



Федеральное государственное бюджетное
учреждение науки Федеральный исследо-
вательский центр «Карельский научный
центр Российской академии наук»

**XI Международная
Петрозаводская конференция**

**Вероятностные методы
в дискретной математике,**

**посвященная 90-летию
со дня рождения В. Ф. Колчина**

27 — 31 мая 2024
г. Петрозаводск, Россия

ТЕЗИСЫ ДОКЛАДОВ

Петрозаводск 2024



Karelian Research Centre
of the Russian Academy of Sciences

**XI International
Petrozavodsk Conference**

**Probabilistic Methods
in Discrete Mathematics**

**dedicated to the 90th anniversary
of V. F. Kolchin**

May 27 — 31, 2024
Petrozavodsk, Russia

ABSTRACTS

Petrozavodsk 2024

ОРГАНИЗАТОРЫ

Математический институт им. В. А. Стеклова Российской академии наук

Математический центр мирового уровня “Математический институт им. В. А. Стеклова Российской академии наук” (МЦМУ МИАН)

Федеральное государственное бюджетное учреждение науки
Федеральный исследовательский центр
“Карельский научный центр Российской академии наук”

Московский физико-технический институт



Мероприятие проводится при финансовой поддержке
Минобрнауки России (грант на создание и развитие
МЦМУ МИАН, соглашение № 75-15-2022-265)
и Московского физико-технического института

Программный комитет

А. М. Зубков (сопредседатель) (Россия)
В. В. Мазалов (сопредседатель) (Россия)
А. М. Райгородский (сопредседатель) (Россия)
В. А. Ватутин (Россия)
А. Krishnamoorthy (Индия)
В. И. Афанасьев (Россия)
Л. А. Петросян (Россия)
Ю. С. Харин (Беларусь)
А. Н. Тихомиров (Россия)
Ю. Л. Павлов (Россия)
А. Н. Реттиева (Россия)
Е. В. Морозов (Россия)
А. С. Румянцев (Россия)

Организационный комитет

В. В. Мазалов (председатель)
А. Н. Реттиева (зам. председателя)
Ю. Л. Павлов (зам. председателя)
И. А. Чеплюкова (ученый секретарь)
В. А. Ватутин
А. С. Румянцев
О. В. Лукашенко
М. М. Лери
Е. В. Хворостянская
Ю. В. Чиркова
А. В. Чирков
А. А. Бедорев

Тематика конференции

Вероятностные и статистические задачи
дискретной математики

Теория случайных графов и гиперграфов

Теория ветвящихся процессов

Прикладная статистика

Теория игр и задачи стохастической оптимизации

Статистическое и имитационное моделирование

Издается в авторской редакции

Содержание

Abebe B.

Elementary probability models for text homogeneity and change point detection..... 11

Абсалямова Н. Д.

Локальные теоремы для возмущенных случайных блужданий..... 14

Алексеев А. К.

Об алгоритмической случайности погрешностей усечения и аппроксимации..... 16

Анохина М. А.

Предельная теорема для момента максимума устойчивого случайного блуждания, достигающего фиксир. уровня... 19

Астафьев С. Н., Морозов Е. В., Румянцев А. С.

Регенеративное оценивание интенсивности выходного потока насыщенной системы для анализа стационарности.. 21

Афанасьев В. И.

Ветвящийся процесс в случайной среде, начинающийся с большого числа частиц..... 23

Бекарева Д. С., Румянцев А. С.

Исследование времени пребывания в многосерверной системе со временами обслуживания фазового типа..... 26

Буитраго Оропеса Х. К.

Максимальные индуцированные деревья в разреженных случайных графах..... 28

Vatutin V. A., Dyakonova E. E., Dong C.

Branching process evolving in extremal random environment 30

Вахрушев С. В., Жуковский М. Е.

Максимальное число расширений в случайном графе.... 33

Волошко В. А.

Локальная информационная геометрия двоичных цепей Маркова высокого порядка..... 36

<i>Головин А. С., Румянцев А. С.</i>	
Расчет стационарного распределения времени обслуживания и длительности периода занятости в односерверной системе.....	39
<i>Демин Д. А., Жуковский М. Е.</i>	
Сложность конечных случайных структур.....	41
<i>Захаров Г. Е.</i>	
Majority dynamics на конечных деревьях.....	43
<i>Зубков А. М.</i>	
Теоремы об итерациях случайных отображений конечных множеств.....	45
<i>Иванов А. В.</i>	
О промежуточных значениях размерностей квантования вероятностных мер.....	46
<i>Ивашко А. А.</i>	
Стратегии оптимальной остановки в игре на разорение ..	48
<i>Итотов А. А.</i>	
An explicit form of the Kolmogorov constant in the theory of continuous-time Markov Stochastic Branching Systems ..	49
<i>Иноземцев Э. Л.</i>	
Нетривиальные и предписанные раскраски кнезеровского графа	52
<i>Колчин А. В.</i>	
Обобщенная схема размещения: некоторые аспекты развития и применения	54
<i>Костина Е. А., Купавский А. Б.</i>	
Рамсеевские множества в пространстве Минковского при радиальных раскрасках.....	57
<i>Кошелев М. М.</i>	
Модулярность в модели посаженного разбиения	59

<i>Курина Г. А., Зиновьева М. Ю.</i>	
Построение нормального распределения по статистическим данным о COVID-19	62
<i>Маслова А. О., Лукашенко О. В.</i>	
Методы Монте-Карло по схеме марковских цепей для обучения факторизационных машин	65
<i>Никитина Н. Н.</i>	
Теоретико-игровые подходы к анализу загруженности городских дорог на примере г. Петрозаводск	67
<i>Pavlov Yu. L.</i>	
Sizes of trees in a Galton-Watson forest in the critical case with infinite variance	70
<i>Перцев Н. В.</i>	
Статистическое моделирование динамики популяций в некоторых задачах иммунологии	72
<i>Печников А. А.</i>	
О некоторых результатах исследования графа цитирования статей российских математических журналов.....	75
<i>Пешкова И. В., Алексеев А. С.</i>	
О моделировании односерверных систем обслуживания с устареванием информации	78
<i>Пешкова И. В., Морозов Е. В., Пагано М.</i>	
Регенеративное моделирование систем обслуживания с суперпозицией входных процессов восстановления	81
<i>Резлер А. В.</i>	
Метод пробных функций для нахождения оценок скорости сходимости однородных цепей Маркова к инвариантному распределению	84
<i>Rettieva A. N.</i>	
Multicriteria dynamic games with asymmetric horizons	85
<i>Рогов А. А., Москвин Н. Д., Воронов Р. В.</i>	
Анализ статистики расстояний на основе максимального общего подграфа для деревьев с упорядоч. вершинами ..	87

<i>Савелов М. П.</i>	
Пределные совместные распределения статистик критериев пакета NIST и их обобщений	90
<i>Саенко В. В.</i>	
Оценка остаточного члена в разложениях строго-устойчивого закона	92
<i>Серкова О. И.</i>	
Нахождение остовных подграфов в случайных графах...	95
<i>Серов А. А.</i>	
Двусторонние оценки квантилей биномиального распределения	98
<i>Соколов Г. М.</i>	
Хроматическое число плоскости в условиях раскраски жордановыми множествами и многоугольниками	101
<i>Sopin I. M.</i>	
Blotto game with testing (The Locks, Bombs and Testing model).....	103
<i>Степанов А. В., Чуновкина А. Г.</i>	
Об оценке параметров асимметричного TSP распределения и его применении.....	106
<i>Тихомиров А. Н.</i>	
Метод Стейна и характеристические функции	109
<i>Fazekas I., Chuprunov A. N.</i>	
Functional limit theorems for purtical numbers of cells in a general allocation scheme	112
<i>Файзуллаев Ш. Ш.</i>	
Пределный процесс для последовательных количеств слов, встретившихся в тексте ровно один раз.....	115
<i>Федотов И. А.</i>	
Вероятностная оценка ширины дерева поиска в задаче изоморфного вхождения подграфа	118

