Math-Net.Ru

Общероссийский математический портал

Г. И. Ивченко, Ю. И. Медведев, Разбиения без малых блоков и r -присоединенные полиномы Белла в параметрической модели: вероятностно-статистический анализ, *Матем. вопр. криптогр.*, 2019, том 10, выпуск 1, 27–40

DOI: <https://doi.org/10.4213/mvk275>

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: 54.237.206.68

26 апреля 2023 г., 17:31:36

**Разбиения без малых блоков и r -присоединенные
полиномы Белла в параметрической модели:
вероятностно-статистический анализ**

Г. И. Ивченко, Ю. И. Медведев

Академия криптографии Российской Федерации, Москва

Получено 18.IV.2018

Аннотация. На множестве всех разбиений n -множества $X_n = \{1, 2, \dots, n\}$ на блоки, размеры которых больше $r \geq 0$, задается вероятностная мера, приписывающая каждому разбиению с k блоками вероятность, пропорциональную θ^k , где $\theta > 0$ — параметр меры. Доказана асимптотическая нормальность общего числа блоков случайного разбиения множества X_n в этой модели и рассчитан статистический критерий проверки гипотезы равномерности $H_0: \theta = 1$ с учетом альтернатив $H_1: \theta \neq 1$.

Ключевые слова: случайные разбиения, распределение числа блоков, r -присоединенные полиномы Белла

**Partitions without small blocks and r -associated Bell polynomials
in a parametric model: probabilistic-statistical analysis**

G. I. Ivchenko, Yu. I. Medvedev

Academy of Cryptography of the Russian Federation, Moscow

Abstract. On the set of all partitions of an n -element set $X_n = \{1, 2, \dots, n\}$ into blocks with sizes exceeding the number $r \geq 0$ a probability measure is defined such that for each partition with k blocks its probability is proportional to θ^k , where $\theta > 0$ is the parameter of the measure. The asymptotic normality of the number of blocks in a random partition of X_n in this model is proved, a statistical test for the uniformity hypothesis $H_0: \theta = 1$ against the alternatives $H_1: \theta \neq 1$ is constructed.

Key words: random partitions, number of blocks distribution, r -associated Bell polynomials

Citation: *Matematicheskie Voprosy Kriptografii*, 2019, v. 10, № 1, pp. 27–40 (Russian)

© Академия криптографии Российской Федерации, 2019 г.

1. Введение

Пусть $A_r = \{i : i > r\}$, $r \geq 0$, — подмножество множества натуральных чисел. Мы рассматриваем A_r -разбиения n -множества $X_n = \{1, 2, \dots, n\}$ на блоки, т. е. разбиения, размеры блоков которых являются элементами A_r . Другими словами, разбиения не должны содержать блоков размеров, не превосходящих r . Анализ разбиений с подобными ограничениями представляет теоретический и практический интерес, в частности, в задачах криптографии.

Обозначим через $\tilde{\sigma}_r(n, k)$ число A_r -разбиений множества X_n на k блоков. Когда $r = 0$, ограничение фактически отсутствует, т. е. мы имеем дело с обычными разбиениями. В этом случае числа $\tilde{\sigma}_r(n, k)$ совпадают с обычными числами Стирлинга 2-го рода $\sigma(n, k)$, определяющими число разбиений множества X_n на k блоков, поэтому, следуя [1, с. 249] и [8, с. 93], где для случая $r = 1$ числа $\tilde{\sigma}_1(n, k)$ названы *присоединенными (associated) числами Стирлинга второго рода*, числа $\tilde{\sigma}_r(n, k)$ (для общего случая $r \geq 1$) будем называть *r -присоединенными числами Стирлинга второго рода*. Свойства этих чисел, которые были введены и исследованы в [4] (см. также [9]), приведены в Приложении. Здесь же отметим лишь тот очевидный факт, что $\tilde{\sigma}_r(n, k) = 0$ для значений $n < (r + 1)k$.

Далее рассматриваются случайные разбиения множества X_n описанного типа, когда на множестве всех A_r -разбиений задается параметрическая вероятностная мера, приписывающая каждому разбиению с k блоками вероятность, пропорциональную θ^k , где $\theta > 0$ — параметр меры. При $\theta = 1$ получается равномерное распределение, приписывающее каждому разбиению одну и ту же вероятность $(\tilde{B}_{n,r}(1))^{-1}$, где $\tilde{B}_{n,r}(1)$ есть общее число разбиений рассматриваемого вида множества X_n (см. ниже формулу (1.3)). Таким образом, рассматриваемая модель включает в себя в качестве частного случая классическую равновероятную модель.

