

Werk

Verlag: Nauka; Российская академия наук санкт-петербургское отделение

Ort: Sankt-Peterburg

Kollektion: RusDML; Mathematica

Digitalisiert: Niedersächsische Staats- und Universitätsbibliothek Göttingen

Werk Id: PPN502905670

PURL: <http://resolver.sub.uni-goettingen.de/purl?PPN502905670>

OPAC: <http://opac.sub.uni-goettingen.de/DB=1/PPN?PPN=502905670>

LOG Id: LOG_0022

LOG Titel: Асимптотика случайных разбиений множеств

LOG Typ: article

Übergeordnetes Werk

Werk Id: PPN496972103

PURL: <http://resolver.sub.uni-goettingen.de/purl?PPN496972103>

OPAC: <http://opac.sub.uni-goettingen.de/DB=1/PPN?PPN=496972103>

Terms and Conditions

The Goettingen State and University Library provides access to digitized documents strictly for noncommercial educational, research and private purposes and makes no warranty with regard to their use for other purposes. Some of our collections are protected by copyright. Publication and/or broadcast in any form (including electronic) requires prior written permission from the Goettingen State- and University Library.

Each copy of any part of this document must contain there Terms and Conditions. With the usage of the library's online system to access or download a digitized document you accept the Terms and Conditions.

Reproductions of material on the web site may not be made for or donated to other repositories, nor may be further reproduced without written permission from the Goettingen State- and University Library.

For reproduction requests and permissions, please contact us. If citing materials, please give proper attribution of the source.

Contact

Niedersächsische Staats- und Universitätsbibliothek Göttingen
Georg-August-Universität Göttingen
Platz der Göttinger Sieben 1
37073 Göttingen
Germany
Email: gdz@sub.uni-goettingen.de

Ю. В. Якубович

АСИМПТОТИКА СЛУЧАЙНЫХ РАЗБИЕНИЙ МНОЖЕСТВ

§1. Постановка задачи и формулировка результатов

Разбиением множества называется его представление в виде объединения попарно непересекающихся подмножеств, которые обычно называют *блоками* разбиения. Если разбиваемое множество конечно, то конечны мощности всех блоков, которые в этом случае обычно называют *длинами блоков*, а также число всех различных разбиений. Множество всех разбиений множества $n = \{1, 2, \dots, n\}$ мы будем обозначать $\mathcal{P}(n)$. Мощность $\mathcal{P}(n)$, то есть число разбиений n , называют n -ным числом Белла и обозначают $B(n)$. Экспоненциальная производящая функция чисел Белла —

$$\sum_{n=0}^{\infty} \frac{B(n)}{n!} z^n = \exp(e^z - 1). \quad (1)$$

Имеет место замечательная формула Добинского

$$B(n) = \frac{1}{e} \sum_{k=0}^{\infty} \frac{k^n}{k!}.$$

Более подробное описание свойств чисел Белла можно найти, например, в [2].

Разбиением числа n называется неупорядоченный набор натуральных чисел $\{y_1, y_2, \dots, y_s\}$ таких, что $\sum_{k=1}^s y_k = n$ (s здесь произвольное натуральное число). Можно задавать разбиение числа n другим способом. Пусть $\mathcal{X}$ обозначает множество финитных последовательностей неотрицательных целых чисел (т.е. последовательностей, имеющих лишь конечное число ненулевых членов). Для $a \in \mathcal{X}$ будем писать $a \vdash n$, если $\sum_{k=1}^{\infty} k a_k = n$. Таким образом, если $a \vdash n$, то a определяет разбиение числа n , а именно

$$n = \underbrace{1 + \dots + 1}_{a_1} + \underbrace{2 + \dots + 2}_{a_2} + \dots + \underbrace{n}_{a_n}.$$

Работа выполнена при поддержке гранта MQV-000 Международного научного фонда (ISF).

Такой вектор $\mathbf{a}$ называется *вектором кратностей* разбиения. Мы в основном будем иметь дело с этим представлением разбиений чисел. Множество всех разбиений числа n обозначим $\mathcal{X}(n)$. В таком случае $\mathcal{X}$ можно себе представить как объединение множеств $\mathcal{X}(n)$ по всем неотрицательным целым n (считая, что существует единственное разбиение нуля на пустое множество слагаемых).

Существует естественная сюръекция π_n из $\mathcal{P}(n)$ на $\mathcal{X}(n)$, сопоставляющая каждому разбиению множества $\mathbf{n}$ разбиение числа n на длины блоков. Рассмотрим теперь равномерную вероятностную меру γ_n на $\mathcal{P}(n)$, сопоставляя каждому разбиению множества $\mathbf{n}$ меру $1/B(n)$. Сюръекция π_n и мера γ_n индуцируют меру μ_n на $\mathcal{X}(n)$. Комбинаторные соображения показывают (см. [4]), что мера μ_n сопоставляет каждому разбиению $\mathbf{a} \vdash n$ число

$$\mu_n(\mathbf{a}) = \frac{n!}{B(n)} \prod_{k=1}^n \frac{1}{(k!)^{a_k} a_k!}. \quad (2)$$

Эти меры называют *белловской статистикой*. А. М. Вершик в работе [6] поставил вопрос об асимптотических свойствах мер μ_n при $n \rightarrow \infty$. Рассмотрим постановку вопроса подробнее.

Упорядочим слагаемые в разбиении $\mathbf{a} \in \mathcal{X}(n)$ по невозрастанию: $n = N_1(\mathbf{a}) + N_2(\mathbf{a}) + \dots$, $N_1(\mathbf{a}) \geq N_2(\mathbf{a}) \geq \dots$. Диаграммой Юнга разбиения $\mathbf{a}$ называется подмножество квадранта $\mathbb{R}_+^2$:

$$\mathcal{D}_{\mathbf{a}} = \left\{ (x, y) : \exists i : 0 \leq x \leq N_i(\mathbf{a}), 0 \leq y \leq i \right\}.$$

Первый вопрос ставится следующим образом: найти такие нормировки по координатным осям, чтобы последовательность нормированных диаграмм, как подмножество $\mathbb{R}_+^2$, имела нетривиальный предел по вероятности при $n \rightarrow \infty$.

Пусть $\mathbf{a} \in \mathcal{X}(n)$. Обозначим r_n положительное решение уравнения

$$re^r = n. \quad (3)$$

(Заметим, что $r_n = \ln n - \ln \ln n + O\left(\frac{\ln \ln n}{\ln n}\right)$ при $n \rightarrow \infty$, см. [2]). Нормируем соответствующую $\mathbf{a}$ диаграмму Юнга $\mathcal{D}_{\mathbf{a}}$, сжав ее в r_n раз по оси Ox и в n/r_n раз по оси Oy . Полученное множество обозначим $\mathcal{D}'_{\mathbf{a}}$. Каждому разбиению $\mathbf{a} \in \mathcal{X}(n)$ сопоставим функции $f_{\mathbf{a}}$ и $g_{\mathbf{a}}$, полученные по правилу

$$\begin{aligned} f_{\mathbf{a}}(x) &= \sup \{ y : (x, y) \in \mathcal{D}'_{\mathbf{a}} \text{ или } y = 0 \}; \\ g_{\mathbf{a}}(y) &= \sup \{ x : (x, y) \in \mathcal{D}'_{\mathbf{a}} \text{ или } x = 0 \}. \end{aligned}$$

По построению обе функции непрерывны слева; мы же сделаем одну из них, например g_a , непрерывной справа, изменив ее значения в точках разрыва, и обозначим эту новую функцию снова g_a . Если нарисовать графики $y = f_a(x)$ и $x = g_a(y)$ на одной плоскости, то они образуют ломаную, являющуюся границей множества D'_a . Заметим, что

$$\int_0^{\infty} f_a(x) dx = \int_0^{\infty} g_a(y) dy = 1 \quad (4)$$

для любых $a \in \mathcal{X}(n)$. Ответ на первый вопрос дает

Теорема 1. Пусть $\varepsilon > 0$. Обозначим

$$T_\varepsilon(n) = \left\{ a(n) \in \mathcal{X}(n) : \forall x \in (0, 1 - \varepsilon) \left| f_{a(n)}(x) - 1 \right| < \varepsilon; \right. \\ \left. \forall y \in (0, 1 - \varepsilon) \left| g_{a(n)}(y) - 1 \right| < \varepsilon \right\}. \quad (5)$$

Тогда $\lim_{n \rightarrow \infty} \mu_n T_\varepsilon(n) = 1$.