<i>Харин Ю. С.</i>	
Критерий отношения правдоподобия проверки сложных гипотез s -мерной равномерности двоичных последовательностей	122
<i>Харламов В. В.</i>	
Переходные явления в ветвящихся процессах в случайной среде: докритический случай	125
<i>Хворостянская Е. В.</i>	
О числе деревьев заданного объема леса Гальтона – Ватсона с ограниченным числом вершин	128
<i>Ходякова М. А.</i>	
Большие отклонения для ветвящихся процессов в случайной среде с замораживанием	130
<i>Чеплюкова И. А.</i>	
Предельные распределения числа вершин заданной степени условного конфигурационного графа	133
<i>Чиняев Б. Б.</i>	
Об асимптотике вероятности связности графов Эрдша-Реньи и неоднородных случайных блужданиях	135
<i>Шабанов Д. А.</i>	
Дробные раскраски случайных гиперграфов	137
<i>Шкляев А. В., Сенько П. И.</i>	
Асимптотика вероятности невырождения критической рекуррентной последовательности	139
<i>Якымив А. Л.</i>	
Асимптотика числа A -отображений с остаточным членом	142

Elementary probability models for text homogeneity and change point detection

Berhane Abebe

Novosibirsk State University, Novosibirsk, Russia

E-mail: b.andemikael@ng.nsu.ru

Numerous studies have been conducted to explore probabilistic models for creating statistical assessments of natural languages, utilizing Zipf's law and its derivatives. In a separate scheme that is constructed similarly to the infinite urn models, words have been chosen from an infinite lexicon. In order to match Zipf's distribution to the word frequency in a text, [1] suggested a straightforward probabilistic model. Likewise, [2] has investigated the statistical test relationship between Zipf's and Heap's rules.

Furthermore, the detection of text change points has gained significant importance in the domains of statistical language models and natural language programming. Changes in word choices, topic, tone, and other elements are all part of the process of identifying change points in text. A variety of techniques for analyzing text change point detection have recently been developed. To mention some [3] used text-tiling method, a new statistical approach of text partitioning by [4] and [5] proposed a probabilistic algorithm using clustering.

A recent work on a text homogeneity test based on urn models using forward and backward counting processes of the number of different words was suggested by [6]. It was extended to text-change point detection. Both models were applied to access text homogeneity and change point detection in the Bible written in Tigrigna, Amharic, and English.

Model Description

Words are chosen from a countably infinite dictionary one by one independent of each other and are numbered 1, 2, Let X_i be the number of the word at i th position, where $1 \leq i \leq n$,

$$\mathbf{P}(X_i = j) = p_j > 0, \quad j \geq 1,$$

$$p_1 + p_2 + \dots = 1,$$

$$p_1 \geq p_2 \geq \dots$$

- n is the number of words in all the concatenated text. We assume $n \rightarrow \infty$ in our theoretical results.
- R_k is the number of different words among k first words of the text.
- R'_k is the number of different words among k last words of the text.

Results

Theorem 1. (for joint distribution)

$$(Z_n, Z'_n) = \{(Z_n(t), Z'_n(t)), 0 \leq t \leq 1\}.$$

If the regularity conditions hold, then it converges weakly in the uniform metrics in $D(0, 1)^2$ to 2-dimensional Gaussian process (Z, Z') with zero expectation and covariance function:

$$\mathbf{E}Z(s)Z(t) = \mathbf{E}Z'(s)Z'(t) = K(s, t), \quad \mathbf{E}Z(s)Z'(t) = K'(s, t),$$

$$K'(s, t) = ((s + t)^\theta - 1)\mathbf{1}(s + t > 1).$$

Corollary.

$$J_n = \frac{\sum_{k=1}^n (R_k - R'_k)}{n\sqrt{R_n}}$$

converges weakly to a centered normal random variable with variance

$$\frac{\theta}{\theta + 2}.$$

To test the homogeneity of a text using any consistent estimate θ^* of parameter θ is given by the p -value is calculated J_{obs} of J_n :

$$p\text{-value} = 2\bar{\Phi} \left(|J_{obs}| \sqrt{1 + 2/\theta^*} \right).$$

Let T is parameter of change point for scaled text and T_n be its consistent estimate. The consistence is impossible without some stochastic assumptions. Hence, let designate

$$\mathbf{P}(X_1^{(1)} = j) = p_j^{(1)}, \quad \mathbf{P}(X_1^{(2)} = j) = p_j^{(2)}.$$

We assume $\{p_j^{(1)}\}_{j \geq 1}$ and $\{p_j^{(2)}\}_{j \geq 1}$ to be descending sequences of positive numbers. We consider two symmetric (and incompatible) assumptions:

$$p_j^{(1)} = j^{-1/\theta} l(j), \quad p_j^{(2)}/p_j^{(1)} \rightarrow 0, \quad (1)$$

$$p_j^{(2)} = j^{-1/\theta} l(j), \quad p_j^{(1)}/p_j^{(2)} \rightarrow 0, \quad (2)$$

$0 < \theta < 1$, l is a slowly varying function.

Theorem 2. *Let*

$$g(t) = \begin{cases} 0, & t < 0, \\ (t/T)^\theta, & 0 \leq t \leq T, \\ 1, & t > T. \end{cases}$$

If (1) holds then for any $t \in [0, 1]$

$$\widehat{R}_n(t) \rightarrow g(t), \quad \widehat{R}'_n(t) \rightarrow g(t - 1 + T) \quad a.s.$$

If (2) holds then for any $t \in [0, 1]$

$$\widehat{R}_n(t) \rightarrow g(t - 1 + T), \quad \widehat{R}'_n(t) \rightarrow g(t) \quad a.s.$$

The work is supported by Mathematical Center in Akademgorodok under agreement no. 075-15-2022-282 with the Ministry of Science and Higher Education of the Russian Federation.

This is a joint work with M. Chebunin and A. Kovalevskii.

References

- [1] *Chakrabarty A., Chebunin M., Kovalevskii A., Pupyshv I., Zakrevskaya N., Zhou Q.* A statistical test for correspondence of texts to the Zipf-Mandelbrot law // *Siberian Electronic Mathematical Reports*. 2020. 17:1959–1974. doi 10.33048/semi.2020.17.13
- [2] *Chebunin M., Kovalevskii A.* Functional central limit theorems for certain statistics in an infinite urn scheme // *Statistics and Probability Letters*. 2016. 119:344–348. doi.org/10.1016/j.spl.2016.08.1099
- [3] *Hearst M. A.* TextTiling: Segmenting text into multi-paragraph subtopic passages // *Computational linguistics*. 1997. 23(1):33–64.
- [4] *Beeferman D., Berger A., Lafferty J. D.* Statistical models for text segmentation. // *Machine learning*. 1999. 34(1-3):177–210.
- [5] *Choi F. Y.Y.* Advances in domain independent linear text segmentation. 2000. arXiv preprint cs/0003083.
- [6] *Abebe B., Chebunin M., Kovalevskii A., Zakrevskaya N.* Statistical tests for text homogeneity: using forward and backward processes of numbers of different words // *Glottometrics*. 2022. 53(1):42–58.

Локальные теоремы для возмущенных случайных блужданий

Наталья Д. Абсалямова

Московский государственный университет им. М. В. Ломоносова,
Москва, Россия

E-mail: natalia.absaliamova@gmail.com

Рассмотрим случайное блуждание $\{S_n\}$, устроенное следующим образом.

- В каждую целочисленную точку числовой прямой положим одно печенье. Через A_i обозначим событие, что в точке с координатой i есть печенье; через B_i обозначим событие, что в точке с координатой i нет печенья.
- В начальный момент времени частица находится в положении $S_0 = 0$.
- Далее в зависимости от наличия печенья в точке происходит следующее:
 - если в точке есть печенье, то частица съедает его и движется направо с вероятностью \tilde{p} и налево с вероятностью $1 - \tilde{p} = \tilde{q}$:
 $\mathbf{P}(S_{n+1} = i + 1 | S_n = i, A_i) = \tilde{p}$, $\mathbf{P}(S_{n+1} = i - 1 | S_n = i, A_i) = 1 - \tilde{p} = \tilde{q}$;
 - если в точке нет печенья, то частица движется направо с вероятностью p и налево с вероятностью $1 - p = q$:
 $\mathbf{P}(S_{n+1} = i + 1 | S_n = i, B_i) = p$, $\mathbf{P}(S_{n+1} = i - 1 | S_n = i, B_i) = 1 - p = q$.

