

УДК 519.214.5

ОЦЕНКА ДЛЯ РАСПРЕДЕЛЕНИЯ ЧИСЕЛ СЕРИЙ В СЛУЧАЙНОЙ ПОСЛЕДОВАТЕЛЬНОСТИ, УПРАВЛЯЕМОЙ СТАЦИОНАРНОЙ ЦЕПЬЮ МАРКОВА

Н. М. Меженная

*Московский государственный технический университет имени Н. Э. Баумана, г. Москва,
Россия*

Проведено исследование асимптотических свойств совместного распределения чисел серий из разных знаков в последовательности случайных величин с полиномиальными распределениями, управляемой стационарной цепью Маркова с конечным числом состояний. Получена оценка расстояния по вариации между распределением случайного вектора из чисел серий заданных знаков и заданной длины в управляемой последовательности и сопровождающим многомерным распределением Пуассона. При доказательстве использованы метод Чена — Стейна, а также оценки расстояния по вариации между смешанным и обычным распределениями Пуассона. Из полученной оценки расстояния по вариации выведены многомерная пуассоновская и нормальная предельные теоремы для указанного случайного вектора.

Ключевые слова: число серий, цепь Маркова, расстояние по вариации, метод Чена — Стейна, смешанное распределение Пуассона, предельная теорема Пуассона, центральная предельная теорема, скрытая марковская модель.

DOI 10.17223/20710410/35/2

ESTIMATOR FOR THE DISTRIBUTION OF THE NUMBERS OF RUNS IN A RANDOM SEQUENCE CONTROLLED BY STATIONARY MARKOV CHAIN

N. M. Mezhenaya

Bauman Moscow State Technical University

E-mail: natalia.mezhenaya@gmail.com

The sequences of random characters from a finite set $\mathcal{A}$ with polynomial distributions controlled by a stationary finite-state Markov chain are considered. For numbers of character runs in them, the asymptotic properties of joint distributions are studied. We deduce an estimate for the total variation distance ρ_{TV} between the distribution of a random vector $\varsigma_{\mathcal{A}}$ with components being numbers of runs in a controlled sequence of an enough length T and accompanying multidimensional Poisson distribution $\text{Pois}(\lambda_{\mathcal{A}})$. The estimate is $\rho_{TV}(\mathcal{L}(\varsigma_{\mathcal{A}}), \text{Pois}(\lambda_{\mathcal{A}})) \leq \gamma(\gamma T(p^*)^{s_*} + 1)$, where $\gamma^2 = |\mathcal{A}|^2(2s_* + 3)(p^*)^{s_*}$, s_* (s^*) is the minimum (maximum) length of run in the set of components of the vector $\varsigma_{\mathcal{A}}$, and p^* is the maximum character probability in distributions given on $\mathcal{A}$. For deriving this estimate, we use the functional variant of Chen — Stein method and an estimation for the total variation distance between the mixed and ordinal Poisson distributions. This estimation is a function of the variance of mixing parameter of mixed Poisson distribution. Using the derived estimate for

the total variation distance ρ_{TV} , we deduce the multidimensional Poisson and normal limit theorems for the random vector ς_A under appropriate conditions for scheme parameters.

Keywords: *number of runs, Markov chain, total variation distance, Chen — Stein method, mixed Poisson distribution, Poisson limit theorem, normal limit theorem, hidden Markov model.*

Введение

Статистические свойства чисел серий в дискретных случайных последовательностях широко используются в задачах передачи информации, контроля качества, анализа генома и т. д. (см. [1, глава 1] и библиографию там же). Вероятностные свойства связанных с ними статистик (чисел серий, длин промежутков между сериями и т. п.) в последовательностях, образованных независимыми или слабо зависимыми случайными величинами, хорошо изучены [1, главы 5, 7; 2]. Имеются как точные, так и асимптотические результаты.

В ряде работ изучены свойства чисел серий и цепочек специального вида из фиксированных знаков в марковских последовательностях. Например, в [3, 4] исследованы распределения времени до появления серии в цепи Маркова с двумя состояниями, а также распределения ряда других случайных величин, связанных с моментами появления серий в такой последовательности. В работах [5–7] найдено распределение чисел серий в марковской случайной последовательности с двумя и тремя состояниями. В [8] получены точные и приближённые формулы для вероятностей появления чисел серий из единиц в отрезках марковских случайных последовательностей с двумя состояниями и выражения для их математического ожидания и дисперсии. В [8, 9] подробно описано использование статистик, связанных с числом серий, для анализа последовательностей из зависимых случайных величин. В [10] получена производящая функция совместных вероятностей появления серий в цепи Маркова с конечным множеством состояний. Она позволяет найти явные выражения для вероятностей и числовых характеристик.

В работах [11–13] рассмотрен вопрос об оценках скорости сходимости для распределения чисел серий и цепочек в марковской цепи к сложному пуассоновскому распределению. Предельное поведение наибольших длин серий изучено в [14, 15], а также в [16, глава 4]. В [17] получены асимптотические формулы для вероятностей цепочек в последовательности независимых случайных величин и проведено обобщение этого результата для марковских последовательностей. В [18, 19] доказаны предельные теоремы Пуассона для числа повторений и кратных повторений (соответственно) цепочек в эргодической марковской цепи с оценками скорости сходимости. В [20] доказана многомерная центральная предельная теорема для чисел серий в марковской случайной последовательности с двумя состояниями.

В последнее время особый интерес представляет анализ свойств скрытых цепей Маркова в связи с задачами распознавания образов, машинного обучения и анализа текста [21, 22]. Основные свойства скрытых марковских моделей подробно описаны в [23, главы 1–3]. Для таких моделей также представляет интерес задача о точном и асимптотическом поведении чисел серий и цепочек как в наблюдаемой последовательности, так и образованной состояниями цепи Маркова. В [24] приведён способ вычисления вероятностей появлений цепочек специального вида в последовательности, образованной состояниями скрытой марковской цепи, по известному участку наблю-

даемой последовательности. Однако вычисления по приведённым в [24] алгоритмам являются весьма трудоёмкими для больших длин последовательностей.

В настоящей работе рассматривается один частный случай скрытой марковской цепи, а именно полиномиальная последовательность, управляемая цепью Маркова. Такая последовательность может трактоваться как последовательность, полученная укрупнением состояний марковской цепи, и как скрытая марковская цепь. В [25] изучено предельное распределение для числа пар совпавших знаков в такой последовательности, когда её длина стремится к бесконечности, а наибольшая вероятность появления любого знака стремится к нулю (согласованным образом). В частности, доказаны пуассоновская и нормальная предельные теоремы (с оценками скорости сходимости) для чисел пар совпадений знаков и ряда связанных с ними случайных величин.

Данная работа посвящена изучению асимптотических свойств совместных распределений чисел серий разных знаков и различных длин в полиномиальной случайной последовательности, управляемой цепью Маркова с конечным числом состояний. Получены предельные теоремы пуассоновского и нормального типа для чисел серий длины не меньше заданной.

1. Оценки расстояния по вариации и предельные теоремы

Пусть $\mathbf{Z} = (Z_0, Z_1, Z_2, \dots, Z_T, \dots)$ — стационарная цепь Маркова с множеством состояний $E_M = \{1, \dots, M\}$ и стационарными вероятностями $\tilde{\pi}_k = \pi_k(n) = \mathbf{P}\{Z_n = k\}$. Обозначим

$$\pi_{kl}^{(m)} = \mathbf{P}\{Z_{t+m} = l | Z_t = k\}, \quad k, l \in E_M,$$

вероятность перехода из состояния k в состояние l за m шагов, $m \in \mathbb{N}$. Для простоты обозначений будем писать π_{kl} вместо $\pi_{kl}^{(1)}$. Известно [26, часть 2, § 3], что существуют константы $C, \alpha > 0$, при которых

$$\max_{l \in E_M} |\pi_{lk}^{(n)} - \tilde{\pi}_k| \leq C \tilde{\pi}_k e^{-\alpha n}. \quad (1)$$

Пусть на множестве $A_N = \{1, \dots, N\}$ заданы M вероятностных распределений $(p_a^{(j)}, a \in A_N, j = 1, \dots, M)$. Рассмотрим последовательность случайных величин $X_0, X_1, \dots, X_T, \dots$, принимающих значения из множества A_N с вероятностями

$$\mathbf{P}\{X_j = k\} = p_k^{(Z_j)}, \quad k \in A_N, \quad j = 0, 1, 2, \dots$$

Пусть $s \geq 1$. Обозначим $\nu_t^a = I\{X_{t-1} \neq a, X_t = \dots = X_{t+s-1} = a\}$ индикатор случайного события, состоящего в том, что в момент t началась серия из знаков a длины не меньше s (для краткости будем писать a -серия). Тогда сумма случайных индикаторов

$$\zeta_s^a = \sum_{t=1}^T \nu_t^a \quad (2)$$

равна числу a -серий длины не меньше s с началом в последовательности $X_1, \dots, X_T$.

