Math-Net.Ru

Общероссийский математический портал

Г. И. Ивченко, Ю. И. Медведев, Параметрические модели случайных r -подстановок и r -разбиений и их вероятностно-статистический анализ, *Матем. вопр. криптогр.*, 2018, том 9, выпуск 1, 47–64

DOI: <https://doi.org/10.4213/mvk247>

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

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

Параметры загрузки:

IP: 3.85.183.62

26 апреля 2023 г., 12:45:17

**Параметрические модели
случайных r -подстановок и r -разбиений
и их вероятностно-статистический анализ**

Г. И. Ивченко, Ю. И. Медведев

Академия криптографии Российской Федерации, Москва

Получено 11.V.2017

Аннотация. Работа посвящена исследованию проблем вероятностной комбинаторики, в которых на множестве рассматриваемых комбинаторных объектов задается некоторая параметрическая вероятностная мера. Проводится сравнительный анализ ряда основных свойств случайных r -подстановок и случайных r -разбиений конечных множеств в однопараметрической модели.

Ключевые слова: случайные комбинаторные объекты, параметрическая модель, r -подстановки, r -разбиения множеств, предельные теоремы, статистические выводы

Parametric models of r -permutations and r -partitions and their probabilistic-statistical analysis

G. I. Ivchenko, Yu. I. Medvedev

Academy of Cryptography of the Russian Federation, Moscow

Abstract. Some problems of the probabilistic combinatorics are investigated when some parametric probabilistic measure on the set of combinatorial objects under consideration is given. Analysis of r -permutations and r -partitions of finite sets in an one-parametric model is performed.

Key words: random combinatorial objects, parametric model, r -permutations, r -partitions, limit theorems, statistical inferences

Citation: *Mathematical Aspects of Cryptography*, 2018, v. 9, № 1, pp. 47–64 (Russian)

© Академия криптографии Российской Федерации, 2018 г.

Введение. Постановка задачи и предварительные соотношения

В множестве классических комбинаторных объектов по теоретико-вероятностным свойствам выделяются три типа: экспоненциальные, биномиальные и отрицательно-биномиальные (см., например, [7, 9, 5, 22]). Важнейшими для криптографических приложений моделями, несомненно, являются подстановки и разбиения конечных множеств, относящиеся к экспоненциальному типу, исследованию которых посвящено большое число работ. При этом традиционным до недавнего времени методом их исследования было введение на множестве рассматриваемых объектов равновероятной меры и последующий анализ различных характеристик этих объектов как случайных величин. Такой вероятностный подход оказался весьма эффективным при решении широкого круга комбинаторных перечислительных задач, когда интересуются числом объектов некоторого множества, обладающих теми или иными заданными свойствами.

Вместе с тем в криптографических приложениях возникают задачи и других типов, для решения которых равномерная модель уже не является адекватной, а именно, когда требуется учитывать различного рода возможные отклонения от равновероятности и исследовать поведение изучаемых характеристик при альтернативных (по отношению к равномерной) мерах.

Общая параметрическая модель для произвольного множества разложимых комбинаторных объектов была предложена в [7, 9, 13], а ее конкретизация для случайных подстановок и разбиений конечных множеств изложена в [11, 12, 14, 16]. Достоинством этой модели является наличие в ней большого числа «свободных» параметров (степеней свободы), варьируя которые, можно конструировать различные альтернативные распределения на множестве изучаемых объектов и тем самым получать возможность проводить их более глубокий не только вероятностный, но и статистический (проверка различных гипотез) анализ. Решению различных задач такого рода для неравновероятных подстановок и разбиений конечных множеств были посвящены работы [6, 7, 10, 11, 16]. В то же время и классические конструкции комбинаторных объектов зачастую оказываются неадекватными моделями для приложений: приходится вводить различного рода усложнения или ограничения на множествах рассматриваемых объектов. Так возникают новые конструкции комбинаторных объектов и новые задачи их анализа.

Настоящая работа посвящена анализу двух таких конструкций, получивших в литературе названия *r-подстановок* и *r-разбиений*, когда r первых элементов множества $X_n = \{1, 2, \dots, r, \dots, n\}$, $n \geq r \geq 0$, должны принадлежать различным циклам перестановки (подстановки) элементов этого множества или, соответственно, различным блокам его разбиения.

Когда $r = 0$ или 1 , такие ограничения фактически отсутствуют, т.е. мы имеем дело с обычными (классическими) подстановками и разбиениями. В этом случае при решении соответствующих задач, как известно, ключевую роль играют числа Стирлинга первого и второго родов, соответственно $[n]_k$ и $\{n\}_k$, определяемые разложениями

$$[x]_n = \sum_{k=0}^n [n]_k x^k \quad \text{и} \quad x^n = \sum_{k=0}^n \{n\}_k (x)_k, \quad (0.1)$$

где $[x]_n = x(x+1) \cdots (x+n-1)$, $n \geq 1$, $[x]_0 = 1$, — возрастающая факториальная функция, а $(x)_n = x(x-1) \cdots (x-n+1)$, $n \geq 1$, $(x)_0 = 1$, — убывающая факториальная функция (разные авторы используют и другие обозначения для этих чисел).

Эквивалентным образом эти числа можно определить также через их экспоненциальные производящие функции (ЭПФ):

$$\sum_{n \geq k} [n]_k \frac{x^n}{n!} = \frac{1}{k!} (-\ln(1-x))^k, \quad \sum_{n \geq k} \{n\}_k \frac{x^n}{n!} = \frac{1}{k!} (e^x - 1)^k. \quad (0.2)$$

При этом, напомним, $[n]_k$ — это число n -подстановок с k циклами, а $\{n\}_k$ — число разбиений n -множества на k непустых блоков. Свойства этих чисел и их применения подробно изложены во многих работах, в частности в отечественной литературе [17–22].

