УДК 519.212.2:681.51

В.И. МАСОЛ, С.В. ПОПЕРЕШНЯК

ПРОВЕРКА СЛУЧАЙНОСТИ РАСПОЛОЖЕНИЯ БИТОВ В ЛОКАЛЬНЫХ УЧАСТКАХ (0, 1)-ПОСЛЕДОВАТЕЛЬНОСТИ

Аннотация. Установлен явный вид совместного распределения числа 2-цепочек и числа 3-цепочек различных фиксированных вариантов в (0, 1)-последовательности длины n, состоящей из нулей и единиц. Предполагается, что элементы (0, 1)-последовательности — это независимые одинаково распределенные случайные величины. Даны таблицы, иллюстрирующие применение установленных формул для (0, 1)-последовательности длины $n = 16$.

Ключевые слова: *s*-цепочки, битовая последовательность, случайность, локальные участки, совместное распределение.

ВВЕДЕНИЕ

Потребность анализа конечных последовательностей, состоящих из двух элементов, с целью выявления присущих им особенностей возникает в различных прикладных задачах. Поиски ответов на вопросы о случайности и взаимосвязи этих элементов (вероятностный аспект [1, 2]), о путях и сложности построения указанных последовательностей (алгоритмический аспект [3]), а также об оценке расстояния между ними (метрический аспект [4]) свидетельствуют о неослабевающем интересе к исследованиям, проводимым в направлении всестороннего изучения двухэлементных упорядоченных конечных совокупностей.

Настоящая статья касается первого аспекта, так как в ней рассматриваются многомерные вероятностные характеристики конечных последовательностей, состоящих из нулей и единиц. Не вызывает сомнения тот факт, что явный вид совместного распределения двух или трех статистик, характеризующих расположение нулей и единиц в (0, 1)-последовательности, позволяет получить не менее точный результат (по сравнению с использованием одномерной статистики) при рассмотрении гипотезы случайности.

Поскольку в тестах на проверку гипотезы случайности используются преимущественно асимптотические оценки распределений одномерных статистик, важным является получение точных совместных распределений вероятностных характеристик конечных (0, 1)-последовательностей. Цель данной статьи — получение точных совместных распределений числа 2-цепочек и числа 3-цепочек заданного вида в (0, 1)-последовательности, образованной конечным случайным количеством нулей и единиц.

ПОСТАНОВКА ЗАДАЧИ

Анализ случайных последовательностей, генерируемых соответствующими конечными абстрактными автоматами, проводится на основании различных статистик (например, [1]). В частности, он может проводиться с одновременным использованием *s*-цепочек и $(s+1)$-цепочек [2, с. 367] или пересекающихся *s*-цепочек [5, с. 162]. Под *s*-цепочкой, $1 \le s \le n$, понимается произвольная подпоследовательность $\gamma_i, \gamma_{i+1}, \dots, \gamma_{i+s-1}$, $i = 1, 2, \dots, n-s+1$, последовательности случайных величин

$$\gamma_1, \gamma_2, \dots, \gamma_n, \tag{1}$$

где γ_k, $k = 1, 2, \ldots, n$, принимает значение из некоторого дискретного множества мощности N. Статистические свойства s-цепочек изучались при различных ограничениях на случайную последовательность (1). Так, в [6] установлена характеристическая функция предельного ($s \to \infty$, $n > s$) совместного распределения частот s-цепочек при условиях, что (1) — это последовательность случайных величин, связанных в простую однородную цепь Маркова и принимающих значения из конечного множества. Дальнейшее развитие относительно оценки точности пуассоновской аппроксимации распределения общего числа серий, образованных k-кратными ($k \ge 2$) повторениями s-цепочек в последовательности (1), которая является простой эргодической цепью Маркова с конечным числом состояний, приведено в работе [7].

Статистическим критериям, построенным по частотам отсутствующих s-цепочек (при $s = \text{const}$ и $\min(n, N) \to \infty$) из некоторого множества, посвящены работы [8, 9], в которых (1) — это последовательность, состоящая из $n > 0$ независимых одинаково распределенных случайных величин.

Для конечных n, N и s в [10, 11] получены неравенства для функции распределения и среднего числа s-цепочек, которые не появились из заданного класса, причем в [11] предполагается, что элементы последовательности (1) — это независимые не обязательно одинаково распределенные случайные величины.

В настоящей статье получено дальнейшее развитие предложенного в [12] подхода к установлению явных совместных распределений числа 2-цепочек и числа 3-цепочек заданного вида.