Математическое ожидание случайной величины ζ_s^a обозначим $\lambda_s^a = \mathbf{E}\zeta_s^a$. Оно определяется формулой $\lambda_s^a = \mathbf{E}\lambda_s^a(\mathbf{Z})$, где

$$\lambda_s^a(\mathbf{Z}) = \mathbf{E}(\zeta_s^a | \mathbf{Z}) = \sum_{t=1}^T (1 - p_a^{(Z_{t-1})}) \prod_{j=t}^{t+s-1} p_a^{(Z_j)}. \quad (3)$$

Пусть $p_* = \min_{a \in A_N} \{p_{a*}\}$, $p^* = \max_{a \in A_N} \{p_a^*\}$, $p_a^* = \max_{j=1, \dots, M} \{p_a^{(j)}\}$, $p_{a*} = \min_{j=1, \dots, M} \{p_a^{(j)}\}$.

Лемма 1. Пусть $T > s \geq 1$, $a \in A_N$ — фиксированный знак. Тогда

$$\mathbf{E}\zeta_s^a = \lambda_s^a = T \sum_{k_0, \dots, k_s \in E_M} (1 - p_a^{(k_0)}) \tilde{\pi}_{k_0} \prod_{j=1}^s p_a^{(k_j)} \pi_{k_{j-1}k_j}; \quad (4)$$

$$\mathbf{D}\lambda_s^a(\mathbf{Z}) \leq \lambda_s^a \left((2s+1)(1 - p_{a*})(p_a^*)^s + \frac{2C\lambda_s^a}{T(e^\alpha - 1)} \right). \quad (5)$$

Замечание 1. Формула, аналогичная (4), приведена в [27] для более общего случая, когда цепь $\mathbf{Z}$ не является стационарной.

Пусть $\mathcal{A} \subset A_N$ — фиксированное подмножество и заданы наборы чисел $(s_a, a \in \mathcal{A})$ и $\lambda_{\mathcal{A}} = (\lambda_{s_a}^a, a \in \mathcal{A})$. Обозначим вектор $\zeta_{\mathcal{A}} = (\zeta_{s_a}^a, a \in \mathcal{A})$. Через $\text{Pois}(\lambda_{\mathcal{A}})$ будем обозначать распределение случайного вектора с независимыми компонентами, каждая из которых распределена по закону Пуассона с параметром $\lambda_{s_a}^a$. Будем также использовать обозначение $\mathcal{L}(X)$ для распределения случайной величины X .

Напомним, что расстояние по вариации между распределениями случайных величин η_1 и η_2 , принимающих значения в счётном множестве $\mathcal{E}$, выражается формулой

$$\rho_{TV}(\mathcal{L}(\eta_1), \mathcal{L}(\eta_2)) = \frac{1}{2} \sum_{a \in \mathcal{E}} |\mathbf{P}\{\eta_1 = a\} - \mathbf{P}\{\eta_2 = a\}|. \quad (6)$$

Пусть $s^* = \max_{a \in \mathcal{A}} \{s_a\}$, $s_* = \min_{a \in \mathcal{A}} \{s_a\}$. Мощность множества $\mathcal{B}$ будем обозначать $|\mathcal{B}|$.

Теорема 1. Пусть $\mathcal{A} \subseteq A_N$, задан набор натуральных чисел $(s_a, a \in \mathcal{A})$, $T \geq \frac{C\lambda_{s_a}^a}{(e^\alpha - 1)(p_a^*)^{s_a}}$. Тогда выполнена оценка

$$\rho_{TV}(\mathcal{L}(\zeta_{\mathcal{A}}), \text{Pois}(\lambda_{\mathcal{A}})) \leq \gamma(\gamma T(p^*)^{s^*} + 1), \quad (7)$$

где $\gamma^2 = |\mathcal{A}|^2(2s^* + 3)(p^*)^{s^*}$.

Замечание 2. Оценка расстояния по вариации, аналогичная оценке (7), для случая одноточечного множества $\mathcal{A}$ получена в [28]. Если $\mathcal{A}$ — одноточечное множество, то вектор $\zeta_{\mathcal{A}}$ представляет собой одномерную случайную величину. В случае одномерного распределения оценка работы [28] точнее, так как при доказательстве использован одномерный вариант метода Чена — Стейна.

Следствие 1. Пусть $\mathcal{A} \subseteq A_N$ — фиксированное множество, $p^* \in (0, 1)$, $T \rightarrow \infty$, а числа $s_a, a \in \mathcal{A}$, меняются так, что $s_a \rightarrow \infty$ и $T(p^*)^{s^*} \rightarrow c > 0$. Тогда при всех $a \in \mathcal{A}$ существуют

$$\lim \mathbf{E}\zeta_{s_a}^a = \lim T \sum_{k_0, \dots, k_s \in E_M} (1 - p_a^{(k_0)}) \tilde{\pi}_{k_0} \prod_{j=1}^s p_a^{(k_j)} \pi_{k_{j-1}k_j} = \lambda_a \geq 0,$$

а компоненты случайного вектора $\zeta_{\mathcal{A}} = (\zeta_{s_a}^a, a \in \mathcal{A})$ асимптотически независимы и имеют в пределе распределения Пуассона с параметрами λ_a соответственно.

Следствие 2. Пусть $\mathcal{A} \subseteq A_N$ — фиксированное множество, $p^* \in (0, 1)$, $T \rightarrow \infty$, а числа $s_a, a \in \mathcal{A}$, меняются так, что $T(p^*)^{s^*} \rightarrow \infty$ и $\gamma T(p^*)^{s^*} \rightarrow 0$. Тогда компоненты случайного вектора $\left(\frac{\zeta_{s_a}^a - \lambda_{s_a}^a}{\sqrt{\mathbf{D}\zeta_{s_a}^a}}, a \in \mathcal{A} \right)$ асимптотически независимы и распределены в пределе по стандартному нормальному закону.

Замечание 3. Следствия 1 и 2 очевидным образом получаются из оценки (7), поэтому подробно на их доказательстве мы не останавливаемся.

2. Вывод оценки расстояния по вариации

Доказательство теоремы 1 начнём с более простой постановки задачи. Пусть случайные величины $Y_1, \dots, Y_T, \dots$ независимы и принимают значения из множества $A_N = \{1, \dots, N\}$, причём

$$\mathbf{P}\{Y_j = k\} = p_k^{(j)}, \quad k \in A_N, \quad j \in \{1, \dots, M\},$$

а наборы $\{p_k^{(j)}\}$ при каждом j удовлетворяют условию $\sum_{k \in A_N} p_k^{(j)} = 1$. Классу таких последовательностей принадлежит последовательность $X_0, \dots, X_T, \dots$, описанная в начале работы, при фиксированной $\mathbf{Z}$.

Пусть $s \geq 1$, $\tilde{\nu}_t^a = I\{Y_{t-1} \neq a, Y_t = \dots = Y_{t+s-1} = a\}$, $\tilde{\zeta}_s^a = \sum_{t=1}^T \tilde{\nu}_t^a$. Случайные величины $\tilde{\nu}_t^a$ и $\tilde{\zeta}_s^a$ аналогичны введённым в начале работы ν_t^a и ζ_s^a (см. (2)). Математическое ожидание случайной величины $\tilde{\zeta}_s^a$ определяется формулой

$$\mathbf{E}\tilde{\zeta}_s^a = \sum_{t=1}^T \mathbf{E}\tilde{\nu}_t^a = \sum_{t=1}^T (1 - p_a^{(t-1)}) \prod_{j=t}^{t+s-1} p_a^{(j)}. \quad (8)$$

Обозначим вектор $\tilde{\zeta}_{\mathcal{A}} = (\tilde{\zeta}_{s_a}^a, a \in \mathcal{A})$. Сопровождающим пуассоновским распределением для вектора $\tilde{\zeta}_{\mathcal{A}}$ будет распределение вектора с независимыми компонентами, каждая из которых распределена по закону Пуассона с параметром $\mathbf{E}\tilde{\zeta}_{s_a}^a$, $a \in \mathcal{A}$, определяемым формулой (8). Для этого распределения будем использовать обозначение $\text{Pois}(\mathbf{E}\tilde{\zeta}_{\mathcal{A}}) = (\text{Pois}(\mathbf{E}\tilde{\zeta}_{s_a}^a), a \in \mathcal{A})$.