Другими словами, теорема утверждает, что для достаточно больших n большинство разбиений таковы, что их диаграмма Юнга, нормированная на r_n по оси Ox и на n/r_n по оси Oy , близка к единичному квадрату. Это означает, что почти все блоки типичного в статистике Белла разбиения имеют мощность, асимптотически эквивалентную r_n . Поэтому естественно рассмотреть более детальные геометрические характеристики разбиений, дающие асимптотику величин таких блоков. Такой вопрос был предложен А. М. Вершиком Д. Александровскому в качестве дипломной работы [7]. Результаты [7] не были опубликованы; мы приводим их и более общие результаты здесь с другими доказательствами.

Уточним вопрос о “втором члене” асимптотик величин блоков. На пространстве $\mathcal{X}(n)$ с вероятностной мерой μ_n введем случайные величины X_k и Y_k : $X_k(a) = a_k$ — число слагаемых k в разбиении a ; $Y_k = \frac{k}{n} X_k$ — доля, приходящаяся на слагаемые k . Каждому разбиению $a \in \mathcal{P}(n)$ сопоставим гистограмму — функцию $\varphi_a: [1, n] \rightarrow \mathbb{R}$, заданную следующим образом: $\varphi_a(x) = Y_{[x]}(a)$. ($[x]$ обозначает целую часть x). Таким образом, для всякого $a \in \mathcal{P}(n)$, φ_a — кусочно-постоянная функция, имеющая промежутки постоянства $[k, k+1)$, где $k \in \mathbb{Z}$. Нас интересует, как нужно одновременно нормировать гистограммы φ_a , чтобы при $n \rightarrow \infty$ нормированные функции мало отличались друг от друга для почти всех a , т. е.

с вероятностью, стремящейся к 1. Другими словами, нужно найти такую нормировку гистограмм, чтобы индуцированная белловской статистикой на $\mathcal{X}(n)$ мера на нормированных функциях имела слабый предел, являющийся δ -мерой. Оказывается, такая нормировка существует, и предельная δ -мера сосредоточена на гауссовой кривой $y = \exp(-\frac{t^2}{2})/\sqrt{2\pi}$. Нормированная функция ψ_a связана с гистограммой φ_a соотношением $\psi_a(x) = \sqrt{r_n}\varphi_a(\sqrt{r_n}x + r_n)$. Вот точная формулировка этого факта.

Теорема 2 (см. также [7]). Пусть a, b – вещественные числа, $a < b$, и $\varepsilon > 0$. Тогда

$$\lim_{n \rightarrow \infty} \mu_n \left\{ a \in \mathcal{X}(n) : \left| \psi_a(t) - \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} \right| < \varepsilon \right\} = 1, \quad (6)$$

причем этот предел равномерен относительно $t \in [a, b]$.

Следствие 2.1. Пусть a, b – вещественные числа, $a < b$, и $\varepsilon > 0$. Тогда

$$\lim_{n \rightarrow \infty} \mu_n \left\{ a \in \mathcal{X}(n) : \left| \sum_{\substack{k: a\sqrt{r_n} \leq \\ k - r_n \leq b\sqrt{r_n}}} Y_k(a) - \frac{1}{\sqrt{2\pi}} \int_a^b e^{-\frac{t^2}{2}} dt \right| < \varepsilon \right\} = 1. \quad (7)$$

Следствие показывает, что при больших n в типичном разбиении множества n доля элементов, попавших в блоки с длинами от $r_n + a\sqrt{r_n}$ до $r_n + b\sqrt{r_n}$ (т.е. вероятность того, что выделенный элемент попадет в эти блоки), примерно равна вероятности того, что случайная величина со стандартным нормальным распределением попадет в отрезок (a, b) .

Автор благодарит А. М. Вершика за постановку задач и помощь в работе.

§2. ОБЗОР БЛИЗКИХ ЗАДАЧ

Метод рассуждений, используемый в настоящей работе, может быть применен не только к мере, определяемой (2), но и к более общей ситуации, под которую попадают некоторые описанные ниже примеры. Повтому опишем сперва эту ситуацию. Рассмотрим последовательность натуральных чисел $\{m_k\}_{k=1}^{\infty}$ и обозначим их экспоненциальную производящую функцию $M(z) = \sum_{k=1}^{\infty} m_k z^k / k!$. Для каждой такой последовательности и натурального числа n

определим меру μ_n на $\mathcal{X}$ следующим образом: при $\mathbf{a} \vdash n$

$$\mu_n(\mathbf{a}) = \frac{N_M(\mathbf{a})}{B_M(n)}, \quad (9)$$

где

$$N_M(\mathbf{a}) = n! \prod_{k=1}^n \frac{m_k^{a_k}}{(k!)^{a_k} a_k!},$$

$$B_M(n) = \sum_{\mathbf{a} \vdash n} N_M(\mathbf{a}),$$

и $\mu_n(\mathbf{a}) = 0$ при $\mathbf{a} \not\vdash n$. В настоящей работе дан критерий, позволяющий решать аналогичные задачи для таких мер μ_n , исследуя асимптотическое поведение функции M при стремлении ее аргумента к радиусу сходимости (Теорема 3).

Приведем некоторые примеры. (Все нижеперечисленные примеры индуцируют статистики, входящие в семейство мер, рассматриваемых в данной работе. Существуют также другие интересные статистики, не принадлежащие этому семейству. См. [1], [3], [6] и дальнейшее обсуждение). В уже рассмотренном случае разбиений конечных множеств, сопоставляя (2) и (9), видим, что $m_k = 1$, и $M(z) = e^z - 1$. Приведем другой пример. Рассмотрим множество всех перестановок множества $\mathbf{n} = \{1, 2, \dots, n\}$. Длины циклов, на которые разлагается перестановка, в сумме дают n и поэтому являются разбиением числа n . В этом случае $m_k = (k-1)!$, и $M(z) = -\ln(1-z)$. Еще два примера. Отображение множества $\mathbf{n}$ в себя разбивает множество на "компоненты связности": аналогично перестановкам определяются циклы, однако они не покрывают все множество $\mathbf{n}$; компонентой связности называется прообраз какого-либо цикла. Также естественным образом определяются компоненты связности помеченного графа с n вершинами. Мощности этих компонент связности также представляют разбиение числа n . В случае отображений $m_k = (k-1)! \sum_{i=1}^{k-1} \frac{k^i}{i!}$, а в случае графов $m_k = C_k^2$, где C_k^2 — биномиальный коэффициент.

Приведем теперь некоторые известные результаты, относящиеся к близким задачам. Аналогичная задача для перестановок с равномерной мерой (мерой Хаара) рассмотрена, в частности, в работе [9]. Пусть $N_i(\mathbf{a})$ — длина i -го по величине цикла в перестановке $g \in S_n$ (S_n здесь симметрическая группа степени n , циклы упорядочены по невозрастанию). На группе S_n задается отобра-

жение в n -мерный симплекс

$$\Sigma^n = \left\{ (x_i)_{i=1}^n : x_1 \geq x_2 \geq \dots \geq x_n \geq 0, \sum_{i=1}^n x_i = 1 \right\}$$

по правилу $\mathbf{a} \mapsto \frac{1}{n}(N_1(\mathbf{a}), \dots, N_n(\mathbf{a}))$. В [9] доказывается, что последовательность дискретных мер на Σ^n , индуцированных этим отображением и мерой Хаара имеет слабый предел — меру τ , для которой найдены все конечномерные распределения. Приведем один из основных результатов этой работы:

$$\lim_{n \rightarrow \infty} \tau \left\{ (x_1, x_2, \dots) : \frac{\ln x_n + n}{\sqrt{n}} \leq b \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^b e^{-\frac{t^2}{2}} dt.$$

Интересно, что в этом случае нет такой нормировки диаграмм Юнга, при которой индуцированная статистикой Хаара мера на нормированных диаграммах имела бы слабым пределом нетривиальную δ -меру. Типичная в статистике Хаара последовательность $N_1, N_2, \dots$, которая и определяет диаграмму Юнга, есть геометрическая прогрессия со знаменателем e^{-1} и первым членом an , где a — случайная величина, нетривиально распределенная на отрезке $[0, 1]$.