Введем теперь соответствующие числа для r -подстановок и r -разбиений. Бродер [24] впервые определил r -число Стирлинга первого рода, обозначая его символом $[n]_r$, как число таких перестановок элементов множества $X_n = \{1, 2, \dots, r, \dots, n\}$, $n \geq r \geq 1$, имеющих $k \geq r$ циклов, что элементы $1, 2, \dots, r$ принадлежат различным циклам, а r -число Стирлинга второго рода $\{n\}_r$ — как число таких разбиений множества X_n на $k \geq r$ блоков, что элементы $1, 2, \dots, r$ попадают в различные блоки. Эти числа определяются разложениями (ср. с (0.1))

$$P_{n,r}(x) = [x+r]_n = \sum_{k=0}^n [n+r]_k x^k, \quad x^n = \sum_{k=0}^n \{n+r\}_k (x-r)_k, \quad (0.3)$$

а их ЭПФ имеют вид

$$\sum_n [n+r]_k \frac{x^n}{n!} = \frac{1}{k!} \left(\frac{1}{1-x} \right)^r \left(\ln \left(\frac{1}{1-x} \right) \right)^k, \quad (0.4)$$

$$\sum_n \{n+r\}_k \frac{x^n}{n!} = \frac{e^{rx}}{k!} (e^x - 1)^k. \quad (0.5)$$

Также имеют место следующие рекуррентные соотношения:

$$\begin{aligned} \left[\begin{matrix} n \\ k \end{matrix} \right]_r &= (n-1) \left[\begin{matrix} n-1 \\ k \end{matrix} \right]_r + \left[\begin{matrix} n-1 \\ k-1 \end{matrix} \right]_r, & n > r, \\ \left[\begin{matrix} n \\ k \end{matrix} \right]_r &= 0, & n < r, \\ \left[\begin{matrix} n \\ k \end{matrix} \right]_r &= \delta_{kr}, & n = r, \\ \left\{ \begin{matrix} n \\ k \end{matrix} \right\}_r &= k \left\{ \begin{matrix} n-1 \\ k \end{matrix} \right\}_r + \left\{ \begin{matrix} n-1 \\ k-1 \end{matrix} \right\}_r, & n > r, \\ \left\{ \begin{matrix} n \\ k \end{matrix} \right\}_r &= 0, & n < r, \\ \left\{ \begin{matrix} n \\ k \end{matrix} \right\}_r &= \delta_{kr}, & n = r, \end{aligned}$$

обобщающие аналогичные соотношения для обычных (при $r = 0$) чисел Стирлинга (здесь δ_{kr} — символ Кронекера).

Введем еще «горизонтальную» производящую функцию

$$B_{n,r}(\theta) = \sum_{k=0}^n \left\{ \begin{matrix} n+r \\ k+r \end{matrix} \right\}_r \theta^k, \quad (0.6)$$

которая называется r -полиномом Белла [23] (для случая $r = 0$ это есть, напомним, обычный полином Белла), а величины

$$B_{n,r} = B_{n,r}(1) = \sum_{k=0}^n \left\{ \begin{matrix} n+r \\ k+r \end{matrix} \right\}_r \quad (0.7)$$

называются r -числами Белла [26].

ЭПФ для r -полиномов Белла, как следует из (0.5), имеет вид

$$\sum_n B_{n,r}(\theta) \frac{z^n}{n!} = \exp \{ r z + \theta (e^z - 1) \},$$

так что

$$B_{n,r}(\theta) = n! [z^n] \exp \{ rz + \theta (e^z - 1) \} \quad (0.8)$$

(здесь и далее используется обозначение $[z^n]f(z) = \text{coef}_{z^n} f(z)$).

Полагая в этих соотношениях $\theta = 1$, получаем соответствующие представления и для r -чисел Белла:

$$B_{n,r} = n! [z^n] \exp \{ e^z - 1 + rz \}. \quad (0.9)$$

Мы ограничиваемся лишь этими сведениями об r -числах Стирлинга, отсылая заинтересованного читателя к работам [24, 25, 29], в которых приведен целый ряд других рекуррентных и комбинаторных соотношений для этих чисел.

З а м е ч а н и е. Еще в 1965 г. в работе авторов [3] были введены числа

$$S(n, k, t) = \frac{1}{k!} \Delta^k t^n, \quad (0.10)$$

где Δ — оператор конечной разности ($\Delta f(x) = f(x+1) - f(x)$), $t \geq 0$ — произвольное действительное число, которые были названы *обобщенными числами Стирлинга второго рода*. Эти числа при целых значениях $t = r$ совпадают с r -числами Стирлинга:

$$S(n, k, r) = \left\{ \begin{matrix} n+r \\ k+r \end{matrix} \right\}_r.$$

Далее мы рассматриваем случайные объекты описанного типа в предположении, что на множестве соответствующих объектов (подстановок или разбиений) задана некоторая вероятностная мера, зависящая от свободного параметра $\theta > 0$, и исследуем распределения общего числа циклов случайной r -подстановки и общего числа блоков случайного r -разбиения в такой параметрической модели. Эти результаты используются затем для построения соответствующих критериев проверки гипотезы $H_0: \theta = 1$ о равновероятности меры с учетом возможных альтернатив. Аналогичный анализ классических объектов (т.е. при $r = 0$) был проведен в предыдущей работе авторов [16].

1. Случайные r -подстановки

Уточним постановку задачи. В этом разделе рассматриваются подстановки на множестве $X_{n+r} = \{1, 2, \dots, r, \dots, n+r\}$, $n \geq 1$, $r \geq 0$, имеющие не менее r циклов и такие, что элементы $1, 2, \dots, r$ принадлежат разным циклам. Будем называть их r -подстановками. Число r -подстановок с $r+k$ циклами есть r -число Стирлинга первого рода $s(n, k, r)$ (мы используем это обозначение, аналогичное классическому, вместо более громоздкого обозначения $\left[\begin{smallmatrix} n+r \\ k+r \end{smallmatrix} \right]_r$).

Далее рассматриваются случайные подстановки описанного типа в случае, когда на множестве всех r -подстановок степени $n+r$ задается вероятностная мера, согласно которой любой подстановке с $r+k$ циклами приписывается вероятность, пропорциональная θ^k , где $\theta > 0$ — свободный параметр меры.