Лемма 2. Пусть $\mathcal{A} \subseteq A_N$, задан набор натуральных чисел $s_a, a \in \mathcal{A}$, $T \geq 1$. Тогда выполнена оценка

$$\rho_{TV}(\mathcal{L}(\tilde{\zeta}_{\mathcal{A}}), \text{Pois}(\mathbf{E}\tilde{\zeta}_{\mathcal{A}})) \leq (2s^* + 3)|\mathcal{A}|(p^*)^{s^*} \Lambda_{\mathcal{A}}, \quad (9)$$

где $\Lambda_{\mathcal{A}} = \sum_{a \in \mathcal{A}} \mathbf{E}\tilde{\zeta}_{s_a}^a$.

Замечание 4. Если цепь Маркова $\mathbf{Z}$ представляет собой последовательность независимых одинаково распределённых случайных величин, то $X_0, X_1, \dots, X_T, \dots$ является стационарной последовательностью с полиномиальным распределением. В этом случае оценка (9) может быть получена из теоремы 2 работы [2].

Доказательство леммы 2 приведено далее в п. 3. Вернёмся к задаче об управляемой последовательности. Запишем неравенство треугольника

$$\rho_{TV}(\mathcal{L}(\zeta_{\mathcal{A}}), \text{Pois}(\lambda_{\mathcal{A}})) \leq \rho_{TV}(\mathcal{L}(\zeta_{\mathcal{A}}), \text{Pois}(\lambda_{\mathcal{A}}(\mathbf{Z}))) + \rho_{TV}(\text{Pois}(\lambda_{\mathcal{A}}(\mathbf{Z})), \text{Pois}(\lambda_{\mathcal{A}})). \quad (10)$$

Здесь обозначение $\text{Pois}(\lambda_{\mathcal{A}}(\mathbf{Z}))$ используется для распределения вектора, компоненты которого имеют смешанные пуассоновские распределения с параметрами $\lambda_{s_a}^a(\mathbf{Z})$. Напомним, что случайная величина X имеет смешанное распределение Пуассона с дискретным случайным параметром Λ , принимающим значения в множестве $\mathcal{E}$, если

$$\mathbf{P}\{X = n\} = \sum_{\lambda \in \mathcal{E}} \frac{\lambda^n}{n!} e^{-\lambda} \mathbf{P}\{\Lambda = \lambda\}.$$

Для оценки первого слагаемого в (10) воспользуемся леммой 2, согласно которой для каждой траектории $\mathbf{z}$ цепи Маркова $\mathbf{Z}$

$$\rho_{TV}(\mathcal{L}(\zeta_{\mathcal{A}}), \text{Pois}(\lambda_{\mathcal{A}}(\mathbf{z}))) \leq (2s^* + 3)|\mathcal{A}|(p^*)^{s^*} \Lambda_{\mathcal{A}}(\mathbf{z}); \quad (11)$$

$$\Lambda_{\mathcal{A}}(\mathbf{z}) = \sum_{a \in \mathcal{A}} \lambda_{s_a}^a(\mathbf{z}) = \sum_{a \in \mathcal{A}} \sum_{t=1}^T (1 - p_a^{(z_{t-1})}) \prod_{j=t}^{t+s_a-1} p_a^{(z_j)} \leq T|\mathcal{A}|(1 - p_*)(p^*)^{s_*} \leq T|\mathcal{A}|(p^*)^{s_*}, \quad (12)$$

а также следующим утверждением.

Лемма 3. Пусть случайная величина Π имеет смешанное дискретное распределение со случайным параметром Θ . Обозначим β не зависящую от Π и Θ случайную величину. Тогда

$$\rho_{TV}(\mathcal{L}(\Pi), \mathcal{L}(\alpha)) \leq \max_{\theta} \rho_{TV}(\mathcal{L}(\Pi|\theta), \mathcal{L}(\beta)), \quad (13)$$

где максимум берется по всем возможным значениям Θ , а через $\mathcal{L}(\Pi|\theta)$ обозначено распределение случайной величины Π при фиксированном значении параметра $\Theta = \theta$.

Из (11)–(13) следует, что

$$\rho_{TV}(\mathcal{L}(\zeta_{\mathcal{A}}), \text{Pois}(\lambda_{\mathcal{A}}(\mathbf{Z}))) \leq 2T(s^* + 1)|\mathcal{A}|^2(p^*)^{2s_*} = \gamma^2 T(p^*)^{s_*}. \quad (14)$$

Для оценки второго слагаемого в правой части (10) понадобится ещё одно утверждение.

Лемма 4. В условиях теоремы 1

$$\rho_{TV}(\text{Pois}(\lambda_{\mathcal{A}}(\mathbf{Z})), \text{Pois}(\lambda_{\mathcal{A}})) \leq \gamma. \quad (15)$$

Подставляя (14) и (15) в (10), получаем

$$\rho_{TV}(\mathcal{L}(\zeta_{\mathcal{A}}), \text{Pois}(\lambda_{\mathcal{A}})) \leq \gamma^2 T(p^*)^{s_*} + \gamma = \gamma(\gamma T(p^*)^{s_*} + 1).$$

Теорема 1 доказана. ■

3. Доказательства лемм

Доказательство леммы 1

Формула (4) в нашем случае очевидна (см. замечание к лемме 1), поэтому сразу перейдём к формуле (5).

Вычислим сначала второй момент случайной величины $\lambda_s^a(\mathbf{Z})$. Согласно (3),

$$\begin{aligned} \mathbf{E}(\lambda_s^a(\mathbf{Z}))^2 &= \mathbf{E} \left(\sum_{t=1}^T (1 - p_a^{(Z_{t-1})}) \prod_{j=t}^{t+s-1} p_a^{(Z_j)} \right)^2 = \\ &= \mathbf{E} \sum_{t=1}^T \sum_{t'=1}^T \left(1 - p_a^{(Z_{t-1})} \right) \prod_{j=t}^{t+s-1} p_a^{(Z_j)} \left(1 - p_a^{(Z_{t'-1})} \right) \prod_{j'=t'}^{t'+s-1} p_a^{(Z_{j'})} = \\ &= \mathbf{E} \sum_{t=1}^T \left(1 - p_a^{(Z_{t-1})} \right)^2 \prod_{j=t}^{t+s-1} \left(p_a^{(Z_j)} \right)^2 + \\ &+ \mathbf{E} \sum_{t=1}^T \sum_{\substack{t'=1, \dots, T; \\ t' \neq t}} \left(1 - p_a^{(Z_{t-1})} \right) \prod_{j=t}^{t+s-1} p_a^{(Z_j)} \left(1 - p_a^{(Z_{t'-1})} \right) \prod_{j'=t'}^{t'+s-1} p_a^{(Z_{j'})}. \end{aligned} \quad (16)$$

Оценим отдельно слагаемые в правой части (16). Для первого слагаемого имеем

$$\begin{aligned} \mathbf{E} \sum_{t=1}^T \left(1 - p_a^{(Z_{t-1})} \right)^2 \prod_{j=t}^{t+s-1} \left(p_a^{(Z_j)} \right)^2 &\leq (1 - p_{a*})(p_a^*)^s \mathbf{E} \sum_{t=1}^T \left(1 - p_a^{(Z_{t-1})} \right) \prod_{j=t}^{t+s-1} p_a^{(Z_j)} \leq \\ &\leq (1 - p_{a*})(p_a^*)^s \lambda_s^a. \end{aligned} \quad (17)$$

