

## ПРЕДЕЛЬНЫЕ ТЕОРЕМЫ ПУАССОНОВСКОГО ТИПА ДЛЯ ЧИСЛА ПАР $H$ -СВЯЗАННЫХ ЦЕПОЧЕК

В. Г. МИХАЙЛОВ

Исследованы свойства распределения числа  $\xi(H)$  пар  $H$ -связанных  $s$ -цепочек в отрезке длины  $n$  последовательности независимых одинаково распределенных случайных величин со счетным множеством значений. С помощью локального варианта метода Чена–Стейна получены оценки точности аппроксимации распределения величины  $\xi(H)$  распределением Пуассона. Выведены вытекающие из этих оценок условия сходимости распределения случайной величины  $\xi(H)$  к распределению Пуассона и нормальному распределению. Для случайной величины  $\xi(S_N)$  (число пар цепочек с одинаковой структурой) в равновероятной случайной последовательности над алфавитом  $\{1, \dots, N\}$  получены новые, более слабые по сравнению с известными, условия сходимости к пуассоновскому распределению.

Пусть

$$X = (X_1, \dots, X_{n+s-1})$$

— отрезок последовательности независимых одинаково распределенных случайных величин, принимающих значения в множестве  $\{1, 2, \dots\}$ ,

$$p_k = \mathbf{P}\{X_1 = k\}, \quad k = 1, 2, \dots, \quad p = \max_k p_k, \quad P = np^s.$$

Пусть задан некоторый конечный набор  $H$  подстановок  $h: \{1, 2, \dots\} \rightarrow \{1, 2, \dots\}$ . Две цепочки  $(x_1, \dots, x_s)$  и  $(y_1, \dots, y_s)$  знаков множества  $\{1, 2, \dots\}$  будем называть  $H$ -связанными, если найдется подстановка  $h \in H$ , для которой  $y_i = h(x_i)$ ,  $i = 1, \dots, s$ . Отношение  $H$ -связанности будем обозначать записью

$$(x_1, \dots, x_s) H (y_1, \dots, y_s).$$

Если  $e \in H$  ( $e$  — тождественная подстановка) и  $h^{-1} \in H$  для всех  $h \in H$ , то отношение  $H$ -связанности является симметричным. Если множество подстановок  $H$  является группой, то определяемое им отношение является отношением эквивалентности. В этом случае связанные отношением цепочки называются  $H$ -эквивалентными (см. [1]). Исследованию  $H$ -эквивалентных цепочек в случайных последовательностях посвящены работы [2–10]. Однако все утверждения и доказательства настоящей работы справедливы без этих предположений, поэтому мы их не используем.

Введем обозначения

$$E_{ij}(h) = \{(h(X_i), \dots, h(X_{i+s-1})) = (X_j, \dots, X_{j+s-1})\},$$

$$E_{ij}(H) = \bigcup_{h \in H} E_{ij}(h) = \{(X_i, \dots, X_{i+s-1}) H (X_j, \dots, X_{j+s-1})\}.$$

Каждой паре  $(i, j)$  из множества  $\{1 \leq i < j \leq n\}$  сопоставим случайную величину

$$W(i, j) = I(E_{ij}(H))$$

(здесь и далее выражение  $I(\dots)$  обозначает индикатор случайного события). Тогда число пар  $H$ -связанных среди  $n$  первых  $s$ -цепочек последовательности  $X$  выразится формулой

$$\xi(H) = \sum_{1 \leq i < j \leq n} W(i, j).$$

Пусть

$$D_k = \mathbf{P} \{E_{1,2s+1}(H) E_{1+k,2s+k+1}(H)\}, \quad k = 0, \dots, s-1.$$

В частности,

$$D_0 = \mathbf{P} \{E_{1,2s+1}(H) E_{1,2s+1}(H)\} = \mathbf{P} \{E_{1,2s+1}(H)\}.$$

Пусть

$$D = D_1 + \dots + D_{s-1}.$$

Распределение Пуассона с параметром  $\lambda$  условимся обозначать  $\text{Po}(\lambda)$ . Запись  $\mathcal{L}(\xi)$  обозначает закон распределения случайной величины  $\xi$ .

**ТЕОРЕМА 1.** Пусть  $n \geq 2s - 1$ ,  $n \rightarrow \infty$ , а параметры  $p_k, s$  меняются так, что выполнено условие  $n^2 D + s|H|P \rightarrow 0$ .

1) Если  $\mathbf{E}\xi(H) \rightarrow \lambda < \infty$ , то  $\mathcal{L}(\xi(H)) \rightarrow \text{Po}(\lambda)$ .

2) Если же  $\mathbf{E}\xi(H) \rightarrow \infty$ , то функция распределения случайной величины  $(\xi(H) - \mathbf{E}\xi(H))(\mathbf{E}\xi(H))^{-1/2}$  сходится к стандартной нормальной функции распределения.

**З а м е ч а н и е 1.** Можно рассматривать  $H$ -связанные цепочки, расположенные в разных последовательностях случайных величин. Пусть

$$X = (X_1, \dots, X_m) \quad \text{и} \quad Y = (Y_1, \dots, Y_n)$$

— две последовательности независимых одинаково (внутри каждой последовательности) распределенных случайных величин, принимающих значения в множестве  $\{1, 2, \dots\}$ . Число пар  $H$ -связанных цепочек из последовательностей  $X$  и  $Y$  выражается формулой

$$N(H) = \sum_{i=1}^m \sum_{j=1}^n I(E(i, j)).$$

В этом определении считается, что

$$E(i, j) = \{(X_i, \dots, X_{i+s-1}) H (Y_j, \dots, Y_{j+s-1})\},$$

а

$$X_{m+i} = X_i, \quad Y_{n+i} = Y_i, \quad i = 1, 2, \dots$$

В работе [9] (см. также [10]) было показано, что при предположениях, аналогичных условиям теоремы 1, случайная величина  $N(H)$  имеет такие же предельные распределения, как и  $\xi(H)$ .