Обозначим через $\xi_{n,r}$ общее число блоков случайного A_r -разбиения множества X_n в такой параметрической модели. Распределение этой случайной величины имеет, очевидно, вид

$$\mathbf{P}_\theta \{ \xi_{n,r} = k \} = \frac{\theta^k \tilde{\sigma}_r(n, k)}{\tilde{B}_{n,r}(\theta)}, \quad k = 1, 2, \dots, \left[\frac{n}{r+1} \right], \quad (1.1)$$

где $[a]$ — целая часть числа a ; нормирующий множитель

$$\tilde{B}_{n,r}(\theta) = \sum_{k=1}^{\left[\frac{n}{r+1} \right]} \theta^k \tilde{\sigma}_r(n, k) \quad (1.2)$$

будем называть r -присоединенным полиномом Белла (свойства этих полиномов приведены в Приложении), а общее число A_r -разбиений множества X_n

$$\tilde{B}_{n,r}(1) = \sum_{k=1}^{\left[\frac{n}{r+1}\right]} \tilde{\sigma}_r(n, k) \quad (1.3)$$

— r -присоединенным числом Белла (для случая $r = 0$, когда ограничение отсутствует, $\tilde{B}_{n,0}(\theta) = B_n(\theta) = \sum_{k=1}^n \theta^k \sigma(n, k)$ есть обычный полином Белла [2], а $B_n(1) = \sum_{k=1}^n \sigma(n, k)$ — число Белла).

В силу (1.1) и (1.2) производящая функция случайной величины $\xi_{n,r}$ есть отношение полиномов (1.2):

$$\mathbf{E}_\theta t^{\xi_{n,r}} = \frac{\tilde{B}_{n,r}(t\theta)}{\tilde{B}_{n,r}(\theta)}, \quad (1.4)$$

и, как следует из [3],

$$\tilde{B}_{n,r}(\theta) = n! [z^n] \exp \left\{ \theta \sum_{j=r+1}^n \frac{z^j}{j!} \right\} = n! [z^n] \exp \{ \theta (e^z - E_r(z)) \}; \quad (1.5)$$

здесь и далее используются обозначения

$$E_r(z) = \sum_{j=0}^r \frac{z^j}{j!} \quad \text{и} \quad [z^n] f(z) = \text{coef}_{2^n} f(z).$$

Из (1.4) следует, что и моменты величины $\xi_{n,r}$ также выражаются через отношения полиномов (1.2). Так, для первых двух моментов имеют место следующие представления (см. Приложение):

$$\mathbf{E}_\theta \xi_{n,r} = \theta \frac{\tilde{B}'_{n,r}(\theta)}{\tilde{B}_{n,r}(\theta)} = \frac{\tilde{B}_{n+1,r}(\theta)}{\tilde{B}_{n,r}(\theta)} - \theta \binom{n}{r} \frac{\tilde{B}_{n-r,r}(\theta)}{\tilde{B}_{n,r}(\theta)}, \quad (1.6)$$

$$\begin{aligned} \mathbf{E}_\theta \xi_{n,r} (\xi_{n,r} - 1) &= \theta^2 \frac{\tilde{B}''_{n,r}(\theta)}{\tilde{B}_{n,r}(\theta)} = \\ &= \frac{\tilde{B}_{n+2,r}(\theta)}{\tilde{B}_{n,r}(\theta)} - \frac{\tilde{B}_{n+1,r}(\theta)}{\tilde{B}_{n,r}(\theta)} - \theta \frac{2n+2-r}{n+1-r} \binom{n}{r} \frac{\tilde{B}_{n+1-r,r}(\theta)}{\tilde{B}_{n,r}(\theta)} + \\ &+ \theta^2 \binom{n}{r} \binom{n-r}{r} \frac{\tilde{B}_{n-2r,r}(\theta)}{\tilde{B}_{n,r}(\theta)}. \end{aligned} \quad (1.7)$$

В настоящей работе исследуется асимптотическое поведение (при $n \rightarrow \infty$ и фиксированных значениях параметров $\theta > 0$ и $r \geq 0$) полиномов $\tilde{B}_{n+k,r}(\theta)$, $k = 0, \pm 1, \pm 2, \dots$, которое используется затем для получения асимптотики распределения случайной величины $\xi_{n,r}$, а также для построения и расчета статистического критерия проверки гипотезы равновероятности $H_0: \theta = 1$ с учетом альтернатив $H_1: \theta \neq 1$.