СОВМЕСТНОЕ РАСПРЕДЕЛЕНИЯЕ ЧИСЛА 2-ЦЕПОЧЕК И ЧИСЛА 3-ЦЕПОЧЕК

Обозначим $\eta(t_1, t_2, \ldots, t_s)$ число всех s-цепочек в последовательности (1), рассматриваемой над дискретным множеством мощности $N = 2$, которые совпадают с фиксированным набором элементов $t_1, t_2, \ldots, t_s$, где $t_i \in \{0, 1\}$, $i = 1, 2, \ldots, s$.

Теорема 1. Пусть выполняются следующие условия:

• последовательность (1) состоит из $n > 0$ независимых случайных величин, имеющих распределение

$$\mathrm{P}\{\gamma_k = 1\} = p,\ \mathrm{P}\{\gamma_k = 0\} = q,\ p + q = 1,\ k = 1, 2, \ldots, n; \tag{2}$$

• k_1, k_2, k_3, t, t_1 — целые числа такие, что $k_1 \ge 0$, $k_2 \ge 0$, $k_3 \ge 0$, $t, t_1 \in \{0, 1\}$.

Тогда имеют место следующие равенства:

$$\mathrm{P}\{\eta(tt) = k_1, \eta(t_1 t t_1^*) = k_2, \eta(t_1 t^* t_1^*) = k_3\} =$$

$$= \sum_{m_1=0}^{n} p^{m_1} q^{m_0} \{C_{m_t - k_1}^{k_2} C_{m_t - k_1}^{k_3} C_{m_{t^*} - m_t + k_1}^{k_3} Z(k_1; k_2) +$$

$$+ C_{m_t - k_1 - 1}^{k_2} C_{m_t - k_1 - 1}^{k_3} C_{m_{t^*} - m_t + k_1 + 1}^{k_3} C_{k_1}^{k_2}\}, \tag{3}$$

где $m_0 = n - m_1$, $t^* = 1 - t$, $Z(a; b) \overset{\text{def}}{=} \begin{cases} C_{a-1}^{b-1}, & \text{если } a \ge b \ge 1, \\ 1, & \text{если } a = b = 0, \\ 0 & \text{в остальных случаях;} \end{cases}$

$$\mathrm{P}\{\eta(tt) = k_1, \eta(t_1 t t_1^*) + \eta(t_1 t^* t_1^*) = k_2\} =$$

$$= \sum_{m_1=0}^{n} p^{m_1} q^{m_0} \sum_{i \in \{k_1, k_1+1\}} \sum C_{m_t - i}^{\delta_t} C_{m_t - i}^{\delta_{t^*}} C_{m_{t^*} - m_t + i}^{m_t - \delta_{t^*} - i} Z(i; m_t - k_1 - \delta_t), \tag{4}$$

где символ $\sum$ обозначает суммирование по всем целым неотрицательным числам δ_t и δ_{t^*} таким, что $\delta_t + \delta_{t^*} = 2(m_t - i) - k_2$;

$$\mathrm{P}\{\eta(tt) = k_1, \eta(t_1 t t_1^*) = k_2\} =$$
$$\sum_{m_1=0}^{n} p^{m_1} q^{m_0} \{C_{m_t-k_1}^{k_2} C_{m_{t^*}}^{m_t-k_1} Z(k_1; k_2) + C_{m_t-k_1-1}^{k_2} C_{m_{t^*}}^{m_t-k_1-1} C_{k_1}^{k_2}\}; \tag{5}$$
$$\mathrm{P}\{\eta(tt) = k_1, \eta(t_1 t^* t_1^*) = k_2\} =$$
$$= \sum_{m_1=0}^{n} p^{m_1} q^{m_0} Z(m_t; m_t - k_1)\{C_{m_t-k_1}^{k_2} C_{m_{t^*}-m_t+k_1}^{k_2} + C_{m_t-k_1-1}^{k_2} C_{m_{t^*}-m_t+k_1+1}^{k_2}\}. \tag{6}$$

Доказательство. Проверим соотношение (4). В силу условия (2) количество единиц (обозначим его ν) в последовательности (1) имеет биноминальное распределение с параметрами (n, p):

$$\mathrm{P}\{\nu = m\} = C_n^m p^m q^{n-m}, \; m = 0, 1, \ldots, n. \tag{7}$$

Используя формулу полной вероятности, получаем

$$\mathrm{P}\{E_1, E_2\} = \sum_{m_1=0}^{n} \mathrm{P}\{\nu = m_1\} \mathrm{P}\{E_1, E_2 / \nu = m_1\}, \tag{8}$$

где $E_1 = \{\eta(tt) = k_1\}$, $E_2 = \{\eta(t_1 t t_1^*) + \eta(t_1 t^* t_1^*) = k_2\}$.