Перейдём к оцениванию второго слагаемого. Пусть $t' \in \{t+1, \dots, t+s\}$. Тогда аналогично предыдущему случаю

$$\begin{aligned} \mathbf{E} \left(1 - p_a^{(Z_{t-1})} \right) \prod_{j=t}^{t+s-1} p_a^{(Z_j)} \left(1 - p_a^{(Z_{t'-1})} \right) \prod_{j'=t'}^{t'+s-1} p_a^{(Z_{j'})} &\leq \\ &\leq \mathbf{E} \left(1 - p_a^{(Z_{t-1})} \right) \prod_{j=t}^{t+s-1} p_a^{(Z_j)} (1 - p_{a*}) (p_a^*)^s. \end{aligned} \quad (18)$$

Пусть $t' \geq t+s+1$. Тогда

$$\begin{aligned} &\mathbf{E} (1 - p_a^{(Z_{t-1})}) \prod_{j=t}^{t+s-1} p_a^{(Z_j)} (1 - p_a^{(Z_{t'-1})}) \prod_{j'=t'}^{t'+s-1} p_a^{(Z_{j'})} = \\ &= \sum_{k_0, \dots, k_s \in E_M} \sum_{l_0, \dots, l_s \in E_M} \mathbf{P} \{ Z_{t-1} = k_0, \dots, Z_{t+s-1} = k_s, Z_{t'-1} = l_0, \dots, Z_{t'+s-1} = l_s \} \times \\ &\quad \times (1 - p_a^{(k_0)}) \prod_{j=1}^s p_a^{(k_j)} (1 - p_a^{(l_0)}) \prod_{j'=1}^s p_a^{(l_{j'})} = \\ &= \sum_{k_0, \dots, k_s \in E_M} \tilde{\pi}_{k_0} (1 - p_a^{(k_0)}) \prod_{j=1}^s p_a^{(k_j)} \pi_{k_{j-1}k_j} \sum_{l_0, \dots, l_s \in E_M} (1 - p_a^{(l_0)}) \pi_{k_s l_0}^{(t'-t-s)} \prod_{j'=1}^s p_a^{(l_{j'})} \pi_{l_{j'-1}l_{j'}} \leq \\ &\leq \sum_{k_0, \dots, k_s \in E_M} \tilde{\pi}_{k_0} (1 - p_a^{(k_0)}) \prod_{j=1}^s p_a^{(k_j)} \pi_{k_{j-1}k_j} \sum_{l_0, \dots, l_s \in E_M} (1 - p_a^{(l_0)}) \tilde{\pi}_{l_0} (1 + C e^{-\alpha(t'-t-s)}) \prod_{j'=1}^s p_a^{(l_{j'})} \pi_{l_{j'-1}l_{j'}} \end{aligned}$$

(в последнем неравенстве мы воспользовались оценкой (1)). Значит,

$$\begin{aligned} &\mathbf{E} \sum_{t=1}^T \sum_{\substack{t'=1, \dots, T; \\ |t'-t| > s}} (1 - p_a^{(Z_{t-1})}) \prod_{j=t}^{t+s-1} p_a^{(Z_j)} (1 - p_a^{(Z_{t'-1})}) \prod_{j'=t'}^{t'+s-1} p_a^{(Z_{j'})} \leq \\ &\leq \sum_{t=1}^T \sum_{\substack{t'=1, \dots, T; \\ |t'-t| > s}} \sum_{k_0, \dots, k_s \in E_M} \tilde{\pi}_{k_0} (1 - p_a^{(k_0)}) \prod_{j=1}^s p_a^{(k_j)} \pi_{k_{j-1}k_j} \times \\ &\quad \times \sum_{l_0, \dots, l_s \in E_M} (1 - p_a^{(l_0)}) \tilde{\pi}_{l_0} (1 + C e^{-\alpha|t'-t-s|}) \prod_{j'=1}^s p_a^{(l_{j'})} \pi_{l_{j'-1}l_{j'}}. \end{aligned} \quad (19)$$

Так как $\sum_{\substack{t'=1, \dots, T; \\ |t'-t| > s}} e^{-\alpha|t'-t-s|} = \frac{1 - e^{\alpha(s+t-T)}}{e^\alpha - 1} + \frac{1 - e^{\alpha(1-(s+t))}}{e^\alpha - 1} \leq \frac{2}{e^\alpha - 1}$, то правая часть (19) оценивается сверху выражением

$$\begin{aligned} &\sum_{t=1}^T \sum_{k_0, \dots, k_s \in E_M} \tilde{\pi}_{k_0} (1 - p_a^{(k_0)}) \prod_{j=1}^s p_a^{(k_j)} \pi_{k_{j-1}k_j} \times \\ &\quad \times \sum_{l_0, \dots, l_s \in E_M} (1 - p_a^{(l_0)}) \tilde{\pi}_{l_0} \left(T + \frac{2C}{e^\alpha - 1} \right) \prod_{j'=1}^s p_a^{(l_{j'})} \pi_{l_{j'-1}l_{j'}} = \\ &= T \sum_{k_0, \dots, k_s \in E_M} \tilde{\pi}_{k_0} (1 - p_a^{(k_0)}) \prod_{j=1}^s p_a^{(k_j)} \pi_{k_{j-1}k_j} \times \\ &\quad \times T \sum_{l_0, \dots, l_s \in E_M} (1 - p_a^{(l_0)}) \tilde{\pi}_{l_0} \left(1 + \frac{2C}{T(e^\alpha - 1)} \right) \prod_{j'=1}^s p_a^{(l_{j'})} \pi_{l_{j'-1}l_{j'}} = \\ &= (\lambda_s^a)^2 \left(1 + \frac{2}{C} T(e^\alpha - 1) \right). \end{aligned} \quad (20)$$

Подставив оценки (17)–(20) в (16), получим

$$\begin{aligned}\mathbf{E}(\lambda_s^a(\mathbf{Z}))^2 &\leq \lambda_s^a \left((2s+1)(1-p_{a*})(p_a^*)^s + \lambda_s^a \left(1 + \frac{2C}{T(e^\alpha - 1)} \right) \right), \\ \mathbf{D}\lambda_s^a(\mathbf{Z}) &\leq \lambda_s^a \left((2s+1)(1-p_{a*})(p_a^*)^s + \frac{2C\lambda_s^a}{T(e^\alpha - 1)} \right).\end{aligned}$$

Лемма 1 доказана. ■

Доказательство леммы 2

Для каждой пары индексов (a, t) , $t \in \{1, \dots, T\}$, $a \in A_N$, определим множество $O(a, t)$ равенством

$$O(a, t) = \{(b, t') : b \in \mathcal{A}, t' \in \{1, \dots, T\} : \max\{1, t - s_b\} \leq t' \leq \min\{T, t + s_a\}\}.$$

Тогда случайный индикатор $\tilde{\nu}_t^a$ и набор случайных индикаторов $(\tilde{\nu}_{t'}^b)$, $(b, t') \notin O(a, t)$, независимы. Согласно [29, глава 10, с. 210, теорема 10.A], расстояние по вариации между распределением случайного вектора $\tilde{\zeta}_A = (\tilde{\zeta}_{s_a}^a, a \in \mathcal{A})$ и сопровождающим пуассоновским распределением $\text{Pois}(\mathbf{E}\tilde{\zeta}_A) = (\text{Pois}(\mathbf{E}\tilde{\zeta}_{s_a}^a), a \in \mathcal{A})$ оценивается как

$$\rho_{TV}(\mathcal{L}(\tilde{\zeta}_A), \text{Pois}(\mathbf{E}\tilde{\zeta}_A)) \leq S_1 + S_2, \quad (21)$$

$$S_1 = \sum_{a \in \mathcal{A}} \sum_{t=1}^T \sum_{(b, t') \in O(a, t)} \mathbf{E}\tilde{\nu}_t^a \mathbf{E}\tilde{\nu}_{t'}^b, \quad S_2 = \sum_{a \in \mathcal{A}} \sum_{t=1}^T \sum_{(b, t') \in O(a, t) \setminus \{(a, t)\}} \mathbf{E}\tilde{\nu}_t^a \tilde{\nu}_{t'}^b.$$

Начнём с оценивания первой суммы S_1 . Так как

$$\mathbf{E}\tilde{\nu}_{t'}^b = (1 - p_b^{(t'-1)}) \prod_{j=t'}^{t'+s_b-1} p_b^{(j)} \leq (1 - p_*) \prod_{j=t'}^{t'+s_b-1} p^* \leq (1 - p_*)(p^*)^{s_*}, \quad (22)$$

то

$$\begin{aligned}S_1 &= \sum_{a \in \mathcal{A}} \sum_{1 \leq t \leq T} \mathbf{E}\tilde{\nu}_t^a \sum_{(b, t') \in O(a, t)} \mathbf{E}\tilde{\nu}_{t'}^b \leq (2s^* + 1)|\mathcal{A}|(1 - p_*)(p^*)^{s_*} \sum_{a \in \mathcal{A}} \sum_{1 \leq t \leq T} \mathbf{E}\tilde{\nu}_t^a = \\ &= (2s^* + 1)|\mathcal{A}|(1 - p_*)(p^*)^{s_*} \sum_{a \in \mathcal{A}} \mathbf{E}\tilde{\zeta}_{s_a}^a = (2s^* + 1)|\mathcal{A}|(1 - p_*)(p^*)^{s_*} \Lambda_A\end{aligned} \quad (23)$$

(в силу формулы (8)).