2. Асимптотика присоединенных полиномов Белла

Из (1.5) по теореме Коши имеем

$$\frac{2\pi i e^\theta}{(n+k)!} \tilde{B}_{n+k,r}(\theta) = \oint \frac{\exp\{\theta(e^z - E_r(z) + 1)\}}{z^{n+k+1}} dz, \quad (2.1)$$

где замкнутый контур интегрирования C обходит начало координат в положительном направлении. Для исследования асимптотического поведения этого интеграла при $n \rightarrow \infty$ мы применим классический метод перевала, следуя выводу в [5, гл. 6] асимптотики чисел Белла $B_n(1)$. При этом вычисления удобнее проводить, 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 известно асимптотическое представление (см. [5, с. 37]):

$$u_n = \ln t - \ln \ln t + \frac{\ln \ln t}{\ln t} + O\left(\left(\frac{\ln \ln t}{\ln t}\right)^2\right) \quad (t \rightarrow \infty). \quad (2.2)$$

Из (2.2) следует, что явная зависимость корня u_n от n и θ имеет вид

$$u_n = \ln n - \ln \ln n - \ln \theta + \frac{\ln \ln n}{\ln n} + O\left(\frac{1}{\ln n}\right). \quad (2.3)$$

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 (детали см. в [6] и в [7]), что при $n \rightarrow \infty$ и любых фиксированных $\theta > 0$ и $k, r \geq 0$ вклад интеграла по этой полуокружности экспоненциально мал, и интеграл в формуле (2.1) можно заменить интегралом по всей прямой $z = u_n + iy$, $-\infty < y < \infty$.

Тем самым задача сводится к анализу интеграла

$$\begin{aligned} I_{n,k,r}(\theta) &= \int_{-\infty}^{\infty} \exp\{-\theta(E_r(u_n+iy)-1) + \theta e^{u_n+iy} - (n+k+1)\ln(u_n+iy)\} dy = \\ &= A_{n,k,r}(\theta) \int_{-\infty}^{\infty} \exp\{\psi_{n,k,r}(y)\} dy, \end{aligned} \quad (2.4)$$

где

$$A_{n,k,r}(\theta) = \exp\{-\theta(E_r(u_n)-1) + \theta e^{u_n} - (n+k+1)\ln u_n\},$$

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

(разность между интегралами (2.1) и (2.4) экспоненциально мала).

Преобразуем функцию $\psi_{n,k,r}(y)$, разложив ее в ряд по степеням y . Имеем:

$$1) \quad -\theta \sum_{j=1}^r \frac{u_n^j}{j!} \sum_{s=1}^j \binom{j}{s} \left(\frac{iy}{u_n}\right)^s = -\theta \sum_{s=1}^r \frac{(iy)^s}{s!} E_{r-s}(u_n),$$

2) замена θe^{u_n} на $(n+1)/u_n$ дает

$$\theta e^{u_n} (e^{iy} - 1) = \frac{n+1}{u_n} \sum_{s=1}^{\infty} \frac{(iy)^s}{s!},$$

3) наконец,

$$-(n+k+1) \ln \left(1 + \frac{iy}{u_n}\right) = (n+k+1) \sum_{s=1}^{\infty} \frac{(-1)^s}{s} \left(\frac{iy}{u_n}\right)^s.$$

Объединяя эти три разложения, получим

$$\psi_{n,k,r}(y) = a_1 iy + \lambda_n \sum_{s=2}^{\infty} a_s (iy)^s, \quad \lambda_n = \frac{n+1}{u_n}, \quad (2.5)$$

где

$$a_s = a_s(n, k, r) = \begin{cases} -\left(\theta E_{r-1}(u_n) + \frac{k}{n}\right), & s=1, \\ \frac{1}{s!} \left[1 + \frac{(-1)^s (s-1)!}{u_n^{s-1}} \frac{n+k+1}{n+1} - \frac{\theta}{\lambda_n} E_{r-s}(u_n)\right], & s=2, 3, \dots \end{cases} \quad (2.6)$$

Проведя в интеграле (2.4) замену переменных по формуле $x = y\sqrt{\lambda_n}$, приходим к представлению

$$I_{n,k,r}(\theta) = \frac{A_{n,k,r}(\theta)}{\sqrt{\lambda_n}} \int_{-\infty}^{\infty} e^{-a_2 x^2} \exp \left\{ \frac{a_1}{\sqrt{\lambda_n}} ix + \sum_{s=3}^{\infty} \frac{a_s}{\sqrt{\lambda_n^{s-2}}} (ix)^s \right\} dx. \quad (2.7)$$

