ON THE LAW OF LARGE NUMBERS FOR NON-HOMOGENEOUS MARKOV CHAINS

1960

SovietRxiv

Abstract

Full Text

MATHEMATICS

M. Rosenblatt-Rot

ON THE LAW OF LARGE NUMBERS FOR NON-HOMOGENEOUS MARKOV CHAINS

(Presented by Academician A. N. Kolmogorov, 6 IV 1960)

Let $(\mathfrak{A}_i, \mathfrak{S}_i, \mu_i)$ be a measure space, $x_i \in \mathfrak{A}_i$, $A_i \in \mathfrak{S}_i$, $i \in I$ (the set of all natural numbers), $\theta = \{i_1, i_2, \dots, i_r\} \subset I$, $(\mathfrak{A}^\theta, \mathfrak{S}^\theta, \mu^\theta) = X_{i\in\theta}(\mathfrak{A}_i, \mathfrak{S}_i, \mu_i)$, $x^\theta = (x_{i_1}, x_{i_2}, \dots, x_{i_r}) = \{x_i,\ i \in \theta\} \in \mathfrak{A}^\theta$, $A^\theta \in \mathfrak{S}^\theta$. Let the transition probability functions $\mathcal{P}_i(x_i, A_{i+1})$ with domains of definition $(\mathfrak{A}_i, \mathfrak{S}_i, \mathfrak{A}_{i+1}, \mathfrak{S}_{i+1})$ $(i \in I)$ define a Markov chain, and let $\alpha_{ij} = \alpha(P_{ij})$ be the coefficient of ergodicity of the transition function $P_{ij}(x_i, A_j)$ over the time interval (i, j) $(1 \leq i < j)$ ([1,2]).

Consider random variables Z_i depending on a finite number $m \geq 1$ of time instants, i.e., respectively on $x^{\theta_i} = \{x_{i-m+1}, x_{i-m+2}, \dots, x_i\} = \{x_k,\ k \in \theta_i\}$, where $\theta_i = \{\tau,\ i-m+1 \leq \tau \leq i\}$, with $DZ_i < \infty^*$ $(i \in I)$. Let $1-\eta_n = O(n^{-\beta})$ $(0 \leq \beta < 1)$; $\eta_n = \max_{1\leq i\leq n-1}(1-\alpha_{i,i+1})$.

Theorem 1. In order that the sequence of random variables Z_i $(i \in I)$, connected into a Markov chain with nonzero coefficients of ergodicity $\alpha_{i,i+1} > 0$, $(i \in I)$, and depending on $m \geq 1$ time instants, should obey the law of large numbers, it is sufficient that

$$\lim_{n\to\infty} \frac{1}{n^{2-\beta}} \sum_{i=1}^{n} DZ_i = 0.$$

If $\alpha_{i,i+1} > \rho > 0$ $(i \in I)$, this condition becomes the condition

$$\lim_{n\to\infty} \frac{1}{n^2} \sum_{i=1}^{n} DZ_i = 0.$$

Theorem 2. If, under the conditions of Theorem 1, $DZ_i \leq C < \infty$ $(i \in I)$, then the sequence Z_i $(i \in I)$ obeys the law of large numbers.

Theorem 3. If, in a discrete Markov chain with nonzero coefficients of ergodicity $\alpha_{i,i+1} > 0$ $(i \in I)$, the probability of occurrence of event i in the k-th trial is $p_k^{(i)}$, and $\mu^{(i)}$ is the number of occurrences of event i in the first n trials, then for any $\varepsilon > 0$

$$\lim_{n\to\infty} P\left\{\left|\frac{\mu^{(i)}}{n} - \frac{p_1^{(i)} + \cdots + p_n^{(i)}}{n}\right| > \varepsilon\right\} = 0.$$

Let X_k be equal to Z_k when $|Z_k| < \delta n$ and equal to zero when $|Z_k| \geq \delta n$; let Y_k be equal to Z_k when $|Z_k| \geq \delta n$ and equal to zero when $|Z_k| < \delta n$ $(k = 1, 2, \dots, n)$.

Theorem 4. If the random variables Z_i $(i \in I)$, connected into a Markov chain with $\alpha_{i,i+1} > \rho > 0$ $(i \in I)$ and depending on $m \geq 1$ time instants, possess finite mathematical expectations $a_i = MZ_i$,

---

* M denotes mathematical expectation, D denotes variance.

moreover, $\mathbf{M}|Z_i| < b < \infty$ $(i \in I)$ and $\mathbf{M}Y_i$ converges uniformly to zero as $n \to \infty$, then for any $\varepsilon > 0$

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} Z_i - \frac{1}{n}\sum_{i=1}^{n} a_i\right| > \varepsilon\right\} = 0.$$