Покажем, что

$$\mathrm{P}\{E_1, E_2 / \nu = m_1\} =$$
$$= (C_n^{m_1})^{-1} \sum_{i \in \{k_1, k_1+1\}} \sum C_{m_t-i}^{\delta_t} C_{m_t-i}^{\delta_{t^*}} C_{m_{t^*}-m_t+i}^{m_t-\delta_{t^*}-i} Z(i; m_t - k_1 - \delta_t). \tag{9}$$

Для этого введем следующие обозначения:

— $\Omega(n, m_1)$ — множество всех попарно различных n-мерных (0, 1)-векторов, каждый из которых содержит m_1 единиц и m_0 нулей, $m_0 + m_1 = n$;

— $D(m_0, m_1, k_1, k_2; t)$ — подмножество множества $\Omega(n, m_1)$, включающие все попарно различные векторы, каждый из которых начинается с элемента t, заканчивается элементом t^*, содержит k_1 (k_2) 2-цепочек вида tt^* (3-цепочек вида $t_1 \alpha t_1^*$, $\alpha \in \{0, 1\}$);

— Q — общее число векторов $\overline{v}$, $\overline{v} \in \Omega(n, m_1)$, каждый из которых имеет k_1 (k_2) 2-цепочек вида tt^* (3-цепочек вида $t_1 \alpha t_1^*$, $\alpha \in \{0, 1\}$).

Используя принятые обозначения, получаем

$$\mathrm{P}\{E_1, E_2 / \nu = m_1\} = Q(|\Omega(n, m_1)|)^{-1}. \tag{10}$$

Определим формулу для нахождения числа Q. Положим $\beta = \{t_1, \text{если } t_1 = t; \; t_1^*, \text{если } t_1 \neq t\}$. Следует отметить, что если некоторый вектор $\overline{v}$, $\overline{v} \in \Omega(n, m_1)$, начинается с t_1, заканчивается t_1^*, имеет k_1 2-цепочек вида tt, то в нем количество 2-цепочек вида $t_1 t_1^*$ равно $m_t - k_1$. Находим

$$Q = \sum_{\nu=0}^{m_{\beta^*}} |D(m_\beta, m_{\beta^*} - \nu, m_t - k_1, k_2; t_1)| +$$
$$+ \sum_{\nu=0}^{m_{\beta^*}} \sum_{\nu'=1}^{m_\beta} |D(m_\beta - \nu', m_{\beta^*} - \nu, m_t - k_1 - 1, k_2; t_1)|. \tag{11}$$

Для произвольного вектора $\overline{v}$, $\overline{v} \in D(m_{t_1}, m_{t_1^*}, m_t - k_1, k_2; t_1)$, выполняется соотношение

$$k_2 = 2(m_t - k_1) - \delta_1^{(t)} - \delta_1^{(t^*)}, \tag{12}$$

где $\delta_1^{(\alpha)}$ — число α-серий ($\alpha \in \{0,1\}$), имеющих длину единица и содержащихся в векторе $\overline{v}$.

Действительно, число всех 3-цепочек вида $t_1 t t_1^*$ ($t_1 t^* t_1^*$) в векторе $\overline{v}$ равно $m_t - k_1 - \delta_1^{(t)}$ ($m_t - k_1 - \delta_1^{(t^*)}$). Отсюда следует (12).

Используя (12), представим множество

$$D(m_{t_1}, m_{t_1^*}, m_t - k_1, k_2; t_1) = \bigcup K(m_{t_1}, m_{t_1^*}, m_t - k_1, k_2, \delta_1^{(t_1)}, \delta_1^{(t_1^*)}; t_1), \tag{13}$$

где объединение осуществляется по всем целым неотрицательным числам $\delta_1^{(t_1)}$ и $\delta_1^{(t_1^*)}$ таким, что $\delta_1^{(t_1)} + \delta_1^{(t_1^*)} = 2(m_t - k_1) - k_2$, множество $K(m_{t_1}, m_{t_1^*}, m_t - k_1, k_2, \delta_1^{(t_1)}, \delta_1^{(t_1^*)}; t_1)$ содержит все те и только те элементы множества $D(m_{t_1}, m_{t_1^*}, m_t - k_1, k_2; t_1)$, каждый из которых имеет $\delta_1^{(t_1)}$ ($\delta_1^{(t_1^*)}$) t_1-цепочек (t_1^*-цепочек) длиной единица.