Пусть $r + \xi_{n,r}^{(u)}$ есть общее число циклов случайной r -подстановки степени $n+r$ в такой параметрической модели. Тогда распределение случайной величины $\xi_{n,r}^{(u)}$ имеет, очевидно, вид

$$\mathbf{P} \left\{ \xi_{n,r}^{(u)} = k \right\} = \frac{\theta^k s(n, k, r)}{P_{n,r}(\theta)}, \quad k = 0, 1, 2, \dots, n, \quad (1.1)$$

где (см. (0.3)) $P_{n,r}(\theta) = [\theta + r]_n = (\theta + r)(\theta + r + 1) \cdots (\theta + r + n - 1)$.

Формула (1.1) дает следующее представление для производящей функции случайной величины $\xi_{n,r}^{(u)}$:

$$\mathbf{E} t^{\xi_{n,r}^{(u)}} = \frac{P_{n,r}(t\theta)}{P_{n,r}(\theta)} = \prod_{j=0}^{n-1} (q_j(r, \theta) + p_j(r, \theta)t), \quad (1.2)$$

где

$$p_j(r, \theta) = 1 - q_j(r, \theta) = \frac{\theta}{\theta + r + j}.$$

Равенство (1.2) означает, что случайная величина $\xi_{n,r}^{(u)}$ может быть представлена в виде суммы $(\xi_{n,r}^{(u)} = X_0 + X_1 + \dots + X_{n-1})$ независимых бернуллиевских случайных величин с распределениями

$$\mathbf{P} \{X_j = 1\} = 1 - \mathbf{P} \{X_j = 0\} = p_j(r, \theta), \quad j = 0, 1, \dots, n-1,$$

откуда следует, что

$$\begin{aligned} \mathbf{E} \xi_{n,r}^{(u)} &= \sum_{j=0}^{n-1} p_j(r, \theta) = \theta \sum_{j=r}^{n+r-1} \frac{1}{\theta+j} = \theta \sum_{j=r}^{n+r-1} \frac{1}{j} - \theta^2 \sum_{j=r}^{n+r-1} \frac{1}{j(\theta+j)}, \\ \mathbf{D} \xi_{n,r}^{(u)} &= \sum_{j=0}^{n-1} p_j(r, \theta) q_j(r, \theta) = \mathbf{E} \xi_{n,r}^{(u)} - \theta^2 \sum_{j=r}^{n+r-1} \frac{1}{(\theta+j)^2}. \end{aligned} \quad (1.3)$$

Исследуем асимптотическое поведение величины $\xi_{n,r}^{(u)}$ при $n \rightarrow \infty$ и фиксированных значениях параметров $\theta > 0$ и $r \geq 1$. Из (1.3) очевидно следует, что в этих условиях имеют место асимптотические соотношения

$$\mathbf{E} \xi_{n,r}^{(u)} \sim \mathbf{D} \xi_{n,r}^{(u)} \sim \theta \ln n. \quad (1.4)$$

Представление (1.2) с использованием стандартных методов позволяет установить для величины $\xi_{n,r}^{(u)}$ как интегральный, так и локальный варианты центральной предельной теоремы с оценкой скорости сходимости в ней.

Теорема 1. При $n \rightarrow \infty$ и фиксированных значениях $\theta > 0$, $r \geq 1$ имеет место соотношение

$$\sup_x \left| \mathbf{P} \left\{ \frac{\xi_{n,r}^{(u)} - \theta \ln n}{\sqrt{\theta \ln n}} < x \right\} - \Phi(x) \right| = O\left(\frac{1}{\sqrt{\ln n}}\right), \quad (1.5)$$

где $\Phi(x)$ — стандартная нормальная функция распределения, а также

$$\max_k \left| \sqrt{\theta \ln n} \mathbf{P} \{ \xi_{n,r}^{(u)} = k \} - \frac{1}{\sqrt{2\pi}} e^{-x^2(k)/2} \right| = O\left(\frac{1}{\sqrt{\ln n}}\right), \quad (1.6)$$

где

$$x(k) = \frac{k - \theta \ln n}{\sqrt{\theta \ln n}}.$$

Доказательство. Условие $\mathbf{D} \xi_{n,r}^{(u)} \rightarrow \infty$, необходимое в центральной предельной теореме, выполнено ввиду (1.4).

Далее, соотношение (1.5) выполняется ввиду очевидного неравенства

$$\sum_{j=0}^{n-1} \mathbf{E} |X_j - \mathbf{E} X_j|^3 \leq \mathbf{D} \xi_{n,r}^{(u)}$$

и неравенства Эссеена.

Наконец, соотношение (1.6) является следствием того факта, что производящая функция (1.2) имеет только действительные корни [2, 27]. $\square$

Теорема 1, в частности, показывает, что параметр r не оказывает сколь-нибудь существенного влияния на асимптотическое поведение рассматриваемой характеристики $\xi_{n,r}^{(u)}$ случайной подстановки, пока он остается конечной величиной: оно оказалось таким же, как и для обычных подстановок, т. е. когда $r = 0$ [4]. Также от параметра r не будет зависеть и вид асимптотического критерия для проверки гипотезы $H_0: \theta = 1$ о равновероятности r -подстановок: он будет таким же, как и для обычных подстановок. Именно (см. [10]), критическая область для гипотезы H_0 при уровне значимости γ и, например, общей двусторонней альтернативе $H_1: \theta \neq 1$ имеет вид

$$X_\gamma(n) = \left\{ \frac{|S_{n,r}^{(u)} - \ln n|}{\sqrt{\ln n}} > u_{\gamma/2} \right\}, \quad \Phi(-u_{\gamma/2}) = \frac{\gamma}{2},$$

где $S_{n,r}^{(u)} + r$ равно наблюдаемому числу циклов r -подстановки.