**З а м е ч а н и е 2.** Изучение числа наборов из  $H$ -связанных ( $H$ -эквивалентных) цепочек для нетривиальных множеств  $H$  началось с работы С. М. Буравлева [2], где рассматривалась последовательность знаков из множества  $\{1, 2, \dots, N\}$ , а в качестве  $H$  бралось множество (группа)  $S_N$  всех подстановок на  $\{1, 2, \dots, N\}$ . Эти исследования были продолжены в работах С. М. Буравлева [3–5] и А. М. Шойтова [6–8], где в качестве  $H$  рассматривались и другие группы. При  $H = \{e\}$  (напомним, что  $e$  — тождественная подстановка)  $H$ -эквивалентными будут только совпавшие цепочки. Этот случай рассматривался в работах [11–13].

Рассмотрим множество

$$\Gamma_0 = \{1 \leq i < j \leq n: j - i \geq 2s\}.$$

Положим

$$\xi_0(H) = \sum_{(i,j) \in \Gamma_0} W(i, j).$$

**ЛЕММА 1.** Пусть  $n \geq 2s - 1$ . Тогда

$$P \{ \xi(H) \neq \xi_0(H) \} \leq E \xi(H) - E \xi_0(H) < 2s |H| P. \quad (1)$$

Доказательство леммы проводится несложными стандартными рассуждениями, и мы его опускаем.

Введем обозначение

$$W_0 = \{W(i, j), (i, j) \in \Gamma_0\}.$$

Пусть

$$\pi_0 = \{\pi(i, j), (i, j) \in \Gamma_0\}$$

— набор независимых случайных величин, имеющих распределения Пуассона с параметрами

$$\mathbf{E} \pi(i, j) = \mathbf{E} W(i, j).$$

Через  $\pi(\lambda)$  обозначим случайную величину, имеющую распределение Пуассона с параметром  $\lambda$ .

Пусть

$$\lambda_0 = \mathbf{E} \xi_0(H).$$

Заметим, что

$$\lambda_0 = |\Gamma_0| \mathbf{E} W(i, j) = C_{n-2s+1}^2 D_0. \quad (2)$$

Обозначим через  $\rho(F, G)$  расстояние по вариации между распределениями  $F$  и  $G$ . Для распределений, сосредоточенных на множестве  $\{0, 1, \dots\}$ , оно выражается формулой

$$\rho(F, G) = \frac{1}{2} \sum_{k=0}^{\infty} |F\{k\} - G\{k\}|.$$

ТЕОРЕМА 2. Пусть  $n \geq 2s - 1$ . Тогда

$$\rho(\mathcal{L}(W_0), \mathcal{L}(\pi_0)) \leq n^2 D + 2(2s - 1) \lambda_0 |H| P + (2s - 1)^2 |H|^2 P^2. \quad (3)$$

Аналогичная оценка для распределения суммы элементов набора  $W_0$  — случайной величины  $\xi_0(H)$  — выглядит следующим образом.

ТЕОРЕМА 3. Пусть  $n \geq 2s - 1$ . Тогда

$$\begin{aligned} \rho(\mathcal{L}(\xi_0(H)), \mathcal{L}(\pi(\lambda_0))) &\leq \\ &\leq \min\{1, \lambda_0^{-1}\} (n^2 D + 2(2s - 1) \lambda_0 |H| P + (2s - 1)^2 |H|^2 P^2). \end{aligned} \quad (4)$$

Из теоремы 3 и леммы 1 следует теорема 1.

Доказательство теоремы 2. Рассмотрим множество

$$\Gamma_{(i,j)}^{\text{ind}} = \{(i', j') \in \Gamma_0: \min\{|i' - i|, |i' - j|, |j' - i|, |j' - j|\} \geq s\}.$$

Нетрудно проверить, что случайная величина  $W(i, j)$  не зависит от совокупности случайных величин

$$\{W(i', j'), (i', j') \in \Gamma_{(i,j)}^{\text{ind}}\}.$$

Положим

$$\Gamma_{(i,j)} = (\Gamma_0 \setminus \{(i, j)\}) \setminus \Gamma_{(i,j)}^{\text{ind}}.$$

Воспользуемся теоремой 10.A из книги [14], из которой следует, что

$$\rho(\mathcal{L}(W_0), \mathcal{L}(\pi_0)) \leq$$

$$\leq \sum_{(i,j) \in \Gamma_0} \mathbf{E}W(i, j)(\mathbf{E}W(i, j) + \mathbf{E}U(i, j)) + \sum_{(i,j) \in \Gamma_0} \mathbf{E}W(i, j)U(i, j), \quad (5)$$

где

$$U(i, j) = \sum_{(i', j') \in \Gamma_{(i,j)}} W(i', j').$$

Выражение типа (5) оценивалось в работе [9]. Мы используем аналогичные рассуждения.