Подчеркнем следующую особенность этого представления. Как следует из (2.6) и (2.3), абсолютная величина a_1 с ростом n ведет себя как $O((\ln n)^{r-1})$, следовательно, $a_1/\sqrt{\lambda_n} = O((\ln n)^{r-1/2}/\sqrt{n})$. Остальные же величины a_s асимптотически имеют вид

$$a_s = \frac{1}{s!} \left(1 + O\left(\frac{1}{(\ln n)^{s-1}}\right) \right),$$

следовательно,

$$\frac{a_s}{\sqrt{\lambda_n^{s-2}}} = O\left(\left(\frac{\ln n}{n}\right)^{(s-2)/2}\right) \quad \text{при всех } s \geq 3.$$

Таким образом, все коэффициенты в показателе второй экспоненты в (2.7) являются бесконечно малыми величинами, когда $n \rightarrow \infty$.

С учетом этого дальнейшая оценка интеграла в (2.7) проводится уже стандартными рассуждениями: используя разложение Тейлора второй экспоненты в нем в точке $x = 0$ и проведя интегрирование с учетом формул (ниже $(2j-1)!! = (2j-1) \cdot (2j-3) \cdot \dots \cdot 3 \cdot 1$)

$$\int_{-\infty}^{\infty} x^{2j} e^{-cx^2} dx = \begin{cases} \sqrt{\frac{\pi}{c}}, & j = 0, \\ \sqrt{\frac{\pi}{c}} \frac{(2j-1)!!}{(2c)^j}, & j = 1, 2, \dots \end{cases}$$

(вклад дают лишь четные степени x), можно получить для него асимптотическое разложение вида

$$\sqrt{\frac{\pi}{a_2}} \left(1 + \sum_{j=1}^{s-1} \frac{b_{n,j}}{\lambda_n^j} + O\left(\frac{(\ln n)^{s(2r-1)}}{n^s}\right) \right),$$

где $b_{n,j}$ — многочлены от a_1 степени $2j$, в частности

$$\begin{aligned} b_{n,1} &= b_{n,1}(k, r) = 3a_4 - \frac{15}{2}a_3^2 + 3a_1a_3 - \frac{1}{2}a_1^2 = \\ &= \frac{\theta^2}{2} E_{r-1}^2(u_n) - \frac{\theta}{2} E_{r-1}(u_n) \left(1 - \frac{2}{\ln^2 n} \right) - \frac{1}{12} + O\left(\frac{1}{\ln^2 n}\right) = \\ &= O((\ln n)^{2(r-1)}). \end{aligned} \quad (2.8)$$

Далее, так как множитель $A_{n,k,r}(\theta)$ в формуле (2.4) с учетом равенства $\theta e^{u_n} = \lambda_n$ можно записать в виде

$$A_{n,k,r}(\theta) = \exp \left\{ -\theta (E_r(u_n) - 1) + \lambda_n - (n+k+1) \ln u_n \right\},$$

получаем для (2.7) асимптотическое представление

$$I_{n,k,r}(\theta) = \exp \left\{ -\theta (E_r(u_n) - 1) + \lambda_n - (n+k+1) \ln u_n \right\} \times \\ \times \sqrt{\frac{\pi}{\lambda_n a_2}} \left(1 + \sum_{j=1}^{s-1} \frac{b_{n,j}}{\lambda_n^j} + O \left(\frac{(\ln n)^{s(2r-1)}}{n^s} \right) \right). \quad (2.9)$$

В итоге, учитывая равенство (2.1), приходим к следующему утверждению.

Теорема 1. Пусть $n \rightarrow \infty$, параметры $\theta > 0$ и $k, r \geq 0$ фиксированны, $u_n = u_n(\theta)$ — положительное решение уравнения $z e^z = (n+1)/\theta$, указанное в (2.3), и $\lambda_n = (n+1)/u_n$. Тогда

$$\tilde{B}_{n+k,r}(\theta) = \frac{(n+k)!}{2\sqrt{\pi \lambda_n a_2}} \exp \left\{ -\theta E_r(u_n) + \lambda_n - (n+k+1) \ln u_n \right\} \times \\ \times \left(1 + \sum_{j=1}^{s-1} \frac{b_{n,j}}{\lambda_n^j} + O \left(\frac{(\ln n)^{s(2r-1)}}{n^s} \right) \right), \quad (2.10)$$

а также

$$\frac{\tilde{B}_{n+k,r}(\theta)}{\tilde{B}_{n,r}(\theta)} = \frac{(n+k)!}{n! u_n^k} \sqrt{\frac{a_2(n, 0, r)}{a_2(n, k, r)}} \times \\ \times \left(1 + \frac{b_{n,1}(k, r) - b_{n,1}(0, r)}{\lambda_n} + O \left(\frac{(\ln n)^{4r-2}}{n^2} \right) \right), \quad (2.11)$$

где величина $a_2 = a_2(n, k, r)$ указана в (2.6), а $b_{n,1} = b_{n,1}(k, r)$ — в (2.8).