Theorem 5. If the random variables Z_i $(i \in I)$ are identically distributed, are connected in a homogeneous Markov chain with ergodicity coefficient $\alpha > 0$, depend on $m \geq 1$ moments of time, and possess a finite expectation $a = \mathbf{M}Z_i$, then for any $\varepsilon > 0$

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} Z_i - a\right| > \varepsilon\right\} = 0.$$

Theorem 6. Let μ_i be the number of occurrences of event i in n successive trials according to the law of a homogeneous Markov chain with ergodicity coefficient $\alpha > 0$, and let p_i be the probability of occurrence of event i in each of the trials. Then for any $\varepsilon > 0$

$$\lim_{n\to\infty} P\left\{\left|\frac{\mu_i}{n} - p_i\right| > \varepsilon\right\} = 0.$$

For the proof of Theorem 1 we need several lemmas.
It is clear that the random variables Z_i are connected in a simple Markov chain for which a moment of time is an interval containing m consecutive ordinary moments of time; denote by $\mathcal{P}^*_{ij}(x^{\theta_i}, A^{\theta_j})$ the transition probability function of this new chain from moment θ_i to moment θ_j, and by $\alpha(\mathcal{P}^*_{ij}) = \alpha^*_{ij}$ its ergodicity coefficient.

Lemma 1. The random variables Z_i $(i \in I)$ are connected in a Markov chain with ergodicity coefficients $\alpha^*_{ij} = \alpha_{i,j-m+1}$ for $i < j - m + 1$, and $\alpha^*_{ij} = 0$ for $j - m + 1 \leq i < j$.

Lemma 2. For sufficiently large n,

$$\mathbf{D}\left(\sum_{i=1}^{n} Z_i\right) \leq C\eta_n^{1-m/2}\left(1-\eta_n^{1/2}\right)^{-1}\sum_{i=1}^{n}\mathbf{D}Z_i,$$

where $C = 16(1+\sqrt{6})$ for $m > 1$, and $C = 8(1+\sqrt{6})$ for $m = 1$.

Let Z', Z'' be two random variables connected in a Markov chain with two moments of time, given by a transition probability function (or stochastic matrix) with ergodicity coefficient α, and suppose that $\mathbf{M}Z' = \mathbf{M}Z'' = 0$, $\mathbf{D}Z' < \infty$, $\mathbf{D}Z'' < \infty$.

Lemma 3. $|\mathbf{M}Z'Z''| \leq K\sqrt{1-\alpha}\,(\mathbf{D}Z' + \mathbf{D}Z'')$, $K = 1+\sqrt{6}$.

The proof of the first part of Theorem 1 follows from Markov' s condition for the law of large numbers, if one observes that when $1-\eta_n = O(n^{-\beta})$ one has

$$\eta_n^{1-\frac{m}{2}}\left(1-\eta_n^{1/2}\right)^{-1} = O(n^{\beta})^*.$$

If $\alpha_{i,i+1} > 0$ $(i \in I)$ and $\beta = 0$, i.e. $1-\eta_n = O(1)$, then $1-\alpha_{i,i+1} < \eta_n < 1-\rho < 1$ $(n \in I,\ \rho > 0)$, whence $\alpha_{i,i+1} > \rho > 0$ $(i \in I)$; conversely, if $\alpha_{i,i+1} > \rho > 0$, it follows that $\beta = 0$, which completes the proof of Theorem 1. We note that the case $\rho = 1$, i.e. $\alpha_{i,i+1} = 1$ $(i \in I)$, corresponds to the case in which the Markov chain degenerates into a sequence-

---

* If $\beta \geq 1$, then from the known inequality

$$\mathbf{D}\left(\sum_{i=1}^{n} Z_i\right) \leq n\sum_{i=1}^{n}\mathbf{D}Z_i$$

for arbitrarily dependent Z_i $(i = 1, \dots, n)$, the result is obtained better than with the aid of Lemma 2.

of independent random variables (see (1), p. 78), i.e., the Markov condition. Let us also note that the conditions of Theorem 1 do not depend on $m \geqslant 1$.

Theorems 2 and 3 are generalizations of the classical theorems of Chebyshev and Poisson for independent random variables and are obtained in the usual way from Theorem 1.

Let us prove Theorem 4. Let

$$P\{Z_k(x^0_{h_k}) < \omega\} = F_k(\omega), \qquad \mathbf{M}X_k = \int_{|\omega|<\delta n}\omega\, dF_k(\omega).$$

Consider the random events

$$A = \left\{ \left| \sum_{k=1}^{n} X_n - \sum_{k=1}^{n} \mathbf{M} Z_k \right| \geqslant n\varepsilon \right\}, \qquad B = \left\{ \left| \sum_{k=1}^{n} X_k - \sum_{k=1}^{n} \mathbf{M} X_k \right| \geqslant n\varepsilon/2 \right\},$$

$$G_k = \{Y_k \neq 0\}, \qquad G = \left\{ \sum_{k=1}^{n} Y_k \neq 0 \right\} \subset \bigcup_{k=1}^{n} G_k,$$

$$E = \left\{ \left| \sum_{k=1}^{n} Z_k - \sum_{k=1}^{n} \mathbf{M} Z_k \right| \geqslant n\varepsilon \right\}.$$

If $\varepsilon > 0,\ \delta > 0$ are given, then, by the conditions of the theorem, there exists a number $n(\varepsilon, \delta)$ such that for $n > n(\varepsilon, \delta)$, $\mathbf{M}|Y_k| < \min(\varepsilon, \delta^2)$ $(k = 1, 2, \dots, n)$.

