Math-Net.Ru

Общероссийский математический портал

А. А. Minakov, Пуассоновская аппроксимация для числа неубывающих серий в цепях Маркова, *Матем. вопр. криптогр.*, 2018, том 9, выпуск 2, 103–116

DOI: <https://doi.org/10.4213/mvk254>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением

<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 3.80.181.102

26 апреля 2023 г., 16:52:39

Poisson approximation for the number of non-decreasing runs in Markov chains

A. A. Minakov

MIREA — Russian Technological University, Moscow

Получено 05.11.2017

Abstract. Let a sequence $X_1, X_2, \dots, X_n$ be a segment of a stationary irreducible and aperiodic Markov chain with state space $\mathcal{A} = \{1, \dots, N\}$, $N \geq 2$. We study the non-overlapping appearances of non-decreasing runs in the sequence $X_1, X_2, \dots, X_n$. By means of Stein method we estimate the total variation distance between the distribution of the number of non-overlapping appearances of non-decreasing monotone runs and the Poisson distribution. As a corollary we prove corresponding limit theorem.

Key words: non-decreasing runs, Poisson approximation, Stein method, Markov chain, Jordan normal form, upper triangular matrix

Пуассоновская аппроксимация для числа неубывающих серий в цепях Маркова

А. А. Минаков

МИРЭА — Российский технологический университет, Москва

Аннотация. Рассматривается последовательность $X_1, X_2, \dots, X_n$ случайных величин, являющаяся отрезком стационарной неприводимой и апериодической цепи Маркова с множеством состояний $\mathcal{A} = \{1, \dots, N\}$, $N \geq 2$. Изучаются непересекающиеся появления неубывающих серий в последовательности $X_1, X_2, \dots, X_n$. С использованием метода Стейна оценивается расстояние по вариации между распределением числа непересекающихся неубывающих серий и пуассоновским распределением. Как следствие доказывается соответствующая предельная теорема.

Ключевые слова: неубывающие серии, пуассоновская аппроксимация, метод Стейна, цепь Маркова, жорданова нормальная форма, верхнетреугольная матрица

Citation: *Mathematical Aspects of Cryptography*, 2018, v. 9, № 2, pp. 103–116 (Russian)

© Академия криптографии Российской Федерации, 2018 г.

1. Introduction

Let $X = (X_1, X_2, \dots, X_n)$ be a segment of stationary irreducible and aperiodic Markov chain $\{X_a\}_{a \in \mathbb{Z}}$ on a finite space $\mathcal{A} = \{1, \dots, N\}$, $N \geq 2$, with transition probability matrix $\mathbb{P} = (p_{i,j})_{i,j \in \mathcal{A}}$ with nonzero elements $p_{i,j} > 0$, $i, j \in \mathcal{A}$, and stationary distribution $\vec{\pi} = (\pi_i)_{i \in \mathcal{A}}$.

Let event $E_{t,s} = \{X_t > X_{t+1} \leq \dots \leq X_{t+s}\}$ denote the appearance of a non-decreasing run of length s ($s \geq 2$) beginning at a local minimum at time $t + 1$.

Let $I\{B\}$ be an indicator of the event B and the random variable

$$\xi_n(s) = \sum_{t=1}^n I\{E_{t,s}\} \quad (1)$$

be the number of non-overlapping appearances of non-decreasing runs of length $s \geq 2$ beginning at local minima of the sequence $X_1, X_2, \dots, X_n$.

In order to avoid edge effects and to facilitate calculations we assume that infinite sequence $\{X_a\}_{a \in \mathbb{Z}}$ is observed. Therefore at the ends of $X_1, X_2, \dots, X_n$ the structure of runs is the same.

In [5, 11] the non-overlapping appearances of increasing runs in random permutations were studied. In [10] the Poisson limit theorem for the number of increasing runs with length larger than or equal to some fixed length was proved. The Poisson limit theorem for increasing runs of fixed length was proved in [4]. In [8] the multidimensional normal theorem for the numbers of monotone tuples with fixed length in a random equiprobable sequence was formulated. In [9] the compound Poisson limit theorem was proved for the number of non-decreasing runs. In all these theorems only sequences of independent identically distributed random variables were considered.

In our paper we construct an estimate for the total variation distance between the distribution of random variable $\xi_n(s)$ and the Poisson distribution for stationary irreducible and aperiodic Markov chain $\{X_a\}_{a \in \mathbb{Z}}$. As a corollary we prove corresponding Poisson limit theorem.

2. Main results

Let us recall some definitions. By $d(\Phi, \Psi)$ we denote the total variation distance between distributions Φ and Ψ . For distributions Φ and Ψ on the set $\{0, 1, \dots\}$ (see, e. g., [3])

$$d(\Phi, \Psi) = \frac{1}{2} \sum_{m=0}^{\infty} |\Psi\{m\} - \Phi\{m\}|.$$

Let $\mathcal{L}(\zeta)$ denote a distribution of random variable ζ and $Po(\lambda)$ denote the Poisson distribution with parameter λ .