Следующая статистика не попадает в класс, определяемый (9), но интересна для нас тем, что нормированные надлежащим образом диаграммы Юнга имеют нетривиальный предел. *Стандартной таблицей Юнга, соответствующей диаграмме Юнга h* называется диаграмма h , заполненная натуральными числами от 1 до n , невозрастающими по столбцам и строкам. Эти объекты возникают в теории представлений симметрических групп. Число стандартных таблиц Юнга, соответствующих диаграмме h , обозначается $\dim h$. Мерой Планшереля называется мера ρ_n на диаграммах Юнга, определяемая формулой

$$\rho_n(h) = \frac{(\dim h)^2}{n!}.$$

Это вероятностная мера в силу соотношения $n! = \sum (\dim h)^2$, где суммирование ведется по всем диаграммам Юнга, соответствующим разбиениям числа n . Эта статистика изучается в работе [10]. Как уже говорилось, нормированные диаграммы Юнга имеют нетривиальный предел. Чтобы описать его, удобно перейти к координатам $X = \frac{1}{2}(x - y)$, $Y = \frac{1}{2}(x + y)$ и затем отождествить диаграмму Юнга h с функцией f_h , подграфиком которой она является.

В [10] доказывается следующая теорема. Пусть $\varepsilon > 0$. Тогда

$$\lim_{n \rightarrow \infty} \rho_n \left\{ \mathfrak{h} : \sup_{s \in \mathbb{R}} \left| \frac{1}{\sqrt{n}} f_{\mathfrak{h}}(s\sqrt{n}) - \Omega(s) \right| \leq \varepsilon \right\} = 1,$$

где

$$\Omega(s) = \begin{cases} \frac{2}{\pi} (s \arcsin s + \sqrt{1-s^2}) & \text{при } |s| < 1 \\ |s| & \text{при } |s| \geq 1. \end{cases}$$

Таким образом видим, что, в отличие от разбиений множеств, в этом случае нормировки диаграмм Юнга по обеим осям одинаковы.

Представляют интерес также многомерные обобщения рассматриваемых задач. Одному из них посвящены работы [11], [12], [13]. Сформулируем один из результатов этих работ. Обозначим CLP_n множество выпуклых ломаных с вершинами на решетке $L_n = (\frac{1}{n}\mathbb{Z})^2$, лежащих внутри $I = [0, 1]^2$ и имеющих начало и конец в точках $(1, 0)$ и $(0, 1)$. (Выпуклость незамкнутой кусочно-гладкой кривой означает, что при выбранной ориентации на плоскости сопровождающий репер всюду невырожден и имеет ту же ориентацию). Оказывается, что равномерная мера на CLP_n имеет слабый предел, который является δ -мерой, сосредоточенной на графике кривой $\sqrt{y} + \sqrt{x} = 1$. Чисто внешняя аналогия между этой и рассматриваемыми нами задачами усиливается, если заметить биекцию между множеством всех выпуклых ломаных с вершинами на решетке L_n , соединяющих точки $(0, 0)$ и $(1, 1)$, и множеством так называемых *строгих разбиений вектора* (n, n) . Строгим разбиением целочисленного вектора называется его разложение в сумму неупорядоченных векторов с неотрицательными целыми координатами, такую, что среди слагаемых нет пропорциональных векторов. Это одно из возможных обобщений на целочисленные векторы понятия разбиения числа. Вышеупомянутая биекция задается следующим образом: строгое векторное разбиение $((n_i, m_i))$, где векторы упорядочены по возрастанию отношений m_i/n_i и $\sum n_i = \sum m_i = n$, переходит в ломаную с вершинами в точках $(1 - \sum_{k=1}^i n_k/n, \sum_{k=1}^i m_k/n)$. Заметим, что в [11] и [13] доказательства используют в основном аналитические методы, тогда как в [12] – вероятностные, точнее, статфизические. Как отмечено А. М. Вершиком, в аналитическом плане оба метода эквивалентны: в одном случае основная теорема вытекает из метода перевала, а в другом – из того же метода, но в форме локальной предельной теоремы как теоремы об эквивалентности большого и малого канонического ансамблей. Подобный “статфизический” метод уже применялся к комбинаторным задачам (автору известны работы

Синая [12], Arratia i Tavaré [1], Fristedt [3], [4]), и настоящая работа содержит некоторое развитие идей этих авторов.

Приведенный выше обзор никоим образом не претендует на полноту и дан лишь затем, чтобы показать в самых общих чертах состояние дел в этой области.

§3. ОБЩАЯ ИДЕЯ РЕШЕНИЯ

Опишем основные идеи используемого метода. В основном мы будем употреблять вероятностную терминологию как наиболее подходящую для описания задач такого рода.

С этого момента мы фиксируем последовательность натуральных чисел m_k . При каждом n меру, которая соответствует ей согласно (9), обозначим μ_n .

Введем семейство функций X_k , $k = 1, 2, \dots$, определяемых по правилу $X_k(a)$ есть число слагаемых k в разбиении $a \in \mathcal{X}$. Допуская вольность обозначений, сужение функции X_k на множество $\mathcal{X}(n)$ мы также будем обозначать X_k . Введем также функцию N — то число, которое мы разбиваем на слагаемые. При всех $a \in \mathcal{X}$ выполнено соотношение $N(a) = \sum_{k=1}^{\infty} k X_k(a)$. Если рассмотреть на $\mathcal{X}$ (или на $\mathcal{X}(n)$) произвольную вероятностную меру ζ , то функции N и X_k становятся случайными величинами. Вектор $X(a) = (X_1(a), X_2(a), \dots)$ является вектором кратностей разбиения a числа $N(a)$. Согласно постановке нашей задачи мы в основном будем изучать поведение вектора X .

Наряду с мерами μ_n , введем на $\mathcal{X}$ семейство вероятностных мер ν_r , зависящее от одного вещественного положительного параметра r (и последовательности $\{m_k\}$), задавая вероятности того, что k -ый член последовательности $x \in \mathcal{X}$ равен a :

$$\nu_r(\{x_k = a\}) = \frac{(m_k r^k / k!)^a}{a!} e^{-m_k r^k / k!}. \quad (10)$$

Значения различных членов последовательности x определяются независимо. Таким образом на $\mathcal{X}$ задается счетное число независимых координатных пуассоновски распределенных случайных величин X_k , k -ая из которых имеет параметр $m_k r^k / (k!)$. Оказывается, при таком определении меры ν_r , выполнено следующее утверждение.

Лемма 1. При любой последовательности натуральных чисел $\{m_k\}$, $k = 1, 2, \dots$, пусть меры μ_n и ν_r на множестве $\mathcal{X}$ определены соотношениями (9) и (10) соответственно. Тогда при всех $n \in \mathbb{N}$ и $r > 0$ мера μ_n совпадает с условной мерой ν_r при условии $N = n$.

Доказательство этой леммы элементарно и может быть найдено в [1].

Таким образом, мы видим, что относительно меры μ_n случайные величины X_k ведут себя как независимые случайные величины, распределенные пуассоновски с параметрами $m_k r^k / k!$, но с наложенным условием

$$\sum_{k=1}^{\infty} k X_k = n. \quad (11)$$

Теперь просматривается план наших действий. Во первых, для каждого n мы можем подбирать параметр r так, чтобы максимизировать вероятность выполнения условия (11). Во-вторых, естественно ожидать, что при больших n малое число слагаемых в сумме (11) несущественно влияет на поведение всей суммы, и поэтому совместное распределение малого числа величин X_k близко к совместному распределению независимых пуассоновских случайных величин с соответствующими параметрами. Строгая формулировка и доказательство этих замечаний и составляет основную часть этой работы.

Прежде всего покажем, как выбирать параметр r . Радиус сходимости функции M обозначим R_M . (Возможно, $R_M = \infty$).

Лемма 2. При всех вещественных $r \in (0, R_M)$ и натуральных n

$$\nu_r \{ N = n \} = \frac{B_M(n) r^n}{n!} e^{-M(r)}. \quad (12)$$

Доказательство. Характеристическая функция ϕ случайной величины N легко находится из условия $N = \sum k X_k$ и имеет вид $\phi(s) = \exp(M(re^{is}) - M(r))$. Применяя формулу обращения (см. [8], гл. ШВ) видим, что

$$\nu_r \{ N = n \} = \frac{1}{2\pi} \int_{-\pi}^{\pi} \exp(M(re^{is}) - M(r) - ins) ds. \quad (13)$$

Известно (см. [1]), что экспоненциальная производящая функция чисел $B_M(n)$ выражается через функцию $M(z)$:

$$\sum_{n=0}^{\infty} \frac{B_M(n) z^n}{n!} = \exp(M(z)) \quad (14)$$

((1) является частным случаем этого отношения). Применяя формулу Коши к равенству (14), видим, что при всех $r \in (0, R_M)$

$$\frac{B_M(n)}{n!} = \frac{1}{2\pi r^n} \int_{-\pi}^{\pi} \exp(M(re^{is}) - ins) ds. \quad (15)$$

Сопоставляя (13) и (15), убеждаемся в справедливости утверждения леммы.