Проверим равенство

$$| K(m_{t_1}, m_{t_1^*}, m_t - k_1, k_2, \delta_1^{(t_1)}, \delta_1^{(t_1^*)}; t_1) | =$$

$$= C_{m_t - k_1}^{\delta_1^{(t)}} C_{m_t - k_1}^{\delta_1^{(t^*)}} C_{k_1 - 1}^{m_t - k_1 - \delta_1^{(t)} - 1} C_{m_{t^*} - m_t + k_1 - 1}^{m_t - k_1 - \delta_1^{(t^*)} - 1}. \tag{14}$$

При этом отметим, что элемент $\overline{v}$, $\overline{v} \in K(m_{t_1}, m_{t_1^*}, m_t - k_1, k_2, \delta_1^{(t_1)}, \delta_1^{(t_1^*)}; t_1)$, определяется однозначно, если зафиксировать:

— одно из $C_{m_t - k_1}^{\delta_1^{(t)}} C_{m_t - k_1}^{\delta_1^{(t^*)}}$ всех возможных расположений t_1-серий и t_1^*-серий, каждая из которых имеет длину единица;

— одно из всех возможных разбиений числа $m_{t_1} - \delta_1^{(t_1)}$ ($m_{t_1^*} - \delta_1^{(t_1^*)}$) на $m_t - k_1 - \delta_1^{(t_1)}$ ($m_t - k_1 - \delta_1^{(t_1^*)}$) слагаемых, значение каждого из которых не меньше двух.

Произведение количества указанных разбиений чисел $m_{t_1} - \delta_1^{(t_1)}$ и $m_{t_1^*} - \delta_1^{(t*)}$ равно $C_{k_1 - 1}^{m_t - k_1 - \delta_1^{(t)} - 1} C_{m_{t^*} - m_t + k_1 - 1}^{m_t - k_1 - \delta_1^{(t^*)} - 1}$. Отсюда следует (14).

Используя (13) и (14), находим

$$| D(m_{t_1}, m_{t_1^*}, m_t - k_1, k_2; t_1) | = \sum C_{m_t - k_1}^{\delta_1^{(t)}} C_{m_t - k_1}^{\delta_1^{(t^*)}} C_{k_1 - 1}^{m_t - k_1 - \delta_1^{(t)} - 1} C_{m_{t^*} - m_t + k_1 - 1}^{m_t - k_1 - \delta_1^{(t^*)} - 1}, \tag{15}$$

где символ $\sum$ обозначает суммирование по всем целым неотрицательным числам $\delta_1^{(t_1)}$ и $\delta_1^{(t_1^*)}$ таким, что $\delta_1^{(t_1)} + \delta_1^{(t_1^*)} = 2(m_t - k_1) - k_2$.

С помощью (10), (11), (15) и равенства $|\Omega(n, m_1)| = C_n^{m_1}$ получаем формулу (9), подстановка которой в (8) завершает (с учетом (7)) проверку соотношения (4).

Доказательство соотношений (3), (5) и (6) проводится аналогично обоснованию равенства (4).

ПРИМЕРЫ К ТЕОРЕМЕ 1

В примерах рассмотрен случай для $p = q = 1/2$ и малой выборки длины n, $n = 16$. (Согласно [13, с. 28] выборка называется малой, если $13 \le n \le 40$.)

Пример использования равенства (3). Табл. 1 иллюстрирует использование соотношения (3) для $n = 16$. В первом, втором и третьем столбцах таблицы помещены все возможные варианты значений k_1, k_2 и k_3, для которых имеем вероятность $\mathrm{P}\,\{\eta(tt) = k_1, \eta(t_1 t t_1^*) = k_2, \eta(t_1 t^* t_1^*) = k_3\} \ge 0.01$. В четвертом столбце табл. 1 даны вероятности (в неубывающем порядке) $\mathrm{P}\,\{\eta(tt) = k_1, \eta(t_1 t t_1^*) = k_2, \eta(t_1 t^* t_1^*) = k_3\}$ для чисел k_1, k_2, k_3. В пятом столбце помещена сумма накопленных вероятностей P_c до реализации события $\{\eta(tt) = k_1, \eta(t_1 t t_1^*) = k_2, \eta(t_1 t^* t_1^*) = k_3\}$ включительно.