Let $\mathbb{M} = (m_{ij})_{i,j \in \mathcal{A}}$ be a matrix such that all entries below the main diagonal are zero and the other entries are equal to corresponding elements of the matrix $\mathbb{P}$ ($m_{ij} = p_{i,j}$ for $i \geq j$). The matrix $\mathbb{M}$ defines transition probabilities from $i \in \mathcal{A}$ to $j \in \mathcal{A}$ such that $i \leq j$.

By the Jordan canonical form theorem (see [6]) there exists an invertible matrix $\mathbb{T} = (t_{ij})_{i,j \in \mathcal{A}}$ (and $\mathbb{T}^{-1} = (t'_{ij})_{i,j \in \mathcal{A}}$) such that

$$\mathbb{M} = \mathbb{T} \cdot \mathbb{J} \cdot \mathbb{T}^{-1}, \quad (2)$$

where

$$\mathbb{J} = \text{Diag}\{\mathbb{J}_{d_1}(p_{i_1, i_1}), \mathbb{J}_{d_2}(p_{i_2, i_2}), \dots, \mathbb{J}_{d_q}(p_{i_q, i_q})\} \quad (3)$$

is a block diagonal matrix and $\mathbb{J}_{d_1}(p_{i_1, i_1}), \dots, \mathbb{J}_{d_q}(p_{i_q, i_q})$ are Jordan blocks of orders $d_1, \dots, d_q$ and $d_1 + \dots + d_q = N$. The matrix $\mathbb{J}$ has q blocks ($q \leq N$). Note that elements $p_{i_1, i_1}, \dots, p_{i_q, i_q}$ are located on the main diagonal of the matrix $\mathbb{P}$.

Choose matrix $\mathbb{T}$ such that the Jordan blocks are arranged from the largest eigenvalue to the smallest one and blocks with the same eigenvalue are arranged from the maximal to the minimal order. Suppose that $b \in \{1, \dots, q\}$ first Jordan blocks have the largest eigenvalue which is equal to p :

$$p_{i_1, i_1} = \dots = p_{i_b, i_b} \equiv p. \quad (4)$$

Let

$$\mathbb{K} = \mathbb{P} - \mathbb{M}. \quad (5)$$

The matrix $\mathbb{K}$ defines transition probabilities from $i \in \mathcal{A}$ to $j \in \mathcal{A}$ such that $i > j$.

Put

$$\vec{\alpha} = (\alpha_1, \dots, \alpha_N) = \vec{\pi} \mathbb{K} \mathbb{M}^{s-1}. \quad (6)$$

The sum of components of vector $\vec{\alpha}$ equals to the probability that a non-decreasing run of length s begins at local minimum at the given moment. Therefore, for any $t \in \{1, \dots, n\}$ and $s \geq 2$ we have

$$\mathbf{P}\{E_{t,s}\} = \alpha_1 + \dots + \alpha_N. \quad (7)$$

Let λ be the expectation of random value $\xi_n(s)$. From (7) it follows that

$$\lambda = \mathbf{E} \xi_n(s) = \sum_{t=1}^n \mathbf{E} \{E_{t,s}\} = n \cdot \mathbf{E} \{E_{t,s}\} = n \cdot (\alpha_1 + \dots + \alpha_N). \quad (8)$$

Let $p_{ij}^{(t)}$ denote the t -step transition probabilities of irreducible and aperiodic Markov chain $\{X_a\}_{a \in \mathbb{Z}}$ and

$$p_{(R)ij}^{(t)} = \pi_j p_{ji}^{(t)} \pi_i^{-1}$$

be the t -step transition probabilities of the reversed Markov chain $\{X_a^R\}_{a \in \mathbb{Z}}$ for $t \in \mathbb{N}$. Following a coupling argument (see [7]) we deduce

$$\max_{x \in \mathcal{A}} \max \left\{ \sum_{y \in \mathcal{A}} |p_{xy}^{(t)} - \pi_y|, \sum_{y \in \mathcal{A}} |p_{(R)xy}^{(t)} - \pi_y| \right\} \leq 2\varrho^t \quad \forall t \geq 1, \quad (9)$$

where

$$\varrho = 1 - \min \left\{ \sum_{y \in \mathcal{A}} \min_{x \in \mathcal{A}} p_{xy}, \sum_{y \in \mathcal{A}} \min_{x \in \mathcal{A}} p_{(R)xy} \right\}. \quad (10)$$

Now we are ready to state the first main result of this paper which is proved by means of the Stein method and results of [1].