В качестве применения этого результата нетрудно получить (с учетом формул (1.6) и (1.7)) асимптотики среднего и дисперсии случайной величины $\xi_{n,r}$ (громоздкие рутинные выкладки опускаем):

$$\mathbf{E}_\theta \xi_{n,r} = \frac{n}{u_n} - \frac{\theta}{r!} u_n^r \left(1 + O \left(\frac{1}{(\ln n)^2} \right) \right), \quad (2.12)$$

$$\mathbf{D}_\theta \xi_{n,r} = \frac{n}{u_n^2} (1 + o(1)). \quad (2.13)$$

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

3. Распределение числа блоков случайного A_r -разбиения

Производящая функция $E_\theta t^{\xi_{n,r}}$ числа блоков случайного A_r -разбиения множества X_n указана в (1.4), и она представляет собой отношение двух r -присоединенных полиномов Белла. Пользуясь полученной в теореме 1 асимптотикой этих полиномов (формула (2.10)), можно установить следующее утверждение об асимптотическом распределении случайной величины $\xi_{n,r}$.

Теорема 2. Пусть $n \rightarrow \infty$, параметры $\theta > 0$ и $r \geq 0$ фиксированны, и $u_n = u_n(\theta)$ — решение уравнения $ze^z = (n+1)/\theta$. Тогда нормированная случайная величина $\tilde{\xi}_{n,r} = (\xi_{n,r} - \mu_n)/\sigma_n$, где $\mu_n = n/u_n$, $\sigma_n = \sqrt{n}/u_n$, имеет в пределе стандартное нормальное распределение $\mathcal{N}(0, 1)$.

Доказательство. Характеристическая функция величины $\tilde{\xi}_{n,r}$, как следует из (1.4), есть

$$E_\theta e^{it\tilde{\xi}_{n,r}} = \frac{e^{-it\sqrt{n}} \tilde{B}_{n,r}(\theta e^{it/\sigma_n})}{\tilde{B}_{n,r}(\theta)},$$

и из (2.10) следует, что для нее справедливо асимптотическое представление

$$E_\theta e^{it\tilde{\xi}_{n,r}} = \exp \left\{ -it\sqrt{n} + n \left(\frac{1}{u_n(\theta e^{it/\sigma_n})} - \frac{1}{u_n(\theta)} \right) - n \ln \frac{u_n(\theta e^{it/\sigma_n})}{u_n(\theta)} \right\} (1+o(1)).$$

Обозначим $\Delta_n = \Delta_n(\theta) = (u_n(\theta e^{it/\sigma_n}) - u_n(\theta))/u_n(\theta)$, тогда

$$\begin{aligned} \ln E_\theta e^{it\tilde{\xi}_{n,r}} &= -it\sqrt{n} - n \frac{\Delta_n}{u_n(\theta e^{it/\sigma_n})} - n \ln(1 + \Delta_n) + o(1) = \\ &= -it\sqrt{n} - n \Delta_n \frac{u_n(\theta e^{it/\sigma_n}) + 1}{u_n(\theta e^{it/\sigma_n})} + \frac{n \Delta_n^2}{2} + O(n \Delta_n^3) + o(1), \end{aligned} \quad (3.1)$$

и дальнейший анализ сводится к исследованию асимптотики величины Δ_n .

В свою очередь, для этого надо рассмотреть разложение при малом x функции $u_n(\theta(1+x))$. По формуле Тейлора имеем:

$$u_n(\theta(1+x)) = u_n(\theta) + \theta x u'_n(\theta) + \frac{(\theta x)^2}{2} u''_n(\theta) + O(x^3 u'''_n(\theta)). \quad (3.2)$$

Производные функции $u_n(\theta)$ находятся дифференцированием тождества $\theta u_n(\theta) e^{u_n(\theta)} \equiv n+1$.