Таблица 1. Значения вероятности осуществления события для чисел k_1, k_2 и k_3 с использованием равенства (3)

k_1	k_2	k_3	P	P_c	k_1	k_2	k_3	P	P_c
3	1	3	0.010483	0.304275	4	1	2	0.018326	0.547851
5	1	2	0.011353	0.314758	4	3	2	0.019455	0.566177
4	2	3	0.012085	0.326111	2	1	3	0.020111	0.585632
4	1	1	0.012665	0.338196	3	2	3	0.020721	0.605743
3	3	2	0.013611	0.350861	6	2	1	0.021286	0.626464
7	2	1	0.013916	0.364472	2	2	3	0.022781	0.64775
1	1	1	0.014099	0.378388	3	2	1	0.025635	0.670531
3	1	1	0.014206	0.392487	5	2	1	0.025696	0.696166
6	3	1	0.014343	0.406693	3	1	2	0.027557	0.721862
0	0	3	0.014771	0.421036	4	2	1	0.028748	0.749419
2	1	1	0.015167	0.435807	5	2	2	0.029388	0.778167
4	3	1	0.015228	0.450974	1	1	3	0.032349	0.807555
5	3	2	0.015747	0.466202	2	1	2	0.03653	0.839904
2	2	1	0.016129	0.481949	2	2	2	0.038864	0.876434
0	0	2	0.016235	0.498078	4	2	2	0.041977	0.915298
6	2	2	0.016479	0.514313	1	1	2	0.042725	0.957275
5	3	1	0.017059	0.530792	3	2	2	0.048935	1

Пример использования равенства (4). Табл. 2 иллюстрирует использование соотношения (4) для $n = 16$. В первом и втором столбцах таблицы даны все возможные варианты значений k_1 и k_2, для которых вероятность $\mathrm{P}\,\{\eta(tt) = k_1, \eta(t_1 t t_1^*) + \eta(t_1 t^* t_1^*) = k_2\} \ge 0.01$. В третьем столбце даны вероятности (в неубывающем порядке) $\mathrm{P}\,\{\eta(tt) = k_1, \eta(t_1 t t_1^*) + \eta(t_1 t^* t_1^*) = k_2\}$ для чисел k_1 и k_2. В четвертом столбце помещена сумма накопленных вероятностей P_c до реализации события $\{\eta(tt) = k_1, \eta(t_1 t t_1^*) + \eta(t_1 t^* t_1^*) = k_2\}$ включительно.

На рис. 1 представлена пузырьковая диаграмма [14] реализации соотношения (4) для $n = 16$. Здесь первый параметр (горизонтальная ось) представляет воз-

Т а б л и ц а 2. Значения вероятности осуществления события для чисел k_1 и k_2 с использованием равенства (4)

k_1	k_2	P	P_c	k_1	k_2	P	P_c
6	2	0.012039	0.176253	6	3	0.031067	0.420851
8	3	0.012695	0.188292	6	4	0.032425	0.451918
6	5	0.01297	0.200987	1	4	0.032654	0.484343
0	3	0.014771	0.213957	4	5	0.033173	0.516997
5	2	0.014862	0.228728	3	5	0.034821	0.55017
0	2	0.016235	0.24359	5	3	0.040771	0.584991
4	2	0.017151	0.259825	1	3	0.04541	0.625762
7	4	0.018005	0.276976	5	4	0.048889	0.671172
1	2	0.018585	0.294981	4	3	0.049805	0.720061
3	2	0.018692	0.313566	2	3	0.053955	0.769866
2	2	0.01944	0.332258	3	3	0.05481	0.823821
7	3	0.020691	0.351698	2	4	0.059052	0.878631
5	5	0.023773	0.372389	4	4	0.062317	0.937683
2	5	0.024689	0.396162	3	4	0.067871	1