Theorem 1. *If $X_1, X_2, \dots, X_n$ is a segment of stationary irreducible and aperiodic Markov chain $\{X_a\}_{a \in \mathbb{Z}}$ on a finite space $\mathcal{A} = \{1, \dots, N\}$, $N \geq 2$, with transition probability matrix $\mathbb{P} = (p_{i,j})_{i,j \in \mathcal{A}}$ with nonzero elements $p_{i,j} > 0$, $i, j \in \mathcal{A}$, and stationary distribution $\vec{\pi} = (\pi_i)_{i \in \mathcal{A}}$, then*

$$\begin{aligned} d(\mathcal{L}(\xi_n(s)), Po(\lambda)) &\leq \frac{(1 - e^{-\lambda})\lambda}{n} \left(2s + 1 + 2(r + 1) \max_{i \in \mathcal{A}} \{\pi_i^{-1}\} \right) + \\ &\quad + 2\lambda N^2 \varrho^{r+1} (N \varrho^{2s+r+1} + 2\varrho^{s+r+1} + \varrho^s + 1), \end{aligned} \quad (11)$$

where λ is defined in (8), ϱ is defined in (10), and parameter r may be equal to integer larger than s .

From Theorem 1 we deduce a limit theorem for a random variable (1).

Theorem 2. Let $X_1, X_2, \dots, X_n$ be a segment of stationary irreducible and aperiodic Markov chain $\{X_a\}_{a \in \mathbb{Z}}$ on a finite space $\mathcal{A} = \{1, \dots, N\}$, $N \in \{2, \dots, s\}$, with transition probability matrix $\mathbb{P} = (p_{i,j})_{i,j \in \mathcal{A}}$ with nonzero elements $p_{i,j} > 0$, $i, j \in \mathcal{A}$, and stationary distribution $\vec{\pi} = (\pi_i)_{i \in \mathcal{A}}$. If

$$n, s \rightarrow \infty, \quad \frac{s}{n} \rightarrow 0 \quad \text{and} \quad np^{s-N} \frac{(s-1)^{N-1}}{(N-1)!} \rightarrow Q \in (0, \infty), \quad (12)$$

then

$$\mathcal{L}(\xi_n(s)) \rightarrow Po(\lambda).$$

3. Proof of Theorems 1 and 2

Let $\Gamma = \{1, \dots, n\}$, $\{Y_i | i \in \Gamma\}$ be a set of random indicators defined on some common probability space and $W = \sum_{i \in \Gamma} Y_i$. For each $i \in \Gamma$ let Γ_i^s be a set of indices $j \in \Gamma \setminus \{i\}$ such that Y_j depends on Y_i strongly. We will use Theorem 2.5 from [2] on the estimate of the total variation distance between the distribution of sum W and the Poisson distribution with parameter $\theta = \mathbf{E} W$. The Theorem 2.5 follows from [1, Th. 2.N].

Theorem 2.5. Let $W = \sum_{i \in \Gamma} Y_i$, where $\{Y_i | i \in \Gamma\}$ are indicator variables. Let for each $i \in \Gamma$ the set $\Gamma \setminus \{i\}$ be splitted into two subsets Γ_i^s and Γ_i^w such that (informally)

$$\Gamma_i^s = \{j \in \Gamma \setminus \{i\} : Y_j \text{ strongly depend on } Y_i\},$$

$$\Gamma_i^w = \Gamma \setminus (\Gamma_i^s \cup \{i\}).$$

Let

$$Z_i = \sum_{j \in \Gamma_i^s} Y_j \quad \text{and} \quad W_i = \sum_{j \in \Gamma_i^w} Y_j.$$

Then

$$\begin{aligned} d(\mathcal{L}(W), Po(\theta)) \leq & \frac{1 - e^{-\theta}}{\theta} \sum_{i \in \Gamma} ((\mathbf{E} Y_i) \mathbf{E} (Y_i + Z_i) + \mathbf{E} (Y_i Z_i)) + \\ & + \max \left\{ 1, \sqrt{\frac{2}{e\theta}} \right\} \sum_{i \in \Gamma} \mathbf{E} |\mathbf{E} Y_i - \mathbf{E} (Y_i | W_i)|. \quad (13) \end{aligned}$$

We apply this estimate to random variable (1). For any $i \in \{1, \dots, n\}$ and $r > s$ we put $\Gamma_i^w = \{1, \dots, i - s - r - 1, i + s + r + 1, \dots, n\}$, $\Gamma_i^s = \{i - s - r, \dots, i + s + r\} \setminus \{i\}$,

$$\begin{aligned} A^* &= \sum_{i=1}^n \sum_{j \in \Gamma_i^s \cup \{i\}} \mathbf{P}\{E_{i,s}\} \mathbf{P}\{E_{j,s}\}, \\ B^* &= \sum_{i=1}^n \sum_{j \in \Gamma_i^s} \mathbf{P}\{E_{i,s}, E_{j,s}\}, \\ C^* &= \sum_{i=1}^n \mathbf{E} \left| \mathbf{P}\{E_{i,s}\} - \mathbf{E} \left(I\{E_{i,s}\} \left| \sum_{j \in \Gamma_i^w} I\{E_{j,s}\} \right| \right) \right|, \end{aligned}$$

then

$$d(\mathcal{L}(\xi_n(s)), Po(\lambda)) \leq \frac{1 - e^{-\theta}}{\theta} (A^* + B^*) + \max \left\{ 1, \sqrt{\frac{2}{e\theta}} \right\} C^*.$$