Теперь перейдём к оцениванию S_2 . Перепишем формулу для S_2 в виде

$$S_2 = \sum_{a, b \in \mathcal{A}} S_2(a, b), \quad S_2(a, b) = \sum_{t=1}^T \sum_{(b, t') \in O(a, t) \setminus \{(a, t)\}} \mathbf{E}\tilde{\nu}_t^a \tilde{\nu}_{t'}^b.$$

Заметим, что события, соответствующие случайным индикаторам $\tilde{\nu}_t^a$ и $\tilde{\nu}_{t'}^a$, $(a, t') \in O(a, t) \setminus \{(a, t)\}$, несовместны. Действительно, при $t' > t$ события $\{Y_{t-1} \neq a, Y_t = \dots = Y_{t+s_a-1} = a\}$ и $\{Y_{t'-1} \neq a, Y_{t'} = \dots = Y_{t'+s_a-1} = a\}$ (отвечающие случайным индикаторам $\tilde{\nu}_t^a$ и $\tilde{\nu}_{t'}^a$ соответственно) зависят от одного и того же случайного знака Y_{t-1} , который в первом событии должен принимать значение a , а во втором — любое значение, отличное от a . Аналогично при $t' < t$. Поэтому $S_2(a, a) = 0$.

Пусть теперь $a \neq b$. События, отвечающие $\tilde{\nu}_t^a$ и $\tilde{\nu}_{t'}^b$, $(b, t') \in O(a, t) \setminus \{(a, t)\}$, совместны, если $t' = t - s_b$ или $t' = t + s_a$ (аналогично оценке для $S_2(a, a)$), и

$$\mathbf{E}\tilde{\nu}_t^a \tilde{\nu}_{t-s_b}^b = \mathbf{E}\tilde{\nu}_t^a \tilde{\nu}_{t+s_a}^b = \frac{1}{\mathbf{P}\{Y_{t-1} \neq a\}} \mathbf{E}\tilde{\nu}_t^a \mathbf{E}\tilde{\nu}_{t-s_b}^b \leq (1 - p_*)^{-1} \mathbf{E}\tilde{\nu}_t^a \mathbf{E}\tilde{\nu}_{t-s_b}^b.$$

Поэтому из (22) имеем

$$S_2(a, b) = 2 \sum_{t=1}^T (1 - p_*)^{-1} \mathbf{E} \tilde{\nu}_t^a \mathbf{E} \tilde{\nu}_{t-s_b}^b \leq 2(p^*)^{s^*} \sum_{1 \leq t \leq T} \mathbf{E} \tilde{\nu}_t^a = 2(p^*)^{s^*} \mathbf{E} \tilde{\zeta}_{s_a}^a.$$

Таким образом,

$$S_2 = \sum_{a, b \in \mathcal{A}} S_2(a, b) \leq 2(p^*)^{s^*} \sum_{a, b \in \mathcal{A}} \mathbf{E} \tilde{\zeta}_{s_a}^a \leq 2|\mathcal{A}|(p^*)^{s^*} \sum_{a \in \mathcal{A}} \mathbf{E} \tilde{\zeta}_{s_a}^a = 2|\mathcal{A}|(p^*)^{s^*} \Lambda_{\mathcal{A}}. \quad (24)$$

Подставляя оценки (23) и (24) в (21), получаем

$$\rho_{TV}(\mathcal{L}(\tilde{\zeta}_{\mathcal{A}}), \text{Pois}(\mathbf{E} \tilde{\zeta}_{\mathcal{A}})) \leq (2s^* + 3)|\mathcal{A}|(p^*)^{s^*} \Lambda_{\mathcal{A}}.$$

Лемма 2 доказана. ■

Доказательство леммы 4

Нам понадобятся ещё одно вспомогательное утверждение и несколько дополнительных определений.

Пусть случайные величины Π_j , $j = 1, \dots, k$, $k \geq 1$, имеют смешанные дискретные распределения с параметрами $\Theta_j = \Theta_j(\xi)$. Будем писать $\mathcal{L}(\Pi_j | \theta_j)$ для обозначения распределения случайной величины Π_j при фиксированном значении параметра $\Theta_j = \theta_j$.

Пусть $\alpha_1, \dots, \alpha_r$ — дискретные случайные величины. Обозначим при каждом фиксированном значении $\Theta_j = \theta_j$

$$r(\mathcal{L}(\Pi_j), \mathcal{L}(\alpha_j)) = \rho_{TV}(\mathcal{L}(\Pi_j | \theta_j), \mathcal{L}(\alpha_j)).$$

Величина $r(\mathcal{L}(\Pi_j), \mathcal{L}(\alpha_j))$ равна расстоянию по вариации (см. (6)) между распределениями $\mathcal{L}(\Pi_j | \theta_j)$ и $\mathcal{L}(\alpha_j)$.

Лемма 5. Пусть ξ — дискретная случайная величина, случайные величины Π_j , $j = 1, \dots, k$, $k \geq 1$, имеют смешанные дискретные распределения с параметрами $\Theta_j = \Theta_j(\xi)$, причём при каждом фиксированном $\xi = x$ случайные величины Π_j , $j = 1, \dots, k$, независимы. Обозначим $\alpha_1, \dots, \alpha_k$ независимые и не зависящие от Π_j , $j = 1, \dots, k$, случайные величины. Тогда

$$\rho_{TV}(\mathcal{L}(\Pi_1, \dots, \Pi_k), \mathcal{L}(\alpha_1, \dots, \alpha_k)) \leq \sum_{j=1}^k \mathbf{E} r(\mathcal{L}(\Pi_j), \mathcal{L}(\alpha_j)). \quad (25)$$

Из (25), подставляя вместо $(\Pi_1, \dots, \Pi_k)$ вектор со смешанным пуассоновским распределением $\text{Pois}(\lambda_{\mathcal{A}}(\mathbf{Z}))$, а вместо $\alpha_1, \dots, \alpha_k$ — вектор с распределением $\text{Pois}(\lambda_{\mathcal{A}})$, получаем

$$\rho_{TV}(\text{Pois}(\lambda_{\mathcal{A}}(\mathbf{Z})), \text{Pois}(\lambda_{\mathcal{A}})) \leq \sum_{a \in \mathcal{A}} \mathbf{E} r(\text{Pois}(\lambda_{s_a}^a(\mathbf{Z})), \text{Pois}(\lambda_{s_a}^a)). \quad (26)$$

Воспользуемся теоремой 1.С из [29, глава 1, с. 12], которая в нашем случае формулируется следующим образом:

$$\rho_{TV}(\text{Pois}(\lambda), \text{Pois}(\mu)) \leq \min \left\{ 1, \frac{1}{\sqrt{\lambda}}, \frac{1}{\sqrt{\mu}} \right\} |\lambda - \mu|.$$

Тогда

$$\begin{aligned} \mathbf{E}r(\text{Pois}(\lambda_{s_a}^a(\mathbf{Z})), \text{Pois}(\lambda_{s_a}^a)) &\leq \min \left\{ 1, \frac{1}{\sqrt{\lambda_{s_a}^a}} \right\} \mathbf{E}|\lambda_{s_a}^a(\mathbf{Z}) - \lambda_{s_a}^a| \leq \\ &\leq \min \left\{ 1, \frac{1}{\sqrt{\lambda_{s_a}^a}} \right\} \sqrt{\mathbf{D}\lambda_{s_a}^a(\mathbf{Z})}. \end{aligned} \quad (27)$$