Чтобы оценить первую сумму в правой части (5), заметим, что

$$|\{(i, j)\} \cup \Gamma_{(i,j)}| = |\Gamma_0 \setminus \Gamma_{(i,j)}^{\text{ind}}| \leq 2(2s-1)(n-2s).$$

Поэтому, используя равенство (2) и неравенство  $nD_0 \leq |H|P$ , получаем

$$\begin{aligned} \sum_{(i,j) \in \Gamma_0} \mathbf{E}W(i, j)(\mathbf{E}W(i, j) + \mathbf{E}U(i, j)) &\leq \\ &\leq 2(2s-1)(n-2s)C_{n-2s+1}^2 D_0^2 \leq 2(2s-1)|H|P\lambda_0. \end{aligned} \quad (6)$$

Оценим вторую сумму в правой части (5). Введем множества

$$\Gamma_{(i,j)}^s = \{(i', j') \in \Gamma_{(i,j)}: -s < i' - i = j' - j < s\},$$

$$\Gamma_{(i,j)}^* = \Gamma_{(i,j)} \setminus \Gamma_{(i,j)}^s,$$

$$\Gamma_{(i,j)}^{**} = \{(i', j') \in \Gamma_{(i,j)}^*: 0 < |i' - i| < s, 0 < |j' - j| < s\}.$$

ЛЕММА 2. Пусть  $(i, j) \in \Gamma_0$ .

(а) При  $(i', j') \in \Gamma_{(i,j)}^* \setminus \Gamma_{(i,j)}^{**}$  выполнены неравенства

$$\mathbf{E}W(i, j)W(i', j') \leq |H|p^s \mathbf{E}W(i, j).$$

(6) При  $(i', j') \in \Gamma_{(i,j)}^{**}$  имеет место неравенство

$$\mathbf{E} W(i, j) W(i', j') \leq |H|^2 p^{2s}.$$

Доказательство леммы проводится по схеме доказательства леммы 1 работы [9] и отличается от последнего лишь в некоторых деталях. Поэтому мы его опускаем.

Из определений следует неравенство

$$\sum_{(i,j) \in \Gamma_0} \sum_{(i',j') \in \Gamma_{(i,j)}^*} \mathbf{E} W(i, j) W(i', j') \leq 2C_{n-2s+1}^2 (D_1 + \dots + D_{s-1}) \leq n^2 D.$$

Используя лемму 2 и оценки

$$|\Gamma_{(i,j)}^* \setminus \Gamma_{(i,j)}^{**}| \leq |\Gamma_{(i,j)}| < 2(2s-1)(n-2s),$$

получаем

$$\begin{aligned} \sum_{(i,j) \in \Gamma_0} \sum_{(i',j') \in \Gamma_{(i,j)}^* \setminus \Gamma_{(i,j)}^{**}} \mathbf{E} W(i, j) W(i', j') &\leq \\ &\leq \lambda_0 |\Gamma_{(i,j)}^* \setminus \Gamma_{(i,j)}^{**}| H p^s < 2(2s-1) |H| P \lambda_0, \end{aligned}$$

$$\begin{aligned} \sum_{(i,j) \in \Gamma_0} \sum_{(i',j') \in \Gamma_{(i,j)}^{**}} \mathbf{E} W(i, j) W(i', j') &\leq \sum_{(i,j) \in \Gamma_0} |\Gamma_{(i,j)}^{**}| H^2 p^{2s} = \\ &= C_{n-2s+1}^2 (2s-1)^2 |H|^2 p^{2s} < \frac{1}{2} ((2s-1) |H| P)^2. \end{aligned}$$

Подставив эти оценки вместе с (6) в (5), приходим к неравенству (3). Теорема 2 доказана.

Доказательство теоремы 3 проводится аналогично; вместо теоремы 10.A из книги [14] надо воспользоваться теоремой 1.A из [14]. Тогда в оценке появится дополнительный множитель  $\min\{1, \lambda_0^{-1}\}$ . Остальные выкладки повторяются дословно, и мы их опускаем.

Воспользуемся теперь теоремой 1 для изучения свойств случайной величины  $\xi(S_N)$  в равновероятной последовательности.

Пусть

$$X = (X_1, \dots, X_{n+s-1})$$

— отрезок последовательности независимых одинаково распределенных случайных величин, принимающих значения в множестве  $\{1, 2, \dots, N\}$  с равными вероятностями  $1/N$ .

В 1999 г. С. М. Буравлевым был получен следующий результат (приводим его в удобном для нас виде).

**ТЕОРЕМА [2].** Пусть случайные величины  $X_t$  независимы и имеют равномерное распределение на множестве  $\{1, \dots, N\}$ . Если  $n, N, s \rightarrow \infty$  так, что

$$\mathbf{E} \xi(S_N) \rightarrow \lambda, \quad 0 < \lambda < \infty, \quad (7)$$

$$sN^{-1} \geq \ln N + 1, \quad (8)$$

то распределение случайной величины  $\xi(S_N)$  сходится к распределению Пуассона с параметром  $\lambda$ .

Мы покажем, что условие (8) можно ослабить.

**ТЕОРЕМА 4.** Пусть случайные величины  $X_t$  независимы и имеют равномерное распределение на множестве  $\{1, \dots, N\}$ , параметры  $n, N, s \rightarrow \infty$  так, что  $n \geq 2s - 1$ , выполнены условие (7) и условие

$$\frac{s}{N} - \frac{1}{4} \ln N \rightarrow \infty. \quad (9)$$

Тогда распределение случайной величины  $\xi(S_N)$  сходится к распределению Пуассона с параметром  $\lambda$ .