The value r in the definition of set Γ_i^s should be selected sufficiently large for $E_{i,s}$ and $E_{j,s}$ to be “weakly” dependent whenever $j \geq i + s + r$ or $i \geq j + s + r$. We note that condition $r > s$ appears due to the Markovian structure of the sequence $\{X_a\}$, which necessitates a larger neighborhood than in the case of independent identically distributed random variables $X_1, \dots, X_n$.

Consider a vector

$$\vec{\beta}^{(h)} = \left(\beta_1^{(h)}, \dots, \beta_N^{(h)} \right) = \vec{\pi} \mathbb{K} \mathbb{M}^{s-1} \mathbb{P}^h \mathbb{K} \mathbb{M}^{s-1}, \quad (14)$$

here the sum of vector components $\beta_1^{(h)} + \dots + \beta_N^{(h)}$ equals the probability of non-overlapping appearances of two non-decreasing runs with length s at the distance $h \geq 1$ between the end of one run and the beginning of other run.

We note that two non-decreasing runs beginning at different local minima have no common elements. In other words, if $I\{E_{i,s}\} = 1$, then $I\{E_{j,s}\} = 0$ for any $j \in \{i - s + 1, \dots, i + s - 1\}$. Hence

$$\mathbf{P}\{E_{i,s}, E_{j,s}\} = 0 \quad (15)$$

for any $j \in \{i - s + 1, \dots, i + s - 1\}$.

If $j < i - s$ or $j > i + s$, and $h = |i - j| - s$ is the distance between the end of one run and the beginning of the other run, and $m_{x,y}, k_{x,y}$ ($x, y \in \mathcal{A}$) are elements of corresponding matrices $\mathbb{M}$ and $\mathbb{K}$, then

$$\begin{aligned}
 \mathbf{P}\{E_{i,s}, E_{j,s}\} &= \beta_1^{(h)} + \dots + \beta_N^{(h)} = \\
 &= \sum_{l_1, \dots, l_{2s+h+2}=1}^N \pi_{l_1} k_{l_1, l_2} m_{l_2, l_3} \cdot \dots \cdot m_{l_s, l_{s+1}} p_{l_{s+1}, l_{s+2}} \cdot \dots \cdot p_{l_{s+h}, l_{s+h+1}} \times \\
 &\quad \times k_{l_{s+h+1}, l_{s+h+2}} m_{l_{s+h+2}, l_{s+h+3}} \cdot \dots \cdot m_{l_{2s+h+1}, l_{2s+h+2}} \leqslant \\
 &\leqslant \sum_{l_1, \dots, l_{s+1}=1}^N \pi_{l_1} k_{l_1, l_2} m_{l_2, l_3} \cdot \dots \cdot m_{l_s, l_{s+1}} \times \\
 &\quad \times \sum_{l_{s+h+1}, \dots, l_{2s+h+2}=1}^N \frac{\pi_{l_{s+h+1}}}{\pi_{l_{s+h+1}}} k_{l_{s+h+1}, l_{s+h+2}} \times \\
 &\quad \times m_{l_{s+h+2}, l_{s+h+3}} \cdot \dots \cdot m_{l_{2s+h+1}, l_{2s+h+2}} \leqslant \\
 &\leqslant \max_{y \in \mathcal{A}} \left\{ \frac{1}{\pi_y} \right\} \mathbf{P}\{E_{i,s}\} \mathbf{P}\{E_{j,s}\}.
 \end{aligned} \tag{16}$$

Using (7), (14), (15), and (16) we get

$$A^* = n(2s+1)(\alpha_1 + \dots + \alpha_N)^2, \tag{17}$$

$$\begin{aligned}
 B^* &= 2n \sum_{h=0}^r \left(\beta_1^{(h)} + \dots + \beta_N^{(h)} \right) \leqslant \\
 &\leqslant 2n(r+1) \max_{y \in \mathcal{A}} \left\{ \frac{1}{\pi_y} \right\} (\alpha_1 + \dots + \alpha_N)^2,
 \end{aligned} \tag{18}$$

where $r > s$ depends on the Markovian structure of the sequence $\{X_a\}$.