Мощность $W_n(\theta_n)$ такого критерия при «близких» альтернативах вида

$$H_{1n}: \theta = \theta_n = 1 + \frac{t}{\sqrt{\ln n}}, \quad t \neq 0,$$

удовлетворяет предельному соотношению

$$W_n(\theta_n) \rightarrow \Phi(t - u_{\gamma/2}) + \Phi(-t - u_{\gamma/2}) \quad (n \rightarrow \infty).$$

Рассмотрим теперь схему серий, т. е. будем считать, что с ростом n может неограниченно возрастать и параметр r . Из (1.3) следует, что в этом случае

$$\mathbf{E} \xi_{n,r}^{(u)} = \theta \ln \frac{n+r-1}{r-1} + o(1) = \theta \ln \left(1 + \frac{n}{r}\right) + o(1). \quad (1.7)$$

Если $r = o(n)$, то $\mathbf{E} \xi_{n,r}^{(u)} \sim \mathbf{D} \xi_{n,r}^{(u)} \rightarrow \infty$, и соотношения (1.5) и (1.6) продолжают оставаться в силе. Если же $r = \alpha n$, $\alpha = \text{const} > 0$, то

$$\mathbf{E} \xi_{n,r}^{(u)} \sim \mathbf{D} \xi_{n,r}^{(u)} \sim \theta \ln \frac{\alpha+1}{\alpha} \equiv \lambda, \quad (1.8)$$

и в этой ситуации имеет место пуассоновская сходимость.

Теорема 2. Пусть при $n, r \rightarrow \infty$ выполнено (1.8). Тогда случайная величина $\xi_{n,r}^{(u)}$ распределена в пределе по закону Пуассона с параметром λ .

Доказательство. Достаточно использовать известную оценку для расстояния по вариации между распределением величины $\xi_{n,r}^{(u)}$ (см. (1.2)) и пуассоновским распределением с параметром λ (см., например, [15, с. 64]):

$$\begin{aligned} \rho_\lambda \left(\xi_{n,r}^{(u)} \right) &= \frac{1}{2} \sum_{k=0}^{\infty} \left| \mathbf{P} \left\{ \xi_{n,r}^{(u)} = k \right\} - e^{-\lambda} \frac{\lambda^k}{k!} \right| \leq \sum_{j=0}^{n-1} \mathbf{P}^2 \{ X_j = 1 \} = \\ &= \theta^2 \sum_{j=r}^{n+r-1} \frac{1}{(\theta+j)^2} \rightarrow 0 \quad (r \rightarrow \infty). \end{aligned} \quad \square$$

2. Случайные r -разбиения

В этом разделе рассматриваются такие разбиения конечного множества X на блоки, что некоторые выделенные элементы X принадлежат различным блокам. Пусть множество $X_{n+r} = \{1, 2, \dots, r, \dots, n+r\}$, $n \geq 1$, $r \geq 0$, разбивается на r или более блоков, причем элементы $1, 2, \dots, r$ принадлежат разным блокам. Такие разбиения будем называть кратко r -разбиениями; число разбиений, имеющих $r+k$ блоков, есть r -число Стирлинга второго рода $S(n, k, r)$ (см. (0.10)).

Далее, по аналогии с разделом 1, рассматриваем случайные r -разбиения множества X_{n+r} , когда на множестве всех таких разбиений задается вероятностная мера, согласно которой любому разбиению с $r+k$ блоками приписывается вероятность, пропорциональная θ^k , где $\theta > 0$ — произвольный параметр. Обозначим $r + \xi_{n,r}^{(6)}$ общее число блоков случайного r -разбиения множества X_{n+r} в такой параметрической модели. Тогда распределение случайной величины $\xi_{n,r}^{(6)} = \xi_{n,r}$ (для краткости верхний индекс далее будем опускать) имеет вид

$$\mathbf{P} \{ \xi_{n,r} = k \} = \frac{\theta^k S(n, k, r)}{B_{n,r}(\theta)}, \quad k = 0, 1, 2, \dots, n, \quad (2.1)$$

где (см. (0.6)) нормирующий множитель

$$B_{n,r}(\theta) = \sum_{k=0}^n S(n, k, r) \theta^k \quad (2.2)$$

есть r -полином Белла (здесь и далее используем обозначение $S(n, k, r)$ вместо более громоздкого обозначения $\left\{ \begin{smallmatrix} n+r \\ k+r \end{smallmatrix} \right\}_r$).

Из (2.1) и (2.2) следует, что производящая функция случайной величины $\xi_{n,r}$ есть

$$\mathbf{E} t^{\xi_{n,r}} = \frac{B_{n,r}(t\theta)}{B_{n,r}(\theta)}. \quad (2.3)$$

В свою очередь, из (2.3) следует, что и моменты величины $\xi_{n,r}$ также выражаются через отношения r -полиномов Белла. Так, имеют место следующие представления для математического ожидания и дисперсии величины $\xi_{n,r}$:

$$\mathbf{E}_{\theta} \xi_{n,r} = \frac{B_{n+1,r}(\theta)}{B_{n,r}(\theta)} - \theta - r, \quad (2.4)$$

$$\mathbf{D}_{\theta} \xi_{n,r} = \frac{B_{n+2,r}(\theta)}{B_{n,r}(\theta)} - \left(\frac{B_{n+1,r}(\theta)}{B_{n,r}(\theta)} \right)^2 - \theta. \quad (2.5)$$

Действительно, дифференцируя соотношение (2.3) в точке $t = 1$, имеем

$$\mathbf{E}_{\theta} \xi_{n,r} = \theta \frac{B'_{n,r}(\theta)}{B_{n,r}(\theta)}, \quad \mathbf{E}_{\theta} \xi_{n,r} (\xi_{n,r} - 1) = \theta^2 \frac{B''_{n,r}(\theta)}{B_{n,r}(\theta)}. \quad (2.6)$$

Далее, воспользовавшись рекуррентой $S(n+1, k, r) = S(n, k-1, r) + k S(n, k, r)$ для r -чисел Стирлинга (см. Введение), для производных r -полиномов Белла нетрудно получить представления

$$\begin{aligned} B'_{n,r}(\theta) &= \theta^{-1} B_{n+1,r}(\theta) - (1 + r\theta^{-1}) B_{n,r}(\theta), \\ B''_{n,r}(\theta) &= \theta^{-2} B_{n+2,r}(\theta) - \theta^{-1} (2(1 + r\theta^{-1}) + \theta^{-1}) B_{n+1,r}(\theta) + \\ &\quad + (r\theta^{-2} + (1 + r\theta^{-1})^2) B_n(\theta). \end{aligned} \quad (2.7)$$

Из (2.6) с учетом (2.7) непосредственно следуют формулы (2.4) и (2.5) (рутинные выкладки опускаем).