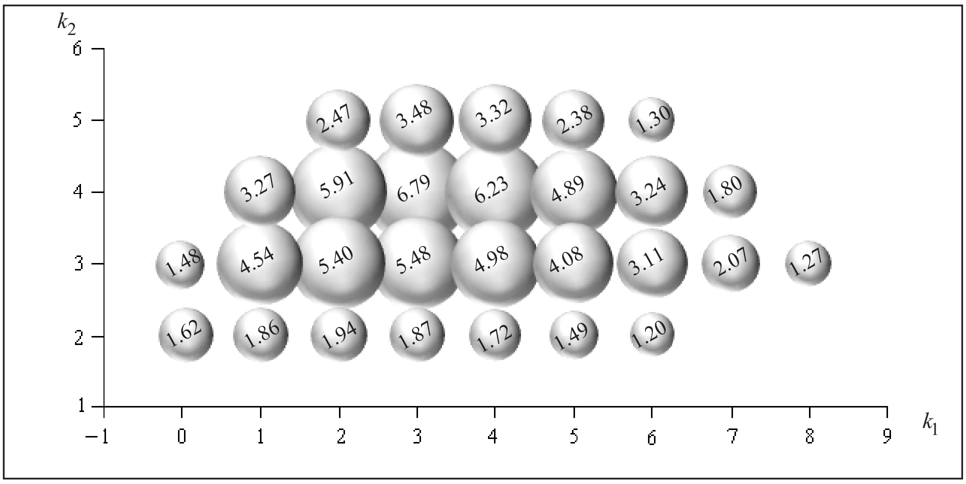

Рис. 1. Диаграмма реализации соотношения (4) для $n = 16$

можные значения k_1, второй параметр (вертикальная ось) представляет возможные значения k_2, третий параметр (размер пузырька) представляет вероятность осуществления события $\{\eta(tt) = k_1, \eta(t_1 t t^*) + \eta(t_1 t^* t_1^*) = k_2\}$, выраженную в процентах. Например, при $k_1 = 4$ и $k_2 = 4$ вероятность осуществления события $\{\eta(tt) = k_1, \eta(t_1 t t_1^*) + \eta(t_1 t^* t_1^*) = k_2\}$ составляет 6.23 %.

Пример использования равенства (5). Табл. 3 иллюстрирует использование соотношения (5) для $n = 16$. Здесь интерпретация столбцов аналогична интерпретации столбцов в табл. 2.

На рис. 2 приведена пузырьковая диаграмма реализации соотношения (5) для $n = 16$. Здесь первый параметр (горизонтальная ось) представляет возможные значения k_1, второй параметр (вертикальная ось) представляет возможные значения k_2, третий параметр (размер пузырька) представляет вероятность осуществления события $\{\eta(tt) = k_1, \eta(t_1 t t_1^*) = k_2\}$, выраженную в процентах.

Т а б л и ц а 3 . Значения вероятности осуществления события для чисел k_1 и k_2 с использованием равенства (5)

k_1	k_2	P	P_c	k_1	k_2	P	P_c
11	2	0.001297	0.104447	9	3	0.002991	0.11319
5	0	0.001358	0.105744	10	2	0.003296	0.116181
10	1	0.001831	0.107102	9	1	0.003326	0.119477
7	4	0.00206	0.108933	3	0	0.003555	0.122803
4	0	0.002197	0.110993	4	4	0.003891	0.126358
6	4	0.004196	0.130249	3	3	0.029266	0.31186
5	4	0.005493	0.134445	4	1	0.036987	0.341126
8	1	0.005722	0.139938	0	0	0.039429	0.378113
2	0	0.005753	0.14566	5	3	0.039459	0.417542
9	2	0.007629	0.151413	4	3	0.042816	0.457001
8	3	0.007904	0.159042	6	2	0.043182	0.499817
1	0	0.009308	0.166946	3	1	0.054016	0.542999
7	1	0.009567	0.176254	5	2	0.064484	0.597015
8	2	0.014847	0.185821	2	1	0.074783	0.661499
6	1	0.015488	0.200668	2	2	0.081757	0.736282
7	3	0.016632	0.216156	4	2	0.086517	0.818039
5	1	0.024338	0.232788	1	1	0.095444	0.904556
7	2	0.026505	0.257126	3	2	0.099152	1
6	3	0.028229	0.283631				