Из (5) и (27) следует, что

$$\mathbf{E}r(\text{Pois}(\lambda_{s_a}^a(\mathbf{Z})), \text{Pois}(\lambda_{s_a}^a)) \leq \sqrt{(2s_a + 1)(1 - p_{a*})(p_a^*)^s + \frac{2C\lambda_{s_a}^a}{T(e^\alpha - 1)}}.$$

Значит, из последней оценки и формулы (26) получаем, что при T , удовлетворяющих неравенству $\frac{C\lambda_{s_a}^a}{T(e^\alpha - 1)} \leq (p_a^*)^{s_a}$, имеет место оценка

$$\rho_{TV}(\text{Pois}(\lambda_{\mathcal{A}}(\mathbf{Z})), \text{Pois}(\lambda_{\mathcal{A}})) \leq \sum_{a \in \mathcal{A}} \sqrt{(2s_a + 3)(p_a^*)^{s_a}} \leq |\mathcal{A}| \sqrt{(2s^* + 3)(p^*)^{s^*/2}} = \gamma.$$

Лемма 4 доказана. ■

Доказательство леммы 5

Последовательно применяя неравенство треугольника, получаем следующую цепочку неравенств:

$$\begin{aligned} \rho_{TV}(\mathcal{L}(\Pi_1, \dots, \Pi_k), \mathcal{L}(\alpha_1, \dots, \alpha_k)) &\leq \rho_{TV}(\mathcal{L}(\Pi_1, \Pi_2, \dots, \Pi_k), \mathcal{L}(\alpha_1, \Pi_2, \dots, \Pi_k)) + \\ &+ \rho_{TV}(\mathcal{L}(\alpha_1, \Pi_2, \dots, \Pi_k), \mathcal{L}(\alpha_1, \dots, \alpha_k)) \leq \rho_{TV}(\mathcal{L}(\Pi_1, \Pi_2, \dots, \Pi_k), \mathcal{L}(\alpha_1, \Pi_2, \dots, \Pi_k)) + \\ &+ \rho_{TV}(\mathcal{L}(\alpha_1, \Pi_2, \Pi_3, \dots, \Pi_k), \mathcal{L}(\alpha_1, \alpha_2, \Pi_3, \dots, \Pi_k)) + \\ &+ \rho_{TV}(\mathcal{L}(\alpha_1, \alpha_2, \Pi_3, \dots, \Pi_k), \mathcal{L}(\alpha_1, \alpha_2, \dots, \alpha_k)) \leq \dots \leq \\ &\leq \sum_{j=1}^k \rho_{TV}(\mathcal{L}(\alpha_1, \dots, \alpha_{j-1}, \Pi_j, \dots, \Pi_k), \mathcal{L}(\alpha_1, \dots, \alpha_{j-1}, \alpha_j, \Pi_{j+1}, \dots, \Pi_k)) \end{aligned} \quad (28)$$

(в формуле (28) считаем, что слагаемое при $j = 1$ в правой части неравенства равно $\rho_{TV}(\mathcal{L}(\Pi_1, \Pi_2, \dots, \Pi_k), \mathcal{L}(\alpha_1, \Pi_2, \dots, \Pi_k))$).

Рассмотрим отдельно первое слагаемое в (28). Согласно определению расстояния по вариации (см. (6)), имеем

$$\begin{aligned} \rho_{TV}(\mathcal{L}(\Pi_1, \Pi_2, \dots, \Pi_k), \mathcal{L}(\alpha_1, \Pi_2, \dots, \Pi_k)) &= \\ &= \frac{1}{2} \sum_{l_1, \dots, l_k=0}^{\infty} |\mathbf{P}\{\Pi_1 = l_1, \Pi_2 = l_2, \dots, \Pi_k = l_k\} - \mathbf{P}\{\alpha_1 = l_1, \Pi_2 = l_2, \dots, \Pi_k = l_k\}|. \end{aligned}$$

Так как параметры смешанных распределений случайных величин $\Pi_1, \Pi_2, \dots, \Pi_k$ зависят от дискретной случайной величины ξ , для вычисления оценки расстояния по вариации воспользуемся формулой полной вероятности:

$$\begin{aligned} \rho_{TV}(\mathcal{L}(\Pi_1, \Pi_2, \dots, \Pi_k), \mathcal{L}(\alpha_1, \Pi_2, \dots, \Pi_k)) &= \frac{1}{2} \sum_{l_1, \dots, l_k=0}^{\infty} \sum_x \mathbf{P}\{\xi = x\} \times \\ &\times |\mathbf{P}\{\Pi_1 = l_1, \Pi_2 = l_2, \dots, \Pi_k = l_k | \xi = x\} - \mathbf{P}\{\alpha_1 = l_1, \Pi_2 = l_2, \dots, \Pi_k = l_k | \xi = x\}| = \\ &= \frac{1}{2} \sum_{l_1, \dots, l_k=0}^{\infty} \sum_x |\mathbf{P}\{\xi = x\} (\mathbf{P}\{\Pi_1 = l_1, \Pi_2 = l_2, \dots, \Pi_k = l_k | \xi = x\} - \\ &- \mathbf{P}\{\alpha_1 = l_1, \Pi_2 = l_2, \dots, \Pi_k = l_k | \xi = x\})|. \end{aligned}$$

Здесь и далее суммирование по индексу x означает, что суммирование ведётся по всем возможным значениям случайной величины ξ .

Так как при фиксированном значении $\xi = x$ случайные величины $\Pi_1, \Pi_2, \dots, \Pi_k$ независимы, продолжая цепочку равенств, получаем

$$\begin{aligned}
& \rho_{TV}(\mathcal{L}(\Pi_1, \Pi_2, \dots, \Pi_k), \mathcal{L}(\alpha_1, \Pi_2, \dots, \Pi_k)) = \\
& = \frac{1}{2} \sum_{l_1, \dots, l_k=0}^{\infty} \left| \sum_x \mathbf{P}\{\xi = x\} (\mathbf{P}\{\Pi_1 = l_1 | \xi = x\} \mathbf{P}\{\Pi_2 = l_2, \dots, \Pi_k = l_k | \xi = x\} - \right. \\
& \quad \left. - \mathbf{P}\{\alpha_1 = l_1\} \mathbf{P}\{\Pi_2 = l_2, \dots, \Pi_k = l_k | \xi = x\}) \right| = \\
& = \frac{1}{2} \sum_{l_1, \dots, l_k=0}^{\infty} \left| \sum_x \mathbf{P}\{\xi = x\} \mathbf{P}\{\Pi_2 = l_2, \dots, \Pi_k = l_k | \xi = x\} \times \right. \\
& \quad \left. \times (\mathbf{P}\{\Pi_1 = l_1 | \xi = x\} - \mathbf{P}\{\alpha_1 = l_1\}) \right| = \\
& = \frac{1}{2} \sum_{l_1, \dots, l_k=0}^{\infty} \left| \sum_x \mathbf{P}\{\xi = x\} \mathbf{P}\{\Pi_2 = l_2, \dots, \Pi_k = l_k | \xi = x\} \times \right. \\
& \quad \left. \times (\mathbf{P}\{\Pi_1 = l_1 | \xi = x\} - \mathbf{P}\{\alpha_1 = l_1\}) \right| \leqslant \\
& \leqslant \frac{1}{2} \sum_{l_1, \dots, l_k=0}^{\infty} \sum_x \mathbf{P}\{\xi = x\} \mathbf{P}\{\Pi_2 = l_2, \dots, \Pi_k = l_k | \xi = x\} \times \\
& \quad \times |(\mathbf{P}\{\Pi_1 = l_1 | \xi = x\} - \mathbf{P}\{\alpha_1 = l_1\})| \leqslant \\
& \leqslant \frac{1}{2} \sum_{l_1=0}^{\infty} \sum_x \mathbf{P}\{\xi = x\} |(\mathbf{P}\{\Pi_1 = l_1 | \xi = x\} - \mathbf{P}\{\alpha_1 = l_1\})| = \mathbf{E}r(\mathcal{L}(\Pi_1), \mathcal{L}(\alpha_1)). \tag{29}
\end{aligned}$$

