ defined respectively in the domains $(\Omega_1, \Sigma_1, \Omega_2, \Sigma_2)$, $(\Omega_1^*, \Sigma_1^*, \Omega_2, \Sigma_2)$ where $\Omega_1^* \subset \Omega_1$, $\Sigma_1^* \subset \Sigma_1$ and let us suppose that for $\omega_1 \varepsilon \Omega_1^*$, $A_2 \varepsilon \Sigma_2$ the functions $P$, $P^*$ coincide; using the Definition 2 from [9], it follows that $\alpha(P) \leqq \alpha(P^*)$, and by means of Lemma 4 of [9] and of (14) we obtain

$$|M(\xi_j\xi_i \mid E_k)| \leqq K\eta_n^{\frac{1}{2}(i-j)}[D(\xi_j \mid E_k) + D(\xi_i \mid E_k)].$$

Therefore $I_k \leqq K(I_k' + I_k'')$, where

$$I_k' = \sum_{i=k+1}^{n}\sum_{j=1}^{k} \eta_n^{\frac{1}{2}(i-j)} D(\xi_i \mid E_k) \leqq \sum_{i=k+1}^{n}\left[\left(\sum_{r=i-k}^{\infty} \eta_n^{\frac{1}{2}r}\right) D(\xi_i \mid E_k)\right]$$
$$\leqq (1 - \eta_n^{\frac{1}{2}})^{-1} \sum_{i=k+1}^{n} D(\xi_i \mid E_k)$$

$$I_k'' = \sum_{j=1}^{k}\sum_{i=k+1}^{n} \eta_n^{\frac{1}{2}(i-j)} D(\xi_j \mid E_k) = \sum_{j=1}^{k}\left[\left(\sum_{r=k+1-j}^{n-j} \eta_n^{\frac{1}{2}r}\right) D(\xi_j \mid E_k)\right]$$
$$\leqq (1 - \eta_n^{\frac{1}{2}})^{-1} \sum_{j=1}^{k} D(\xi_j \mid E_k).$$

Using the relation (18) from [9] and the known relation

$$D_1(\xi \mid E) = \sum_{k=0}^{n} P(E_k) D(\xi \mid E_k) < D\xi$$

where $E = (E_0, E_1, \cdots, E_n)$, it is easy to obtain

$$\sum_{k=1}^{n} I_k P(E_k) \leqq 2K[\alpha^{(n)}]^{-1} \sum_{i=1}^{n} D_1(\xi_i \mid E) < 2K[\alpha^{(n)}]^{-1} D_n. \tag{15}$$

By means of (3) and (15), (11) follows from (13). Also, because $\alpha_i > \rho > 0$ $(i \varepsilon I)$ and $\alpha^{(n)} > \rho > 0$ $(n \varepsilon I)$ are equivalent, (12) follows.

2. *Proof of Lemma* 2. First we shall consider that all the random events

$E_i(i \varepsilon I)$ belong to the same measurable space $(\mathfrak{A}, \Sigma)$. The event $X$ which consists in the occurrence of an infinite number of events $E_n(n \varepsilon I)$, can be expressed by

$$X = \bigcap_{j=1}^{\infty} \bigcup_{i>j} E_i .$$

If $P(E_i) = 1$ for a finite set of indices $i_1 < i_2 < \cdots < i_s$ we may study the problem for $i > i_s$ only; if $P(E_i) = 1$ for an infinite set of indices, it follows evidently $P(X) = 1$. Thus, we may consider only the case $P(E_i') > 0$ $(i \varepsilon I)$, i.e. the case for which the conditional probabilities $P(E_i \mid E_{i-1}')$ exist and are determined for all $i \varepsilon I$; in this last case, we have

$$P(X) = \lim_{j\to\infty} P\left(\bigcup_{i>j} E_i\right) = 1 - \lim_{j\to\infty} P\left(\bigcap_{i>j} E_i'\right)$$
$$= 1 - \lim_{j\to\infty} \left\{P(E_{j+1}') \prod_{i=j+1}^{\infty} [1 - P(E_i \mid E_{i-1}')]\right\}.$$

If the series (10) diverges, the infinite product diverges to zero, i.e. $P(X) = 1$.