Случай $\theta = 1$ соответствует равновероятной модели, в которой любое r -разбиение имеет вероятность появления $B_{n,r}^{-1}$ (см. (0.7)). Для этого случая распределение случайной величины $\xi_{n,r}$ при $n \rightarrow \infty$ и фиксированном значении $r \geq 1$ асимптотически нормально (это эквивалентно асимптотической нормальности r -чисел Стирлинга второго рода [30]). Такой же характер асимптотического поведения распределения случайной величины $\xi_{n,r}$ сохраняется и в общей параметрической модели (при произвольном значении параметра θ), о чем будет идти речь ниже. Сначала исследуем асимптотическое (при $n \rightarrow \infty$ и фиксированных значениях параметров $\theta > 0$ и $k, r \geq 0$) поведение r -полиномов Белла $B_{n+k,r}(\theta)$, которое будет использовано затем для получения асимптотики распределения случайной величины $\xi_{n,r}$, а также для построения и расчета статистического критерия проверки гипотезы $H_0: \theta = 1$ о равновероятности разбиений с учетом возможных альтернатив.

3. Асимптотическое поведение r -полиномов Белла

Экспоненциальная производящая функция для r -полиномов Белла (см. (2.2)) указана в (0.8), откуда по теореме Коши имеем

$$\frac{2\pi i e^\theta}{(n+k)!} B_{n+k,r}(\theta) = \oint \exp\{rz + \theta e^z\} z^{-n-k-1} dz, \quad (3.1)$$

где замкнутый контур интегрирования C обходит начало координат в положительном направлении. Для исследования асимптотического поведения этого интеграла, когда $n \rightarrow \infty$, применим метод перевала, следуя выводу асимптотики чисел Белла [1]. При этом вычисления удобнее производить, используя точку перевала, соответствующую случаю $k = r = 0$.

Итак, в нашем случае точка перевала есть решение уравнения $ze^z = t$ при $t = (n+1)/\theta$. Это уравнение имеет единственный положительный корень $u_n = u_n(\theta)$, для которого нетрудно получить следующую асимптотику:

$$u_n = \ln t - \ln \ln t + \frac{\ln \ln t}{\ln t} + O\left(\left(\frac{\ln \ln t}{\ln t}\right)^2\right) \quad (t \rightarrow \infty). \quad (3.2)$$

Из (3.2) следует, что явная зависимость корня u_n от n и θ имеет вид

$$u_n = \ln n - \ln \ln n - \ln \theta + \frac{\ln \ln n}{\ln n} + O\left(\frac{1}{\ln n}\right). \quad (3.3)$$

Далее, контур интегрирования C трансформируется следующим образом: он состоит из большого отрезка вертикальной прямой $z = u_n + iy$, $-R < y < R$, $R > u_n$, и полуокружности с центром $(u_n, 0)$ и радиусом R , охватывающей начало координат. Как и для случая $\theta = 1$, показывается, что при $n \rightarrow \infty$ и любых фиксированных $\theta > 0$ и $k, r \geq 0$ вклад интеграла по этой полуокружности экспоненциально мал, поэтому интеграл в формуле (3.1) может быть заменен интегралом по всей прямой $z = u_n + iy$, $-\infty < y < \infty$. Тем самым задача сводится к анализу интеграла

$$\begin{aligned} I_{n,k,r}(\theta) &= \int_{-\infty}^{\infty} \exp\{r(u_n + iy) + \theta e^{u_n + iy} - (n+k+1) \ln(u_n + iy)\} dy = \\ &= \exp\{ru_n + \theta e^{u_n} - (n+k+1) \ln u_n\} \int_{-\infty}^{\infty} e^{\psi_{n,k,r}(y)} dy, \end{aligned} \quad (3.4)$$

где

$$\psi_{n,k,r}(y) = iry + \theta e^{u_n} (e^{iy} - 1) - (n+k+1) \ln\left(1 + \frac{iy}{u_n}\right)$$

(разность между интегралами (3.1) и (3.4) при $n \rightarrow \infty$ экспоненциально мала).

Заменив здесь θe^{u_n} на $(n+1)/u_n$, разложив функцию $\psi_{n,k,r}(y)$ в ряд:

$$\psi_{n,k}(y) = \left(r - \frac{k}{u_n}\right)iy - \lambda_{n,k} \frac{y^2}{2} + \lambda_{n,k} \sum_{m=3}^{\infty} a_{m,k} (iy)^m,$$

где

$$\lambda_{n,k} = \frac{(n+1)(u_n+1)+k}{u_n^2} = \frac{n+1}{u_n} \left(1 + \frac{n+k+1}{u_n(n+1)}\right),$$

$$a_{m,k} = \left(1 + \frac{n+k+1}{u_n(n+1)}\right)^{-1} \left(\frac{1}{m!} + (-1)^m \frac{n+k+1}{m u_n^{m-1}(n+1)}\right), \quad m \geq 3,$$

произведя в интеграле (3.4) замену переменных по формуле $x = y \sqrt{\lambda_{n,k}}$, приходим к представлению

$$I_{n,k,r}(\theta) = A_{n,k,r}(\theta) \int_{-\infty}^{\infty} e^{-x^2/2} \exp \left\{ \left(r - \frac{k}{u_n}\right) \frac{ix}{\sqrt{\lambda_{n,k}}} + \sum_{m=3}^{\infty} \frac{a_{m,k} (ix)^m}{\sqrt{\lambda_{n,k}^{m-2}}} \right\} dx, \quad (3.5)$$