Случайное блуждание $\{S_n\}$, обладающее тремя свойствами, описанными выше, называется возмущенным случайным блужданием. Модель была впервые рассмотрена в работе [1] Бенджамина и Вилсона, в которой изучался случай многомерного блуждания на \mathbb{Z}^d , $d \geq 1$ с несколькими одинаковыми печеньями в каждой точке. В работе [2] была рассмотрена связь данного блуждания с ветвящимися процессами.

Мы рассматриваем случай несимметричного случайного блуждания ($p > 1/2$) с одним печеньем, усиливающим притяжение направо ($\tilde{p} > p$).

В такой ситуации частица почти наверное пройдет через некоторые целочисленные точки один раз за всю историю, что будет являться моментами регенерации для выбранного процесса. В работе используется метод, предложенный в [3] для процессов, обладающих свойством регенерации, учитывая модификации, описанные в [4]. С его помощью получена теорема следующего вида.

Теорема.

$$P(S_n = k) \sim \frac{C(k/n)}{\sqrt{2\pi n}} \exp\left(-L\left(\frac{k}{n}\right)n\right), \quad n \rightarrow \infty,$$

где $C(k/n)$ зависит от k/n , $k/n \in [\theta_1, \theta_2] \subset (0, 1)$.

Для функции $L(\theta)$ найдена явная формула.

Список литературы

- [1] *Benjamini I., Wilson D.* Excited random walk // *Electron. Commun. Probab.* 2003. 8:86–92.
- [2] *Kosygina E., Zerner M.* Positively and negatively excited random walks on integers, with branching processes // *Electron. J. Probab.* 2008. 13:1952–1979.
- [3] *Бакай Г. А.* О характеристике вероятностей больших уклонений для регенерирующих последовательностей // *Труды Математического института имени В. А. Стеклова.* 2022. 316:47–63.
- [4] *Бакай Г. А.* О больших уклонениях момента достижения далекого уровня случайным блужданием в случайной среде // *Дискретная математика.* 2022. 34(4):3–13.

Об алгоритмической случайности погрешностей усечения и аппроксимации

Алексей К. Алексеев

Московский физико-технический институт, Долгопрудный, Россия

E-mail: aleksey.k.alekseev@gmail.com

При аппроксимации нестационарных систем ЧДУ ошибки дискретизации сеточной функции определяются формальным рядом (аппроксимирующим операторы сдвига) $u_{i+1}^n = u_i^n + \sum_{j=1}^{\infty} \frac{h^j}{j!} \frac{\partial^j u(t^n, x_i)}{\partial x^j}$ [1] (n - номер шага по времени, i -по координате). Этот ряд определяет ошибки усечения δu , которые в свою очередь определяют ошибки аппроксимации Δu . Не смотря на то, что этот ряд де-факто сходится, на практике, как правило, используется только его младший член (в простейшем случае для конечных разностей $(u_{i+1}^n - u_i^n)/h$ он имеет вид $\delta u_i^n \approx h \frac{\partial^2 u(t^n, x_i)}{\partial x^2}$). В некоторых простейших случаях оценка ошибки усечения получается по аналитическим формулам [2], но чаще с помощью действия специального постпроцессора [3]. Он используется для оценки погрешности ценных функционалов [3, 4] либо для построения решения повышенной точности с помощью коррекции погрешности [5]. Тем не менее, наличие этого ряда в ошибках усечения (а в дальнейшем его суммы в ошибках аппроксимации) вызывает серьезное беспокойство как с точки зрения существования и возможности определения высших производных, так и с точки зрения наличия бесконечного числа членов.

В качестве альтернативы естественно использовать разложение Тейлора в форме с Лагранжевым остатком $u_{i+1}^n = u_i^n + h \frac{\partial u(t^n, x_i + \beta_i^n h)}{\partial x}$. Этот вариант позволяет уйти от бесконечного ряда ценой введения неизвестного параметра $\beta_i^n \in [0, 1]$. Его свойства и предоставляемые им возможности обсуждаются в данном докладе. Сравнив $u_{i+1}^n = u_i^n + h \frac{\partial u(t^n, x_i + \beta_i^n h)}{\partial x}$ и $u_{i+1}^n = u_i^n + \sum_{j=1}^{\infty} \frac{h^j}{j!} \frac{\partial^j u(t^n, x_i)}{\partial x^j}$ получим $\frac{\partial u(t^n, x_i + \beta_i^n h)}{\partial x} = \sum_{j=1}^{\infty} \frac{h^{j-1}}{j!} \frac{\partial^j u(t^n, x_i)}{\partial x^j}$. В первом порядке β_i^n можно выразить как $\beta_i^n \approx \sum_{j=2}^{\infty} \frac{h^{j-2}}{j!} \frac{\partial^j u(t^n, x_i)}{\partial x^j} \cdot \left(\frac{\partial^2 u(t^n, x_i)}{\partial x^2} \right)^{-1}$.

“Программа”, позволяющая рассчитывать β_i^n явно длиннее самой переменной (за исключением тривиального случая нулевых производных, не реализуемого в практически интересных нелинейных случаях, таких

как течения с разрывами) и поэтому кажется естественным, что Лагранжев коэффициент β_i^n может считаться алгоритмически случайной величиной (по Колмогорову, [6, 7]). Погрешность усечения используется при реализации метода коррекции погрешности [5] $(u_i^n)^{m+1} = (u_i^n)^m + \beta_i^{n,m} h^k (L_{Low}^h)^{-1} \partial_x^{k+1} (u_i^n)^m$ ($\beta_i^{n,m} \in [0, 1]$) в форме итераций (m -номер итерации). Последовательный выбор случайной $\beta_i^{n,m}$ объясняет наблюдаемую на первых итерациях метода немонотонность сходимости. Набор итераций позволяет найти ансамбль $\beta_i^{n,m}$ и изучить свойства Лагранжевых коэффициентов. Если предположить, что алгоритмическая случайность погрешности дискретизации $\Delta u^{(k)} \in R^N$ обеспечивает произвольность выбора, то принадлежность погрешности к пространствам очень большой размерности (что естественно для численных решений в двумерном и трехмерном случаях) дает нестандартные возможности для определения нормы погрешности и положения истинного решения. Известно, что в пространствах достаточно большой размерности N расстояние $d_{1,2}$ между двумя произвольно выбранными векторами $v^{(1)} \in R^N$ и $v^{(2)} \in R^N$ “с вероятностью 1” больше длины этих векторов $\|v^{(i)}\| \leq d_{1,2}$. Это вызвано тем обстоятельством, что хорда в этих пространствах “почти всегда больше радиуса” [8], а два произвольно выбранных единичных вектора с большой вероятностью ортогональны [9], а именно $P\{(u^{(1)}, u^{(2)}) > \delta\} < \sqrt{\pi/2} e^{-\delta^2 N/2}$. В итоге, для двух численных решений, полученных независимыми методами, расстояние между ними “с вероятностью 1” определяет верхнюю грань нормы погрешности дискретизации (расчета) $u^{(1)} - u^{(2)} = u^{(1)} - \tilde{u} - u^{(2)} + \tilde{u} = \Delta u^{(1)} - \Delta u^{(2)}$, $d(u^{(i)}, \tilde{u}) = \|\tilde{u} - u^{(i)}\| \leq d_{1,2} = \|\Delta u^{(1)} - \Delta u^{(2)}\|$. К сожалению численные расчеты [10] показывают, что неравенство $\|\tilde{u} - u^{(i)}\| \leq \|\Delta u^{(1)} - \Delta u^{(2)}\|$ выполняется не всегда и для надежной оценки погрешности необходима дополнительная информация (в простейшем случае увеличение ансамбля анализируемых решений, полученных независимыми алгоритмами). Алгоритмически случайный выбор в большом пространстве сеточных функций должен обеспечивать ортогональность погрешности аппроксимации, чего не наблюдается. Тем не менее в расчетах наблюдается наличие существенных по величине ортогональных компонент в погрешностях аппроксимации для независимых алгоритмов, что дает возможность апостериорной оценки погрешности [10]. В работе представлены результаты численных расчетов иллюстрирующие свойства Лагранжева коэффициента и его возможности при апостериорной оценке погрешности аппроксимации.