If the random events $E_i(i \varepsilon I)$ belong to different measurable spaces $(\mathfrak{A}_i, \Sigma_i)$ $(i \varepsilon I)$, we obtain the same result by applying the obtained result to the cylinders $E_i \times \mathsf{X}_{n\neq i}\mathfrak{A}_n$ of the infinite cartesian product of all $(\mathfrak{A}_i, \Sigma_i)(i \varepsilon I)$.

3. *Proof of Theorem* 1. Let us consider the random events

$$A_m = \{\max_{s \leqq n < ls} n^{-1}|S_n| \geqq \epsilon\}$$
$$B_m = \{\max_{s \leqq n < ls} |S_n| \geqq \epsilon s\} \supset A_m , \qquad s = l^m.$$

Using (11) and (4)

$$\sum_{m=0}^{\infty} P(A_m) \leqq \sum_{m=0}^{\infty} P(B_m) \leqq 20K\epsilon^{-2} \sum_{m=0}^{\infty} \left\{(s^2\alpha^{(ls)})^{-1} \sum_{j=1}^{ls} D\xi_j\right\} \tag{16}$$
$$\leqq 20K\epsilon^{-2} \sum_{i=1}^{\infty} \omega_i^{-2} D\xi_i .$$

If (6) is fulfilled, for each $i$ which verifies (4) we deduce

$$\omega_i^2 \geqq (Al^{-2} - l^{-4})i^2\alpha^{(i)}. \tag{17}$$

From (16), (17) follows our theorem.

4. *Theorem* 3 follows from the definition of the strong stability, using Lemma 2.

5. *Proof of Theorem* 5. Let us define the auxiliary random variables $\xi_n'$ equal to $\xi_n$ if $|\xi_n| \leqq n$ and to zero if $|\xi_n| > n$; obviously $\xi_n(n \varepsilon I)$ is also a Markov chain with the same transition functions as $\xi_n(n \varepsilon I)$.

(a) *The* $\xi_n$ , $\xi_n'(n \varepsilon I)$ *are equivalent.* Let us set $\varphi(k) = P\{|\xi_n| \varepsilon u_k\} = P(|\xi| \varepsilon u_k)$, $F_n(x) = P\{\xi_n < x\}$, $F(x) = P\{\xi < x\}$. We have

$$J_n = \sum_{i=n}^{\infty} P(\xi_i \neq \xi_i') = \sum_{i=n}^{\infty} (i - n + 1)\varphi(i) < \sum_{i=n}^{\infty} i\varphi(i) < \int_{|x| \geqq n} |x| \, dF(x).$$

From the inequality $(M|\xi|)^{1+\beta} \leqq M|\xi|^{1+\beta}$, $(\beta \geqq 0)$, and the conditions of our theorem, it follows that $M|\xi|$ is finite and therefore $J_n \to 0$ $(n \to \infty)$ i.e. $\xi_n$, $\xi'_n(n \varepsilon I)$ are equivalent.

(b) *The* $\xi'_n(n \varepsilon I)$ *verifies the* N.S.L. We have

$$D\xi'_n \leqq M\xi'^2_n = \sum_{k=0}^{n-1} \int_{u_k} x^2 \, dF_n(x) \leqq \sum_{k=0}^{n} (k+1)^2 \varphi(k). \tag{18}$$

In the conditions of our theorem we have

$$\omega_i^2 > cl^{-2}(1 - l^{\beta-2})i^{2-\beta}, \qquad s \leqq i < ls, \qquad s = l^m \tag{19}$$

so that it follows that

$$\begin{aligned} J = \sum_{i=1}^{\infty} \omega_i^{-2} D\xi'_i &\leqq c^{-1}l^2(1 - l^{\beta-2})^{-1} \sum_{i=1}^{\infty} \left[ i^{\beta-2} \sum_{k=0}^{i} (k+1)^2 \varphi(k) \right] \\ &= c^{-1}l^2(1 - l^{\beta-2})^{-1} \sum_{k=0}^{\infty} \left[ (k+1)^2 \varphi(k) \sum_{i=k}^{\infty} i^{\beta-2} \right]. \end{aligned} \tag{20}$$