Аналогично для слагаемых в (28) при $j \geqslant 2$:

$$\begin{aligned}
& \rho_{TV}(\mathcal{L}(\alpha_1, \dots, \alpha_{j-1}, \Pi_j, \dots, \Pi_k), \mathcal{L}(\alpha_1, \dots, \alpha_{j-1}, \alpha_j, \Pi_{j+1}, \dots, \Pi_k)) = \\
& = \frac{1}{2} \sum_{l_1, \dots, l_k=0}^{\infty} |\mathbf{P}\{\alpha_1 = l_1, \dots, \alpha_{j-1} = l_{j-1}, \Pi_j = l_j, \dots, \Pi_k = l_k\} - \\
& \quad - \mathbf{P}\{\alpha_1 = l_1, \dots, \alpha_{j-1} = l_{j-1}, \alpha_j = l_j, \Pi_{j+1} = l_{j+1}, \dots, \Pi_k = l_k\}| = \\
& = \frac{1}{2} \sum_{l_1, \dots, l_k=0}^{\infty} |\mathbf{P}\{\alpha_1 = l_1, \dots, \alpha_{j-1} = l_{j-1}\} \mathbf{P}\{\Pi_j = l_j, \dots, \Pi_k = l_k\} - \\
& \quad - \mathbf{P}\{\alpha_1 = l_1, \dots, \alpha_{j-1} = l_{j-1}\} \mathbf{P}\{\alpha_j = l_j\} \mathbf{P}\{\Pi_{j+1} = l_{j+1}, \dots, \Pi_k = l_k\}| \leqslant \\
& \leqslant \frac{1}{2} \sum_{l_1, \dots, l_k=0}^{\infty} |\mathbf{P}\{\Pi_j = l_j, \dots, \Pi_k = l_k\} - \mathbf{P}\{\alpha_j = l_j\} \mathbf{P}\{\Pi_{j+1} = l_{j+1}, \dots, \Pi_k = l_k\}|.
\end{aligned}$$

Дальше достаточно применить доказанную ранее оценку (29):

$$\rho_{TV}(\mathcal{L}(\alpha_1, \dots, \alpha_{j-1}, \Pi_j, \dots, \Pi_k), \mathcal{L}(\alpha_1, \dots, \alpha_{j-1}, \alpha_j, \Pi_{j+1}, \dots, \Pi_k)) \leqslant \mathbf{E}r(\mathcal{L}(\Pi_j), \mathcal{L}(\alpha_j)).$$

Тогда из (28) и последнего неравенства получим

$$\rho_{TV}(\mathcal{L}(\Pi_1, \dots, \Pi_k), \mathcal{L}(\alpha_1, \dots, \alpha_k)) \leqslant \sum_{j=1}^k \mathbf{E}r(\mathcal{L}(\Pi_j), \mathcal{L}(\alpha_j)).$$

Лемма 5 доказана. ■

Доказательство леммы 3

Воспользуемся формулой (25) при $k = 1$:

$$\rho_{TV}(\mathcal{L}(\Pi_1), \mathcal{L}(\alpha_1)) \leqslant \mathbf{E}r(\mathcal{L}(\Pi_1), \mathcal{L}(\alpha_1)).$$

Так как $\mathbf{E}r(\mathcal{L}(\Pi_1), \mathcal{L}(\alpha_1)) \leqslant \max_x r(\mathcal{L}(\Pi_1 | x_j), \mathcal{L}(\alpha_1))$, из двух последних формул получаем (13). ■

Заключение

Проведено исследование свойств распределения чисел серий в последовательности случайных величин с полиномиальными распределениями, управляемой стационарной цепью Маркова с конечным числом состояний. Использованы функциональный вариант метода Чена — Стейна, а также оценки для аппроксимации смешанного пуассоновского распределения простым распределением Пуассона. В работе получены:

- 1) оценка для расстояния по вариации между распределением случайного вектора из чисел серий заданных знаков и заданной длины в управляемой полиномиальной последовательности и сопровождающим многомерным распределением Пуассона;
- 2) выведены многомерные пуассоновская и нормальная предельные теоремы для указанного случайного вектора, когда длина T наблюдаемой последовательности стремится к бесконечности, а максимальная вероятность появления каждого знака — к нулю (согласованным образом с T).

Автор выражает глубокую признательность рецензенту за ценные замечания и рекомендации, а также В. Г. Михайлову за постановку задачи, советы и внимание к работе.

ЛИТЕРАТУРА

1. *Balakrishnan N. and Koutras M. V.* Runs and Scans with Applications. N.Y.: John Wiley & Sons Inc., 2002. 452 p.
2. *Михайлов В. Г.* Об асимптотических свойствах числа серий событий // Тр. по дискр. матем. 2006. Т. 9. С. 152–163.
3. *Aki S. and Hirano K.* Discrete distributions related to succession events in two-state Markov chain // Statistical Science and Data Analysis / eds. K. Matusita, M. L. Puri, T. Hayakawa. Zeist: VSP International Science Publishers, 1993. P. 467–474.
4. *Aki S. and Hirano K.* Sooner and later waiting time problems for runs in Markov dependent bivariate trials // Ann. Inst. Stat. Math. 1999. V. 51. P. 17–29.
5. *Han Q. and Aki S.* Formulae and recursions for the joint distributions of success runs of several lengths in a two-state Markov chain // Stat. Probab. Lett. 1998. V. 40. No. 3. P. 203–214.
6. *Савельев Л. Я., Балакин С. В.* Совместное распределение числа единиц и числа 1-серий в двоичных марковских последовательностях // Дискретная математика. 2004. Т. 16. № 3. С. 43–62.
7. *Савельев Л. Я., Балакин С. В.* Комбинаторное вычисление моментов характеристик серий в троичных марковских последовательностях // Дискретная математика. 2011. Т. 23. № 2. С. 76–92.
8. *Савельев Л. Я.* Распределения числа состояний в двоичных марковских стохастических моделях // Сиб. журн. вычисл. матем. 2015. Т. 18. № 2. С. 191–200.
9. *Савельев Л. Я., Балакин С. В.* Некоторые применения стохастической теории серий // Сиб. журн. индустр. матем. 2012. Т. 15. № 3. С. 111–123.
10. *Shinde R. L. and Kotwal K. S.* On the joint distribution of runs in the sequence of Markov-dependent multi-state trials // Stat. Probab. Lett. 2006. V. 76. No. 10. P. 1065–1074.
11. *Geske M. X., Godbole A. P., Schaffner A. A., et al.* Compound Poisson approximations for word patterns under Markovian hypotheses // J. Appl. Probab. 1995. V. 32. P. 877–892.
12. *Erhardsson T.* Compound Poisson approximation for Markov chains using Stein's method // Ann. Probab. 1999. V. 27. No. 1. P. 565–596.
13. *Chrysaphinou O. and Vaggelatou E.* Compound Poisson approximation for multiple runs in a Markov chain // Ann. Inst. Stat. Math. 2002. V. 54. No. 2. P. 411–424.

14. *Fu J. C., Wang L., and Lou W. Y. W.* On exact and large deviation approximation for the distribution of the longest run in a sequence of two-state Markov dependent trials // *J. Appl. Probab.* 2003. V. 40. No. 2. P. 346–360.
15. *Eryilmaz S.* Some results associated with the longest run statistic in a sequence of Markov dependent trials // *Appl. Math. Comput.* 2006. V. 175. No. 1. P. 119–130.
16. *Pinsky M. A. and Karlin S.* The long run behavior of Markov chains // *An Introduction to Stochastic Modeling. Fourth Edition* / eds. M. A. Pinsky, S. Karlin. Boston: Elsevier, 2011. P. 165–222.
17. *Fu J. C. and Johnson B. C.* Approximate probabilities for runs and patterns in i.i.d. and Markov-dependent multistate trials // *Adv. Appl. Probab. Appl. Probab. Trust.* 2009. V. 41. No. 1. P. 292–308.
18. *Михайлов В. Г., Шойтов А. М.* О длинных повторениях цепочек в цепи Маркова // *Дискретная математика.* 2014. Т. 26. № 3. С. 79–89.
19. *Михайлов В. Г.* Оценки точности пуассоновской аппроксимации для распределения числа серий повторений длинных цепочек в цепи Маркова // *Дискретная математика.* 2015. Т. 27. № 4. С. 67–78.
20. *Mytalis G. C. and Zazanis M. A.* Central Limit Theorem approximations for the number of runs in Markov-dependent binary sequences // *J. Stat. Plan. Inference.* 2013. V. 143. No. 2. P. 321–333.
21. *Mahmoudzadeh E., Montazeri M. A., Zekri M., and Sadri S.* Extended hidden Markov model for optimized segmentation of breast thermography images // *Infrared Phys. Technol.* 2015. V. 72. P. 19–28.
22. *Yang W., Tao J., and Ye Z.* Continuous sign language recognition using level building based on fast hidden Markov model // *Pattern Recognit. Lett.* 2016. V. 78. P. 28–35.
23. *Elliott R. J., Aggoun L., and Moore J. B.* Hidden Markov Models. N. Y.: Springer, 1995. V. 29. 382 p.
24. *Aston J. A. D. and Martin D. E. K.* Distributions associated with general runs and patterns in hidden Markov models // *Ann. Appl. Stat.* 2007. V. 1. No. 2. P. 585–611.
25. *Меженная Н. М.* О числе совпадений знаков в дискретной случайной последовательности, управляемой цепью Маркова // *Сибирские электронные математические известия.* 2016. Т. 13. С. 305–317.
26. *Розанов Ю. А.* Случайные процессы. Краткий курс. М.: Наука, 1979. 184 с.
27. *Mezhenneya N. M.* On the distribution of the number of runs in polynomial sequence controlled by Markov chain // *OP&PM Surv. Appl. Ind. Math.* 2016. V. 23. No. 2. P. 186–187.
28. *Меженная Н. М.* О предельном распределении числа серий в полиномиальной последовательности, управляемой цепью Маркова // *Вестник УдГУ.* 2016. Т. 26. № 3. С. 324–335.
29. *Barbour A. D., Holst L., and Janson S.* Poisson Approximation. Oxford: Oxford Univ. Press, 1992. 277 p.