Список литературы

- [1] *Шокин Ю. И., Яненко Н. Н.* Метод дифференциального приближения. Применение к газовой динамике, Наука, 1985. 209 с.
- [2] *Alekseev A. K., Navon I. M.* A posteriori pointwise error estimation for compressible fluid flows using adjoint parameters and Lagrange remainder // *Int. J. Numer. Meth. Fluids.* 2005. 47:45–74.
- [3] *Alekseev A. K., Navon I. M.* A Posteriori Error Estimation by Postprocessor Independent of Flowfield Calculation Method // *Computers & Mathematics with Applications.* 2006. 51:397–404.
- [4] *Алексеев А. К., Махнев И. Н.* Использование Лагранжевых коэффициентов при апостериорной оценке погрешности расчета // *Сибирский журнал вычислительной математики.* 2009. 12(4):375–388.
- [5] *Diskin B., Thomas J. L.* Solving Upwind-biased Discretizations: Defect-correction Iterations // *ICASE Report No. 99-14.*
- [6] *Колмогоров А. Н.* Три подхода к определению понятия “количество информации” // *Проблемы передачи информации.* 1965. 1(1):3–11.
- [7] *Martin-Lof P.* Algorithms and Randomness // *Review of International Statistical Inst.* 1969. 37(3).
- [8] *Зорич В. А.* Многомерная геометрия, функции очень многих переменных и вероятность // *ТВИ.* 2014. 59(3):436–451.
- [9] *Sidiropoulos P.* N-sphere chord length distribution // *arXiv:1411.5639v1.* 2014.
- [10] *Alekseev A. K., Bondarev A. E.* On a posteriori error estimation using distances between numerical solutions and angles between truncation errors // *Mathematics and Computers in Simulation.* 2012. 190:892–904.

Предельная теорема для момента максимума устойчивого случайного блуждания, достигающего фиксированного уровня

Мария А. Анохина

Московский государственный университет им. М. В. Ломоносова,
Москва, Россия

E-mail: anokhina.mary1@gmail.com

Рассмотрим случайное блуждание $S_n = X_1 + \dots + X_n$, $n \in \mathbb{N}$, $S_0 = 0$, где X_1, X_2, \dots – н. о. р. (независимые одинаково распределенные) случайные величины (сл. в.). Пусть $\mathcal{A} := \{0 < \alpha < 1; |\beta| < 1\} \cup \{1 < \alpha < 2; |\beta| \leq 1\} \cup \{\alpha = 1, \beta = 0\} \cup \{\alpha = 2, \beta = 0\} \subset \mathbb{R}^2$. Для $(\alpha, \beta) \in \mathcal{A}$ и сл. в. X будем говорить, что $X \in \mathcal{D}(\alpha, \beta)$, если распределение X принадлежит области притяжения устойчивого закона с характеристической функцией

$$G_{\alpha, \beta}(t) := \exp \left\{ -c|t|^\alpha \left(1 - i\beta \frac{t}{|t|} \tan \frac{\pi\alpha}{2} \right) \right\}, \quad c > 0,$$

и, к тому же, $\mathbf{E}X = 0$, если оно существует.

Если существует $\rho \in (0, 1)$ такое, что $\mathbf{P}(S_n > 0) \rightarrow \rho$, $n \rightarrow \infty$, то известен обобщенный закон арксинуса для момента достижения максимума блужданием S_n

$$\mathbf{P} \left(\frac{\tau_M}{n} \leq x \right) \rightarrow \frac{\sin(\pi\rho)}{\pi} \int_0^x t^{-\rho} (1-t)^{\rho-1} dt, \quad n \rightarrow \infty, \quad x \in [0, 1].$$

где τ_M – момент первого достижения максимума блужданием.

Нас интересует исследование асимптотического поведения при $n \rightarrow \infty$ вероятности

$$\mathbf{P} \left(\frac{\tau_M}{n} \leq x \mid M_n \in [x_n, x_n + \Delta) \right), \quad x \in [0, 1],$$

где $x_n \in \mathbb{R}^+$, $\Delta > 0$. Исследование асимптотического поведения такой вероятности связано с исследованием вероятности $\mathbf{P}(S_i > 0, i \leq n \mid S_n \in [x_n, x_n + \Delta))$. В зоне нормальных уклонений, то есть когда $x_n \approx \sqrt{n}$, $n \rightarrow \infty$, в арифметическом случае асимптотическое поведение интересующей нас вероятности можно получить напрямую, а в нерешетчатом случае уже требуются некоторые дополнительные оценки, полученные в работах [1] и [2]. Сначала сформулируем более общий результат.

© М. А. Анохина, 2024

Теорема 1. Пусть X_i , $i = 1, \dots, n$, – н. о. р. нерешетчатые сл. в. и $X_1 \in \mathcal{D}(\alpha, \beta)$. Пусть $k \in (0, \infty) \cap \mathbb{Z}$, $L(n)$ – медленно меняющаяся на бесконечности функция, $c_n = n^{1/\alpha} L(n)$, $\{x_n\}$ – некоторая последовательность: $x_n/c_n \rightarrow y > 0$, $n \rightarrow \infty$, $0 < a < b < 1$, $\Delta > 0$. Тогда при $n \rightarrow \infty$

$$\mathbf{P} \left(\frac{\tau_M}{n} \in [a, b] \mid M_n \in [x_n, x_n + \Delta) \right) \rightarrow \frac{\int_0^b p_{\alpha, \beta} \left(yt^{-1/\alpha} \right) t^{\rho-1-1/\alpha} (1-t)^{-\rho} dt}{\int_0^1 p_{\alpha, \beta} \left(yt^{-1/\alpha} \right) t^{\rho-1-1/\alpha} (1-t)^{-\rho} dt},$$

где $p_{\alpha, \beta}(x)$ – плотность соответствующего меандра Леви, а

$$\rho = \begin{cases} 1/2, & \alpha = 1; \\ 1/2 + \arctan(\beta \tan(\pi\alpha/2))/(\pi\alpha), & \text{иначе.} \end{cases}$$

В частности, при $\alpha = 2$ и $\beta = 0$ получаем предельную теорему в случае нормального закона.

Теорема 2. Пусть X_i , $i = 1, \dots, n$, – н. о. р. нерешетчатые сл. в. с $\mathbf{E}X_1 = 0$ и $\mathbf{E}X_1^2 = \sigma^2 \in (0, \infty)$, $\{x_n\}$ – некоторая последовательность: $x_n/\sqrt{n} \rightarrow y > 0$, $n \rightarrow \infty$, $0 < a < b < 1$, $\Delta > 0$. Тогда при $n \rightarrow \infty$

$$\mathbf{P} \left(\frac{\tau_M}{n} \in [a, b] \mid M_n \in [x_n, x_n + \Delta) \right) \rightarrow \frac{y}{\sqrt{2\pi}\sigma} \int_a^b \frac{e^{-\frac{y^2}{2\sigma^2} \left(\frac{1}{t}-1\right)}}{t\sqrt{t(1-t)}} dt.$$

Аналогичные результаты верны и в арифметическом случае для $\mathbf{P}(\tau_M/n \in [a, b] \mid M_n = k_n)$, где $k_n \in \mathbb{N}$.

Список литературы

- [1] *Vatutin V. A., Wachtel V.* Local probabilities for random walks conditioned to stay positive // *Probability Theory Related Fields*. 2009. 143:177–217.
- [2] *Wachtel V.* Local limit theorem for the maximum of asymptotically stable random walks // *Probability Theory Related Fields*. 2012. 152:407–424.

Регенеративное оценивание интенсивности выходного потока насыщенной системы для анализа стационарности

¹Сергей Н. Астафьев, ²Евсей В. Морозов,
²Александр С. Румянцев

¹Институт прикладных математических исследований КарНЦ РАН,
Петрозаводск, Россия

²Новгородский государственный университет им. Ярослава Мудрого,
Великий Новгород, Россия

E-mail: seryu@mail.ru; emorozov@karelia.ru;
ar0@krc.karelia.ru

Первым шагом в анализе систем обслуживания, как правило, является установление условий стационарности, при которых распределения исследуемых характеристик сходятся к стационарным. Эта задача особенно сложна в условиях, когда исследуемые случайные процессы являются немарковскими. В этом случае, как правило, требуются специальные свойства процесса, например, наличие регенераций [1], либо стохастическая монотонность по параметру входного потока и т. п. Кроме того, исследование стационарности часто удобно вести с помощью вспомогательной *насыщенной* системы, интенсивность выходного потока в которой может быть использована как верхняя граница интенсивности входного потока сохраняющая стационарность исходной системы [2].