Next, we consider the term C^* and estimate the value

$$\mathbf{E} \left| \mathbf{P}\{E_{i,s}\} - \mathbf{E} \left(I\{E_{i,s}\} \left| \sum_{j \in \Gamma_i^w} I\{E_{j,s}\} \right. \right) \right|.$$

Using the Markov property and stationarity of chain $\{X_a\}_{a \in \mathbb{Z}}$, we obtain

$$\begin{aligned}
& \left| \mathbf{E} \left\{ \mathbf{P} \{E_{i,s}\} - \mathbf{E} \left(I \{E_{i,s}\} \middle| \sum_{j \in \Gamma_i^w} I \{E_{j,s}\} \right) \right\} \right| \leq \\
& \leq \left| \mathbf{E} \left\{ \mathbf{P} \{E_{i,s}\} - \mathbf{E} \left(I \{E_{i,s}\} \middle| X_1, \dots, X_{i-s-r-1}, X_{i+s+r+1}, \dots, X_n \right) \right\} \right| = \\
& = \left| \mathbf{E} \left\{ \mathbf{P} \{E_{i,s}\} - \mathbf{E} \left(I \{E_{i,s}\} \middle| X_{i-s-r-1}, X_{i+s+r+1} \right) \right\} \right| = \\
& = \sum_{z, w \in \mathcal{A}} \left| \mathbf{P} \{E_{i,s}\} \mathbf{P} \{X_{i-s-r-1} = z, X_{i+s+r+1} = w\} - \right. \\
& \quad \left. - \mathbf{P} \{I \{E_{i,s}\} = 1, X_{i-s-r-1} = z, X_{i+s+r+1} = w\} \right|.
\end{aligned}$$

We have

$$\mathbf{P} \{X_{i-s-r-1} = z, X_{i+s+r+1} = w\} = \pi_z p_{zw}^{(2s+2r+2)},$$

and if $m_{x,y}, k_{x,y}$ ($x, y \in \mathcal{A}$) are elements of corresponding matrices $\mathbb{M}$ and $\mathbb{K}$, then

$$\begin{aligned}
& \left| \mathbf{P} \{I \{E_{i,s}\} = 1, X_{i-s-r-1} = z, X_{i+s+r+1} = w\} - \right. \\
& \quad \left. - \mathbf{P} \{E_{i,s}\} \mathbf{P} \{X_{i-s-r-1} = z, X_{i+s+r+1} = w\} \right| = \\
& = \left| \sum_{l_1, \dots, l_{s+1}=1}^N \pi_{l_1} k_{l_1, l_2} m_{l_2, l_3} \cdot \dots \cdot m_{l_s, l_{s+1}} \left(p_{(R)l_1 z}^{(s+r+1)} p_{l_{s+1} w}^{(r+1)} - \pi_z p_{zw}^{(2s+2r+2)} \right) \right|
\end{aligned}$$

where $p_{(R)xz}^{(t)} = \pi_z p_{zx}^{(t)} / \pi_x^{-1}$ is the t -step transition probability of the reversed Markov chain. Furthermore, if we set

$$\varepsilon_{xy}^{(t)} = |p_{xy}^{(t)} - \pi_y| \quad \text{and} \quad \varepsilon_{(R)xy}^{(t)} = |p_{(R)xy}^{(t)} - \pi_y|,$$

then

$$\begin{aligned}
& \left| \mathbf{P} \{I \{E_{i,s}\} = 1, X_{i-s-r-1} = z, X_{i+s+r+1} = w\} - \right. \\
& \quad \left. - \mathbf{P} \{E_{i,s}\} \mathbf{P} \{X_{i-s-r-1} = z, X_{i+s+r+1} = w\} \right| \leq \\
& \leq \sum_{l_1, \dots, l_{s+1}=1}^N \pi_{l_1} k_{l_1, l_2} m_{l_2, l_3} \cdot \dots \cdot m_{l_s, l_{s+1}} \sum_{x, y \in \mathcal{A}} \left| p_{(R)xz}^{(s+r+1)} p_{yw}^{(r+1)} - \pi_z p_{zw}^{(2s+2r+2)} \right| \leq \\
& \leq \mathbf{P} \{E_{i,s}\} \sum_{x, y \in \mathcal{A}} \left(\varepsilon_{(R)xz}^{(s+r+1)} \varepsilon_{yw}^{(r+1)} + \pi_w \varepsilon_{(R)xz}^{(s+r+1)} + \pi_z \varepsilon_{yw}^{(r+1)} + \pi_z \varepsilon_{zw}^{(2s+2r+2)} \right).
\end{aligned}$$

Substituting this estimate into definition C^* , we get

$$C^* \leq \sum_{i=1}^n \mathbf{P}\{E_{i,s}\} \left(\sum_{x,y,z,w \in \mathcal{A}} \varepsilon_{(R)xz}^{(s+r+1)} \varepsilon_{yw}^{(r+1)} + N \sum_{x,z \in \mathcal{A}} \varepsilon_{(R)xz}^{(s+r+1)} + \right. \\ \left. + N \sum_{y,w \in \mathcal{A}} \varepsilon_{yw}^{(r+1)} + N^2 \sum_{z,w \in \mathcal{A}} \pi_z \varepsilon_{zw}^{(2s+2r+2)} \right).$$

