A LOCAL LIMIT THEOREM FOR THE NUMBERS OF TRANSITIONS IN A MARKOV CHAIN AND ITS APPLICATIONS

MATHEMATICS

1966

SovietRxiv

Abstract

Full Text

UDC 519.217

MATHEMATICS

Corresponding Member of the Academy of Sciences of the USSR N. V. SMIRNOV,
O. V. SARMANOV, V. K. ZAKHAROV

A LOCAL LIMIT THEOREM FOR THE NUMBERS OF TRANSITIONS IN A MARKOV CHAIN AND ITS APPLICATIONS

A simple homogeneous Markov chain with $s+1$ states $E_i,\ i=1,2,\dots,s+1$, and a positive matrix of transition probabilities $\{p_{ij}\}$, $p_{ij}>0,\ i,j=1,2,\dots,s+1$, is considered. By $p_i^{(1)}$ we denote the initial probabilities of E_i and suppose that $p_i^{(1)}>0$. According to (1), the matrix $\{p_{ij}\}$ will be simple and regular, and the chain will be ergodic; in this case there exists a unique set of positive final probabilities $p_i,\ i=1,2,\dots,s+1$, independent of the initial probabilities $p_i^{(1)}$;

$$\sum_{i=1}^{s+1} p_i^{(1)} = \sum_{i=1}^{s+1} p_i = 1.$$

Let m_i denote the number of occurrences of E_i in n trials, and let m_{ij} denote the number of transitions from state E_i to state E_j in the same n trials.

1. Let us compute the number K of distinct chains of length n, composed of $s+1$ states, having a prescribed set of numbers of transitions m_{ij} from E_i to E_j, beginning with state E_{i_0} and ending with state E_{j_0}.

The numbers m_i and m_{ij} satisfy the conditions

$$\sum_{i=1}^{s+1} m_i = n, \qquad \sum_{i,j=1}^{s+1} m_{ij} = n-1$$

and are connected by the relations

$$\sum_{j=1}^{s+1} m_{ij} = m_i \quad \text{for } i \neq j_0; \qquad \sum_{j=1}^{s+1} m_{j_0 j} = m_{j_0} - 1; \tag{1}$$

$$\sum_{i=1}^{s+1} m_{ij} = m_j \quad \text{for } j \neq i_0; \qquad \sum_{i=1}^{s+1} m_{ii_0} = m_{i_0} - 1,$$

in consequence of which, among the $(s+1)(s+2)$ variables m_i and m_{ij}, only s^2+s+1 will be independent; as independent variables we shall choose $m_i,\ i = 1, 2, \dots, s+1$, and $m_{ij},\ i, j = 1, 2, \dots, s$.

We shall carry out the computation of K by induction on the number of states of the chain. Let

$$K^{(s)}_{i_0 j_0}(m_1, m_2, \dots, m_s; k_{ij};\ i, j = 1, 2, \dots, s-1)$$

be the number, known to us, of chains from s states with fixed numbers m_i of occurrences of all $E_i,\ i = 1, 2, \dots, s$, and numbers of transitions $k_{ij},\ i, j = 1, 2, \dots, s-1$ (k_{is} and k_{sj} are uniquely expressed in terms of only the indicated parameters). From these chains we construct new chains into which the state E_{s+1} will enter m_{s+1} times and in which a fixed set of transitions $m_{ij},\ i, j = 1, 2, \dots, s$, from state E_i to state E_j, will be obtained.