В то же время, модели многих современных телекоммуникационных систем являются слишком сложными для аналитического исследования, включая анализ стационарности, и в этом случае необходим численный анализ, который опирается на результаты имитационного моделирования [3]. К полученным на основе результатов моделирования оценкам необходимо применять методы статистической проверки стационарности. Такие методы разработаны, в частности, для широкого класса марковских моделей стохастических систем обслуживания в [4].

В данной работе предложен подход к построению доверительного интервала для максимальной интенсивности входного потока, гарантирующей стационарность немарковской системы обслуживания. Подход основан на регенеративном оценивании интенсивности выходного потока

насыщенной системы по моментам так называемой искусственной регенерации. Такие моменты могут быть построены, если распределение времени обслуживания обладает так называемым тяжелым хвостом (например, распределение Парето). В этом случае для получения моментов регенерации применяется экспоненциальное расщепление, а затем используется регенеративная версия центральной предельной теоремы для построения интервальной оценки требуемой характеристики системы. В данной работе на примере модели системы с многосерверными клиентами (модели вычислительного кластера) дано сравнение данного подхода с методом, предложенным в работе [4]. Показано, что предложенный метод имеет ряд преимуществ при нахождении условий стационарности данной модели.

Работа поддержана РНФ, проект 21-71-10135.

Список литературы

- [1] *Morozov E., Steyaert B.* Stability Analysis of Regenerative Queueing Models Mathematical Methods and Applications. Springer, Cham, 2021. 185 p.
- [2] *Baccelli F., Foss S.* On the saturation rule for the stability of queues // Journal of Applied Probability. 1995. 2(32):494–507.
- [3] *Asmussen S., Glynn P.* Stochastic simulation: algorithms and analysis. New York: Springer, 2007. 476 p.
- [4] *Mandjes M., Patch B., Walton N.* Detecting Markov Chain Instability: A Monte Carlo Approach. 2016. doi: 10.48550/arXiv.1608.03257

Ветвящийся процесс в случайной среде, начинающийся с большого числа частиц

Валерий И. Афанасьев

Математический институт им. В. А. Стеклова РАН, Москва, Россия

E-mail: viafan@mi-ras.ru

Пусть $(\Omega, \mathcal{F}, \mathbf{P})$ – исходное вероятностное пространство и Δ – пространство вероятностных мер на $\mathbf{N}_0 := \{0, 1, \dots\}$ с метрикой полной вариации. Рассмотрим случайные элементы Q_1, Q_2, \dots , отображающие $(\Omega, \mathcal{F}, \mathbf{P})$ в Δ . Это означает, что Q_n при каждом $n \in \mathbf{N}$ является вероятностной мерой на \mathbf{N}_0 . Последовательность $\Pi = \{Q_1, Q_2, \dots\}$ называется *случайной средой*.

Последовательность неотрицательных целочисленных случайных величин $\{Z_n, n \in \mathbf{N}_0\}$ называется *ветвящимся процессом в случайной среде*, если $Z_0 = 1$ и

$$Z_{n+1} = \sum_{i=1}^{Z_n} \xi_i^{(n)}, \quad n \in \mathbf{N}_0,$$

где предполагается, что при фиксированной случайной среде Π случайные величины $\{\xi_i^{(n)} : n \in \mathbf{N}_0, i \in \mathbf{N}\}$ независимы, причем при фиксированном $n \in \mathbf{N}_0$ величины $\xi_1^{(n)}, \xi_2^{(n)}, \dots$ одинаково распределены с распределением Q_{n+1} . На языке ветвящихся процессов Z_n – численность частиц n -го поколения; $\xi_i^{(n)}$ – число непосредственных потомков i -ой частицы из n -го поколения.

Другими словами, случайный процесс $\{Z_n, n \in \mathbf{N}_0\}$ при фиксированной случайной среде Π является (неоднородным) ветвящимся процессом Гальтона–Ватсона; при этом для каждого $n \in \mathbf{N}$ закон размножения частиц n -го поколения есть Q_{n+1} . Сопоставим (случайному) распределению Q_n при $n \in \mathbf{N} := \{1, 2, \dots\}$ производящую функцию $f_n(\cdot)$. Предполагается, что случайные элементы Q_1, Q_2, \dots независимы и одинаково распределены.

Положим

$$X_n = \ln f_n'(1), \quad \eta_n = \frac{f_n''(1)}{(f_n'(1))^2}, \quad n \in \mathbf{N}$$

(предполагается, что п.н. $f'_1(1), f''_n(1) \in (0, +\infty)$). Заметим, что случайные векторы $(X_1, \eta_1), (X_2, \eta_2), \dots$ являются независимыми и одинаково распределенными. Введем *сопровождающее случайное блуждание*

$$S_0 = 0, \quad S_n = \sum_{i=1}^n X_i, \quad n \in \mathbf{N}.$$

В дальнейшем предполагается, что

$$\mathbf{E}X_1 = 0, \quad 0 < \mathbf{E}X_1^2 := \sigma^2 < +\infty; \quad (1)$$

кроме того,

$$\mathbf{E} \ln^{2+q} (\eta_1 \vee 1) < +\infty \quad (2)$$

при некотором $q > 0$.

Теперь в отличие от стандартной ситуации, когда процесс $\{Z_n\}$ начинается с одной частицы, т.е. $Z_0 = 1$, рассмотрим ситуацию, когда $Z_0 = m_n(x)$, где $m_n(x) \in \mathbf{N}$ и для некоторого фиксированного $x > 0$ при $n \rightarrow \infty$

$$\ln m_n(x) \sim \sigma \sqrt{n}x. \quad (3)$$

Положим при $t \geq 0$

$$Y_n(t) = \frac{Z_{\lfloor nt \rfloor}}{m_n(x) \exp S_{\lfloor nt \rfloor}}.$$

Вместо обозначений $\{Z_i \mid Z_0 = m_n(x)\}$ и $\{Y_n(t) \mid Z_0 = m_n(x)\}$ будем использовать $Z_i^{(n,x)}$ и $Y_n^{(n,x)}(t)$ соответственно. Введем момент выживания процесса $\{Z_i^{(n,x)}\}$, т.е. момент

$$T^{(n,x)} = \min \{i : Z_i^{(n,x)} = 0\}.$$

Пусть $\{W(t), t \geq 0\}$ – стандартное броуновское движение и τ_a – момент первого достижения точки $a \neq 0$ этим процессом. Для $x > 0$ введем случайный процесс $Y^{(x)} = \{Y^{(x)}(t), t \geq 0\}$: $Y^{(x)}(t) = 1$ при $t \leq \tau_{-x}$, $Y^{(x)}(t) = 0$ при $t > \tau_{-x}$. Пусть символ \xrightarrow{D} означает сходимость случайных элементов (величин, векторов, процессов) по распределению в том или ином метрическом пространстве. В некоторых случаях это пространство указывается.

Теорема 1. Пусть выполнены условия (1)-(3), тогда при $n \rightarrow \infty$

$$\frac{T^{(n,x)}}{n} \xrightarrow{D} \tau_{-x}.$$

Теорема 2. Пусть выполнены условия (1)-(3), тогда при $n \rightarrow \infty$

$$\left\{ Y_n^{(n,x)}, t \geq 0 \right\} \Rightarrow Y^{(x)},$$

где символ \Rightarrow означает сходимость в смысле конечномерных распределений.

Теорема 3. Пусть выполнены условия (1)-(3), тогда при $n \rightarrow \infty$

$$\left\{ \frac{\ln \left(Z_{[nt]}^{(n,x)} + 1 \right)}{\sigma \sqrt{n}}, t \geq 0 \right\} \xrightarrow{D} \{x + W(t \wedge \tau_{-x}), t \geq 0\}.$$

Здесь символ \xrightarrow{D} означает сходимость по распределению в пространстве $D[0, +\infty)$ с топологией Скорохода.

Отметим, что теория критических ветвящихся процессов в случайной среде, начинающихся с одной частицы и рассматриваемых при условии невырождения, изложена в работах [1] и [2]. Классические критические ветвящиеся процессы Гальтона – Ватсона, начинающиеся с большого числа частиц, рассмотрены в работе [3].

Исследование выполнено за счет гранта Российского научного фонда № 24-11-00037 (<https://rscf.ru/project/24-11-00037/>).