REFERENCES

1. *Balakrishnan N. and Koutras M. V.* Runs and Scans with Applications. N. Y., John Wiley & Sons Inc., 2002, 452 p.
2. *Mikhaylov V. G.* Ob asimptoticheskikh svoystvakh chisla seriy sobytiy [On asymptotic properties of the number of runs of events]. *Tr. Diskr. Mat.*, 2006, vol. 9, pp. 152–163. (in Russian)
3. *Aki S. and Hirano K.* Discrete distributions related to succession events in two-state Markov chain. *Statistical Science and Data Analysis.* Eds. K. Matusita, M. L. Puri, T. Hayakawa. Zeist, VSP International Science Publishers, 1993, pp. 467–474.

4. *Aki S. and Hirano K.* Sooner and later waiting time problems for runs in Markov dependent bivariate trials. *Ann. Inst. Stat. Math.*, 1999, vol. 51, pp. 17–29.
5. *Han Q. and Aki S.* Formulae and recursions for the joint distributions of success runs of several lengths in a two-state Markov chain. *Stat. Probab. Lett.*, 1998, vol. 40, no. 3, pp. 203–214.
6. *Savel'ev L. Ja. and Balakin S. V.* The joint distribution of the number of ones and the number of 1-runs in binary Markov sequences. *Discr. Math. Appl.*, 2004, vol. 14, no. 4, pp. 353–372.
7. *Savel'ev L. Ja. and Balakin S. V.* A combinatorial approach to calculation of moments of characteristics of runs in ternary Markov sequences. *Discr. Math. Appl.*, 2011, vol. 21, no. 1, pp. 47–67.
8. *Savel'ev L. Ja.* Raspredeleniya chisla sostoyaniy v dvoichnykh markovskikh stokhasticheskikh modelyakh [Calculation of the number of states in binary Markov stochastic models]. *Sib. Zh. Vychisl. Mat.*, 2015, vol. 18, no. 2, pp. 191–200. (in Russian)
9. *Savel'ev L. Ja. and Balakin S. V.* Nekotorye primeneniya stokhasticheskoy teorii seriy [Some applications of the stochastic theory of runs]. *Sib. Zh. Ind. Mat.*, 2012, vol. 15, no. 3, pp. 111–123. (in Russian)
10. *Shinde R. L. and Kotwal K. S.* On the joint distribution of runs in the sequence of Markov-dependent multi-state trials. *Stat. Probab. Lett.*, 2006, vol. 76, no. 10, pp. 1065–1074.
11. *Geske M. X., Godbole A. P., Schaffner A. A., et al.* Compound Poisson approximations for word patterns under Markovian hypotheses. *J. Appl. Probab.*, 1995, vol. 32, pp. 877–892.
12. *Erhardsson T.* Compound Poisson approximation for Markov chains using Stein's method. *Ann. Probab.*, 1999, vol. 27, no. 1, pp. 565–596.
13. *Chryssaphinou O. and Vaggelatos E.* Compound Poisson approximation for multiple runs in a Markov chain. *Ann. Inst. Stat. Math.*, 2002, vol. 54, no. 2, pp. 411–424.
14. *Fu J. C., Wang L., and Lou W. Y. W.* On exact and large deviation approximation for the distribution of the longest run in a sequence of two-state Markov dependent trials. *J. Appl. Probab.*, 2003, vol. 40, no. 2, pp. 346–360.
15. *Eryilmaz S.* Some results associated with the longest run statistic in a sequence of Markov dependent trials. *Appl. Math. Comput.*, 2006, vol. 175, no. 1, pp. 119–130.
16. *Pinsky M. A. and Karlin S.* The long run behavior of Markov chains. *An Introduction to Stochastic Modeling. Fourth Edition*, eds. M. A. Pinsky, S. Karlin. Boston, Elsevier, 2011, pp. 165–222.
17. *Fu J. C. and Johnson B. C.* Approximate probabilities for runs and patterns in i.i.d. and Markov-dependent multistate trials. *Adv. Appl. Probab. Appl. Probab. Trust*, 2009, vol. 41, no. 1, pp. 292–308.
18. *Mikhailov V. G. and Shoitov A. M.* On repetitions of long tuples in a Markov chain. *Discr. Math. Appl.*, 2015, vol. 25, no. 5, pp. 295–303.
19. *Mikhailov V. G.* Estimates of accuracy of the Poisson approximation for the distribution of number of runs of long string repetitions in a Markov chain. *Discr. Math. Appl.*, 2016, vol. 26, no. 2, pp. 105–113.
20. *Mytalis G. C. and Zazanis M. A.* Central Limit Theorem approximations for the number of runs in Markov-dependent binary sequences. *J. Stat. Plan. Inference*, 2013, vol. 143, no. 2, pp. 321–333.
21. *Mahmoudzadeh E., Montazeri M. A., Zekri M., and Sadri S.* Extended hidden Markov model for optimized segmentation of breast thermography images. *Infrared Phys. Technol.*, 2015, vol. 72, pp. 19–28.
22. *Yang W., Tao J., and Ye Z.* Continuous sign language recognition using level building based on fast hidden Markov model. *Pattern Recognit. Lett.*, 2016, vol. 78, pp. 28–35.

23. *Elliott R. J., Aggoun L., and Moore J. B.* Hidden Markov Models. N.Y., Springer, 1995, vol. 29, 382 p.
24. *Aston J. A. D. and Martin D. E. K.* Distributions associated with general runs and patterns in hidden Markov models. *Ann. Appl. Stat.*, 2007, vol. 1, no. 2, pp. 585–611.
25. *Mezhennaya N. M.* O chisle sovpadeniy znakov v diskretnoy sluchaynoy posledovatel'nosti, upravlyаемой tsep'yu Markova [On the number of characters matchings in discrete random sequence controlled by Markov chain]. *Sib. Elektron. Mat. Izv.*, 2016, vol. 13, pp. 305–317. (in Russian)
26. *Rozanov Yu. A.* Sluchaynye protsessy. Kratkiy kurs [Random processes. Short course]. Moscow, Nauka Publ., 1979, 184 p. (in Russian)
27. *Mezhennaya N. M.* On the distribution of the number of runs in polynomial sequence controlled by Markov chain. *OP&PM Surv. Appl. Ind. Math.*, 2016, vol. 23, no. 2, pp. 186–187.
28. *Mezhennaya N. M.* О предельном распределении числа серий в полиномиальной последовательности, управляемой цепью Маркова [On the limit distribution of a number of runs in polynomial sequence controlled by Markov chain]. *Vestn. Udmurtsk. Univ. Mat. Mekh. Komp. Nauki*, 2016, vol. 26, iss. 3, pp. 324–335. (in Russian)
29. *Barbour A. D., Holst L., and Janson S.* Poisson Approximation. Oxford, Oxford Univ. Press, 1992, 277 p.