Consider an arbitrary chain composed of s states. Among its elements we must in a certain way place series of states E_{s+1}. Among all pairs of adjacent states $E_i E_j,\ i, j = 1, 2, \dots, s$, and the number of such pairs is k_{ij}, choose γ_{ij} pairs which we shall separate by a series of states E_{s+1}; the number of such choices is equal to $C_{k_{ij}}^{\gamma_{ij}}$. Thus, m_{s+1} occurrences of the state E_{s+1} must be divided into

$$\gamma = \sum_{i,j=1}^{s} \gamma_{ij}$$

series,

which can be done in a number of ways equal to $C_{m_{s+1}-1}^{\gamma-1}$. Note that when γ_{ij} series of states E_{s+1} are placed between the components of the pairs $E_i E_j$, the number of these pairs will decrease by γ_{ij}, and precisely this new number of remaining pairs $k_{ij} - \gamma_{ij}$ must be equal to m_{ij} in the new chain; hence

$$k_{ij} = m_{ij} + \gamma_{ij}, \qquad i, j = 1, 2, \dots, s. \tag{2}$$

Since

$$\sum_{i,j=1}^{s} k_{ij} = \sum_{i=1}^{s} m_i - 1,$$

we have

$$\gamma = \sum_{i=1}^{s} m_i - 1 - \sum_{i,j=1}^{s} m_{ij} = \text{const.} \tag{3}$$

For fixed γ_{ij}, the number of new chains obtained from the old ones by adding m_{s+1} states E_{s+1} will be equal to

$$K_{i_0 j_0}^{(s)}(m_1, m_2, \dots, m_s;\, m_{ij} + \gamma_{ij};\, i,j = 1,2,\dots,s-1) C_{m_{s+1}-1}^{\gamma-1} \prod_{i,j=1}^{s} C_{m_{ij}+\gamma_{ij}}^{\gamma_{ij}}.$$

Summing over all γ_{ij}, we obtain the recurrence formula

$$K_{i_0 j_0}^{(s+1)}(m_1, m_2, \dots, m_{s+1};\, m_{ij};\, i,j = 1,2,\dots,s) =$$

$$= C_{m_{s+1}-1}^{\gamma-1} \sum_{\substack{\gamma_{ij} \\ i,j=1,2,\dots,s}} K_{i_0 j_0}^{(s)}(m_1, m_2, \dots, m_s;\, m_{ij}+\gamma_{ij};\, i,j = 1,2,\dots,s-1) \times \prod_{i,j=1}^{s} C_{m_{ij}+\gamma_{ij}}^{\gamma_{ij}}. \tag{4}$$

Let us also show how $K_{i_0 j_0}^{(2)}$ is found for two states E_{i_0}, E_{j_0} (with $i_0 \neq j_0$).

We divide the m_{i_0} states E_{i_0} and the m_{j_0} states E_{j_0} into the same number $m_{i_0} - m_{i_0 i_0}$ of ordered series (with such a number of series there will be exactly $m_{i_0 i_0}$ transitions from E_{i_0} to E_{i_0}, while the numbers of transitions $m_{i_0 j_0}, m_{j_0 i_0}$, and $m_{j_0 j_0}$ are fixed automatically). The ordered series, without disturbing their order, are placed into one chain, putting the series from E_{i_0} in the odd positions and the series from E_{j_0} in the even positions; then the chains will begin with E_{i_0} and end with E_{j_0}.

Thus the number of all such chains is

$$K_{i_0 j_0}^{(2)} = C_{m_{i_0}-1}^{m_{i_0} - m_{i_0 i_0} - 1} C_{m_{j_0}-1}^{m_{i_0} - m_{i_0 i_0} - 1}. \tag{5}$$

The quantity $K_{i_0 i_0}^{(2)}$ is found analogously.

By direct substitution one can verify that the following expression satisfies the recurrence formula (4) (for $i_0 \neq j_0$):