где

$$A_{n,k,r}(\theta) = \frac{e^{r u_n + (n+1)/u_n - (n+k+1) \ln u_n}}{\sqrt{\lambda_{n,k}}}.$$

Оценка интеграла в (3.5) проводится далее уже стандартными рассуждениями: используя разложение Тейлора в точке $x = 0$ второй экспоненты в нем и произведя интегрирование (вклад дают лишь четные степени x), можно получить для него асимптотическое разложение по степеням малого параметра $\lambda_{n,k}^{-1} \sim (\ln n)/n$ с любой степенью точности. Ограничиваясь двумя главными членами этого разложения, получаем представление

$$I_{n,k,r}(\theta) = A_{n,k,r}(\theta) \sqrt{2\pi} \left[1 + \frac{b_{n,k,r}}{\lambda_{n,k}} + O\left(\frac{1}{\lambda_{n,k}^2}\right) \right],$$

где

$$\begin{aligned} b_{n,k,r} &= 3a_{4,k} - \frac{15}{2}a_{3,k}^2 + 3\left(r - \frac{k}{u_n}\right)a_{3,k} - \frac{1}{2}\left(r - \frac{k}{u_n}\right)^2 = \\ &= -\frac{1+6r(r-1)}{12} + \frac{7+12k(2r-1)}{24 \ln n} + O\left(\frac{\ln \ln n}{\ln^2 n}\right). \end{aligned} \quad (3.6)$$

В итоге приходим к следующему утверждению.

Теорема 3. Пусть $n \rightarrow \infty$, параметры $\theta > 0$ и $k, r \geq 0$ фиксированы, $u_n = u_n(\theta)$ — положительное решение уравнения $ze^z = (n+1)/\theta$, указанное в (3.3), и

$$\lambda_{n,k} = \frac{(n+1)(u_n+1)+k}{u_n^2}.$$

Тогда

$$B_{n+k,r}(\theta) = \frac{(n+k)!}{\sqrt{2\pi}} \frac{e^{ru_n+(n+1)/u_n-(n+k+1)\ln u_n-\theta}}{\sqrt{\lambda_{n,k}}} \left[1 + \frac{b_{n,k,r}}{\lambda_{n,k}} + O\left(\left(\frac{\ln n}{n}\right)^2\right) \right], \quad (3.7)$$

где величина $b_{n,k,r}$ указана в (3.6), а также

$$\frac{B_{n+k,r}(\theta)}{B_{n,r}(\theta)} = \frac{(n+k)!}{n! u_n^k} \sqrt{\frac{\lambda_{n,0}}{\lambda_{n,k}}} \left[1 + \frac{b_{n,k,r}}{\lambda_{n,k}} - \frac{b_{n,0,r}}{\lambda_{n,0}} + O\left(\left(\frac{\ln n}{n}\right)^2\right) \right]. \quad (3.8)$$

В качестве применения этого результата легко получаем асимптотические разложения для среднего (2.4) и дисперсии (2.5) случайной величины $\xi_{n,r}$:

$$\mathbf{E}_\theta \xi_{n,r} = \frac{n}{\ln n} \left(1 + \frac{\ln \ln n}{\ln n} + \frac{\ln \theta}{\ln n} + O\left(\left(\frac{\ln \ln n}{\ln n}\right)^2\right) \right), \quad (3.9)$$

$$\mathbf{D}_\theta \xi_{n,r} = \frac{n}{(\ln n)^2} \left(1 + 2 \frac{\ln \ln n}{\ln n} + 2 \frac{\ln \theta}{\ln n} + O\left(\left(\frac{\ln \ln n}{\ln n}\right)^2\right) \right). \quad (3.10)$$

Отметим, что, как следует из этих соотношений, параметры θ и r весьма слабо влияют на асимптотическое поведение моментов величины $\xi_{n,r}$.

4. Распределение числа блоков случайного r -разбиения

Производящая функция $\mathbf{E}_\theta t^{\xi_{n,r}} = P_{n,r,\theta}(t)$ уменьшенного на r числа блоков случайного r -разбиения множества X_{n+r} указана в (2.3), и она представляет собой отношение двух r -полиномов Белла. Известно [30], что все корни полинома $B_{n,r}(\theta)$ различные, действительные и неположительные. Это позволяет провести исследование распределения случайной величины $\xi_{n,r}$ по той же схеме, что и в разделе 1 при исследовании числа циклов случайной r -подстановки. Именно, обозначив ненулевые корни полинома $B_{n,r}(\theta)$ через $-\alpha_1, -\alpha_2, \dots, -\alpha_{n-1}$, можно записать для него представление

$$B_{n,r}(\theta) = \theta(\theta + \alpha_1) \dots (\theta + \alpha_{n-1}).$$

Отсюда следует, что

$$P_{n,r,\theta}(t) = t \prod_{j=1}^{n-1} \left(\frac{t\theta}{\theta + \alpha_j} + \frac{\alpha_j}{\theta + \alpha_j} \right). \quad (4.1)$$

Но это означает, что случайная величина $\xi_{n,r}$ может быть представлена в виде суммы независимых бернуллиевских случайных величин:

$$\xi_{n,r} = \xi_{n,r,1} + \xi_{n,r,2} + \dots + \xi_{n,r,n-1} + 1, \quad (4.2)$$

где

$$p_{n,r,j}(\theta) = \mathbf{P}\{\xi_{n,r,j} = 1\} = 1 - \mathbf{P}\{\xi_{n,r,j} = 0\} = \frac{\theta}{\theta + \alpha_j}, \quad j = 1, 2, \dots, n-1.$$

Поэтому можно записать новые представления для среднего значения и дисперсии величины $\xi_{n,r}$ через корни $\alpha_j, j = 1, \dots, n-1$:

$$\begin{aligned} \mathbf{E}_\theta \xi_{n,r} &= 1 + \sum_{j=1}^{n-1} p_{n,r,j}(\theta) = 1 + \theta \sum_{j=1}^{n-1} \frac{1}{\theta + \alpha_j}, \\ \mathbf{D}_\theta \xi_{n,r} &= \sum_{j=1}^{n-1} p_{n,r,j}(\theta)(1 - p_{n,r,j}(\theta)) = \theta \sum_{j=1}^{n-1} \frac{\alpha_j}{(\theta + \alpha_j)^2}, \end{aligned} \quad (4.3)$$

а также установить для нее как интегральный, так и локальный варианты центральной предельной теоремы с оценкой скорости сходимости. Необходимым условием для этого является условие $\mathbf{D}_\theta \xi_{n,r} \rightarrow \infty$, которое выполняется при любых конечных $\theta > 0$ и $r \geq 0$ (см. (3.10)).