Чтобы максимизировать вероятность события $N = n$, продифференцируем правую часть (12) по r . Необходимым условием максимума является уравнение

$$rM'(r) = n, \quad (16)$$

которое при достаточно больших n имеет единственное вещественное решение, которое мы будем обозначать r_n . Заметим, что r_n стремится к радиусу сходимости функции M .

Символами E_r и D_r будем обозначать соответственно функционалы математического ожидания и дисперсии, взятые относительно меры ν_r . Математическое ожидание случайной величины N может быть легко вычислено (например, из формулы для характеристической функции): $E_r N = rM'(r)$. Из (16) видим, что $E_{r_n} N = n$.

Как уже говорилось, мы будем изучать совместные распределения малого (в некотором смысле) количества величин X_k . Особый интерес представляет случай, когда их количество растет с ростом n . Поэтому возникает необходимость использовать какие-то метрики на множестве мер.

Как и Арратиа и Таварé [1], мы будем оценивать расстояние между распределениями, используя метрику полной вариации. Напомним, что по определению расстояние полной вариации между дискретными мерами τ и ζ есть

$$d_{TV}(\tau, \zeta) = \frac{1}{2} \sum_s |\tau(s) - \zeta(s)|.$$

При таком определении расстояние полной вариации между вероятностными мерами есть вещественное число из промежутка $[0, 1]$. Если $d_{TV}(\xi_n, \xi)$, где ξ_n и ξ – распределения случайных величин, стремится к нулю, то это равносильно сходимости соответствующих случайных величин по вероятности.

Чтобы доказать теорему 2, мы введем параметр $t \in [a, b]$ и покажем, что равномерно по t имеем место (6). Однако в проводимых

рассуждениях не играет роли тот факт, что $t \in [a, b]$, и мы будем считать, что параметр t принадлежит произвольному множеству. Это окажется полезным в других приложениях рассматриваемого метода, см. теорему 4 ниже.

Пусть $K_n^{(t)}$ – семейство подмножеств $\mathbb{N}$, зависящее от натурального числа n и произвольного параметра t . Пусть $\mathbb{Z}_+^{K_n^{(t)}}$ обозначает множество неотрицательных целых чисел, возведенное в декартову степень, равную мощности множества $K_n^{(t)}$. Индексами вектора из $\mathbb{Z}_+^{K_n^{(t)}}$ удобно считать элементы из множества $K_n^{(t)}$. Существует естественная проекция $P_{K_n^{(t)}}$ множества $\mathcal{X}$ на $\mathbb{Z}_+^{K_n^{(t)}}$: $P_{K_n^{(t)}}(a_1, a_2, \dots) = (a_k, k \in K_n^{(t)})$. (Заметим, что $\mathcal{X}$ в этих обозначениях есть $\mathbb{Z}_+^{\mathbb{N}}$). Если ζ – мера на $\mathcal{X}$, то символом $\zeta|_{K_n^{(t)}}$ мы будем обозначать меру на $\mathcal{X}$, полученную следующим образом. Распределение векторной величины $\mathbf{X}_{K_n^{(t)}} = (X_k, k \in K_n^{(t)})$ будем обозначать $\tau_{K_n^{(t)}}$; мера $\tau_{K_n^{(t)}}$ и проекция $P_{K_n^{(t)}}$ индуцируют на $\mathcal{X}$ меру, которую мы и будем обозначать $\zeta|_{K_n^{(t)}}$. Другими словами, можно задать меру $\zeta|_{K_n^{(t)}}$ значениями на одноэлементных множествах:

$$\zeta|_{K_n^{(t)}}(\mathbf{a}) = \sum_{\mathbf{a}': X_k(\mathbf{a}') = X_k(\mathbf{a}), k \in K_n^{(t)}} \zeta(\mathbf{a}').$$

Согласно такому определению, мера $\zeta|_{K_n^{(t)}}$ полностью определяет совместное распределение случайного вектора $\mathbf{X}_{K_n^{(t)}}$ относительно меры ζ . Мы хотим оценивать расстояния между мерами $\mu_n|_{K_n^{(t)}}$ и $\nu_{r_n}|_{K_n^{(t)}}$. Для этого будем использовать следующую лемму. (Аналогичный результат может быть найден в [3]).

Лемма 3. Пусть при всех n множество $B_n^{(t)} \subset \mathbb{Z}_+^{K_n^{(t)}}$ такое, что

$$\nu_{r_n}((P_{K_n^{(t)}})^{-1}(B_n^{(t)})) \rightarrow 1, \quad (17)$$

$$\frac{\nu_{r_n}(N = n \mid P_{K_n^{(t)}} = \mathbf{a}_n^{(t)})}{\nu_{r_n}(N = n)} \rightarrow 1 \quad (18)$$

равномерно по всем t и по всем последовательностям $\mathbf{a}_n^{(t)} \in B_n^{(t)}$. Тогда расстояние полной вариации между мерами $\mu_n|_{K_n^{(t)}}$ и $\nu_{r_n}|_{K_n^{(t)}}$ стремится к нулю при $n \rightarrow \infty$.

Доказательство. Нетрудно видеть, что

$$d_{TV}(\mu_n|_{K_n^{(t)}}, \nu_{r_n}|_{K_n^{(t)}}) \leq \sup_{B \subset \mathbb{Z}_+^{K_n^{(t)}}} \left(\nu_{r_n}((P_{K_n^{(t)}})^{-1}(B)) - \frac{\nu_{r_n}((P_{K_n^{(t)}})^{-1}(B) \cap \{a : N(a) = n\})}{\nu_{r_n}(N = n)} \right). \quad (19)$$

Для каждого $B \subset \mathbb{Z}_+^{K_n^{(t)}}$ выражение, стоящее под знаком супремума в (19), ограничено сверху величиной

$$\begin{aligned} & \nu_{r_n}((P_{K_n^{(t)}})^{-1}(B \setminus B_n^{(t)})) + \sum_{a_n^{(t)} \in B \cap B_n^{(t)}} \left(\nu_{r_n}((P_{K_n^{(t)}})^{-1}(a_n^{(t)})) - \frac{\nu_{r_n}((P_{K_n^{(t)}})^{-1}(a_n^{(t)}) \cap \{a : N(a) = n\})}{\nu_{r_n}(N = n)} \right) \\ &= \tau_{K_n^{(t)}}(B \setminus B_n^{(t)}) + \sum_{a_n^{(t)} \in B \cap B_n^{(t)}} \tau_{K_n^{(t)}}(a_n^{(t)}) \left(1 - \frac{\nu_{r_n}(N = n \mid P_{K_n^{(t)}} = a_n^{(t)})}{\nu_{r_n}(N = n)} \right) \\ &\leq \tau_{K_n^{(t)}}(\mathbb{Z}_+^{K_n^{(t)}} \setminus B_n^{(t)}) + \sum_{a_n^{(t)} \in B_n^{(t)}} \tau_{K_n^{(t)}}(a_n^{(t)}) \left| 1 - \frac{\nu_{r_n}(N = n \mid P_{K_n^{(t)}} = a_n^{(t)})}{\nu_{r_n}(N = n)} \right|. \end{aligned}$$

Последнее выражение не зависит от B и, если выполнены условия леммы, стремится к нулю равномерно по t при $n \rightarrow \infty$, что и завершает доказательство.