Because $f(x) = x^{\beta-2}$ is non-negative and monotonically decreasing to zero,

$$\sum_{i=k}^{\infty} f(i) < f(k) + \int_k^{\infty} f(x) \, dx$$

from which follows

$$\sum_{i=k}^{\infty} i^{\beta-2} < 2(1-\beta)^{-1} k^{\beta-1}. \tag{21}$$

From (20), (21) and the obvious inequality $(k+1)^2 k^{\beta-1} \leqq 4k^{1+\beta}$ it follows that

$$J \leqq L \sum_{k=0}^{\infty} k^{1+\beta} \varphi(k) \leqq LM|\xi|^{1+\beta}, \text{ where } L = 8c^{-1}l^2(1-\beta)^{-1}(1 - l^{\beta-2})^{-1}.$$

Therefore, from Theorem 1 we obtain our proof, because equivalent sequences verify simultaneously the N.S.L. ([7], Theorem 4).

**3. Proof of Theorem 2.**

1. The first part of the theorem follows from Theorem 1. We shall prove the second part. Let $I_1$ be the set of $r \varepsilon I$ for which $b_r < r^2$ and $I_2$ its complementary set in $I$; also let $I_{ij}$ be the set of $r \varepsilon I_j$ for which $r - 1 \varepsilon I_i (i = 1, 2)$.

2. *The auxiliary sequence of constants* $\delta_r$. Let us set

$$a_r = \tfrac{1}{2} \min (r^{-2} b_r ; \ 1 - r^{-2} b_r) \leqq \tfrac{1}{4}, \qquad r \varepsilon I_1$$

and let us define a strictly monotonically decreasing sequence of positive constants $\delta_r(r \varepsilon I)$ for which $\delta_r < a_r$ if $r \varepsilon I_1$ and $\delta_r < \frac{1}{4}$ for $r \varepsilon I_2$. For instance we may take $0 \leqq \delta_1 < \min (a_1, \frac{1}{4})$ for $1 \varepsilon I_1$ and $0 \leqq \delta_1 < \frac{1}{4}$ for $1 \varepsilon I_2$ and also $0 \leqq \delta_r < \min (a_r ; \delta_{r-1})$ for $1 < r \varepsilon I_1$ and $0 \leqq \delta_r < \delta_{r-1}$ for $1 < r \varepsilon I_2$.

3. *The stochastic matrices.* We shall consider a chain with three states

$\omega_i^{(r)}(i = 1, 2, 3)$ in the instants $r \varepsilon I_1$ and only two states $\omega_i^{(r)}(i = 1, 2)$ in the instants $r \varepsilon I_2$.

We shall use the auxiliary constants $\pi_1 = -1, \pi_2 = 0, \pi_3 = 1, \theta_1 = -1, \theta_2 = 1$. Denote the transition probability $P(\omega_j^{(r)} \mid \omega_i^{(r-1)})$ from $\omega_i^{(r-1)}$ to $\omega_j^{(r)}$ by $p_{ij}^{(r)}(r \varepsilon I)$ [where: $i, j = 1, 2, 3$ in $I_{11}$; $i = 1, 2, 3, j = 1, 2$ in $I_{12}$; $i = 1, 2, j = 1, 2, 3$ in $I_{21}$; $i, j = 1, 2$ in $I_{22}$] and define

$$I_{11} : p_{i1}^{(r)} = p_{i3}^{(r)} = \tfrac{1}{2}r^{-2}b_r + \pi_i\delta_r \qquad (i = 1, 2, 3)$$

$$I_{12} : p_{i1}^{(r)} = \tfrac{1}{2} + \pi_i\delta_r \qquad (i = 1, 2, 3)$$

$$I_{21} : p_{i1}^{(r)} = p_{i3}^{(r)} = \tfrac{1}{2}r^{-2}b_r + \theta_i\delta_r \qquad (i = 1, 2)$$