При доказательстве теоремы 4 нам понадобится

**ЛЕММА 3.** Пусть  $\mu_0(s, N)$  — число пустых ячеек в равновероятной схеме размещения  $s$  частиц по  $N$  ячейкам.

Тогда

$$\mathbf{P} \{E_{1,2s+1}(S_N)\} = \frac{N!}{N^s} \sum_{k=0}^{N-1} \frac{1}{k!} \mathbf{P} \{\mu_0(s, N) = k\}. \quad (10)$$

**Доказательство.** Как показано в [2],

$$\mathbf{P} \{E_{1,2s+1}(S_N)\} = N^{-2s} \sum_{m=1}^{\min\{s, N\}} \sigma(s, m) ((N)_m)^2,$$

где

$$(N)_m = N(N-1)\dots(N_m+1),$$

а  $\sigma(s, m)$  — числа Стирлинга второго рода (см., например, [15, с. 127]). Аналогичные рассуждения дают формулу

$$\mathbf{P}\{E_{1,2s+1}(S_N) \cap \{\mu(1) = m\}\} = N^{-2s} \sigma(s, m) ((N)_m)^2.$$

Здесь  $\mu(1)$  — число различных значений, принятых случайными величинами  $X_1, \dots, X_s$ . Так как элементы последовательности  $X$  имеют равномерное распределение на множестве  $\{1, \dots, N\}$ , то

$$\mathbf{P}\{\mu(1) = m\} = \sigma(s, m) (N)_m N^{-s}, \quad m = 1, \dots, s.$$

Поэтому

$$\mathbf{P}\{E_{1,2s+1}(S_N) \cap \{\mu(1) = N - r\}\} = \frac{N!}{N^s r!} \mathbf{P}\{\mu(1) = N - r\}. \quad (11)$$

Учитывая равенство

$$\mu(1) = N - \mu_0(s, N)$$

(в смысле равенства распределений), из (11) получаем (10). Лемма доказана.

**З а м е ч а н и е 4.** Формулу (10) можно переписать в виде

$$\mathbf{P}\{E_{1,2s+1}(S_N)\} = N^{-s} \mathbf{E}(N)_{\mu(s, N)}, \quad (12)$$

где  $\mu(s, N)$  — число занятых ячеек при равновероятном размещении  $s$  частиц по  $N$  ячейкам.

**Доказательство теоремы 4.** Мы хотим воспользоваться теоремой 1. Поэтому наша задача состоит в проверке условия

$$n^2 D \rightarrow 0, \quad s|H|P = snN!N^{-s} \rightarrow 0. \quad (13)$$

Следующее утверждение указывает условия, при которых выполняется вторая часть (13).

**ЛЕММА 4.** *Существует такая константа  $\beta < \infty$ , что при  $n, N, s \rightarrow \infty$  и условиях*

$$s \geq \beta N, \quad (14)$$

$$n^2 \frac{N!}{N^s} \sum_{k=r}^{N-1} \frac{1}{k!} \mathbf{P}\{\mu_0(s, N) = k\} = O(1) \quad (15)$$

*имеет место сходимость  $snN!N^{-s} \rightarrow 0$ .*

Отметим сразу, что условия леммы 4 вытекают непосредственно из условий теоремы 4.

Доказательство леммы. Положим

$$\alpha = sN^{-1}, \quad \Lambda = Ne^{-\alpha}.$$

Сначала рассмотрим случай, когда  $\Lambda^{-1} = O(1)$ . Воспользуемся тем, что при  $s, N \rightarrow \infty$

$$E\mu_0(s, N) = N \left(1 - \frac{1}{N}\right)^s = \Lambda(1 + O(N^{-1} \ln N)),$$

а расстояние по вариации между  $\mathcal{L}(\mu_0(s))$  и  $\text{Po}(\Lambda)$  оценивается согласно теореме 2 из [16] как  $O((\alpha + 1)e^{-\alpha})$  при  $n, N \rightarrow \infty$  и любом поведении параметра  $\alpha$ . Поэтому

$$\begin{aligned} \sum_{k=0}^{N-1} \frac{1}{k!} \mathbf{P} \{ \mu_0(s, N) = k \} &\geq \frac{1}{([\Lambda])!} \mathbf{P} \{ 0 \leq \mu_0(s, N) \leq \Lambda \} \geq \\ &\geq \frac{1}{[\Lambda]!} (\mathbf{P} \{ 0 \leq \pi(\Lambda) \leq \Lambda \} - C(\alpha + 1)e^{-\alpha}), \end{aligned} \quad (16)$$

где  $\pi(\Lambda)$  — случайная величина, имеющая распределение Пуассона с параметром  $\Lambda$ ,  $[\Lambda]$  — целая часть числа  $\Lambda$ , а  $C < \infty$ . Заметим, что из условий леммы 4 следует, что

$$(\alpha + 1)e^{-\alpha} \leq (\beta + 1)e^{-\beta},$$

и выбором значения  $\beta$  это выражение можно сделать сколь угодно малым. Кроме того,  $\Lambda^{-1} = O(1)$ . Поэтому из (16) получаем, что для некоторого положительного числа  $\varepsilon$  найдется такое число  $\beta(\varepsilon)$ , что при  $s \geq \beta(\varepsilon)N$  и достаточно больших  $n, N$  выполнено неравенство

$$\sum_{k=0}^{N-1} \frac{1}{k!} \mathbf{P} \{ \mu_0(s, N) = k \} \geq \frac{\varepsilon}{[\Lambda]!}.$$

Вместе с (15) это дает оценку

$$n^2 = O\left(\frac{N^s [\Lambda]!}{N!}\right). \quad (17)$$

Таким образом,

$$snN!N^{-s} = O(s(N!N^{-s}[\Lambda]!)^{1/2}).$$

Возьмем

$$\beta = \max\{\beta(\varepsilon), z_0 + 1\},$$

где  $z_0$  — корень уравнения  $1 + e^{-z} = z$ , и покажем, что при  $n, N, s \rightarrow \infty$ ,  $s \geq \beta N$

$$s^2 N! N^{-s} [\Lambda]! \rightarrow 0. \tag{18}$$

Этим соотношением  $snN!N^{-s} \rightarrow 0$  будет доказано.