§4. АНАЛИТИЧЕСКАЯ ЧАСТЬ

Чтобы показать (18), нам понадобится следующая чисто аналитическая

Лемма 4. Пусть $S(r, x)$ — вещественная функция, определенная на прямоугольнике $(0, R) \times [c, d]$ ($0 < R \leq \infty$, $c < d$). Пусть $f_t(r, x)$ — семейство вообще говоря комплекснозначных функций двух вещественных переменных $r \in (0, R)$ и $x \in [c, d]$, зависящее от произвольного параметра t . Пусть выполнены следующие условия:

(1) существует $R_0 < R$ такое, что при всех $r \in (R_0, R)$ функция $S(r, x)$ имеет глобальный максимум по x в точке $x = x_0 \in (c, d)$;

(2) $S(r, x_0) \rightarrow \infty$ при $r \rightarrow R$;

(3) $f_t(r, x) = o(\exp(S(r, x_0)))$ равномерно по $x \in [c, d]$ и по t при $r \rightarrow \infty$.

(4) существует такая функция $\delta(r) \rightarrow 0$ при $r \rightarrow R$, что

(i) $\delta(r) \sqrt{-S''_{xx}(r, x_0)} \rightarrow \infty$ при $r \rightarrow R$;

(ii) $S''_{xx}(r, x) = S''_{xx}(r, x_0)(1 + o(1))$ равномерно по $x \in [-\delta(r), \delta(r)]$ при $r \rightarrow \infty$;

(iii) $f_t(r, x) = f_t(r, x_0)(1 + o(1))$ равномерно по $x \in [-\delta(r), \delta(r)]$ и по t при $r \rightarrow \infty$;

Тогда

(А) если найдутся константы $C > 0$ и $R_0 < R$ такие, что при всех t и $r \in (R_0, R)$ $f_t(r, x_0) \notin [-C, C]$, то равномерно по t при $r \rightarrow R$

$$\int_c^d e^{S(r, x)} f_t(r, x) dx = \sqrt{-\frac{2\pi}{S''_{xx}(r, x_0)}} e^{S(r, x_0)} f_t(r, x_0)(1 + o(1)); \quad (20)$$

(Б) если найдется $R_0 < R$ такая, что $f_t(r, x_0) = 0$ при всех t и $r \in (R_0, R)$, то

$$\int_c^d e^{S(r, x)} f_t(r, x) dx = o\left(\sqrt{-\frac{2\pi}{S''_{xx}(r, x_0)}} e^{S(r, x_0)}\right). \quad (21)$$

Замечание. Если $f_t(r, x_0) \not\equiv 0$ в некоторой левой окрестности точки $r = R$, но $f_t(r, x_0) \rightarrow 0$ при $r \rightarrow R$, то лемма не дает ответа на вопрос об асимптотике рассматриваемого интеграла (требуется более глубокий анализ). Как мы увидим, в применении, рассматриваемом нами, такой случай не возникает.

Доказывается эта лемма применением обобщенного метода Лапласа к интегралу, стоящему в левой части (20). Напомним, что идея доказательства заключается в том, что интеграл по отрезку $[c, d]$ асимптотически эквивалентен интегралу по $\delta(r)$ -окрестности точки x_0 . Аналогичные оценки проведены в книге М. В. Фёдорюка [5]. Не составляет труда перенести доказательство со случая, рассмотренного в [5], на наш случай. Мы не будем проводить доказательство целиком, а поясним только, что означает каждое из условий. Условия (1) и (2) требуются, чтобы заменить интеграл от e^S по отрезку $[c, d]$ интегралом по окрестности точки x_0 . Условие (3) гарантирует, что умножение экспоненты S на функцию f_t не влияет на возможность замены области интегрирования; это условие можно существенно ослабить. Выполнение условия (4i) позволяет осуществить замену переменных в интеграле по $\delta(r)$ -окрестности точки x_0 , сводящую подынтегральную функцию к функции, асимптотически эквивалентной гауссовой кривой $\exp(-\frac{x^2}{2})$ и делающую область интегрирования

расширяющейся к $(-\infty, \infty)$. Наконец, условия (4ii) и (4iii) требуются, чтобы обосновать переход от эквивалентности под знаком интеграла к эквивалентности интегралов.

Вернемся к нашей задаче. Чтобы показать выполнение (18), нужно, во-первых, выбрать подходящее множество $B_n^{(t)}$ и, во-вторых, показать, что дробь в (18) равномерно стремится к единице. Лемма 2 дает выражение для знаменателя. Чтобы получить явное выражение для числителя, применим следующие рассуждения. Поскольку при рассмотрении на $\mathcal{X}$ меры ν_{r_n} координаты X_k независимы, то

$$\nu_{r_n}(N = n \mid P_{K_n^{(t)}} = a_n^{(t)}) = \nu_{r_n} \left(\sum_{k \notin K_n^{(t)}} k X_k = n - \sum_{k \in K_n^{(t)}} k a_k \right) \quad (22)$$

Обозначим $N_n^{(t)} = \sum_{k \notin K_n^{(t)}} k X_k$ и $\phi_n^{(t)}$ — ее характеристическую функцию (когда рассматриваемая на $\mathcal{X}$ мера есть ν_r). Тогда

$$\phi_n^{(t)}(s) = \exp \left(\left(M(r e^{is}) - \sum_{k \in K_n^{(t)}} \frac{m_k r^k e^{i k s}}{k!} \right) - \left(M(r) - \sum_{k \in K_n^{(t)}} \frac{m_k r^k}{k!} \right) \right).$$

Применим теперь формулу обращения. Имеем

$$\nu_{r_n} \left(N_n^{(t)} = n - \sum_{k \in K_n^{(t)}} k a_k \right) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi_n^{(t)}(s) \exp \left(-i \left(n - \sum_{k \in K_n^{(t)}} k a_k \right) s \right) ds$$

Принимая во внимание (12), (22) и проводя некоторые сокращения, получаем

$$\begin{aligned} & \frac{\nu_{r_n}(N = n \mid P_{K_n^{(t)}} = a_n^{(t)})}{\nu_{r_n}(N = n)} \\ &= \frac{n!}{B_M(n) r_n^n} \int_{-\pi}^{\pi} \exp \left(M(r_n e^{is}) - i n s \right. \\ & \quad \left. + \sum_{k \in K_n^{(t)}} \left(\frac{m_k r_n^k (1 - e^{i k s})}{k!} + i k a_k s \right) \right) ds. \end{aligned} \quad (23)$$

Итак, пусть мы выбрали семейство множеств $K_n^{(t)}$ и хотим найти множества $B_n^{(t)}$, для которых равномерно по t выполняется (17), с тем чтобы затем проверить (18). Рассмотрим множества

$$B_n^{(t)} = \left\{ a_n^{(t)} = (a_{k,n}^{(t)}, k \in K_n^{(t)}) : \left| \sum_{k \in K_n^{(t)}} \frac{m_k r_n^k}{(k-1)!} - \sum_{k \in K_n^{(t)}} k a_{k,n}^{(t)} \right| \leq b_n \right\},$$

где b_n – какие-то числа. Поскольку первая сумма в определении $B_n^{(t)}$ есть ничто иное, как $E_{r_n} \sum_{k \in K_n^{(t)}} k X_k$, то применение неравенства Чебышева показывает, что при $n \rightarrow \infty$

$$\nu_{r_n}((P_{K_n^{(t)}})^{-1}(\mathbb{R}^d \setminus B_n^{(t)})) \leq \frac{1}{b_n^2} \sum_{k \in K_n^{(t)}} \frac{k m_k r_n^k}{(k-1)!} \rightarrow 0$$