$$K = \frac{(m_{j_0} - 1)\prod_{i=1}^{s+1} m_i!}{m_{j_0} m_{s+1} \prod_{i,j=1}^{s+1} m_{ij}!} \begin{vmatrix} 1 - \frac{m_{11}}{m_1} & \dots & -\frac{m_{1i_0}}{m_1} & \dots & -\frac{m_{1j_0}}{m_1} & \dots & -\frac{m_{1s}}{m_1} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ -\frac{m_{i_01}}{m_{i_0}} & \dots & 1 - \frac{m_{i_0i_0}}{m_{i_0}} & \dots & -\frac{m_{i_0j_0}}{m_{i_0}} & \dots & -\frac{m_{i_0s}}{m_{i_0}} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ -\frac{m_{j_01}}{m_{j_0}-1} & \dots & -\frac{m_{j_0i_0}+1}{m_{j_0}-1} & \dots & 1 - \frac{m_{j_0j_0}-1}{m_{j_0}-1} & \dots & -\frac{m_{j_0s}}{m_{j_0}-1} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ -\frac{m_{s1}}{m_s} & \dots & -\frac{m_{si_0}}{m_s} & \dots & -\frac{m_{sj_0}}{m_s} & \dots & 1 - \frac{m_{ss}}{m_s} \end{vmatrix}. \tag{6}$$

For $i_0 = j_0$, the row of the determinant with this number takes the form

$$-\frac{m_{j_01}}{m_{j_0} - 1}, \dots, 1 - \frac{m_{j_0}}{m_{j_0} - 1}, \dots, -\frac{m_{j_0s}}{m_{j_0} - 1},$$

and the remaining rows remain unchanged.

Remark. An expression equivalent to (6) was obtained by elementary methods in (2); the formula for K given in (3) is erroneous.

2. Let $P_{i_0j_0}^{(n)}(\{m_{ij}\})$ be the probability of observing numbers of transitions forming the matrix $\{m_{ij}\}$, $i, j = 1, 2, \dots, s+1$, in n trials beginning with the occurrence of E_{i_0} and ending with the occurrence of E_{j_0}. Then

$$P_{i_0j_0}^{(n)}(\{m_{ij}\}) = K p_{i_0}^{(1)} \prod_{i,j=1}^{s+1} p_{ij}^{m_{ij}}, \tag{7}$$

where K is determined by formula (6). Putting

$$\begin{aligned} m_{ij} &= np_ip_{ij} + \xi_{ij}\sqrt{np_ip_{ij}}, && i, j = 1, 2, \dots, s+1, \\ m_i &= np_i + z_i\sqrt{np_i}, && i \neq j_0, \\ m_{j_0} - 1 &= np_{j_0} + z_{j_0}\sqrt{np_{j_0}}, && \\ \Delta\xi_{ij} &= 1/\sqrt{np_ip_{ij}}, && i, j = 1, 2 \dots, s, \\ \Delta z_i &= 1/\sqrt{np_i}, && i = 1, 2, \dots, s. \end{aligned} \tag{8}$$

we find, as a consequence of (1),

$$\sum_{j=1}^{s+1} \xi_{ij}\sqrt{p_{ij}} = z_i, \qquad i = 1, 2, \dots, s+1. \tag{9}$$

Moreover, the following limiting relations (obtained as $n \to \infty$) are valid:

$$\sum_{i=1}^{s+1} \xi_{ij} \sqrt{p_i p_{ij}} = z_j \sqrt{p_j}, \qquad j = 1, 2, \dots, s+1,$$

$$\sum_{i=1}^{s+1} z_i \sqrt{p_i} = 0. \tag{9'}$$

Regarding the variables ξ_{ij} and z_i as varying within finite limits, with the aid of Stirling’ s formula and taking (8) and (9) into account, we bring (7) to the form

$$P_{i_0 j_0}^{(n)}(\{m_{ij}\}) = p_{i_0}^{(1)} p_{j_0} C e^{-Q/2} \prod_{i,j=1}^{s} \Delta\xi_{ij} \prod_{i=1}^{s} \Delta z_i (1 + O(n^{-1/2})), \tag{10}$$

where

$$C = \frac{1}{(2\pi)^{(s^2+s)/2}} \left(\frac{\prod_{i=1}^{s} p_i}{p_{s+1}^{s+2} \prod_{i=1}^{s} p_{i,s+1} \prod_{j=1}^{s+1} p_{s+1,j}} \right)^{1/2} \left| \begin{matrix} 1 - p_{11} & -p_{12} & \dots & -p_{1s} \\ -p_{21} & 1 - p_{22} & \dots & -p_{2s} \\ \cdot & \cdot & \cdot & \cdot \\ -p_{s1} & -p_{s2} & \dots & 1 - p_{ss} \end{matrix} \quad . \right|, \tag{11}$$

$$Q = \sum_{i,j=1}^{s+1} \xi_{ij}^2 - \sum_{i=1}^{s+1} z_i^2 = \sum_{i,j=1}^{s+1} \left(\xi_{ij} - z_i \sqrt{p_{ij}} \right)^2. \tag{12}$$