Используя свойство  $\Lambda^{-1} = O(1)$  и вытекающее из (14) неравенство  $\Lambda \leq Ne^{-\beta}$ , оцениваем логарифм левой части соотношения (18) сверху выражением

$$\begin{aligned} N(\ln N - 1) + (Ne^{-\beta} + r)(\ln N - \beta - 1) - \beta N \ln N + O(\ln N) = \\ = (1 + e^{-\beta} - \beta) N \ln N + O(N), \end{aligned}$$

которое стремится к  $-\infty$ , так как

$$1 + e^{-\beta} - \beta < 0 \quad \text{при} \quad \beta > z_0.$$

Значит, соотношение (18) выполнено.

Пусть теперь  $\Lambda \rightarrow 0$ . Тогда

$$\sum_{k=0}^{N-1} \frac{1}{k!} \mathbf{P} \{ \mu_0(s, N) = k \} \geq \mathbf{P} \{ \mu_0(s, N) = 0 \} \geq 1 - \Lambda.$$

Поэтому, повторив приведенные выше рассуждения, приходим к выводу, что в этом случае вместо (18) достаточно проверить соотношение

$$s^2 N! N^{-s} \rightarrow 0.$$

Оно выполнено при  $s \geq N$  и тем более при  $s \geq \beta N$ .

Теперь из доказанного для случаев  $\Lambda^{-1} = O(1)$  и  $\Lambda \rightarrow 0$  стандартными рассуждениями выводим соотношение  $snN!N^{-s} \rightarrow 0$  при любом допустимом условиями леммы поведении  $\Lambda$ . Лемма 4 доказана.

Для проверки первой части условия (13) нам понадобятся следующие оценки величин  $D_m$ ,  $m = 1, \dots, s - 1$ .

ЛЕММА 5. При  $\varepsilon < 1$  и  $N \geq (1 - \varepsilon^2)^{-1}$

$$D_m < \left( \frac{N}{1 - \varepsilon} \left( \frac{1}{N} + \varepsilon^2 \right)^2 \right)^m D_0 + \frac{N!}{N^{s-m} [\varepsilon N]!}. \tag{19}$$

ЛЕММА 6. Если  $m \leq N$ , то

$$D_m \leq D_0 \exp \left\{ -\frac{m(m-1)}{2N} \right\}. \tag{20}$$

Если же  $m \geq N$ , то

$$D_m \leq D_0 N^{-m} N!. \tag{21}$$

Доказательство леммы 5. Нам надо оценить вероятность

$$D_m = \mathbf{P} \{E_{1,2s+1}(S_N) E_{1+m,2s+m+1}(S_N)\}, \quad (22)$$

где  $m = 1, \dots, s-1$  и

$$E_{ij}(S_N) = \{(X_i, \dots, X_{i+s-1}) S_N(X_j, \dots, X_{j+s-1})\}.$$

Указанное в правой части (22) событие влечет за собой событие

$$E(m, S_N) = \{(X_{m+1}, \dots, X_s) S_N(X_{2s+m+1}, \dots, X_{3s})\}$$

(оно относится к величинам  $X_t$  из общих участков цепочек  $(X_1, \dots, X_s)$  и  $(X_{m+1}, \dots, X_{s+m})$  и общих участков цепочек  $(X_{2s+1}, \dots, X_{3s})$  и  $(X_{2s+m+1}, \dots, X_{3s+m})$ ) и связывает определенным образом остальные величины  $X_t$ , участвующие в записи (22).

Заметим, что событие  $E(m, S_N)$  влечет наличие случайного обратимого отображения  $\varphi$ , которое удовлетворяет условиям

$$\varphi(X_t) = X_{2s+t}, \quad t = m+1, \dots, s$$

и переводит множество

$$M_X = \bigcup_{t=m+1}^s \{X_t\} \text{ в множество } \varphi(M_X) = \bigcup_{t=m+1}^s \{X_{2s+t}\}.$$

Случайное отображение  $\varphi$  определяется этим правилом лишь на случайном множестве  $M_X$ , но при выполнении событий  $E_{1,2s+1}(S_N)$  и  $E_{1+m,2s+m+1}(S_N)$  доопределяется (вообще говоря, разными способами) до подстановок на множестве  $\{1, \dots, N\}$ , реализующих указанные события.

Для подмножества  $I_1 \subseteq \{1, \dots, m\}$  положим  $J_1 = \{1, \dots, m\} \setminus I_1$ . В предположении, что выполнено событие  $E(m, S_N)$  и, следовательно, задано отображение  $\varphi$ , определим событие

$$A(I_1) = \{X_i \in M_X, X_{2s+i} = \varphi(X_i), i \in I_1\} \cap \\ \cap \{X_j \notin M_X, X_{2s+j} \notin \varphi(M_X), j \in J_1\}.$$

Для  $I_2 \subseteq \{s+1, \dots, s+m\}$  положим  $J_2 = \{s+1, \dots, s+m\} \setminus I_2$  и определим по тому же правилу событие  $A(I_2)$ .