равномерно по t , если только последовательность b_n удовлетворяет условию

$$D_{r_n} \sum_{k \in K_n^{(t)}} k X_k = o(b_n^2). \quad (24)$$

Выбирая таким образом множества $B_n^{(t)}$, мы хотим найти некие общие условия на множества $K_n^{(t)}$, последовательность b_n и функцию M , чтобы выполнялись предположения леммы 2. Заметим, что одно такое условие, а именно (24), мы уже наложили. Осталось найти предположения, при которых правая часть (23) стремится к единице равномерно по t и по $a_n^{(t)} \in B_n^{(t)}$. Поскольку мы а priori знаем, что интеграл в правой части (23) имеет вещественное значение, будем считать его вещественную часть, чтобы применить лемму 4. Обозначим $U(r, s) = \operatorname{Re} M(re^{is})$, $V(r, s) = \operatorname{Im} M(re^{is})$. Нетрудно проверить, что $U(r, \cdot)$ есть четная, а $V(r, \cdot)$ – нечетная функции. Чтобы не иметь зависимости от параметров $a_n^{(t)}$ и t функции S из леммы 4, разобьем экспоненту в интеграле (23) на произведение двух экспонент: первую – от $M(re^{is}) - ins$, и вторую – от суммы по k . Имеем

$$\begin{aligned} & \frac{\nu_{r_n}(N = n \mid P_{K_n^{(t)}} \circ X = a_n^{(t)})}{\nu_{r_n}(N = n)} \frac{B_M(n) r_n^n}{n!} \\ &= \int_{-\pi}^{\pi} \operatorname{Re} (\exp (M(r_n e^{is}) - ins)) \\ & \times \operatorname{Re} \left(\exp \left(\sum_{k \in K_n^{(t)}} \left(\frac{m_k r_n^k (1 - e^{iks})}{k!} + ika_k s \right) \right) \right) ds \\ & - \int_{-\pi}^{\pi} \operatorname{Im} (\exp (M(r_n e^{is}) - ins)) \\ & \times \operatorname{Im} \left(\exp \left(\sum_{k \in K_n^{(t)}} \left(\frac{m_k r_n^k (1 - e^{iks})}{k!} + ika_k s \right) \right) \right) ds \end{aligned}$$

$$\begin{aligned}
&= \int_{-\pi}^{\pi} e^{U(r_n, s)} \cos(V(r_n, s) - ns) \exp\left(\sum_{k \in K_n^{(t)}} \frac{m_k r_n^k (1 - \cos(ks))}{k!}\right) (*) \\
&\quad \times \cos\left(\sum_{k \in K_n^{(t)}} \left(-\frac{m_k r_n^k \sin(ks)}{k!} + k a_k s\right)\right) ds \\
&- \int_{-\pi}^{\pi} e^{U(r_n, s)} \sin(V(r_n, s) - ns) \exp\left(\sum_{k \in K_n^{(t)}} \frac{m_k r_n^k (1 - \cos(ks))}{k!}\right) (**) \\
&\quad \times \sin\left(\sum_{k \in K_n^{(t)}} \left(-\frac{m_k r_n^k \sin(ks)}{k!} + k a_k s\right)\right) ds.
\end{aligned}$$

Мы хотим, используя лемму 4, получить асимптотические оценки интегралов (*) и (**). Рассмотрим $S(r, x) = U(r, x)$,

$$\begin{aligned}
f_{t, a_n^{(t)}} &= g(V(r_n, s) - ns) \exp\left(\sum_{k \in K_n^{(t)}} \frac{m_k r_n^k (1 - \cos(ks))}{k!}\right) \\
&\quad \times g\left(\sum_{k \in K_n^{(t)}} \left(-\frac{m_k r_n^k \sin(ks)}{k!} + k a_k s\right)\right) ds,
\end{aligned}$$

где $g = \cos$ в (*) и $g = \sin$ в (**). Заметим, что параметр t в формулировке леммы 4 может быть из произвольного множества, и то, что мы заменили его парой параметров $(t, a_n^{(t)})$, не делает лемму неприменимой.

Проверим выполнение предположений леммы 4 для таких функций S и f . Прежде всего заметим, что $U(r_n, s)$ достигает глобального максимума по s в точке $s = 0$, т. к. выполнены условия $m_k > 0$ при всех $k \geq 1$. (См., например, [5]). Тем самым для обоих интегралов выполнено условие (1) леммы 4. Как мы уже отмечали, $r_n \rightarrow R_M$ при $n \rightarrow \infty$. Поэтому $U(r_n, 0) \rightarrow \infty$, т. е. для обоих интегралов выполнено условие (2) леммы 4. Нетрудно видеть, что

$$\ln|f_t(r_n, s)| \leq 2 \sum_{k \in K_n^{(t)}} \frac{m_k r_n^k}{k!}.$$

Чтобы удовлетворить условию (3) леммы 4 для обоих интегралов, потребуем выполнения соотношения

$$\sum_{k \in K_n^{(t)}} \frac{m_k r^k}{k!} = o(M(r)) \quad (25)$$

при $r \rightarrow R_M$ равномерно по t , и оценим по модулю единицей все синусы и косинусы (включая функцию g). Поскольку $U''_{ss}(r, 0) \rightarrow -\infty$ при $r \rightarrow R_M$, то существует функция $\delta(r) \rightarrow 0$ такая, что выполнено условие (4i) леммы 4. Разлагая функцию $U''_{ss}(R, \cdot)$ в ряд Тейлора с остаточным членом в форме Лагранжа, видим, что $U''_{ss}(r, s) = U''_{ss}(r, 0) + sU'''_{sss}(r, \bar{s})$, где $\bar{s} \in [-\delta(r), \delta(r)]$. Поскольку $\delta(r)$ стремится к нулю, а $U'''_{sss}(r, 0) = 0$ (т. к. U — четная функция s), то выполнено условие (4ii). Условие (4iii) будем проверять отдельно для каждого из трех сомножителей, входящих в функцию $f_{t, a_n^{(t)}}$. Условия Коши–Римана в полярных координатах имеют вид

$$\begin{cases} rU'_r(r, s) = V'_s(r, s) \\ rV'_r(r, s) = -U'_s(r, s). \end{cases}$$

Разлагая $V(r, s)$ в ряд Тейлора по s , применяя первое из этих уравнений и (16), и замечая, что $M'(r) = U'_r(r, 0)$, видим, что первое произведение принимает вид $g(V''_{ss}(r_n, \bar{s}))$, где $\bar{s} \in [-\delta(r), \delta(r)]$. Поскольку $\delta(r)$ стремится к нулю, а $V''_{ss}(r, 0) = 0$ (т. к. V — нечетная функция s), то для первого сомножителя в функции $f_{t, a_n^{(t)}}$ условие (4iii) выполнено. Для экспоненты от суммы по k , разлагая ее аргумент в ряд Тейлора и оценивая косинус по модулю единицей, видим, что

$$\left| \sum_{k \in K_n^{(t)}} \frac{m_k r_n^k (1 - \cos(ks))}{k!} \right| \leq s^2 \sum_{k \in K_n^{(t)}} \frac{k m_k r_n^k}{(k-1)!}.$$

Налагая условие, что равномерно по t

$$\sum_{k \in K_n^{(t)}} \frac{k m_k r_n^k}{(k-1)!} = o \left(\left(r \frac{d}{dr} \right)^2 M(r) \right), \quad r \rightarrow R_M \quad (27)$$

также убеждаемся в справедливости условия (4iii). Наконец, для получения оценки третьего сомножителя вспомним, что $a_n^{(t)} \in B_n^{(t)}$. Имеем

$$\left| \sum_{k \in K_n^{(t)}} \left(-\frac{m_k r_n^k \sin(ks)}{k!} + k a_k s \right) \right| \leq s b_n + \left| \sum_{k \in K_n^{(t)}} \frac{m_k r_n^k k s^2 \sin(k\bar{s})}{2(k-1)!} \right|,$$

где $\bar{s} \in [-\delta(r_n), \delta(r_n)]$. Налагая еще одно условие

$$b_n^2 = o(D_{r_n} N) \quad (28)$$

получаем, что и для третьего сомножителя условие (4iii) выполнено. Соответственно, при выполнении (24), (27) и (28), это условие выполнено для функции $f_{t, a_n^{(t)}}$.

Теперь мы можем применить лемму 4. Для интеграла (*) имеет место случай (А), а для интеграла (**) — случай (В). Поэтому видим, что при $n \rightarrow \infty$

$$\frac{\nu_{r_n}(N = n \mid P_{K_n^{(t)}} \circ X = a_n^{(t)})}{\nu_{r_n}(N = n)} = \frac{n!}{B_M(n)r_n^n} \sqrt{-\frac{2\pi}{U''_{ss}(r_n, 0)}} e^{M(r_n)}(1+o(1)) \quad (29)$$

равномерно по t и по $a_n^{(t)} \in B_n^{(t)}$. Осталось лишь заметить, что асимптотика дроби $r_n^n B_M(n)/n!$ вычисляется в точности по такому же алгоритму, но только вместо интеграла (23) оценивается интеграл

$$\frac{r_n^n B_M(n)}{n!} = \int_{-\pi}^{\pi} \exp(M(r_n e^{is}) - ins) \, ds.$$

Повторяя предыдущие рассуждения, которые лишь упрощаются в силу более простого вида функции f_t (например, она не зависит от параметров), убеждаемся, что при $n \rightarrow \infty$

$$\frac{B_M(n)r_n^n}{n!} = \sqrt{-\frac{2\pi}{U''_{ss}(r_n, 0)}} e^{M(r_n)}(1+o(1)).$$

Сравнивая это равенство с равенством (29), видим, что, при сделанных предположениях, (18) выполнено равномерно по t и по $a_n^{(t)} \in B_n^{(t)}$.