Список литературы

- [1] Afanasyev V. I., Geiger J., Kersting G., Vatutin V. A. Criticality for branching processes in random environment // Ann. Probab. 2005. 33(2):645–673. doi: 10.1214/009117904000000928
- [2] Kersting G., Vatutin V. Discrete time branching processes in random environment. London: John Viley and Sons, Inc., 2017. 286 p.
- [3] Lindvall T. Convergence of critical Galton – Watson branching processes // J. Appl. Probab. 1972. 9(2):445–450. doi: 10.2307/3212815

Исследование времени пребывания в многосерверной системе со временами обслуживания фазового типа

Диана С. Бекарева, Александр С. Румянцев

Институт прикладных математических исследований КарНЦ РАН,
Петрозаводск, Россия

E-mail: abcdi_do@mail.ru; ar0@krc.karelia.ru

Повсеместное использование многопроцессорных вычислительных систем, позволяющих обслуживать большие объемы запросов клиентов за приемлемое время, требует проведения анализа их качественных характеристик для обеспечения возможности оценивания и контроля качества обслуживания и надежности.

При организации многосерверных вычислительных систем широко используются технологии параллельных, распределенных и облачных вычислений, позволяющие снижать среднее время отклика. В то же время, их применение создает определенные риски, связанные как с отдельными вычислительными узлами, так и с каналами передачи данных. Для повышения надежности системы применяется дублирование заданий, называемое *репликацией*, а также согласование результатов вычислений, называемое *кворумом*. Выбор подходящих параметров репликации и кворума важен для обеспечения приемлемой надежности без существенной потери скорости работы системы.

Для исследования многопроцессорных вычислительных систем, в связи со случайным характером поступающей работы, применяются методы теории массового обслуживания [1]. С точки зрения моделей, исследуемый класс относится к многосерверным системам обслуживания с разделением-слиянием заявок [2, 3]. При этом каждая из заявок на обслуживание разделяется на r копий, при достижении результата в $q \leq r$ копий заявка считается обслуженной. Параметры r, q задают репликацию и кворум. Выбор параметров r и q влияет не только на надежность системы, но и на среднее время пребывания заявок в системе, являющееся одним из наиболее важных показателей ее работы.

Для анализа систем разделения-слияния с sr серверами, в которых заявка занимает r серверов и для завершения обслуживания требуется кворум q , могут применяться эквивалентные системы с s серверами, в

которых время обслуживания заявки распределено как q -я порядковая статистика выборки из r с.в. (соответствующих временам обслуживания). Для таких систем в общем случае удастся установить монотонность времени времени пребывания заявки в системе по параметрам r и q , а также получить явные выражения для времени пребывания в системах типа $M/M/c$ (при $q = 1$), а также если распределения имеют фазовый тип, либо тяжелый хвост (распределения Парето, Вейбулла) для односервеного случая [4].

При сравнении указанных систем была показана монотонность их характеристик по параметрам репликации и кворума, однако число физических серверов (равное cr) в исходных системах различно. В этой связи, в данном исследовании решается задача, в которой число серверов исходной системы фиксировано. Исследуется время пребывания $V_{q,r}$ в системе в зависимости от параметров репликации r и кворума q .

Работа поддержана РНФ, проект 21-71-10135.

Список литературы

- [1] *Asmussen S., Glynn P.* Stochastic simulation: algorithms and analysis. New York: Springer, 2007. 476 p.
- [2] *Fiorini P. M., Lipsky L.* Exact Analysis of Some Split-Merge Queues // ACM SIGMETRICS Performance Evaluation Review. 2015. 2(43):51–53.
- [3] *Joshi G., Soljanin E., Wornell G.* Efficient Redundancy Techniques for Latency Reduction in Cloud Systems // ACM Transactions on Modeling and Performance Evaluation of Computing Systems. 2017. 2(2):1–30.
- [4] *Rumyantsev A. S., Dolgaleva D. S., Golovin A. S.* Steady-state performance analysis of multiserver queueing models with redundancy // Program Systems: Theory and Applications. 2023. 14(1):55–94.

Максимальные индуцированные деревья в разреженных случайных графах

Хуан Карлос Буитраго Оропеса

Московский физико-технический институт, Долгопрудный, Россия

E-mail: buitrago.okh@phystech.edu

В этой работе мы обсудим некоторые характеристики биномиального случайного графа $G(n, p)$ (см. [1–6]), где $p = p(n)$ является функцией от n . Рассмотрим набор помеченных вершин $[n] := \{1, \dots, n\}$. В $G(n, p)$ каждая из $\binom{n}{2}$ пар вершин соединяется ребром независимо с вероятностью p (более формально, $G(n, p)$ – это случайный элемент, принимающий значения во множестве всех графов на $\{1, \dots, n\}$, с распределением $P(G(n, p) = H) = p^{e(H)}(1-p)^{\binom{n}{2}-e(H)}$, где $e(H)$ – количество ребер в H).

Напомним, что множество $U \subset \{1, \dots, n\}$ графа, в котором никакие две вершины не смежны, называется *независимым множеством*. Размер наибольшего независимого множества называется *числом независимости графа*. Распределение размера самого большого независимого множества в случайном графе тщательно изучено. В [7, 8] было доказано, что число независимости $G(n, p = \text{const})$ с вероятностью, стремящейся к 1, принимает одно из двух значений, $f_0(n)$ и $f_0(n) + 1$, где

$$f_0(n) = \left\lfloor 2 \log_b n - 2 \log_b \log_b n + 2 \log_b \frac{e}{2} + 0.9 \right\rfloor, \quad b = 1/(1-p).$$

Некоторые уточнения этого результата можно найти в [9]. В таких случаях говорят, что число независимости *сконцентрировано в двух точках*.

В дальнейшем концентрация в двух точках была доказана и для других характеристик случайного графа. Кроме того, вопрос ставился и при заданных структурных свойствах индуцированного подграфа. Так, в [10] изучен максимальный размер индуцированного дерева в $G(n, p = \text{const})$ и доказана концентрация в двух точках.

Теорема 1 (D. Kamaldinov, A. Skorkin, M. Zhukovskii, 2021). Пусть $f_\varepsilon(n) = \lfloor 2 \log_{1/(1-p)}(enp) + 2 + \varepsilon \rfloor$. Тогда найдется $\varepsilon > 0$, при котором с вероятностью, стремящейся к 1, максимальный размер индуцированного дерева в $G(n, p = \text{const})$ принимает одно из двух значений, $f_\varepsilon(n)$ и $f_\varepsilon(n) + 1$.

Предыдущий результат применимый к модели случайного графа, где вероятность появления ребра является постоянной величиной. В настоящей работе нам удалось доказать, что результат теоремы 1 о максимальном размере индуцированного дерева в $G(n, p)$ можно расширить для достаточно широкого диапазона $p = o(1)$.

Теорема 2. Пусть $\varepsilon > 0$, $n^{-\frac{\varepsilon-2}{3\varepsilon-2}+\varepsilon} \leq p = o(1)$. Тогда найдется такое $\delta > 0$, что с вероятностью, стремящейся к 1, максимальный размер индуцированного дерева в $G(n, p)$ принимает одно из двух значений, $g(n)$ и $g(n) + 1$, где

$$g(n) = \lfloor 2 \log_{1/(1-p)}(enp) + \delta \rfloor.$$

Список литературы

- [1] *Bollobás B.* Random Graphs. 2nd ed. Cambridge University Press; 2001.
- [2] *Janson S., Luczak T., Ruciński A.* Random Graphs. New York: Wiley; 2000.
- [3] *Жуковский М. Е., Райгородский А. М.* Случайные графы: модели и предельные характеристики // Успехи математических наук. 2015. 70(1):35–88.
- [4] *Дервянко Н. М., Киселев С. Г.* Числа независимости случайных подграфов некоторого дистанционного графа // Проблемы передачи информации. 2017. 53(4):307–318.
- [5] *Егорова А. Н., Жуковский М. Е.* Опровержение закона нуля или единицы для экзистенциальных монадических свойств разреженного биномиального случайного графа // Доклады Академии наук. 99(1):68–70.
- [6] *Ostrovsky L. B., Zhukovskii M. E.* Monadic second-order properties of very sparse random graphs // Annals of Pure and Applied Logic. 2017. 168(11):2087–2101.
- [7] *Bollobás B., Erdős P.* Cliques in random graphs // Math. Proc. Camb. Phil. Soc. 1976. 80:419–427.
- [8] *Matula D.* The largest clique size in a random graph. Tech. Rep. Dept. Comp. Sci., Southern Methodist University, Dallas, Texas; 1976.
- [9] *Krivelevich M., Sudakov B., Vu V. H., Wormald N. C.* On the probability of independent sets in random graphs // Random Structures & Algorithms. 2003. 22(1):1–14.
- [10] *Kamaldinov D., Skorkin A., Zhukovskii M.* Maximum sparse induced subgraphs of the binomial random graph with given number of edges // Discrete Mathematics. 2021. 344(2):112205. ISSN 0012-365X.