Заметим, что событие  $\bigcup_{I_1 \subseteq \{1, \dots, m\}} A(I_1)$  (при выполнении события  $E(m, S_N)$ ) необходимо для возможности доопределения отображе-

ния  $\varphi$  до подстановки  $h \in S_N$ , для которой  $X_{2s+t} = h(X_t)$ ,  $t = 1, \dots, s$ , т.е. до подстановки, реализующей событие  $E_{1,2s+1}(S_N)$ . Событие  $\bigcup_{I_2 \subseteq \{s+1, \dots, s+m\}} A(I_2)$  аналогичным образом связано с событием  $E_{1+m,2s+m+1}(S_N)$ . Поэтому, рассматривая события как множества элементарных исходов, можем записать соотношение

$$E_{1,2s+1}(S_N) \cap E_{1+m,2s+1+m}(S_N) \subseteq E(m, S_N) \cap \bigcup_{I_1 \subseteq \{1, \dots, m\}} \bigcup_{I_2 \subseteq \{s+1, \dots, s+m\}} A(I_1) \cap A(I_2). \quad (23)$$

При  $u = 1, 2$  и любом заданном множестве  $M \subseteq \{1, \dots, N\}$  условная вероятность

$$P\{A(I_u) \mid E(m, S_N), M_X = M\}$$

зависит лишь от множества  $M$  (точнее, от числа его элементов  $|M|$ ) и выполнены равенства

$$\begin{aligned} P\{A(I_u) \mid E(m, S_N), M_X = M\} &= \\ &= \prod_{i \in I_u} (P\{X_i \in M\} P\{X_{i+2s} = \varphi(X_i) \mid E(m, S_N), M_X = M\}) \times \\ &\times \prod_{j \in J_u} (P\{X_j \notin M\} P\{X_{j+2s} \notin \varphi(M) \mid E(m, S_N), M_X = M\}) = \\ &= \frac{|M|^{|I_u|}}{N^{2|I_u|}} \left(1 - \frac{|M|}{N}\right)^{2|J_u|}. \quad (24) \end{aligned}$$

В этих выкладках существенно использована равновероятность случайных величин  $X_t$ , при которой выполнены равенства

$$P\{X_{i+2s} = \varphi(X_i) \mid E(m, S_N), M_X = M\} = \frac{1}{N},$$

$$P\{X_{j+2s} \notin \varphi(M) \mid E(m, S_N), M_X = M\} = 1 - \frac{|M|}{N}.$$

В ином случае эти вероятности, вообще говоря, зависят от случайного отображения  $\varphi$ .

Для сокращения записей введем обозначения

$$\Sigma_1 = \sum_{I_1 \subseteq \{1, \dots, m\}} P\{A(I_1) \mid E(m, S_N), M_X = M\},$$

$$\Sigma_2 = \sum_{I_2 \subseteq \{s+1, \dots, s+m\}} \mathbf{P} \{A(I_2) \mid E(m, S_N), M_X = M\}.$$

Согласно (24)

$$\Sigma_1 = \Sigma_2 = \left( \frac{|M|}{N^2} + \left(1 - \frac{|M|}{N}\right)^2 \right)^m.$$

Используя эти формулы и соотношения (22) и (23), получаем

$$\begin{aligned} D_m &= \mathbf{P} \{E_{1,2s+1}(S_N) \cap E_{1+m,2s+1+m}(S_N)\} \leq \\ &\leq \sum_{M \subseteq \{1, \dots, N\}} \mathbf{P} \{E(m, S_N), M_X = M\} \Sigma_1 \Sigma_2 = \\ &= \sum_{M \subseteq \{1, \dots, N\}} \mathbf{P} \{E(m, S_N), M_X = M\} \left( \frac{|M|}{N^2} + \left(1 - \frac{|M|}{N}\right)^2 \right)^{2m} = \\ &= \sum_{k=0}^{N-1} \mathbf{P} \{E(m, S_N), |M_X| = N - k\} \left( \frac{1}{N} + \frac{k(k-1)}{N^2} \right)^{2m}. \end{aligned}$$

Согласно формулам (11) и (10)

$$\mathbf{P} \{E(m, S_N), |M_X| = N - k\} = \frac{N!}{N^{s-m} k!} \mathbf{P} \{\mu_0(s-m, N) = k\}.$$

Поэтому из предыдущих выкладок вытекает оценка

$$D_m \leq \frac{N!}{N^{s-m}} \sum_{k=0}^{N-1} \frac{1}{k!} \mathbf{P} \{\mu_0(s-m, N) = k\} \left( \frac{1}{N} + \frac{k(k-1)}{N^2} \right)^{2m}. \quad (25)$$

Заметим теперь, что при  $k \leq \varepsilon N$

$$\begin{aligned} \mathbf{P} \{\mu_0(s, N) = k\} &= \mathbf{P} \{\mu_0(s-1, N) = k\} \left(1 - \frac{k}{N}\right) + \\ &+ \mathbf{P} \{\mu_0(s-1, N) = k+1\} \frac{k+1}{N} \geq (1-\varepsilon) \mathbf{P} \{\mu_0(s-1, N) = k\} \geq \dots \\ &\dots \geq (1-\varepsilon)^m \mathbf{P} \{\mu_0(s-m, N) = k\} \geq \\ &\geq (C(\varepsilon))^{-m} \mathbf{P} \{\mu_0(s-m, N) = k\} \left( \frac{1}{N} + \frac{k(k-1)}{N^2} \right)^{2m}, \end{aligned}$$