Так, из тождества $u_n(\theta) e^{u_n(\theta)} + \theta u'_n(\theta) e^{u_n(\theta)} + \theta u_n(\theta) u'_n(\theta) e^{u_n(\theta)} \equiv 0$ находим, что

$$\theta u'_n(\theta) = -\frac{u_n(\theta)}{u_n(\theta) + 1}. \quad (3.3)$$

Далее, дифференцируя (3.3) и заменяя возникающие при этом первые производные по полученной формуле, находим, что

$$\theta^2 u_n''(\theta) = \frac{u_n(\theta)}{u_n(\theta)+1} \left(1 + \frac{1}{(u_n(\theta)+1)^2} \right). \quad (3.4)$$

Из соотношений (3.3), (3.4) и (2.3) следует, что при $n \rightarrow \infty$ первые две производные функции $u_n(\theta)$ имеют порядок $O(1)$. Такое же заключение справедливо и для производных более высоких порядков функции $u_n(\theta)$. Заменяя теперь в (3.2) x на

$$e^{it/\sigma_n} - 1 = \frac{it}{\sigma_n} - \frac{t^2}{2\sigma_n^2} + O\left(\frac{1}{\sigma_n^3}\right),$$

получаем для Δ_n асимптотическое представление

$$\Delta_n = -\frac{it}{\sqrt{n}} \frac{u_n}{u_n+1} + O\left(\frac{1}{\sigma_n^3 u_n}\right). \quad (3.5)$$

Теперь в (3.1), с учетом (3.5),

$$\begin{aligned} -it\sqrt{n} - n\Delta_n \frac{u_n(\theta e^{it/\sigma_n})+1}{u_n(\theta e^{it/\sigma_n})} &= it\sqrt{n} \left(\frac{u_n(\theta e^{it/\sigma_n})+1}{u_n(\theta e^{it/\sigma_n})} \frac{u_n(\theta)}{u_n(\theta)+1} - 1 \right) = \\ &= it\sqrt{n} \frac{\Delta_n}{u_n(\theta)} (1+o(1)) + o(1) = o(1), \end{aligned}$$

а $n\Delta_n^2 = -t^2(1+o(1))$. Это в итоге дает нужное соотношение $\ln \mathbf{E}_\theta e^{it\tilde{\xi}_{n,r}} \rightarrow -t^2/2$. Теорема доказана. $\square$

Дополним эту теорему некоторыми комментариями.

1) Параметры нормировки в теореме 2 можно заменить их асимптотическими значениями $\mu_n \sim n/\ln n$, $\sigma_n \sim \sqrt{n}/\ln n$, поэтому число блоков $\xi_{n,r}$ случайного A_r -разбиения множества X_n асимптотически нормально с параметрами $(n/\ln n, n/(\ln n)^2)$ для любых фиксированных значений параметров $\theta > 0$ и $r \geq 0$, т.е. асимптотическое распределение величины $\xi_{n,r}$ от этих параметров не зависит. Для случая $r = 0$ (ограничения отсутствуют) асимптотическая нормальность с параметрами $(n/\ln n, n/(\ln n)^2)$ числа блоков ξ_n в случайном равновероятном разбиении n -множества, т.е. при гипотезе $H: \theta = 1$, впервые была доказана в [6] и впоследствии неоднократно передоказывалась (см., например, [1, с. 308–313]), а скорость сходимости с оценкой вида $O\left(\frac{\ln n}{\sqrt{n}}\right)$ в нормальной предельной теореме получена в [7].

Таким образом, утверждение теоремы 2 является новым и для случая $r = 0$, так как оно обобщает классический результат на случай неравновероятных моделей случайных разбиений. Тем самым теорема 2 позволяет построить и рассчитать статистический критерий проверки гипотезы $H_0: \theta = 1$ с учетом альтернатив $H_1: \theta \neq 1$. Подобные критерии могут представлять интерес в ситуациях, когда при разбиении множества одни типы разбиения являются более предпочтительными, другие — менее. Например, если более предпочтительным является появление большего числа блоков, то следует выбирать параметрическую модель с $\theta > 1$, и, наоборот, если предпочтение отдается малому числу блоков, то следует выбирать модель с параметром $\theta < 1$ (см. следующий раздел).

2) Теорема 2 доказана методом прямого асимптотического анализа характеристической функции величины $\xi_{n,r}$ с использованием асимптотической формулы для r -присоединенных полиномов Белла $\tilde{B}_{n,r}(\theta)$, установленной в теореме 1. В классической модели (когда $\theta = 1$ и $r = 0$) аналогичная теорема может быть доказана также с использованием того факта, что обычный полином Белла $B_n(\theta)$ имеет различные действительные корни, следствием чего является представление числа блоков ξ_n разбиения n -множества в виде суммы независимых бернуллиевских случайных величин. В свою очередь, такое представление с использованием стандартных методов позволяет установить для числа блоков ξ_n как интегральный, так и локальный варианты центральной предельной теоремы с оценками скорости сходимости. Известно также, что при некоторых (рассматривавшихся в литературе) обобщениях классических полиномов Белла отмеченное выше свойство их корней сохраняется (корни продолжают оставаться действительными и различными). Нам представляется вполне реалистичной гипотеза о том, что таковым свойством обладают и рассматриваемые в настоящей работе r -присоединенные полиномы Белла $\tilde{B}_{n,r}(\theta)$. В таком случае будет справедлив следующий, более сильный вариант теоремы 2:

при $n \rightarrow \infty$ имеет место соотношение

$$\sup_x \left| \mathbf{P}_\theta \left\{ \frac{\xi_{n,r} - \mathbf{E}_\theta \xi_{n,r}}{\sqrt{\mathbf{D}_\theta \xi_{n,r}}} < x \right\} - \Phi(x) \right| \leq \frac{C}{\sqrt{\mathbf{D}_\theta \xi_{n,r}}}, \quad (3.6)$$

где $\Phi(x)$ — стандартная нормальная функция распределения и $C > 0$ — некоторая константа, а также

$$\max_k \left| \sqrt{\mathbf{D}_\theta \xi_{n,r}} \mathbf{P}_\theta \{ \xi_{n,r} = k \} - \frac{1}{\sqrt{2\pi}} e^{-x^2(k)/2} \right| \leq \frac{C}{\sqrt{\mathbf{D}_\theta \xi_{n,r}}}, \quad (3.7)$$

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

4. Статистический критерий

Как отмечено выше, теорема 2 позволяет построить и рассчитать статистический критерий проверки гипотезы $H_0: \theta = 1$ с учетом альтернатив $H_1: \theta \neq 1$. Рассмотрим наиболее интересный для практики случай «близких» (при $n \rightarrow \infty$) альтернатив вида

$$H_{1n}: \theta = \theta_n = 1 + \delta_n, \quad \delta_n \rightarrow 0, \quad (4.1)$$

когда величина δ_n (порядок отклонения от равновероятности) выбирается так, чтобы при заданном уровне значимости критерия α асимптотическое значение его мощности было бы отделено как от α , так и от 1 (так называемые «пороговые» альтернативы). Такие альтернативы, как известно, определяются условием

$$\frac{\mu_n(1+\delta_n) - \mu_n(1)}{\sigma_n(1)} \rightarrow \text{const} \neq 0.$$

Но из разложения (2.3) легко следует, что в нашем случае

$$\frac{\mu_n(1+\delta_n) - \mu_n(1)}{\sigma_n(1)} = \delta_n \frac{\sqrt{n}}{\ln n} (1 + o(1)),$$

так что «пороговые» альтернативы определяются значениями

$$\delta_n = t \frac{\ln n}{\sqrt{n}}, \quad t \neq 0.$$

Пусть, для определенности, «близкие» альтернативы (4.1) задаются параметром $t > 0$. Тогда, основываясь на теореме 2, стандартными рассуждениями получаем, что критерий уровня значимости α для проверки гипотезы H_0 против таких альтернатив задается асимптотически критической областью

$$X_\alpha^+(n) = \left\{ \xi_{n,r} > \frac{n}{\ln n} + z_\alpha \frac{\sqrt{n}}{\ln n} \right\}, \quad z_\alpha = \Phi^{-1}(1-\alpha), \quad (4.2)$$

и предельное значение его мощности при альтернативах вида

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

равно $\Phi(t - z_\alpha)$.

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

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

асимптотически имеет вид

$$X_\alpha^-(n) = \left\{ \xi_{n,r} < \frac{n}{\ln n} - z_\alpha \frac{\sqrt{n}}{\ln n} \right\}, \quad (4.3)$$

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

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 (4.2) и (4.3) и имеет вид

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

В таком случае предельное значение мощности равно $\Phi(t - z_{\alpha/2}) + \Phi(-t - z_{\alpha/2})$.

5. Приложение

1. r -присоединенные числа Стирлинга

Из (1.2) и (1.5) следует представление

$$\tilde{\sigma}_r(n, k) = \frac{n!}{k!} [z^n] (e^z - E_r(z))^k, \quad n \geq (r+1)k, \quad (5.1)$$

так что $\tilde{\sigma}_r(n, k) = 0$ для значений $n < (r+1)k$.