Using (9), we obtain

$$C^* \leq \sum_{i=1}^n \mathbf{P}\{E_{i,s}\} (4N^2 \varrho^{s+2r+2} + 2N^2 \varrho^{s+r+1} + 2N^2 \varrho^{r+1} + 2N^3 \varrho^{2s+2r+2}) = \\ = \sum_{i=1}^n \mathbf{P}\{E_{i,s}\} 2N^2 \varrho^{r+1} (2\varrho^{s+r+1} + \varrho^s + 1 + N\varrho^{2s+r+1}) = \\ = 2\lambda N^2 \varrho^{r+1} (N\varrho^{2s+r+1} + 2\varrho^{s+r+1} + \varrho^s + 1), \quad (19)$$

where λ was defined in (8).

In the next lemma the formula for the expectation of random variable $\xi_n(s)$ is given.

Lemma 1. *We have*

$$\lambda = \mathbf{E} \xi_n(s) = n \cdot \sum_{l=1}^N \sum_{k=1}^{N-1} w_{k,l} \sum_{t=k+1}^N \pi_t p_{t,k}, \quad (20)$$

where for any $k \leq l$

$$w_{k,l} = \sum_{f=1}^q \sum_{j=1+d_1+\dots+d_{f-1}}^{d_1+\dots+d_f} t'_{j,l} \sum_{i=1}^j t_{k,i} p_{i_f, i_f}^{s-j+i-1} \binom{s-1}{s-j+i-1},$$

where $t_{k,i}$ and $t'_{j,l}$ are elements of the invertible matrices $\mathbb{T} = (t_{ij})_{i,j \in \mathcal{A}}$ and $\mathbb{T}^{-1} = (t'_{ij})_{i,j \in \mathcal{A}}$ in (2), $d_1, \dots, d_q$ are orders of Jordan blocks $\mathbb{J}_{d_1}(p_{i_1, i_1}), \dots, \mathbb{J}_{d_q}(p_{i_q, i_q})$ in (3), $d_0 = 0$, and $p_{i,j}$ ($i, j \in \{1, \dots, N\}$) are elements of transition probability matrix $\mathbb{P}$, and for any $k > l$

$$w_{k,l} = 0.$$

Proof of Lemma 1 will be given in the Appendix.

If we combine (17), (18), (19), and the results of Lemma 1 with (13), we get the estimate (11). This completes the proof of Theorem 1. $\square$

Let us prove the upper bound for λ under the conditions of Theorem 2. Using lemma 1 we get

$$\begin{aligned} \lambda &= n \cdot \sum_{l=1}^N \sum_{k=1}^{N-1} \sum_{t=k+1}^N \pi_t p_{t,k} \sum_{f=1}^q \sum_{j=1+d_1+\dots+d_{f-1}}^{d_1+\dots+d_f} t'_{j,l} \sum_{i=1}^j t_{k,i} p_{i_f,i_f}^{s-j+i-1} \binom{s-1}{s-j+i-1} \leq \\ &\leq n \binom{s-1}{s-N} \cdot \sum_{f=1}^q p_{i_f,i_f}^{s-N} \cdot \sum_{l=1}^N \sum_{k=1}^{N-1} \sum_{t=k+1}^N \pi_t p_{t,k} \sum_{j=1+d_1+\dots+d_{f-1}}^{d_1+\dots+d_f} t'_{j,l} \sum_{i=1}^j t_{k,i} = \\ &= n \binom{s-1}{s-N} \cdot \sum_{f=1}^q p_{i_f,i_f}^{s-N} \cdot G_f \equiv \lambda', \end{aligned}$$

where we have denoted

$$G_f \equiv \sum_{l=1}^N \sum_{k=1}^{N-1} \sum_{t=k+1}^N \pi_t p_{t,k} \sum_{j=1+d_1+\dots+d_{f-1}}^{d_1+\dots+d_f} t'_{j,l} \sum_{i=1}^j t_{k,i}.$$

Remind that in (4) we supposed that $b \in \{1, \dots, q\}$ first Jordan blocks have the largest eigenvalue which is equal to p . Then in the formula for λ' we isolate summands which depend on largest eigenvalue p :

$$\lambda' = n \binom{s-1}{s-N} \left(p^{s-N} \sum_{f=1}^b G_f + \sum_{f=b+1}^q p_{i_f,i_f}^{s-N} G_f \right).$$

Note that under conditions $s \rightarrow \infty$ and $G = \text{const} > 0$ we have

$$\binom{s+G}{s} = \binom{s+G}{G} = \frac{(s+G)^G}{G!} (1 + o(1)). \quad (21)$$

Using (21) and conditions (12) we obtain

$$\begin{aligned} \lambda' &= n \cdot \frac{(s-1)^{N-1}}{(N-1)!} \left(p^{s-N} \sum_{f=1}^b G_f + \sum_{f=b+1}^q p_{i_f,i_f}^{s-N} G_f \right) (1 + o(1)) \rightarrow \\ &\rightarrow Q \cdot \sum_{f=1}^b G_f \in (0, \infty) \end{aligned}$$

then $\lambda \in (0, \infty)$.