From

$$\left| \sum_{k=1}^{n} X_k - \sum_{k=1}^{n} \mathbf{M} Z_k \right| \leqslant \left| \sum_{k=1}^{n} X_k - \sum_{k=1}^{n} \mathbf{M} X_k \right| + \sum_{k=1}^{n} \mathbf{M}|Y_k|$$

it is seen that $A \subset B$ for $n > n(\varepsilon, \delta)$; on the other hand,

$$\mathbf{D} X_k < \mathbf{M} X_k^2 < \delta n \, \mathbf{M}|X_k| < \delta n b,$$

and, taking Lemma 2 into account, we obtain

$$P(A) \leqslant P(B) < \frac{1}{n^2} \mathbf{D} \left(\sum_{k=1}^{n} X_k \right) < \frac{C}{n^2} \sum_{k=1}^{n} \mathbf{D} X_k < C\delta b.$$

For $n > n(\varepsilon, \delta)$ it is easily obtained that

$$P(G) \leqslant \sum_{k=1}^{n} P(G_k) \leqslant \frac{1}{\delta n} \sum_{k=1}^{n} \mathbf{M}|Y_k| < \delta,$$

so that from

$$\left| \sum_{k=1}^{n} Z_k - \sum_{k=1}^{n} \mathbf{M} Z_k \right| \leqslant \left| \sum_{k=1}^{n} X_k - \sum_{k=1}^{n} Z_k \right| + \sum_{k=1}^{n} |Y_k|$$

it follows that $E \subset A \cup G$, $\quad P(E) \leqslant P(A) + P(G) < \delta(Cb + 1)$, which proves the theorem.

If the number of states l is finite, then from ([2]), (3.16), p. 371, it is seen that, in order that $\alpha = 0$, it is necessary and sufficient that there exist at least two states i, j $(i \neq j)$ such that for any $k = 1, 2, \ldots, l$ the equality $p_{ik} p_{jk} = 0$ holds, i.e., that in the matrix formed only by the rows i, j, in each column there be at least one zero.

Consequently, in the Markov case studied in ([4,5]), when b_r $(r = 1, 2, \ldots, l)$ are the possible values of the random variable Z_i $(i \in I)$ and in gen-

transition matrix there exists at least one column k without zeros, there is an $\alpha > 0$ such that the result mentioned is contained as a special case in Theorem 4, with $\alpha_{i,i+1} = \alpha$ $(i \in I)$, $Y_k = 0$ for sufficiently large n,

$$b = \sum_{r=1}^{l} |b_r|.$$

(In (4) the existence of the variance Z_i is required, which is not needed here.)

Theorem 5 is a special case of Theorem 4, and both generalize the well-known theorem of A. Ya. Khinchin ([3]) for independent random variables. Theorem 6 is a special case of Theorem 5 and generalizes the classical Bernoulli theorem for independent random variables.

Theorem 1 was obtained by the author earlier in ([8]) and was used in the proof of Theorem 6 from ([6]) and Theorem 4 from ([7]).

Faculty of Mathematics and Physics
Parhon University

Mathematical Institute
of the Romanian Academy of Sciences
Bucharest, Romania

Received
19 XII 1959

REFERENCES

1. R. L. Dobrushin, *Theory of Probability and Its Applications*, **1**, no. 1, 72 (1956).

2. R. L. Dobrushin, *Theory of Probability and Its Applications*, **1**, no. 4, 365 (1956).

3. A. Khintchine, C. R., **188**, 477 (1929).

4. A. A. Markov, "Extension of the law of large numbers to quantities depending on one another," in: *Selected Works*, Publishing House of the

Academy of Sciences of the USSR, 1951, pp. 339–361.

5. N. Sapogov, "On the paper 'Extension of the law of large numbers to quantities depending on one another,' " in: A. A. Markov, *Selected Works*, Publishing House of the Academy of Sciences of the USSR, 1951, pp. 660–662.

6. M. Rosenblatt-Roth, DAN, **112**, no. 1, 16 (1957).

7. M. Rosenblatt-Roth, DAN, **112**, no. 2, 202 (1957).

8. M. Rosenblatt-Roth, Candidate dissertation, Moscow State University, 1956.

Note: Figure translations are in progress. See original paper for figures.