Покажем теперь, что выполнения (27) достаточно, чтобы были выполнены и (28), и (25). Если предел левой части (27) при $r \rightarrow R_M$ конечен, то предел левой части (25) при $r \rightarrow R_M$ также конечен, а правой — бесконечен, и (25) выполнено. Если же предел левой части (27) бесконечен, то дважды воспользуемся правилом Лопиталья для дробного отношения, входящего в определение $o(\cdot)$. Это показывает, что отношение левой части (25) и аргумента $o(\cdot)$ равномерно по t стремится к нулю, т. е. (25) действительно выполнено. Что касается (28), то из (27) прямо следует, что последовательность b_n , удовлетворяющая (28) и (24), найдется.

Тем самым мы доказали следующую теорему.

Теорема 3. Пусть для семейства множеств $K_n^{(t)}$, зависящего от натурального числа n и произвольного параметра t , выполняется (27), т. е. при $r \rightarrow R_M$

$$\sum_{k \in K_n^{(t)}} \frac{km_k r^k}{(k-1)!} = o\left(\left(r \frac{d}{dr}\right)^2 M(r)\right)$$

равномерно по t . Тогда, в обозначениях второго параграфа,

$$d_{TV}(\mu_n|_{K_n^{(t)}}, \nu_{r_n}|_{K_n^{(t)}}) \rightarrow 0, \quad n \rightarrow \infty.$$

Замечание. Хотя теорема и дает ответ на вопрос, какие наборы величин X_k можно совместно приближать соответствующим набором независимых случайных величин, она не дает ответа на важный вопрос о скорости сходимости расстояния полной вариации между распределениями этих наборов. В ряде частных случаев ответ можно найти в статье [1].

§5. ПРИМЕНЕНИЕ К СТАТИСТИКЕ РАЗБИЕНИЙ МНОЖЕСТВ

Большой круг задач, связанных со случайными разбиениями множеств, рассмотрен в книге Сачкова [2]. Однако его результаты носят в основном “локальный” характер, в том смысле, что он рассматривает совместное поведение конечного числа координатных переменных X_k . Для рассмотрения совместного распределения растущего с ростом n числа переменных X_k , а также для решения задач, поставленных в первом параграфе, оказалась полезной изложенная выше техника. Некоторые результаты об асимптотике статистики Белла можно найти в [4].

Напомним, что для статистики Белла, индуцированной разбиениями множеств, $m_k = 1$, откуда $M(z) = e^z - 1$. Отсюда мы видим, что r_n есть решение уравнения (3) (ср. (16)). Условие (27) приобретает вид

$$\sum_{k \in K_n^{(t)}} \frac{k r_n^k}{(k-1)!} = o(r_n^2 e^{r_n}) \quad (30)$$

равномерно по t при $n \rightarrow \infty$.

Теорема 4. Пусть r_n есть решение уравнения $te^{r_n} = n$, и последовательности натуральных чисел b_n и c_n удовлетворяют условиям $(r_n - b_n)/\sqrt{r_n} \rightarrow \infty$ и $(c_n - r_n)/\sqrt{r_n} \rightarrow \infty$ при $n \rightarrow \infty$. Пусть $K_{1,n} = \{1, 2, \dots, b_n\} \cup \{c_n, c_n + 1, \dots, n\}$, а $K_{2,n} \subset \mathbb{N} \setminus K_{1,n}$, причем мощность $\#K_{2,n} = o(\sqrt{r_n})$ при $n \rightarrow \infty$. Тогда при $i = 1, 2$

$$d_{TV}(\mu_n|_{K_{i,n}}, \nu_{r_n}|_{K_{i,n}}) \rightarrow 0, \quad n \rightarrow \infty. \quad (31)$$

Доказательство. В силу теоремы 3 для доказательства достаточно проверить выполнение (30) для множеств вида $K_{1,n}$ и $K_{2,n}$. Для $K_{1,n}$ свойство (30) следует из нормального приближения пуассоновского распределения (см. [8]). Для $K_{2,n}$ воспользуемся

формулой Стирлинга. Представим $k \in K_{2,n}$ в виде $k = r_n + t_n \sqrt{r_n}$, где t_n — вещественное число. Тогда

$$\frac{k^2 r_n^k}{k!} \sim \frac{(r_n + t_n \sqrt{r_n})^2 r_n^{r_n + t_n \sqrt{r_n}} e^{r_n + t_n \sqrt{r_n}}}{\sqrt{2\pi(r_n + t_n \sqrt{r_n})} (r_n + t_n \sqrt{r_n})^{r_n + t_n \sqrt{r_n}}}.$$

Нетрудно проверить, что если умножить последнюю дробь на $\#K_{2,n} = o(\sqrt{r_n})$ и поделить на $r_n^2 e^{r_n}$, то полученное выражение будет стремиться к нулю при $n \rightarrow \infty$, что и доказывает справедливость теоремы.

Следствие 4.1. Пусть t — вещественное число. Обозначим $K_n^{(t)}$ множество, состоящее из одного элемента — целой части числа $r_n + t\sqrt{r_n}$. Тогда при любых вещественных $a < b$ имеем

$$d_{TV}(\mu_n|_{K_n^{(t)}}, \nu_{r_n}|_{K_n^{(t)}}) \rightarrow 0$$

равномерно по $t \in [a, b]$ при $n \rightarrow \infty$.

Доказательство. Собственно, это не есть прямое следствие Теоремы 4, а скорее следствие ее доказательства. Нужно опять проверить выполнение (30) для данного в условии следствия семейства множеств $K_n^{(t)}$, что непосредственно следует из доказательства Теоремы 4.

Теперь мы имеем возможность доказать Теорему 2.

Доказательство теоремы 2. Согласно следствию 4.1 математические ожидания и дисперсии случайных величин $\psi_a(t)$ относительно меры μ_n , равномерно по t асимптотически эквивалентны при $n \rightarrow \infty$ соответствующим математическим ожиданиям и дисперсиям относительно меры ν_{r_n} . Повтому

$$\begin{aligned} E\psi_a(t) &\sim \frac{r_n^{[r_n + t\sqrt{r_n}]}}{\sqrt{r_n} [r_n + t\sqrt{r_n}]!} e^{-r_n}; \\ D\psi_a(t) &\sim \frac{r_n^{[r_n + t\sqrt{r_n}] - 1}}{[r_n + t\sqrt{r_n}]!} e^{-2r_n} \end{aligned}$$

равномерно по $t \in [a, b]$ при $n \rightarrow \infty$. Применим к случайной величине $\psi_a(t)$ неравенство Чебышева:

$$\mu_n \left\{ a \in \mathcal{X}(n) : |\psi_a(t) - E\psi_a(t)| < \varepsilon \right\} > 1 - \frac{1}{\varepsilon^2} D\psi_a(t). \quad (32)$$

Применяя нормальное приближение пуассоновски распределенных случайных величин (см. [8]), видим, что математическое ожидание $\psi_a(t)$ стремится к $e^{-t^2/2}/\sqrt{2\pi}$, а ее же дисперсия стремится

к нулю равномерно по $t \in [a, b]$ при $n \rightarrow \infty$. Переходя в (32) к пределу при $n \rightarrow \infty$, убеждаемся в справедливости теоремы.

Доказательство следствия 2.1. Этот факт следует из равномерности предела (6) относительно t и равенства

$$\lim_{n \rightarrow \infty} \left(\sum_{k: a\sqrt{r_n} \leq k - r_n \leq b\sqrt{r_n} Y_k(a(n)) - \int_a^b \psi_{a(n)}(x) dx \right) = 0$$

для любой последовательности $a(n) \in \mathcal{X}(n)$, в справедливости которого нетрудно убедиться, рассматривая интеграл как предел сумм Римана.