$$I_{22} : p_{i1}^{(r)} = \tfrac{1}{2} + \theta_i\delta_r \qquad (i = 1, 2).$$

4. *The initial and the absolute probabilities.* If $P_i^{(r)}$ is the absolute (for $r = 1$' initial) probability of $\omega_i^{(r)}$ and we take $P_1^{(1)} = P_3^{(1)} = \tfrac{1}{2}\mathfrak{b}_1$, for $1 \varepsilon I_1$ and $P_1^{(1)} = P_2^{(1)} = \tfrac{1}{2}$ for $1 \varepsilon I_2$ it is easy to prove by induction that $P_1^{(r)} = P_3^{(r)} = \tfrac{1}{2}r^{-2}\mathfrak{b}_r$ for $r \varepsilon I_1$ and $P_1^{(r)} = P_2^{(r)} = \tfrac{1}{2}$ for $r \varepsilon I_2$.

5. *The ergodic coefficients.* From the Definition 6 in [9] it is easy to obtain that $\alpha_r$ is equal to $1 - 4\delta_r$ in the cases $I_{11}$, $I_{21}$, i.e. for $r \varepsilon I_1$ and to $1 - 2\delta_r$ in the cases $I_{12}$, $I_{22}$, i.e. for $r \varepsilon I_2$; then $\alpha^{(n)} > 1 - 4\delta_1(n \varepsilon I)$.

6. *The random variables* $\xi_r$ can be defined by $\xi_r(\omega_i^{(r)}) = \pi_i r$ $(i = 1, 2, 3)$ for $r \varepsilon I_1$ and by $\xi_r(\omega_i^{(r)}) = \theta_i(b_r)^{\frac{1}{2}}(i = 1, 2)$ for $r \varepsilon I_2$. Using the absolute probabilities $P_i^{(r)}(r \varepsilon I)$, it is easy to obtain $M\xi_r = 0, D\xi_r = b_r(r \varepsilon I)$.

7. *The* $\xi_r(r \varepsilon I)$ *does not verify the* N.S.L. If we set $E_r = \{|\xi_r - M\xi_r| > \epsilon r\}$, obviously, in the case $I_{11}$, $E_r = \omega_1^{(r)} \cup \omega_3^{(r)}$, $E'_{r-1} = \omega_2^{(r-1)}$, $P(E_r \mid E'_{r-1}) = r^{-2}b_r$, and in the case $I_{12}$, $E_r = \mathfrak{A}_r$, $E'_{r-1} = \omega_2^{(r-1)}$, $P(E_r \mid E'_{r-1}) = 1$.

In the cases $I_{21}$, $I_{22}$, because $E_{r-1} = \mathfrak{A}_{r-1}$, $P(E_r \mid E'_{r-1})$ is not defined. For $b_r(r \varepsilon I)$ only one of three cases can occur: (1) $I_{11}$ is an infinite set, (2) $I'_{11}$ is a finite set, (3) $I'_{22}$ is a finite set.

In the first case the series (10) contains an infinite set of elements equal to 1; in the second case, the series (10) contains all the elements of the divergent series (9) except a finite set of them. Thus, in these two cases, from Lemma 2, it follows that the $\xi_r(r \varepsilon I)$ does not verify the N.S.L. In the third case, for all $r \varepsilon I$, except a finite set, we have $P(E_r) = 1$; in the proof of Lemma 2 we have seen that with probability 1 there occurs an infinity of events from the sequence $E_r(r \varepsilon I)$, i.e. $\xi_r(r \varepsilon I)$ does not verify the N.S.L.

**4. Proof of Theorem 4.**

1. *The first part of the proof.* If $\xi_n(n \varepsilon I)$ verifies the S.L. there exists some sequence of constants $c_n(n \varepsilon I)$, so that if we denote $q_n = c_n - n^{-1}(n - 1)c_{n-1}$ and we consider the random variables $\sigma'_n = \sigma_n - c_n$, $\lambda_n = n^{-1}\xi_n$, $\sigma'_n - n^{-1}(n - 1)\sigma'_{n-1} = \lambda_n - q_n$ and the random events

$$A = \{\sigma'_n \to 0, \quad n \to \infty\}; \qquad A_1 = \{\sigma'_{n-1} \to 0, \quad n \to \infty\},$$

$$B = \{\sigma'_n - n^{-1}(n - 1)\sigma'_{n-1} \to 0, \quad n \to \infty\},$$

$$E_i = \{|\lambda_i - {}_m\lambda_i| > \epsilon\} = \{|\xi_i - {}_m\xi_i| > i\epsilon\}$$

we deduce $A \cap A_1 \subset B$, $P(A) = P(A_1) = 1$. It follows that $P(B) = 1$, i.e. the Markov chain $\lambda_n(n \varepsilon I)$ verifies the S.L. From Lemma 2 follows the convergence of the series (10).