Branching process evolving in extremal random environment

Vladimir A. Vatutin¹, Elena E. Dyakonova¹, C. Dong²

¹Steklov Mathematical Institute of Russian Academy of Sciences, Moscow;

²Xidian University, Xi'an, China

E-mail: vatutin@mi-ras.ru; elena@mi-ras.ru; czdong@xidian.edu.cn

Let $F(s)$ be a random variable taking values in the space \mathcal{M} of all probability generating and let $\{F_n(s), s \in [0, 1], n = 1, 2, \dots\}$ be a sequence of independent copies of F . The infinite sequence $\mathcal{E} = \{F_n, n \in \mathbb{N}\}$ is called a random environment.

A sequence of nonnegative random variables $\mathcal{Z} = \{Z_n, n = 0, 1, \dots\}$ specified on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ is called a branching process in random environment (BPRE), if Z_0 is independent of \mathcal{E} and, given \mathcal{E} the process \mathcal{Z} is a Markov chain with

$$\mathcal{L}(Z_n | Z_{n-1} = z_{n-1}, \mathcal{E} = (f_1, f_2, \dots)) = \mathcal{L}(\xi_{n1} + \dots + \xi_{ny_{n-1}})$$

for all $n = 1, 2, \dots$, $z_{n-1} \in \{0, 1, \dots\}$ and $f_1, f_2, \dots \in \mathcal{M}$, where $\xi_{n1}, \xi_{n2}, \dots$ is a sequence of i. i. d. random variables with distribution f_n . Thus, Z_{n-1} is the $(n-1)$ th generation size of the population of the branching process and f_n is the offspring distribution of an individual at generation $n-1$.

We impose the following restrictions on the properties of the BPRE.

Let (α, β) be a pair of nonnegative numbers belonging to the union of the sets

$$\{0 < \alpha < 1; |\beta| < 1\} \cup \{1 < \alpha < 2; |\beta| < 1\} \cup \{\alpha = 1, \beta = 0\} \cup \{\alpha = 2, \beta = 0\}.$$

For a random variable X , we write $X \in \mathcal{D}(\alpha, \beta)$ if the distribution of X belongs (without centering) to the domain of attraction of a stable law with density $g(x) = g_{\alpha, \beta}(x)$, $x \in (-\infty, +\infty)$, and the characteristic function

$$G(w) = \int_{-\infty}^{+\infty} e^{iwx} g(x) dx = \exp \left\{ -c|w|^\alpha \left(1 - i\beta \frac{w}{|w|} \tan \frac{\pi\alpha}{2} \right) \right\}, \quad c > 0.$$

Consider a random walk

$$S_0 = 0, \quad S_n = X_1 + \dots + X_n, \quad n \geq 1,$$

with independent and identically distributed increments. We assume that the random walk $S = \{S_n, n = 0, 1, \dots\}$ meets the following conditions.

Hypothesis A1. *The increments $X_n, n = 1, 2, \dots$ of S belong to $D(\alpha, \beta)$ with $|\beta| < 1$.*

This means, in particular, that there is an increasing sequence of positive numbers

$$a_n := n^{1/\alpha} \ell(n), n = 1, 2, \dots, \quad (1)$$

where $\ell(1), \ell(2), \dots$ is a slowly varying sequence, such that, as $n \rightarrow \infty$

$$\mathcal{L} \left\{ \frac{S_{[nt]}}{a_n}, t \geq 0 \right\} \Longrightarrow \mathcal{L} \{Y_t, t \geq 0\}, \quad (2)$$

where the symbol \Longrightarrow stands for the convergence in distribution in the space $D[0, +\infty)$ with Skorokhod topology and the process $\mathcal{Y} = \{Y_t, t \geq 0\}$ is strictly stable with index $\alpha \in (0, 2]$ and positivity parameter $\rho = \mathbb{P}(Y_1 > 0) \in (0, 1)$.

According to the classification of BPRES (see, for instance, [1] and [2]), Condition A1 means that we consider the critical BPRES's.

Hypothesis A2. *We assume that either of the following assumptions hold:*

- *((1; 0)-lattice case) The law of X_1 under \mathbb{P} is supported on the lattice $\{\dots, -2, -1, 0, 1, 2, \dots\}$, with maximal span 1.*
- *(absolutely continuous case) The law of X_1 under \mathbb{P} is absolutely continuous with respect to the Lebesgue measure on \mathbb{R} , and there exists $n \in \mathbb{N}$ such that the density $f_n(x) := \mathbb{P}(S_n \in dx)/dx$ of S_n is essentially bounded (therefore, $f_n(x) \in L^\infty$).*

Our third assumption on the environment concerns reproduction laws of particles. Set

$$\gamma(b) = \frac{\sum_{k=b}^{\infty} k^2 F(\{k\})}{\left(\sum_{i=b}^{\infty} i F(\{i\})\right)^2}.$$

Hypothesis A3. *There exist $\varepsilon > 0$ and $b \in \mathbb{N}$ such that*

$$\mathbb{E}[(\log^+ \gamma(b))^{\alpha+\varepsilon}] < \infty,$$

where $\log^+ x = \log(\max(x, 1))$.

Our main result looks as follows.

Theorem. *Let Conditions A1–A3 be valid, $\min(m, n) \rightarrow \infty$ and $m = o(n)$. Then*

1) *if $\varphi(n) \rightarrow \infty$ as $n \rightarrow \infty$ in such a way that $\varphi(n) = o(a_m)$, then for any $z \in (0, \infty)$*

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{1}{a_m} \log Z_{n-m} \leq z | S_n \leq \varphi(n), Z_n > 0 \right) = \frac{\int_0^z w^{\alpha\rho} g^-(w) dw}{\int_0^\infty w^{\alpha\rho} g^-(w) dw},$$

where $g^-(w)$ is the density of the meander of the process $-Y$;

2) if $\varphi(n) \rightarrow \infty$ as $n \rightarrow \infty$ in such a way that $\varphi(n) \sim T a_m, T \in (0, \infty)$, then for any $z \in [0, \infty)$

$$\begin{aligned} \lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{1}{a_m} \log Z_{n-m} \leq z | S_n \leq \varphi(n), Z_n > 0 \right) \\ = \frac{\alpha\rho + 1}{T^{\alpha\rho+1}} \int_0^z w^{\alpha\rho} dw \int_0^T g(t-w) C(w, t) dt, \end{aligned}$$

where

$$C(a, b) = \mathbb{P} \left(\inf_{0 \leq s \leq 1} Y_s \geq 0 | Y_0 = a, Y_1 = b \right);$$

3) if $a_m = o(\varphi(n)), \varphi(n) = o(a_n)$, then for any $z \in [0, \infty)$

$$\lim_{n \rightarrow \infty} \mathbb{P} \left(\frac{1}{a_m} (\log Z_{n-m} - S_n) \leq z | S_n \leq \varphi(n), Z_n > 0 \right) = \mathbb{P}(Y_1 \leq z).$$

The work of E. E. Dyakonova and V. A. Vatutin was performed at the Steklov International Mathematical Center and supported by the Ministry of Science and Higher Education of the Russian Federation (agreement no. 075-15-2022-265). The research of C. Dong and V. A. Vatutin was also supported by the Ministry of Science and Technology of PRC, project G2022174007L.

References

- [1] *Afanasyev V. I., Beinghoff Ch., Kersting G., Vatutin V. A.* Criticality for branching processes in random environment // *Ann. Probab.* 2005. 33(2):645–673.
- [2] *Kersting G., Vatutin V.* *Discrete Time Branching Processes in Random Environment.* Wiley, London, 2017.
- [3] *Vatutin V. A., Dong C., Dyakonova E. E.* Random walks conditioned to stay nonnegative and branching processes in an unfavourable environment // *Sb. Math.* 2023. 214(11):1501–1533.