In the definition of Γ_i we choose r such that $r/s \rightarrow 1$ under $s \rightarrow \infty$ (for example, let $r = s + 1$), and $n, s \rightarrow \infty, s/n \rightarrow 0$. Therefore using the estimate (11) and the inequality $\varrho < 1$, we obtain

$$\begin{aligned} d(\mathcal{L}(\xi_n(s)), Po(\lambda)) &\leq \\ &\leq \frac{(1-e^{-\lambda})\lambda}{n} \left(2s + 1 + 2(r+1) \max_{i \in \mathcal{A}} \{\pi_i^{-1}\} \right) + \\ &\quad + 2\lambda N^2 \varrho^{r+1} (N \varrho^{2s+r+1} + 2\varrho^{s+r+1} + \varrho^s + 1) \leq \\ &\leq \frac{2(1-e^{-\lambda'}) \max_{i \in \mathcal{A}} \{\pi_i^{-1}\} (s+r+2) \lambda'}{n} + 2\lambda' N^2 \varrho^{r+1} (N+4) = \\ &= O\left(\frac{s+r}{n} + \varrho^r\right) = o(1). \end{aligned}$$

The proof of Theorem 2 is complete. $\square$

References

- [1] Barbour A. D., Holst L., Janson S., *Poisson approximation*, Oxford: Oxford Univ. Press, 1992, 280 pp.
- [2] *An Introduction to Stein's Method*, eds A. D. Barbour, L. H. Y. Chen, Singapore: Singapore Univ. Press, 2005, 238 pp.
- [3] Billingsley P., *Convergence of Probability Measures*, 2, New York: John Wiley & Sons, Inc., 1999, 296 pp.
- [4] Chrysaphinou O., Papastavridis S., Vaggelatou E., "Poisson limit theorems for the appearances of attributes". In: *Stein's method and applications*, eds A. D. Barbour, L. H. Y. Chen, Singapore: Singapore Univ. Press, 2005, 19–35.
- [5] Goncharov V. L., "From the domain of combinatorics", *Izv. Akad. Nauk SSSR, Ser. Mat.*, **8**:1 (1944), 3–48. (In Russian.)
- [6] Horn R. A., Johnson C. R., *Matrix Analysis*, Cambridge: Cambridge Univ. Press, 1985, 656 pp.
- [7] Lindvall T., *Lectures on the Coupling Method*, New York: John Wiley & Sons, Inc., 1992, 270 pp.
- [8] Mezhenyaya N. M., "Multidimensional normal theorem for number of monotone tuples with fixed length in a random equiprobable sequence", *Obozrenie prikl. i promyshl. matem.*, **14**:3 (2007), 503–505. (In Russian.)
- [9] Minakov A. A., "Compound Poisson approximation for the distribution of the number of monotone tuples in random sequence", *Prikl. Diskr. Mat.*, № 2 (28) (2015), 21–29. (In Russian.)
- [10] Pittel B. G., "Limiting behavior of a process of runs", *Ann. Probab.*, **9** (1981), 119–129.
- [11] Wolfowitz J., "Asymptotics distribution of runs up and down", *Ann. Math. Statist.*, **15**:2 (1944), 163–172.

4. Appendix

Proof of Lemma 1.

First we calculate components of vector (6) and then substitute them into (8).

From (2) and properties of Jordan matrices the following equalities follows

$$\mathbb{M}^{s-1} = \mathbb{T} \cdot \mathbb{J}^{s-1} \cdot \mathbb{T}^{-1}$$

and

$$\mathbb{J}^{s-1} = \text{Diag}\{\mathbb{J}_{d_1}^{s-1}(p_{i_1, i_1}), \mathbb{J}_{d_2}^{s-1}(p_{i_2, i_2}), \dots, \mathbb{J}_{d_q}^{s-1}(p_{i_q, i_q})\}.$$

For brevity denote $\mathbb{J}^{s-1}$ by

$$\mathbb{Y} = (y_{i,j})_{i,j \in \mathcal{A}}.$$

Then multiply $\mathbb{Y}$ by $\mathbb{T}$, and denote the result by $\mathbb{F}$:

$$\mathbb{T} \cdot \mathbb{Y} = (f_{k,l})_{k,l \in \mathcal{A}} \equiv \mathbb{F}.$$

Taking into account that $\mathbb{Y}$ is upper triangular matrix, we get the expression for (k, l) -th element of $\mathbb{F}$:

$$f_{k,l} = \sum_{i=1}^l t_{k,i} y_{i,l}.$$

Now multiplying $\mathbb{F}$ by $\mathbb{T}^{-1}$ we obtain the matrix $\mathbb{M}^{s-1}$:

$$\mathbb{M}^{s-1} = \mathbb{F} \cdot \mathbb{T}^{-1}.$$

Denote $\mathbb{M}^{s-1}$ by $\mathbb{W}$. Using the fact that $\mathbb{W}$ is upper triangular matrix, we get (k, l) -th element of $\mathbb{W}$ for $k \leq l$:

$$w_{k,l} = \sum_{j=1}^N t'_{j,l} \sum_{i=1}^j t_{k,i} y_{i,j}, \quad (22)$$

and for $k > l$:

$$w_{k,l} = 0.$$

Now let's find the formula for each Jordan block of the matrix $\mathbb{Y}$.