2. *The second part of the proof.* Let us consider a real-valued positive function $\varphi(n) = o(n)$ and the event $E_r = \{|\xi_r - M\xi_r| > \varphi(r)\}$; we shall construct a Markov chain $\xi_r(r \varepsilon I)$ which verifies the N.S.L. but the series (10) diverges.

Obviously it is sufficient to consider $\varphi(n)$ monotonically increasing to infinity, $n^{-1}\varphi(n)$ monotonically decreasing to zero for $n \to \infty$ and $\varphi(1) = 1$. From the relation

$$0 < n^{-1}\varphi(n) - (n+1)^{-1}\varphi(n+1) < n^{-1}\varphi(n) \to 0 \quad (n \to \infty) \tag{22}$$

it follows that there exist some values of $n$ so that for a given $s \varepsilon I$, we have

$$s^{-2} < n^{-1}\varphi(n) < 3s^{-2}. \tag{23}$$

For each $s \varepsilon I$, let $n_s$ be the least integer in $I$ for which these inequalities are true. From (22) it follows that there exists an $s_0 \varepsilon I$, so that for $s \geqq s_0$ the correspondence between $s$ and $n_s$ is one-to-one. If $s$ runs over $I$, we denote by $I_1$ the subset of $I$, containing all the integers $n_s$ and by $I_2$ its complement in $I$. We define now the function $g_n$ equal to $n^{-2}$ for $n \varepsilon I_2$ and to $s^{-2}$ for $n = n_s \varepsilon I_1$, and also the function

$$v_r = \tfrac{1}{2} r g_r [\varphi(r)]^{-1}.$$

From (23), $2v_r < 1$. We consider the constants $\delta_r(r \varepsilon I)$ strictly monotonically decreasing to zero, for which

$$0 \leqq \delta_r < \min(v_r, \tfrac{1}{2} - v_r) \leqq \tfrac{1}{4}.$$

3. *The construction of the Markov chain.* We define the Markov chain $\xi_r(r \varepsilon I)$ with three states $\omega_i^{(r)}(i = 1, 2, 3)$ for which $\xi_r(\omega_i^{(r)}) = c\pi_i\varphi(r)$ $(i = 1, 2, 3)$ and the transition probabilities $p_{i1}^{(r)} = p_{i3}^{(r)} = v_r + \pi_i\delta_r (i = 1, 2, 3)$.

4. *The sequence $\xi_r(r \varepsilon I)$ verifies the* N.S.L. Denote by $P_i^{(r)}$ the absolute (for $r = 1$, initial) probability of $\omega_i^{(r)}(i = 1, 2, 3)$. If we suppose that $P_1^{(1)} = P_3^{(1)} = v_1$, by induction it is easy to obtain $P_1^{(r)} = P_3^{(r)} = v_r(r \varepsilon I)$. Using the Definition 6 from [9] we obtain that $\alpha_r = 1 - 4\delta_r$ so that $\alpha^{(n)} = 1 - 4\delta_1 > 0$ $(n \varepsilon I)$. It is easy to see that $M\xi_r = 0$, $D\xi_r = cr^2 g_r\varphi(r)$ and therefore

$$\sum_{r=1}^{\infty} r^{-2}D\xi_r = c^2 \sum_{r=1}^{\infty} g_r r^{-1}\varphi(r) < c^2 \sum_{r=1}^{\infty} g_r = c^2\Big(\sum_{r\varepsilon I_1} + \sum_{r\varepsilon I_2}\Big)g_r \leqq 2c^2 \sum_{r=1}^{\infty} r^{-2} = c^2\pi^2/3.$$

The desired result follows from Theorem 2.