Доказательство теоремы 1. Доказательство проще провести не прямо, а как следствие из теоремы 2. Прежде всего заметим, что для любых $x_1 < x_2$ выражение

$$\int_{f_a(x_2)}^{f_a(x_1)} g_a(y) dy$$

обозначает деленную на n сумму слагаемых разбиения a , больших или равных $x_1 r_n$ и меньших $x_2 r_n$. Таким образом, при любых $b > 0$ и достаточно больших n

$$\sum_{k: -b\sqrt{r_n} \leq k - r_n < b\sqrt{r_n} Y_k(a) = \int_{f_a(1+b/\sqrt{r_n})}^{f_a(1-b/\sqrt{r_n})} g_a(y) dy. \quad (33)$$

При $\delta > 0$ и $b > 0$ обозначим

$$S_{\delta, b} = \left\{ a \in \mathcal{X}(n) : \left| \int_{f_a(1+b/\sqrt{r_n})}^{f_a(1-b/\sqrt{r_n})} g_a(y) dy - \frac{1}{\sqrt{2\pi}} \int_{-b}^b e^{-\frac{t^2}{2}} dt \right| < \delta \right\} = 1.$$

Принимая во внимание (33) видим, что в этих обозначениях следствие 2.1 утверждает, что при любых $\delta > 0$ и $b > 0$ при $n \rightarrow \infty$ имеет место $\mu_n S_{\delta, b}(n) \rightarrow 1$. (Нетрудно показать, что замена нестрогого неравенства в (7) на строгое в (33) ни на что не влияет). Выберем $\epsilon > 0$ и покажем, что можно так подобрать δ и b , что при достаточно больших n множества $S_{\delta, b}(n)$ содержатся в $T_\epsilon(n)$. Это включение и доказывает теорему.

Для краткости обозначим $\Phi(b) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^b \exp(-\frac{t^2}{2}) dt$ — функцию распределения нормального закона. Тогда нетрудно проверить, что условие $a \in S_{\delta,b}(n)$ равносильно

$$\int_{f_a(1+b/\sqrt{r_n})}^{f_a(1-b/\sqrt{r_n})} g_a(y) dy \in (2\Phi(b) - 1 - \delta, 2\Phi(b) - 1 + \delta). \quad (34)$$

Заметим также, что по определению функций f_a и g_a

$$\begin{aligned} \text{если } y \in \left(f_a \left(1 + \frac{b}{\sqrt{r_n}} \right), f_a \left(1 - \frac{b}{\sqrt{r_n}} \right) \right), \\ \text{то } 1 - \frac{b}{\sqrt{r_n}} \leq g_a(y) \leq 1 + \frac{b}{\sqrt{r_n}}. \end{aligned} \quad (35)$$

Отсюда видим, что

$$\int_{f_a(1+b/\sqrt{r_n})}^{f_a(1-b/\sqrt{r_n})} g_a(y) dy \geq \left(1 - \frac{b}{\sqrt{r_n}} \right) \left(f_a \left(1 - \frac{b}{\sqrt{r_n}} \right) - f_a \left(1 + \frac{b}{\sqrt{r_n}} \right) \right).$$

Это не есть точная оценка, но при достаточно больших n к ней можно приблизиться сколь угодно близко. Поэтому, принимая во внимание (34), видим, что при $a \in S_{\delta,b}(n)$

$$\left(1 - \frac{b}{\sqrt{r_n}} \right) \left(f_a \left(1 - \frac{b}{\sqrt{r_n}} \right) - f_a \left(1 + \frac{b}{\sqrt{r_n}} \right) \right) \geq 2\Phi(b) - 1 - \delta.$$

Отсюда

$$\left(f_a \left(1 - \frac{b}{\sqrt{r_n}} \right) - f_a \left(1 + \frac{b}{\sqrt{r_n}} \right) \right) \geq \Phi(b) - \frac{\delta}{2}$$

при достаточно больших n . Из этого неравенства и из (35) заключаем, что на отрезке длины, не меньшей, чем $\Phi(b) - \frac{\delta}{2}$ функция g принимает значения, большие $1 - \frac{b}{\sqrt{r_n}}$. Пусть этот отрезок содержит отрезок $(\alpha, \alpha + \Phi(b) - \frac{\delta}{2})$ при некотором $\alpha \geq 0$. Покажем, что $\alpha = 0$. Действительно, если $\alpha > 0$, то $\int_0^\infty g_a(y) dy \geq (\alpha + \Phi(b) - \frac{\delta}{2})(1 - \frac{b}{\sqrt{r_n}})$ при $a \in S_{\delta,b}(n)$. Множество $S_{\delta,b}$ непусто при любых $\delta > 0$ и $b > 0$, а правая часть неравенства может быть сделана строго больше единицы выбором δ и b . Это противоречит (4), которое выполнено при любых $a \in \mathcal{X}$. Тем самым для того, чтобы при любых

$y \in (0, 1 - \varepsilon)$ и при любых $a \in S_{\delta, b}(n)$ для достаточно больших n выполнялось $g_a(y) \in (1 - \varepsilon, 1 + \varepsilon)$, нужно выбрать δ и b такими, чтобы $\Phi(b) - \frac{\delta}{2} > 1 - \varepsilon$.

Покажем теперь, что при таком выборе δ и b при любых $a \in S_{\delta, b}(n)$ для достаточно больших n выполнено $f_a(y) \in (1 - \varepsilon, 1 + \varepsilon)$. Поскольку при таких δ и b $g_a(1 - \varepsilon) > 1 - \varepsilon$, то, по определению f_a , при всех $x \in (0, 1 - \varepsilon)$ имеет место $f_a(x) > 1 - \varepsilon$. Предположим, что найдется такой $x_0 > 0$, что $f_a(x_0) > 1 + \varepsilon$. Тогда $\int_0^\infty f_a(x) dx \geq x_0(1 + \varepsilon) + (1 - b/\sqrt{\tau_n} - x_0)(\Phi(b) - \frac{\delta}{2})$. Выбором n , δ и b правая часть этого неравенства может быть сделана строго больше единицы, причем δ и b можно выбрать так, чтобы $\Phi(b) - \frac{\delta}{2} > 1 - \varepsilon$, а это противоречит (4). Тем самым, при $\Phi(b) - \frac{\delta}{2} > 1 - \varepsilon$ и при достаточно больших n имеет место включение $T_\varepsilon(n) \supset S_{\delta, b}(n)$, и теорема доказана.

ЛИТЕРАТУРА

1. R. Arratia, S. Tavaré *Independent process approximation for random combinatorial structures*. — *Advances in Math.* (to appear).
2. В. Н. Сачков *Вероятностные методы в комбинаторном анализе*. М., Наука, 1978.
3. B. Fristedt *The structure of random partitions of large integers*. — *Transact. Amer. Math. Soc.* **337** no. 2 (1993).
4. B. Fristedt *The structure of random partitions of large sets*. — *University of Minnesota Mathematics report* (May 1987), 86–154.
5. М. В. Федорюк *Метод перевала*. М., Наука, 1977.
6. А. М. Вершик *Статистика на разбиениях натуральных чисел и предельные формы случайных диаграмм*.
7. Д. Александровский *Асимптотика статистики Белла на разбиениях*. Дипломная работа, ЛГУ, мат.-мех. ф-т, 1984.
8. В. Феллер *Введение в теорию вероятностей и ее приложения*, т. 2. "Мир", 1967.
9. А. М. Вершик, А. А. Шмидт *Предельные меры, возникающие в асимптотической теории симметрических групп*. — *Теория вероятностей и ее применения*. Ч. 1 (1977) вып. 1; Ч. 2 (1978) вып. 2.
10. А. М. Вершик, С. В. Керов *Асимптотика меры Планшереля симметрической группы и предельная форма диаграмм Юнга*. — *ДАН СССР* (1977).
11. А. М. Вершик *Предельная форма выпуклых целочисленных многоугольников и близкие вопросы*. — *Функц. анализ и его прил.* **28** вып. 1 (1994), 16–25.
12. Я. Г. Синай *Вероятностный подход к анализу статистик выпуклых ломаных*. — *Функц. анализ и его прил.* **28** Вып. 2 (1994), 41–48.
13. I. Bárány *The limit shape theorem for convex lattice polygons* (to appear).

Yakubovich Yu. V. Asymptotics of random partitions of a set.

The paper contains two results on asymptotic behavior of uniform probability measure on partitions of a finite set as its cardinality tends to

infinity. The first one states that there exists a normalization of the corresponding Young diagrams such that the induced measure has a weak limit. This limit is shown to be δ -measure supported by the unit square (Theorem 1). It implies that the majority of partition blocks have approximately the same length. Theorem 2 clarifies the limit distribution of these blocks.

The techniques used can also be useful for deriving a range of analogous results.

С.-Петербургский государственный
университет

Поступило 15 января 1995 г.