Далее, дифференцируя экспоненциальную производящую функцию

$$f_{k,r}(z) = \sum_{n \geq (r+1)k} \tilde{\sigma}_r(n, k) \frac{z^n}{n!} = \frac{1}{k!} (e^z - E_r(z))^k,$$

получим соотношение

$$f'_{k,r}(z) = \sum_n \tilde{\sigma}_r(n+1, k) \frac{z^n}{n!} = k f_{k,r}(z) + f_{k-1,r}(z) \frac{z^r}{r!}.$$

Отсюда, приравнявая соответствующие коэффициенты, имеем следующее рекуррентное соотношение для чисел $\tilde{\sigma}_r(n, k)$:

$$\tilde{\sigma}_r(n+1, k) = k \tilde{\sigma}_r(n, k) + \binom{n}{r} \tilde{\sigma}_r(n-r, k-1). \quad (5.2)$$

2. r -присоединенные полиномы Белла

Используя соотношение (5.2), имеем

$$\begin{aligned} \tilde{B}'_{n,r}(\theta) &= \sum_{k=1}^{\left[\frac{n}{r+1}\right]} \theta^{k-1} k \tilde{\sigma}_r(n, k) = \\ &= \sum_{k=1}^{\left[\frac{n}{r+1}\right]} \theta^{k-1} \tilde{\sigma}_r(n+1, k) - \binom{n}{r} \sum_{k=1}^{\left[\frac{n}{r+1}\right]} \theta^{k-1} \tilde{\sigma}_r(n-r, k-1), \end{aligned}$$

откуда следует рекуррента

$$B'_{n,r}(\theta) = \frac{1}{\theta} \tilde{B}_{n+1,r}(\theta) - \binom{n}{r} \tilde{B}_{n-r,r}(\theta). \quad (5.3)$$

Далее, дифференцируя последнее соотношение и заменяя возникающие при этом первые производные по формуле (5.3), имеем:

$$\begin{aligned} \tilde{B}''_{n,r}(\theta) &= \theta^{-1} \tilde{B}'_{n+1,r}(\theta) - \binom{n}{r} \tilde{B}'_{n-r,r}(\theta) - \frac{1}{\theta^2} \tilde{B}_{n+1,r}(\theta) = \\ &= \frac{1}{\theta^2} \tilde{B}_{n+2,r}(\theta) - \frac{1}{\theta^2} \tilde{B}_{n+1,r}(\theta) - \\ &\quad - \frac{1}{\theta} \frac{2n+2-r}{n+1-r} \binom{n}{r} \tilde{B}_{n+1-r,r}(\theta) + \binom{n}{r} \binom{n-r}{r} \tilde{B}_{n-2r,r}(\theta). \end{aligned} \quad (5.4)$$

С учетом соотношений (5.3)–(5.4) и следующих из (1.4) формул

$$\begin{aligned} \mathbf{E}_\theta \xi_{n,r} &= \theta \frac{\tilde{B}'_{n,r}(\theta)}{\tilde{B}_{n,r}(\theta)}, \\ \mathbf{E}_\theta \xi_{n,r}(\xi_{n,r}-1) &= \theta^2 \frac{\tilde{B}''_{n,r}(\theta)}{\tilde{B}_{n,r}(\theta)}, \end{aligned}$$

получаем указанные в (1.6) и (1.7) представления для первых двух моментов случайной величины $\xi_{n,r}$.

Авторы признательны А. М. Зубкову и В. Г. Михайлову за внимательное прочтение рукописи и высказанные замечания, способствовавшие улучшению изложения ее результатов.

Список литературы

- [1] Сачков В. Н., *Введение в комбинаторные методы дискретной математики*, 2-е изд., М.: МЦНМО, 2004, 424 с.
- [2] Bell E. T., “Exponential polynomials”, *Ann. Math.*, **35** (1934), 258–277.
- [3] Ивченко Г. И., Медведев Ю. И., “Случайные комбинаторные объекты”, *Доклады РАН*, **396**:2 (2004), 151–154.
- [4] Charalambides Ch., “The generalized Stirling and C-numbers”, *Sankhya (Ser. A)*, **36** (1974), 419–436.
- [5] де Брёйн Н. Г., *Асимптотические методы в анализе*, М.: ИЛ, 1961, 248 с.
- [6] Harper L., “Stirling behavior is asymptotically normal”, *Ann. Math. Stat.*, **38**:2 (1961), 410–414.
- [7] Hwang H. K., “On convergence rates in the central limit theorems for combinatorial structures”, *European J. Combin.*, **19**:3 (1998), 329–343.
- [8] Риордан Дж., *Введение в комбинаторный анализ*, М.: ИЛ, 1963, 288 с.
- [9] Moll V. H., Ramirez J. L., Villamizar D., “Combinatorial and arithmetical properties of the restricted and associated Bell and factorial numbers” (2017), arXiv: 1707.08138v2.