5. *The series* (10) *diverges.* Easily we obtain $E_r = \omega_1^{(r)} \cup \omega_3^{(r)}$, $E'_{r-1} = \omega_2^{(r-1)}$, $P(E_r \mid E'_{r-1}) = 2v_r$ and

$$\sum_{r=1}^{\infty} P(E_r \mid E'_{r-1}) > 2 \sum_{r\varepsilon I_1} v_r = 2 \sum_{s=1}^{\infty} s^{-2} n_s [\varphi(n_s)]^{-1}$$

which diverges because of (23).

**5. Remarks.**

A. *Remarks concerning Theorem* 1.

$A_1$. *The Condition* (6) *is verified for*

$$\alpha_i = i^{-\beta} \prod_{j=1}^{n} [f(p_j\,; i)]^{-\gamma_j}, \beta + \sum_{j=1}^{n} \gamma_j < 2 \qquad i \,\varepsilon\, I$$

*where* $\beta > 0, \gamma_j > 0$, $p_j \,\varepsilon\, I$ *and* $f(p_j\,; i)$ *is the iterated logarithm of the order* $p_j$ *of* $i$ *in a basis greater than* 1. Let us set $u = (m + 1) \log l$, $v = m \log l$, $F(m, p) = f(p; ls)\, (f(p; s))^{-1} = f(p - 1; u)\, (f(p - 1; v))^{-1}$, $s = l^m$. For each $p \,\varepsilon\, I$ there exists a $m_0(p) \,\varepsilon\, I$ so that for $m \geqq m_0$, $F(m, p) \leqq l$. Indeed, for $p = 1$, $F(m, 1) = uv^{-1} = m^{-1}(m + 1) < 2 \leqq l^2$; if $F(m, p) \leqq l$, i.e. $f(p - 1; u) \leqq lf(p - 1; v) < (f(p - 1; v))^l$, it follows that $f(p; u) \leqq lf(p; v)$; i.e. $F(m, p + 1) \leqq l$. Now if we set $\gamma = \sum_{i=1}^{n} \gamma_i$ we obtain

$$\alpha^{(ls)} = \alpha^{(s)}\, l^{-\beta} \prod_{i=1}^{n} [F(m, p_i)]^{-\gamma_i} \geqq l^{-(\beta+\gamma)}\, \alpha^{(s)}$$

which concludes our proof.

$A_2$. *The Condition* (6) *generalizes the conditions in the papers* [12], [13]. Indeed, the conditions in these references require the existence of a constant $k (0 < k < 1)$ for which

$$l^2 k \alpha^{(ls)} \geqq \alpha^{(s)}, \qquad m \,\varepsilon\, I, \quad s = l^m. \tag{24}$$

But, if (6) is satisfied with a determined constant $A > l^{-2}$, then (24) is also satisfied with any $k$ for which $(Al^2)^{-1} \leqq k < 1$; conversely, if (24) is verified with a determined $k (0 < k < 1)$, it follows that (6) is also verified with $A \leqq (l^2 k)^{-1}$.

$A_3$. *If* $n\alpha^{(n)}$ *is monotonically increasing to infinity, then* (6) *is satisfied with* $A = l^{-1}$; *in this particular case Theorem* 1 *can be easily obtained.* Indeed if $y_n = [n\alpha^{(n)}]^{-1} D\xi_n$, $K_1 = (20K)^{-1}$, $\epsilon_1 = K_1 \epsilon n$, from (11) it follows that

$$P(\mathfrak{U}_n > n\epsilon) < (K_1\epsilon)^{-2}\, n^{-1} \sum_{i=1}^{n} [n\alpha^{(n)}]^{-1} D\xi_i < (K_1\epsilon)^{-2}\, n^{-1} \sum_{i=1}^{n} y_i\,.$$

Because

$$\sum_{n=1}^{\infty} n^{-1} y_n = \sum_{n=1}^{\infty} [n^2 \alpha^{(n)}]^{-1} D\xi_n$$

by means of ([3], 47, Chapter IX, Theorem 3), it follows from (7) that $n^{-1} \sum_{i=1}^{n} y_i \to 0\ (n \to \infty)$, i.e. the N.S.L. holds.