Summing (10) over i_0 and j_0 from 1 to $s+1$, we obtain the probability of observing the transitions $\{m_{ij}\}$ in an arbitrary segment of n observations:

$$P^{(n)}(\{m_{ij}\}) = C e^{-Q/2} \prod_{i,j=1}^{s} \Delta\xi_{ij} \prod_{i=1}^{s} \Delta z_i (1 + O(n^{-1/2})). \tag{13}$$

Thus, we have proved the

Local limit theorem. *The numbers of transitions $\{m_{ij}\}$, under the normalization (8), have a joint asymptotically normal distribution with density $Ce^{-Q/2}$ with $s^2 + s$ normally correlated variables ξ_{ij}, $i, j = 1, 2, \dots, s$, z_i, $i = 1, 2, \dots, s$.*

* In (3), in the expression for Q, p_i is printed instead of p_{ij}.

3. The frequencies m_i/n and m_{ij}/m_i are naturally to be regarded as statistical estimates, respectively, of the final probabilities p_i and of the transition probabilities p_{ij}, obtained by the maximum-likelihood method.

Let us note that the random variable (12) has a χ^2 distribution with s^2+s degrees of freedom.

From (12) and (8) it follows that

$$Q = \sum_{i,j=1}^{s+1} \frac{(m_{ij} - m_i p_{ij})^2}{n p_i p_j} = \sum_{i,j=1}^{s+1} \left(\frac{m_{ij}}{m_i} - p_{ij} \right)^2 \frac{n p_i}{p_{ij}} + \eta_n = Q_1 + \eta_n, \qquad (14)$$

where η_n, as $n \to \infty$, tends to zero with probability one.

Considering an inequality of the form $Q_1 < L$ as a criterion for testing, from observations, the simple hypothesis H_0 fixing the probabilities $\{p_{ij}\}$, we conclude on the basis of (14) that, for large n, this criterion is asymptotically distributed according to the χ^2 law with s^2+s degrees of freedom. As always, the critical region is the region of large values of Q_1.

If, along with the hypothesis H_0, we allow the possibility of an alternative hypothesis H_1, close to H_0, under which the probabilities p_{ij} receive a certain displacement $c_{ij}/\sqrt{n}$ ($\sum_{j=1}^{s+1} c_{ij} = 0$), then, under the hypothesis H_1, the criterion Q_1 asymptotically follows a noncentral χ^2-distribution with s^2+s degrees of freedom and noncentrality parameter $\sum c_{ij}^2$. This circumstance can be used to estimate the power of the criterion and to estimate a sufficient number of observations for distinguishing the hypotheses.

Remark. One could also have refrained from assuming the positivity of all p_{ij}. The main conclusions of the paper remain valid if the matrix $\{p_{ij}\}$ is indecomposable and regular (1), since then there exists a unique set of final probabilities. The equality to zero of individual p_{ij} leads only to a corresponding reduction in the number of degrees of freedom.

V. A. Steklov Mathematical Institute
Academy of Sciences of the USSR

Received
8 XII 1965

REFERENCES

1. S. N. Bernstein, *Collected Works*, 4, Moscow, 1964, p. 455.

2. P. Whittle, *J. Roy. Stat. Soc.*, ser. B, **17**, 235 (1955).

3. N. V. Smirnov, *Vestn. LGU*, No. 11, 47 (1955).

4. A. N. Kolmogorov, *Izv. AN SSSR, ser. matem.*, **13**, 281 (1949).