Consider the $(s-1)$ th power of the f th Jordan block

$$\mathbb{J}_{d_f}^{s-1}(p_{i_f, i_f}) = (p_{i_f, i_f} \mathbb{E}_{d_f} + \mathbb{H}_{d_f})^{s-1}, \quad f \in \{1, \dots, q\},$$

where $\mathbb{E}_{d_f}$ is the identity matrix of size $d_f \times d_f$, and $\mathbb{H}_{d_f}$ is the matrix of size $d_f \times d_f$ such that superdiagonal of $\mathbb{H}_{d_f}$ consists of 1 and all other entries are zeros:

$$\mathbb{H}_{d_f} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ & & & \ddots & & \\ 0 & 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & 0 & \dots & 0 & 0 \end{pmatrix}.$$

It follows from [6, § 3.2.5] that

$$\mathbb{J}_{d_f}^{s-1}(p_{i_f, i_f}) = \sum_{h=s-d_f}^{s-1} \binom{s-1}{h} p_{i_f, i_f}^h \mathbb{H}_{d_f}^{s-h-1}.$$

Therefore the (i, j) th $(i, j \in \{1, \dots, d_f\}, i \leq j)$ element of f -th Jordan block $\mathbb{J}_{d_f}^{s-1}(p_{i_f, i_f})$ equals to $p_{i_f, i_f}^{s-j+i-1} \binom{s-1}{s-j+i-1}$. Noting that the matrix $\mathbb{Y}$ has also block structure, for $f \in \{1, \dots, q\}$, $i, j \in \{1+d_1+\dots+d_{f-1}, \dots, d_1+\dots+d_f\}$, $i \leq j$ we have

$$y_{i,j} = p_{i_f, i_f}^{s-j+i-1} \binom{s-1}{s-j+i-1}.$$

Next we substitute the formula for $y_{i,j}$ into (22). For $k \leq l$ we have:

$$\begin{aligned} w_{k,l} &= \sum_{j=1}^{d_1} t'_{j,l} \sum_{i=1}^j t_{k,i} y_{i,j} + \sum_{j=1+d_1}^{d_1+d_2} t'_{j,l} \sum_{i=1}^j t_{k,i} y_{i,j} + \dots + \\ &\quad \dots + \sum_{j=1+d_1+\dots+d_{q-1}}^{d_1+\dots+d_q} t'_{j,l} \sum_{i=1}^j t_{k,i} y_{i,j} = \\ &= \sum_{f=1}^q \sum_{j=1+d_1+\dots+d_{f-1}}^{d_1+\dots+d_f} t'_{j,l} \sum_{i=1}^j t_{k,i} y_{i,j} = \\ &= \sum_{f=1}^q \sum_{j=1+d_1+\dots+d_{f-1}}^{d_1+\dots+d_f} t'_{j,l} \sum_{i=1}^j t_{k,i} p_{i_f, i_f}^{s-j+i-1} \binom{s-1}{s-j+i-1}. \end{aligned}$$

Now we multiply the vector $\vec{\pi}$ by $\mathbb{K}$ defined in (5):

$$\vec{\pi} \cdot \mathbb{K} = \left(\sum_{l=2}^N \pi_l p_{l,1}, \sum_{l=3}^N \pi_l p_{l,2}, \dots, \pi_N p_{N,N-1}, 0 \right).$$

And multiply the vector $\vec{\pi}\mathbb{K}$ by $\mathbb{W}$:

$$\vec{\alpha} = \vec{\pi}\mathbb{K}\mathbb{W} = \sum_{k=1}^{N-1} \vec{W}_k \sum_{l=k+1}^N \pi_l p_{l,k}.$$

Now let us write the formula for the entries of $\vec{\alpha}$:

$$\alpha_l = \sum_{k=1}^{N-1} w_{k,l} \sum_{t=k+1}^N \pi_t p_{t,k}, \quad (23)$$

where $l \in \{1, \dots, N\}$.

Using (8) and (23), we get:

$$\lambda = n \cdot \sum_{l=1}^N \sum_{k=1}^{N-1} w_{k,l} \sum_{t=k+1}^N \pi_t p_{t,k}.$$

This completes the proof of Lemma 1.