$A_4$. *If* $n\alpha^{(n)} \to \infty\ (n \to \infty)$ *and* $D\xi_i \leqq c < +\infty\ (i \,\varepsilon\, I)$, *then, instead of* (5) *and* (7) *we may take respectively*

$$\sum_{i=1}^{\infty} \omega_i^{-2} < +\infty, \qquad \sum_{n=1}^{\infty} [n^2 \alpha^{(n)}]^{-1} < +\infty. \tag{25}$$

If $\alpha_i > \rho > 0$ $(i \varepsilon I)$ (e.g. the chain is homogeneous $\alpha_i = \rho > 0$, $(i \varepsilon I))$, $D\xi_i \leqq c < +\infty$ $(i \varepsilon I)$, because $\omega_i^2 \geqq \rho l^{-2}(1 - l^{-2})^{-1} i^2 \geqq \rho l^{-2} i^2$ $(i \varepsilon I)$ follows the N.S.L. If the chain is discrete, $n\alpha^{(n)} \to \infty$ $(n \to \infty)$ (e.g., $\alpha_i > \rho > 0$, $i \varepsilon I$) $p_k^i$ the probability of the appearance of the event $i$ in the $k$th trial, $\mu_i$ the number of occurrences of $i$ in the first $n$ trials and (25) is verified, then $n^{-1} \mu_i - n^{-1} \sum_{k=1}^{n} p_k^i$ converges to zero with probability 1; if $\alpha_i = \rho > 0$ $(i \varepsilon I)$, $p_k^i = p^i$ then $n^{-1} \mu_i$ converges to $p^i$ with probability 1. This follows from Theorem 1 taking $\xi_k$ equal to the number of occurrences of $i$ in the $k$th trial.

B. *Remarks concerning Theorem* 2.

$B_1$. *Let* $1 < l \varepsilon I$, $s = l^r$, $\zeta_r = s^{-1}(S_{ls} - S_{s+1})$ $(r \varepsilon I)$. *If* $\alpha_i > \rho > 0$ $(i \varepsilon I)$, *then the condition* $\sum_{i=1}^{\infty} D\zeta_i < \infty$ *is sufficient for the* N.S.L. *of* $\xi_i(i \varepsilon I)$. *This condition is the best in terms of* $\zeta_i(i \varepsilon I)$, *in the sense that it is possible to construct a non degenerated Markov chain* $\xi_i(i \varepsilon I)$ *for which this series diverges,* $\alpha_i > \rho > 0$ $(i \varepsilon I)$ *and which does not verify the* N.S.L. From (3), it follows that

$$C' \sum_{i=s+1}^{ls} i^{-2} D\xi_i \leqq C' \sum_{i=s+1}^{ls} (s+1)^{-2} D\xi_i \leqq D\zeta_r \leqq Cl^2 \sum_{i=s+1}^{ls} i^{-2} D\xi_i$$

$$C' \sum_{i=l+1}^{\infty} i^{-2} D\xi_i \leqq \sum_{r=1}^{\infty} D\zeta_r \leqq Cl^2 \sum_{i=l+1}^{\infty} i^{-2} D\xi_i \tag{26}$$

and by means of Theorem 2 we obtain the first result.

For the second part we may use here the Markov chain constructed in the proof of Theorem 2. Indeed, it does not satisfy the N.S.L. Since in the proof of Theorem 2 we have proved that the series (8) diverges, then the series here diverges also because of (26).

$B_2$. *If the sequence of arbitrarily dependent random variables* $\xi_i(i \varepsilon I)$ *verifies the* S.L. *then* $\zeta_i(i \varepsilon I)$ *is strongly stable.* Set $s = l^r$, $s' = ls$, $a_r = s^{-1}(s' m\sigma_{s'} - sm\sigma_s)$, $\tau_r = \sigma_s - m\sigma_s$, $\tau_{r+1} = \sigma_{s'} - m\sigma_{s'}$, $R_1 = \{\tau_r \to 0, r \to \infty\} = \{\tau_r \to 0, \tau_{r+1} \to 0, r \to \infty\}$, $R_2 = \{l\tau_{r+1} - \tau_r \to 0, r \to \infty\}$. Under our conditions, $\sigma_n(n \varepsilon I)$ defined by (2) is strongly stable; i.e., $P(\sigma_n - m\sigma_n \to 0, n \to \infty) = 1$ and therefore $P(R_1) = 1$; because $R_1 \subset R_2$, $l\tau_{r+1} - \tau_r = \zeta_r - a_r$ it follows that $P(R_2) = 1$, i.e. $\zeta_i(i \varepsilon I)$ is strongly stable.