где

$$C(\varepsilon) = (1 - \varepsilon)^{-1} \left( \frac{1}{N} + \varepsilon^2 \right)^2.$$

Поэтому при  $N \geq (1 - \varepsilon^2)^{-1}$

$$\begin{aligned} \sum_{k=0}^{N-1} \frac{1}{k!} \mathbf{P} \{ \mu_0(s-m, N) = k \} \left( \frac{1}{N} + \frac{k(k-1)}{N^2} \right)^{2m} &< \\ &< \sum_{k=0}^{[\varepsilon N]} \frac{1}{k!} \mathbf{P} \{ \mu_0(s-m, N) = k \} \left( \frac{1}{N} + \frac{k(k-1)}{N^2} \right)^{2m} + \\ &+ \sum_{k=[\varepsilon N]+1}^N \frac{1}{k!} \mathbf{P} \{ \mu_0(s-m, N) = k \} < \\ &< (C(\varepsilon))^m \sum_{k=0}^{N-1} \frac{1}{k!} \mathbf{P} \{ \mu_0(s, N) = k \} + \frac{1}{[\varepsilon N]!}. \quad (26) \end{aligned}$$

Используя теперь (10), (25) и (26), получаем неравенство (19). Лемма 5 доказана.

Доказательство леммы 6. Рассмотрим событие

$$F(m, S_N) = \{ (X_{s+1}, \dots, X_{s+m}) S_N (X_{3s+1}, \dots, X_{3s+m}) \}.$$

Очевидно, что события  $E_{1,2s+1}(S_N)$  и  $F(m, S_N)$  независимы и

$$E_{1,2s+1}(S_N) \cap E_{1+m,2s+1+m}(S_N) \subseteq E_{1,2s+1}(S_N) \cap F(m, S_N).$$

Поэтому

$$\begin{aligned} D_m = \mathbf{P} \{ E_{1,2s+1}(S_N) \cap E_{1+m,2s+1+m}(S_N) \} &\leq \\ &\leq \mathbf{P} \{ E_{1,2s+1}(S_N) \} \mathbf{P} \{ F(m, S_N) \} = D_0 \mathbf{P} \{ F(m, S_N) \}. \quad (27) \end{aligned}$$

Согласно формуле (12) и неравенству  $\mu(m, N) \leq \min \{ m, N \}$  получаем оценки

$$\begin{aligned} \mathbf{P} \{ F(m, S_N) \} = N^{-m} \mathbf{E}(N)_{\mu(m, N)} &\leq N^{-m} (N)_m = \\ &= \prod_{k=0}^{m-1} \left( 1 - \frac{k}{N} \right) \leq \exp \left\{ -\frac{m(m-1)}{2N} \right\} \quad (28) \end{aligned}$$

при  $m \leq N$  и оценки

$$\mathbf{P} \{F(m, S_N)\} \leq N^{-m} N! \quad (29)$$

при  $m \geq N$ . Из (27), (28) и (29) следуют (20) и (21). Лемма 6 доказана.

Перейдем к проверке первого из условий (13), которое мы разделим на три соотношения

$$S_1 = n^2 \sum_{1 \leq m \leq \sqrt{N} \ln N} D_m \rightarrow 0, \quad (30)$$

$$S_2 = n^2 \sum_{\sqrt{N} \ln N < m < N-1} D_m \rightarrow 0, \quad (31)$$

$$S_3 = n^2 \sum_{m=N}^{s-1} D_m \rightarrow 0. \quad (32)$$

Для проверки соотношения (30) воспользуемся оценкой (19) при

$$\varepsilon = N^{-1/4}(\varepsilon(s, N) + (\ln N)^{-1}), \quad \varepsilon(s, N) = N^{1/4} e^{-s/N}.$$

По условию (9) при переходе к пределу  $\varepsilon(s, N) \rightarrow 0$ . Получаем, что при достаточно больших  $N$

$$\begin{aligned} S_1 &< n^2 D_0 \sum_{1 \leq m \leq \sqrt{N} \ln N} \left( \frac{N}{1-\varepsilon} \left( \frac{1}{N} + \varepsilon^2 \right)^2 \right)^m + \frac{n^2 N!}{N^s [\varepsilon N]!} \sum_{1 \leq m \leq \sqrt{N} \ln N} N^m < \\ &< 2n^2 D_0 \left( \varepsilon(s, N) + \frac{1}{\ln N} \right)^4 + \frac{2n^2 N!}{N^s [N^{3/4}(\varepsilon(s, N) + (\ln N)^{-1})]} N^{\sqrt{N} \ln N}. \quad (33) \end{aligned}$$

Сначала изучим второе слагаемое в правой части (33). По формуле (17)

$$\frac{n^2 N!}{N^s [N^{3/4}(\varepsilon(s, N) + (\ln N)^{-1})]!} N^{\sqrt{N} \ln N} = O \left( \frac{[\Lambda]! N^{\sqrt{N} \ln N}}{[N^{3/4}(\varepsilon(s, N) + (\ln N)^{-1})]!} \right),$$

а логарифм выражения под знаком « $O$ -большое» равен

$$\Lambda(\ln[\Lambda] - 1) - \frac{3}{4} N^{3/4} \varepsilon(s, N) \ln N - \frac{3}{4} N^{3/4} + o(N^{3/4}). \quad (34)$$