Итак, в рассматриваемом случае справедливо следующее утверждение.

Теорема 4. При $n \rightarrow \infty$ имеет место соотношение

$$\sup_x \left| \mathbf{P}_\theta \left\{ \frac{\xi_{n,r} - \mathbf{E}_\theta \xi_{n,r}}{\sqrt{\mathbf{D}_\theta \xi_{n,r}}} < x \right\} - \Phi(x) \right| = O\left(\frac{1}{\sqrt{\mathbf{D}_\theta \xi_{n,r}}} \right), \quad (4.4)$$

а также

$$\max_k \left| \sqrt{\mathbf{D}_\theta \xi_{n,r}} \mathbf{P}_\theta \{\xi_{n,r} = k\} - \frac{1}{\sqrt{2\pi}} e^{-x^2(k)/2} \right| = O\left(\frac{1}{\sqrt{\mathbf{D}_\theta \xi_{n,r}}} \right), \quad (4.5)$$

где

$$x(k) = \frac{k - \mathbf{E}_\theta \xi_{n,r}}{\sqrt{\mathbf{D}_\theta \xi_{n,r}}}.$$

Дополним эту теорему некоторыми комментариями.

Из теоремы 4 и формул (3.9), (3.10) следует, что при любых конечных $\theta > 0$, $r \geq 0$ и $n \rightarrow \infty$ число блоков в случайном r -разбиении асимптотически нормально с параметрами

$$\left(\frac{n}{\ln n}, \frac{n}{(\ln n)^2} \right),$$

что для случая обычных разбиений ($r = 0$) отмечалось в предыдущей работе авторов [16]; при этом остается прежней и скорость сходимости с оценкой вида $O\left(\frac{\ln n}{\sqrt{n}}\right)$ в нормальной предельной теореме, как и в классическом случае равновероятных разбиений [28].

Далее, из соотношений (4.3) получаем, что при «близких» к гипотезе $H_0: \theta = 1$ альтернативах вида

$$H_{1n}: \theta = \theta_n = 1 + \delta_n, \quad \delta_n \rightarrow 0 \quad (n \rightarrow \infty), \quad (4.6)$$

справедливы представления

$$\begin{aligned} \mathbf{E}_{\theta_n} \xi_{n,r} &= 1 + \theta_n \sum_{j=1}^{n-1} \frac{1}{\theta_n + \alpha_j} = \mathbf{E}_1 \xi_{n,r} + \delta_n \mathbf{D}_1 \xi_{n,r} (1 + O(\delta_n)), \\ \mathbf{D}_{\theta_n} \xi_{n,r} &\sim \mathbf{D}_1 \xi_{n,r} \sim \frac{n}{(\ln n)^2}. \end{aligned}$$

Таким образом, утверждения (4.4) и (4.5) обобщают классические результаты на случай неравновероятных моделей случайных разбиений с ограничениями. Тем самым они позволяют построить и рассчитать для таких ситуаций статистический критерий проверки гипотезы $H_0: \theta = 1$ с учетом, в частности, альтернатив H_{1n} .

5. Статистический критерий

Пусть, для определенности, «близкие» альтернативы (4.6) задаются параметром $\delta_n < 0$. Тогда, основываясь на теореме 4, стандартными рассуждениями получаем, что критерий уровня значимости α для проверки гипотезы H_0 при таких альтернативах H_{1n} задается асимптотически критической областью

$$X_{\alpha}^{-}(n) = \left\{ S_n < \frac{n}{\ln n} - u_{\alpha} \frac{\sqrt{n}}{\ln n} \right\}, \quad u_{\alpha} = \Phi^{-1}(1 - \alpha), \quad (5.1)$$

где S_n — наблюдаемое значение случайной величины $\xi_{n,r}$.

«Пороговыми» альтернативами (т. е. для которых асимптотическое значение мощности критерия отделено как от α , так и от 1) в рассматриваемом случае являются альтернативы вида

$$H_{1n}^- : \theta = \theta_n = 1 - t \frac{\ln n}{\sqrt{n}}, \quad t > 0,$$

и при таких альтернативах предельное значение мощности критерия равно $\Phi(t - u_\alpha)$.

Аналогично, критическая область критерия уровня значимости α для проверки гипотезы H_0 против правосторонних альтернатив вида

$$H_{1n}^+ : \theta = \theta_n = 1 + t \frac{\ln n}{\sqrt{n}}, \quad t > 0,$$

асимптотически имеет вид

$$X_\alpha^+(n) = \left\{ S_n > \frac{n}{\ln n} + u_\alpha \frac{\sqrt{n}}{\ln n} \right\}, \quad (5.2)$$

и предельное значение мощности также равно $\Phi(t - u_\alpha)$.

Если же гипотеза H_0 проверяется против двусторонних альтернатив

$$H_{1n} = H_{1n}^- \cup H_{1n}^+ : \theta = \theta_n = 1 + t \frac{\ln n}{\sqrt{n}}, \quad t \neq 0,$$

то соответствующая критическая область строится с помощью объединения двух односторонних областей вида (5.1) и (5.2):

$$X_\alpha(n) = X_{\alpha/2}^-(n) \cup X_{\alpha/2}^+(n) = \left\{ \left| S_n - \frac{n}{\ln n} \right| \frac{\sqrt{n}}{\ln n} > u_{\alpha/2} \right\}.$$

В таком случае предельное значение мощности равно

$$\Phi(t - u_{\alpha/2}) + \Phi(-t - u_{\alpha/2}).$$

В заключение отметим, что в данной работе мы ограничились анализом «степенной» модели, когда случайный объект (подстановка или разбиение) с характеристикой k наблюдается с вероятностью, пропорциональной θ^k .