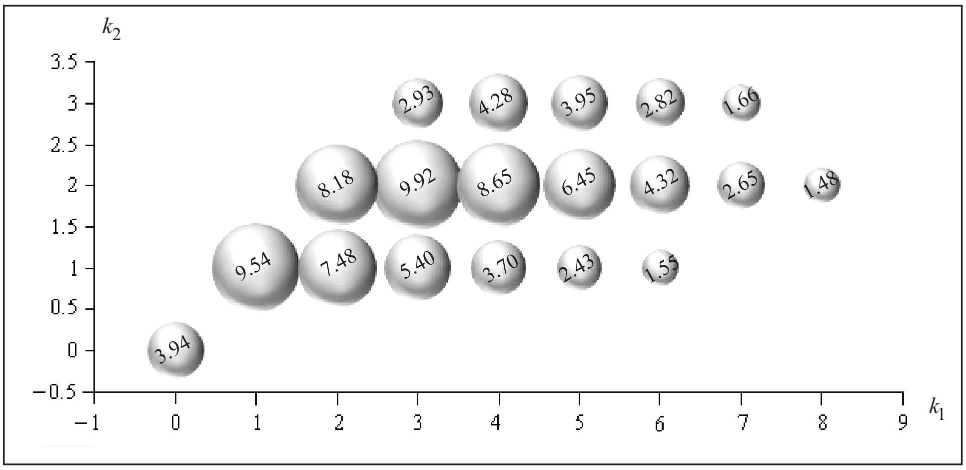

Рис. 2. Диаграмма реализации соотношения (5) для $n = 16$

Пример использования равенства (6). Табл. 4 иллюстрирует использование соотношения (6) для $n = 16$. Интерпретация столбцов аналогична интерпретации столбцов табл. 2.

На рис. 3 приведена пузырьковая диаграмма реализации соотношения (6) для $n = 16$. Здесь первый параметр (горизонтальная ось) представляет возможные значения k_1, второй параметр (вертикальная ось) представляет возможные значения k_2, третий параметр (размер пузырька) представляет вероятность осуществления события $\{\eta(tt) = k_1, \eta(t_1 t^* t_1^*) = k_2\}$, выраженную в процентах.

Таблица 4. Значения вероятности осуществления события для чисел k_1 и k_2 с использованием равенства (6)

k_1	k_2	P	P_c	k_1	k_2	P	P_c
5	3	0.01004	0.168058	1	3	0.035034	0.405852
6	0	0.010071	0.188169	3	3	0.037537	0.440886
7	0	0.010071	0.178098	2	3	0.044189	0.478423
5	0	0.010254	0.19824	6	1	0.046158	0.522612
0	3	0.014771	0.208494	1	2	0.047211	0.56877
7	2	0.014954	0.223265	3	1	0.049347	0.615981
1	1	0.015808	0.238219	5	1	0.056305	0.665328
0	2	0.016235	0.254027	5	2	0.05864	0.721633
8	1	0.01767	0.270262	4	1	0.059052	0.780273
4	3	0.022736	0.287932	2	2	0.078278	0.839325
7	1	0.029709	0.310668	4	2	0.082397	0.917603
2	1	0.032684	0.340377	3	2	0.091919	1
6	2	0.032791	0.373061				

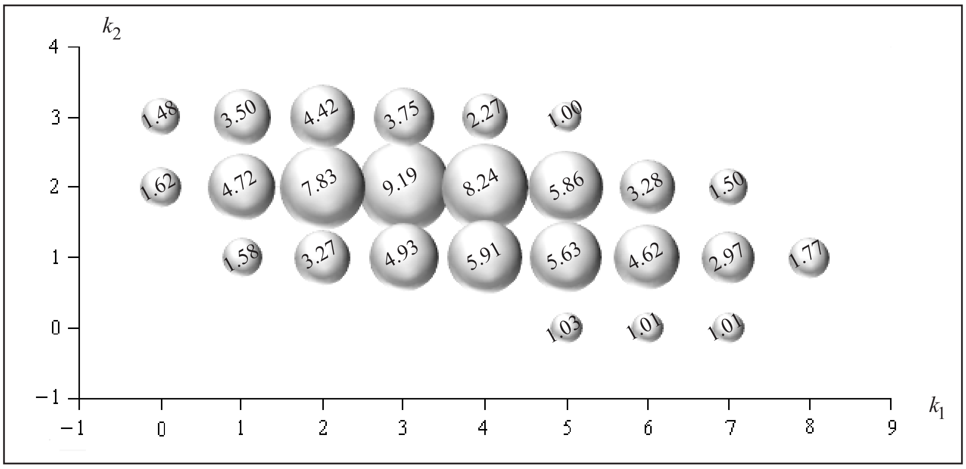

Рис. 3. Диаграмма реализации соотношения (6) для $n = 16$

ЗАКЛЮЧЕНИЕ

В настоящей статье установлены совместные распределения числа 2-цепочек и числа 3-цепочек фиксированного вида случайной битовой последовательности конечной длины. Даны примеры использования этих распределений. Возможным применением полученных формул может быть проверка гипотезы случайности расположения нулей и единиц в локальных участках (0, 1)-последовательности.