Так как

$$\Lambda = Ne^{-s/N} = N^{3/4}\varepsilon(s, N),$$

то

$$\Lambda(\ln[\Lambda] - 1) < N^{3/4}\varepsilon(s, N)\left(\frac{3}{4}\ln N - 1\right).$$

Поэтому выражение (34) стремится к  $-\infty$ . Следовательно,

$$\frac{2n^2 N!}{N^s [N^{3/4}\varepsilon(s, N)]!} N^{\sqrt{N} \ln N} \rightarrow 0. \quad (35)$$

Перейдем к первому слагаемому в (33). Нам достаточно показать, что

$$n^2 D_0 = O(1). \quad (36)$$

Для этого заметим, что при предельном переходе в теореме 4

$$E\xi(S_N) \geq E\xi_0(S_N) = \frac{n^2}{2} D_0(1 + o(1)).$$

Поэтому соотношение (36) вытекает из условия (7).

Следовательно, первое слагаемое в правой части (33) стремится к нулю. Вместе с (35) это дает соотношение  $S_1 \rightarrow 0$ .

При проверке соотношения (31) воспользуемся неравенством (20). Получаем

$$\begin{aligned} S_2 &\leq n^2 D_0 \sum_{\sqrt{N} \ln N < m \leq N-1} e^{-m(m-1)/(2N)} \leq \\ &\leq n^2 D_0 \sum_{\sqrt{N} \ln N < m \leq N-1} e^{-m(\sqrt{N} \ln N - 1)/(2N)} = \\ &= n^2 D_0 O\left(\frac{\sqrt{N}}{\ln N} e^{-\ln^2 N/2}\right) = n^2 D_0 O((\ln N)^{-1} e^{(1-\ln N)/2}). \end{aligned}$$

Как уже было показано,  $n^2 D_0 = O(1)$ . Поэтому  $S_2 \rightarrow 0$  при  $N \rightarrow \infty$ .

Проверка условия (32) проводится аналогично с использованием оценки (21). Получаем, что

$$S_3 = O(N! N^{-N}) \rightarrow 0 \quad \text{при} \quad N \rightarrow \infty.$$

Этим проверка условий (30)–(32) и (13) завершена.

Теперь можно воспользоваться утверждением 1) теоремы 1. Теорема 4 доказана.

Автор признателен А. М. Зубкову, Г. И. Ивченко и А. М. Шойтову за полезные замечания.

## СПИСОК ЛИТЕРАТУРЫ

1. Сачков В. Н. Введение в комбинаторные методы дискретной математики. — М.: Наука, 1982.
2. Буравлев С. М. Повторения с точностью до перестановки в последовательности независимых испытаний. — Дискрет. матем., 1999, т. 11, в. 2, с. 53–75.
3. Буравлев С. М. Предельные теоремы для случайных величин от числа пар отрезков, связанных подстановками из латинского прямоугольника. — В сб.: Третья Всероссийская школа-коллоквиум по стохастическим методам. Тезисы докладов. — М.: ТВП, 1996, с. 37–40.
4. Буравлев С. М. Локальная асимптотическая нормальность числа пар отрезков последовательности, связанных подстановками из латинского прямоугольника. — Обзор. прикл. промышл. матем., 1998, т. 5, в. 2, с. 205–207.
5. Буравлев С. М. Уточнение одного предельного распределения. — Обзор. прикл. промышл. матем., 1999, т. 6, в. 1, с. 127–128.
6. Шойтов А. М. Предельные распределения числа наборов  $H$ -эквивалентных отрезков в последовательности независимых испытаний. — Обзор. прикл. промышл. матем., 1999, т. 6, в. 1, с. 218–220.
7. Шойтов А. М. Связь отрезков отношением  $H$ -эквивалентности в схеме серий. — Обзор. прикл. промышл. матем., 1999, т. 6, в. 1, с. 220–221.
8. Шойтов А. М. Предельные распределения числа наборов  $H$ -эквивалентных отрезков в равновероятной полиномиальной схеме серий. — Дискрет. матем., 2002, т. 14, в. 1, с. 82–98.
9. Михайлов В. Г. Об асимптотических свойствах распределения числа пар  $H$ -связанных цепочек. — Дискрет. матем., 2002, т. 14, в. 3, с. 122–129.
10. Михайлов В. Г. Предельная теорема Пуассона для числа  $H$ -совпадений цепочек. — Обзор. прикл. промышл. матем., 2000, т. 7, с. 123–124.
11. Зубков А. М., Михайлов В. Г. Предельные распределения случайных величин, связанных с длинными повторениями в последовательности независимых испытаний. — Теор. вероятн. примен., 1974, т. 19, в. 1, с. 173–181.
12. Зубков А. М., Михайлов В. Г. О повторениях  $s$ -цепочек в последовательности независимых величин. — Теор. вероятн. примен., 1979, т. 24, в. 2, с. 267–279.
13. Михайлов В. Г. Оценка точности сложной пуассоновской аппроксимации для распределения числа совпадающих цепочек. — Теор. вероятн. примен., 2001, т. 46, в. 4, с. 713–723.

14. *Barbour A. D., Holst L., Janson S.* Poisson Approximation. — Oxford: Oxford Univ. Press, 1992.
15. *Сачков В. Н.* Комбинаторные методы дискретной математики. — М.: Наука, 1977.
16. *Михайлов В. Г.* О пуассоновской аппроксимации для распределения числа пустых ячеек в неоднородной схеме размещения. — Теор. вероятн. примен., 1997, т. 42, в. 1, с. 184–189.