В то же время для теории и практики представляет несомненный интерес анализ и других моделей параметрического типа — «экспоненциальных», «логарифмических», моделей с несколькими параметрами и т. д. Надеемся, что такой анализ будет проведен специалистами, которые проявят интерес к этой проблематике.

Выражаем глубокую благодарность А. М. Зубкову за ценные замечания, способствовавшие устранению некоторых неточностей в изложении результатов статьи.

Список литературы

- [1] де Брёйн Н. Г., *Асимптотические методы в анализе*, М. : ИЛ, 1961, 248 с.
- [2] Ватутин В. А., Михайлов В. Г., “Предельные теоремы для числа пустых ячеек в равновероятной схеме размещения частиц комплектами”, *Теория вероятн. и ее примен.*, **27**:4 (1982), 684–692.
- [3] Медведев Ю. И., Ивченко Г. И., “Асимптотические представления конечных разностей от степенной функции в произвольной точке”, *Теория вероятн. и ее примен.*, **10**:1 (1965), 151–156.
- [4] Ивченко Г. И., Медведев Ю. И., “О случайных подстановках”, *Труды по дискретной математике*, **5**, 2002, 73–92.
- [5] Ивченко Г. И., Медведев Ю. И., “Метод В. Л. Гончарова и его развитие в анализе различных моделей случайных подстановок”, *Теория вероятн. и ее примен.*, **47**:3 (2002), 558–566.
- [6] Ивченко Г. И., Медведев Ю. И., “Об одном классе неравновероятных подстановок случайной степени”, *Труды по дискретной математике*, **7**, 2003, 75–88.
- [7] Ивченко Г. И., Медведев Ю. И., “Неравновероятные меры на множествах разложимых комбинаторных объектов”, *Обозрение прикл. и промышл. матем.*, **10**:2 (2003), 348–349.
- [8] Ивченко Г. И., Медведев Ю. И., Сачков В. Н., “Некоторые проблемы вероятностной комбинаторики”, В сб.: *Математика и безопасность информационных технологий*, Матер. конф. в МГУ 23.10.2003, М. : МЦНМО, 2004, 45–52.
- [9] Ивченко Г. И., Медведев Ю. И., “Случайные комбинаторные объекты”, *Доклады АН*, **396**:2 (2004), 151–154.
- [10] Ивченко Г. И., Медведев Ю. И., “Статистика параметрической модели случайных подстановок”, *Труды по дискретной математике*, **8**, 2004, 116–127.
- [11] Ивченко Г. И., Медведев Ю. И., “Статистические выводы для случайных подстановок по неполным данным”, *Труды по дискретной математике*, **9**, 2006, 66–76.
- [12] Ивченко Г. И., Медведев Ю. И., “Случайные подстановки: общая параметрическая модель”, *Дискретная математика*, **18**:4 (2006), 105–112.
- [13] Ивченко Г. И., Медведев Ю. И., “Случайные комбинаторные объекты в общей параметрической модели”, *Труды по дискретной математике*, **10**, 2007, 73–86.
- [14] Ивченко Г. И., Соболева М. В., “Некоторые неравновероятные модели случайных подстановок”, *Дискретная математика*, **23**:3 (2011), 23–31.
- [15] Ивченко Г. И., Медведев Ю. И., *Дискретные распределения. Вероятностно-статистический справочник: Одномерные распределения*, М. : ЛЕНАНД, 2015, 256 с.
- [16] Ивченко Г. И., Медведев Ю. И., “Параметрические модели случайных комбинаторных объектов экспоненциального типа и вопросы их вероятностно-статистического анализа”, *Математические вопросы криптографии*, **8**:3 (2017), 41–56.
- [17] Сачков В. Н., “Случайные разбиения множеств”, *Теория вероятн. и ее примен.*, **19**:1 (1974), 178–194.
- [18] Сачков В. Н., “Случайные разбиения”, *Вопросы кибернетики*, **16**, Тр. II Всесоюз. сем. по комбин. матем., Ч. 1, М. : Сов. радио, 1975, 88–94.
- [19] Сачков В. Н., *Комбинаторные методы дискретной математики*, М. : Наука, 1977, 320 с.
- [20] Сачков В. Н., *Вероятностные методы в комбинаторном анализе*, М. : Наука, 1978, 288 с.

- [21] Сачков В. Н., “Случайные разбиения множеств”, Математические вопросы кибернетики, **8**, 1999, 33–54.
- [22] Сачков В. Н., *Введение в комбинаторные методы дискретной математики*. 2-е изд., М. : МЦНМО, 2004, 424 с.
- [23] Bell E. T., “Exponential polynomials”, *Ann. Math.*, **35** (1934), 258–277.
- [24] Broder A. Z., “The r -Stirling numbers”, *Discrete Math.*, **49**:3 (1984), 241–259.
- [25] Carlitz L., “Weighted Stirling numbers of the first and second kind. I”, *Fibonacci Quart.*, **18** (1980), 147–162.
- [26] Corcino C. B., Corcino R. B., “An asymptotic formula for r -Bell numbers with real arguments”, *ISRN Discrete Mathematics* (2013), Article ID 274697, 7 с.
- [27] Harper L., “Stirling behavior is asymptotically normal”, *Ann. Math. Stat.*, **38**:2 (1961), 410–414.
- [28] Hwang H. K., “On convergence rates in the central limit theorems for combinatorial structures”, *Eur. J. Comb.*, **19**:3 (1998), 329–343.
- [29] Koutras M., “Non-central Stirling numbers and some applications”, *Discrete Math.*, **42** (1982), 73–89.
- [30] Mezö I., “On the maximum of r -Stirling numbers”, *Adv. Appl. Math.*, **41**:3 (2008), 293–306.