СПИСОК ЛИТЕРАТУРЫ

1. Rukhin A., Soto J., Nechvatal J., Smid M., Barker E., Leigh S., Levenson M., Vangel M., Banks D., Heckert A., Dray J., Vo S. A statistical test suite for random and pseudorandom number generators for cryptographic applications. National Institute of Standards and Technology. Special Publication 800-22 revision 1a, 2010. 131 p.
2. Ивченко Г.И., Медведев Ю.И. Введение в математическую статистику. Москва: Изд-во ЛКИ, 2010. 600 с.
3. Верещагин Н.К., Успенский В.А., Шень А. Колмогоровская сложность и алгоритмическая случайность. Москва: МЦНМО, 2013. 576 с.

4. Котов В.Н. Применение теории измерений в биологических исследованиях. Киев: Наук. думка, 1985. 100 с.
5. Харин Ю.С., Берник В.И., Матвеев Г.В., Агиевич С.В. Математические и компьютерные основы криптологии. Минск: Новое знание, 2003. 382 с.
6. Беляев П.Ф. О совместном распределении частот длинных s-цепочек в простых однородных цепях Маркова с конечным множеством исходов. *Труды по дискретной математике*. 1997. Т. 1. С. 19–42.
7. Mikhaylov V.G. Estimates of accuracy of the Poisson approximation for the distribution of number of runs of long string repetitions in a Markov chain. *Discrete Math. Appl.* 2016. Vol. 26, N 2. P. 105–113.
8. Tikhomirova M.I., Chistyakov V.P. On the asymptotic behaviour of moments of the number of absent s-tuples. *Discrete Math. Appl*. 1997. Vol. 7, N 1. P. 13–31.
9. Чистяков В.П., Тихомирова М.И. Статистические критерии, построенные по частотам s-грамм из некоторого множества. *Труды по дискретной математике*. 2000. Т. 3. С. 295–302.
10. Михайлов В.Г. Некоторые неравенства для функции распределения числа непоявившихся s-цепочек. *Труды по дискретной математике*. 1997. Т. 1. С. 221–226.
11. Михайлов В.Г. Неравенства для среднего числа повторений m-цепочек и для среднего числа непоявившихся m-цепочек из заданного класса. *Труды по дискретной математике*. 2000. Т. 3. С. 147–154.
12. Масол В.И. О распределении некоторых статистик (0, 1)-вектора. *Исследование операций и АСУ*. 1987. Вып. 29. С. 23–27.
13. Гайдышев И.П. Программное обеспечение анализа данных AtteStat. Руководство пользователя. Версия 13, 2012. 505 с.
14. Hu Y., Polk T., Yang J., Zhao Y., Liu Sh. Spot-tracking lens: A zoomable user interface for animated bubble charts. *2016 IEEE Pacific Visualization Symposium (PacificVis)*, Taipei, Taiwan. 2016. Vol. 1. P. 16–23.

Надійшла до редакції 26.06.2019

В.І. Масол, С.В. Поперешняк

ПЕРЕВІРКА ВИПАДКОВОСТІ РОЗТАШУВАННЯ БІТІВ У ЛОКАЛЬНИХ ДІЛЯНКАХ (0, 1)-ПОСЛІДОВНОСТІ

Анотація. Встановлено явний вигляд сумісного розподілу кількості 2-ланцюжків і кількості 3-ланцюжків різних фіксованих варіантів в (0, 1)-послідовності довжини n, що складається з нулів і одиниць. Вважається, що елементи (0, 1)-послідовності — це незалежні однаково розподілені випадкові величини. Наведено таблиці, що ілюструють застосування встановлених формул для (0, 1)-послідовності довжини $n = 16$.

Ключові слова: s-ланцюжки, бітова послідовність, випадковість, локальні ділянки, сумісний розподіл.

V. Masol, S. Popereshnyak

CHECKING THE RANDOMNESS OF BITS DISPOSITION IN LOCAL SEGMENTS OF THE (0, 1)-SEQUENCE

Abstract. An explicit form of the joint distribution of the number of 2-chains and the number of 3-chains of various fixed variants in a (0, 1)-sequence of length n consisting of zeros and ones is established. It is assumed that the elements of (0, 1)-sequences are independent identically distributed random variables. Tables illustrating the application of the established formulas for a (0, 1)-sequence of length $n = 16$ are given.

Keywords: s-chains, bit sequence, randomness, local segments, joint distribution.

Масол Владимир Иванович,
доктор физ.-мат. наук, профессор Киевского национального университета имени Тараса Шевченко, e-mail: vimasol@ukr.net.

Поперешняк Светлана Владимировна,
кандидат физ.-мат. наук, доцент Киевского национального университета имени Тараса Шевченко, e-mail: spopereshnyak@gmail.com.