C. *Remarks concerning Theorem* 5.

$C_1$. If the variables $\xi_i(i \varepsilon I)$ are $\Omega$-identically distributed, $\alpha_i > \rho > 0$ $(i \varepsilon I)$ [e.g. (1) independent variables, (2) the $\xi_i(i \varepsilon I)$ are identically distributed and the chain is homogeneous] in order that they verify the N.S.L. it is necessary and sufficient that in the same class with this sequence there exists some variable $\xi$ having a finite expectation. This follows from Theorem 5 if $\beta = 0$.

$C_2$. In Theorem 5 and remark $C_1$, we do not assume the existence of $D\xi_i(i \varepsilon I)$.

D. *General remarks.*

$D_1$. We have supposed everywhere in this paper that $n\alpha^{(n)} \to \infty$ $(n \to \infty)$, because in the other case the obtained results are not interesting (see [9], p. 444, Remark 3).

$D_2$. Theorems 1, 2, 5, Lemma 1 and some facts in the Remark $A_4$ generalize the results of Kolmogorov ([4]–[6]), Theorems 3, 4 and the Remarks $B_1$, $B_2$ those of Prokhorov, and Lemma 2 and some facts in the Remark $A_4$ those of Borel and Cantelli.

$D_3$. The totality of homogeneous chains with $\alpha > 0$ (the end of remark $A_4$) intersects with the totalities of chains which satisfy the conditions from ([8], p. 364; [14], p. 59, 61, 166).

## REFERENCES


[1] Dobrushin, R. L. (1956). The central limit theorem for non-homogeneous Markov chains. *Teor. Veroyatnost. i Primenen.* **1** 72–89, 365–425.

[2] Dynkin, E. B. (1954). On some limit theorems for Markov chains. *Ukrain. Mat. Ž.* **6** 21–29.

[3] Halmos, P. R. (1950). *Measure Theory.* Van Nostrand, New York.

[4] Kolmogorov, A. N. (1930). Sur la loi forte des grands nombres. *C. R. Acad. Sci. Paris* **191** 910–913.

[5] Kolmogorov, A. N. (1928 and 1929). Ueber die Summen durch den Zufall bestimmter unabhängiger Grössen. *Math. Ann.* **99** 309–319; **102** 484–488.

[6] Kolmogorov, A. N. (1933). *Grundbegriffe der Wahrscheinlichkeitsrechnung.* Ergebnisse der Mathematik.

[7] Prokhorov, Iu V. (1950). On the strong law of large numbers. *Izv. Akad. Nauk SSSR.* **14** 523–536.

[8] Romanovski, V. I. (1949). *Discrete Markov Chains.* Moscow.

[9] Rosenblatt-Roth, M. (1963). Some theorems concerning the law of large numbers for non-homogeneous Markov chains. *Wahrscheinlichkeitstheorie und angewandte Gebiete* **1** 433–445.

[10] Rosenblatt-Roth, M. (1963). Sur la dispersion des sommes de variables aléatoires enchaînées. *C. R. Acad. Sci. Paris* **256** 5499–5501.

[11] Rosenblatt-Roth, M. (1961). On the strong law of large numbers for non-homogeneous Markov chains. *Dokl. Akad. Nauk. SSSR.* **141** 1310–1312.

[12] Rosenblatt-Roth, M. (1962). On the law of large numbers and on the strong law of large numbers for non-homogeneous Markov chains. *Dokl. Akad. Nauk. SSSR.* **147** 1294–1295.

[13] Rosenblatt-Roth, M. On some problems of random variables connected in a Markov chain. *Transactions of the Third Prague Conference on Information Theory, Statistical Decision Functions, Random Processes.* (In print.)

[14] Sarymsakov, T. A. (1954). *The Fundaments of the Theory of Markov Processes.* Moscow.