Максимальное число расширений в случайном графе

Степан В. Вахрушев, Максим Е. Жуковский

Московский физико-технический институт, Москва, Россия

E-mail: vakhrushev.sv@phystech.edu

Изучаются числа расширений в плотном биномиальном случайном графе, а именно нахождение асимптотического распределения максимального числа расширений. Под расширением заданного множества вершин (корней) V (небольшого размера) подразумевается непересекающееся с этим множеством множество U , ребра в котором (а также между U и V) должны быть проведены в соответствии с некоторым шаблоном – корневым графом (R, H) , в котором зафиксировано множество корней R . Так, например, степень вершины – это количество расширений, для которых шаблоном является ребро, одна из вершин которого является корнем. Другими примерами количеств расширений являются число треугольников, содержащих заданную вершину, количество общих соседей заданного набора вершин, количество путей заданной длины между двумя фиксированными вершинами и т. д.

Ряд работ, восходящих к Боллобашу и Спенсеру – [2, 8, 9], посвящены изучению распределения максимального числа расширений (по всем способам выбрать корни). Недавно Игорь Родионов и Максим Жуковский [6] смогли обобщить данный результат для максимального числа общих соседей k вершин $\Delta_{k,n}$, показав, что после масштабирования этой случайной величины с другими константами будет, как и у Боллобаша, наблюдаться сходимость к распределению Гумбеля. В настоящей работе авторам обобщить все известные результаты (см. также [4, 7]): доказано существование предельного распределения (и явно предъявлено) для широкого класса *симметричных расширений*. Единственное ограничение, накладываемое на шаблон, выглядит следующим образом: если v_1, v_2 – два корня шаблона, а u_1, u_2 – две некорневые вершины, такие, что u_1 – общий сосед вершин v_1, v_2 , а u_2 смежна с v_1 , то u_2 должна быть смежна и с v_2 . Это задаёт разбиение корней на классы: разбиение “первого уровня” представляется в виде объединения полных двудольных подграфов.

Сформулируем необходимые обозначения и полученный общий результат работы. Рассмотрим симметричный корневой граф H с h вершинами и f рёбрами. Разобьём его множество корней R на классы эквивалентности (в соответствии с определением симметричных расширений

выше) так, что для любого $i \in [r]$, существует в точности m_i классов размера k_i (здесь $k_1 < \dots < k_r$ – все возможные размеры классов). Оказывается, что предельное распределение (но не константы масштабирования) зависит только от двудольного корневого подграфа H , состоящего из того же множества корней в качестве одной доли, вершин, которые смежны с корнями в H в качестве другой доли, и рёбер между этими долями. Этот подграф можно задать вектором $W(H) := ((m_1, k_1), (m_2, k_2), \dots, (m_r, k_r))$ и дополнительно константами $g_{i,j}$, $i \in [r]$, $j \in [m_i]$, являющимися числами общих соседей для j -ого класса корней размера k_i в корневом графе. Без потери общности предположим, что $g_{i,1} \geq \dots \geq g_{i,m_i}$ для каждого $i \in [r]$. Также обозначим за $g_i := \sum_{j=1}^{m_i} g_{i,j}$ общее количество вершин, инцидентных классам корней размера k_i , а за $g := \sum_{i=1}^r g_i$ общее количество вершин, смежных хотя бы с одним корнем. Наконец, пусть $s \geq 0$ – общее количество вершин, несмежных с корнями. Очевидно, что

$$|R| + g + s = h .$$

Теорема. *Используя обозначения выше, определим*

$$\begin{aligned} a_n &= \frac{n^{s+g-1} p^f}{g_{1,1}! g_{1,2}! \dots g_{r,m_r}!} \left[n p^{\sum_{i=1}^r k_i g_i} + \sqrt{2n \ln n} \times \right. \\ &\times \left. \left(\sum_{i=1}^r g_i p^{k_i (g_i - 1)} \sqrt{k_i p^{k_i} (1 - p^{k_i})} \left(1 - \frac{\ln(k_i!)}{2k_i \ln n} - \frac{\ln[4\pi k_i \ln n]}{4k_i \ln n} \right) \right) \right], \\ b_n &= \frac{n^{s+g-1} p^f}{g_{1,1}! g_{1,2}! \dots g_{r,m_r}!} \sqrt{\frac{n}{2 \ln n}} p^{\sum_{i=1}^r k_i g_i} . \end{aligned}$$

Тогда

$$\frac{\max_T X(T) - a_n}{b_n} \xrightarrow{d} \sum_{i=1}^r \sqrt{\frac{1 - p^{k_i}}{k_i p^{k_i}}} \sum_{j=1}^{m_i} g_{i,j} \eta_{i,j},$$

где случайные вектора $\eta_i = (\eta_{i,j}, j \in [m_i])$ независимы в совокупности и имеют плотность

$$p_{\eta_i}(x_1, x_2, \dots, x_{m_i}) = e^{-x_1} \cdot e^{-x_2} \cdot \dots \cdot e^{-x_{m_i}} \cdot e^{-e^{-x_{m_i}}} \cdot I(x_1 \geq x_2 \geq \dots \geq x_{m_i}).$$

В частности, в общем случае предельное распределение не является распределением Гумбеля!

Работа выполнена при поддержке российского научного фонда, грант 22-11-00131.

Список литературы

- [1] *Bohman T., Frieze A., Lubetzky E.* Random triangle removal // *Adv. Math.* 2015. 280:379–438.
- [2] *Bollobás B.* The distribution of the maximum degree of a random graph // *Discrete Mathematics.* 1980. 32:201–203.
- [3] *Gnedenko B.* Sur La Distribution Limite Du Terme Maximum D'Une Serie Aleatoire // *Annals of Mathematics.* 1943. 44(3):423–453.
- [4] *Isaev M., Rodionov I., Zhang R., Zhukovskii M.* Extremal independence in discrete random systems // *Annales de l'Institut Henri Poincaré (B)* (to appear).
- [5] *Ivchenko G. I.* On asymptotic behaviour of the degrees of vertices in a random graph // *Teor. Veroyatnost. i Primenen.* 1973. 18(1):195–203.
- [6] *Rodionov I., Zhukovskii M.* The distribution of the maximum number of common neighbors in the random graph // *European Journal of Combinatorics.* 2023. 107:103602.
- [7] *Šileikis M., Warnke L.* Counting extensions revisited // *Random Struct. Algorithms.* 2022. 61:3–30.
- [8] *Spencer J. H.* Threshold functions for extension statements // *J. of Comb. Th. Ser A.* 1990. 53:286–305.
- [9] *Spencer J. H.* Counting extensions // *J. of Comb. Th. Ser A.* 1990. 55:247–255.

Локальная информационная геометрия двоичных цепей Маркова высокого порядка

Валерий А. Волошко^{1,2}

¹НИИ прикладных проблем математики и информатики,

²Белорусский государственный университет,

Минск, Белоруссия

E-mail: valoshka@bsu.by

Рассмотрим двоичную стационарную цепь Маркова $x_t \in \{0, 1\}$, $t \in \mathbb{Z}$, порядка $s \in \mathbb{N}_0 = \{0, 1, \dots\}$, определяемую набором невырожденных (не равных нулю или единице) переходных вероятностей $\mathbf{p}(q)$ в единичное значение для всех предысторий $q = (q_i)_{i=1}^s \in \{0, 1\}^s$ длины s (вероятность перехода в нулевое значение из условия нормировки равна $1 - \mathbf{p}(q)$):

$$\mathbf{p}(q) ::= \mathbb{P} \{x_t = 1 | x_{t-1} = q_1, \dots, x_{t-s} = q_s\} \in (0, 1).$$

Семейство распределений вероятностей описанных цепей Маркова на бесконечных двоичных последовательностях обозначим $\text{MC}(s)$. Нулевой порядок $s = 0$ отвечает схеме независимых испытаний Бернулли с одним параметром – вероятностью успеха $\mathbf{p} \in (0, 1)$. При $\mathbf{p} = 1/2$ получается равномерная вероятностная мера на множестве бесконечных двоичных последовательностей, определяемая свойством:

$$\mathbb{P} \{x_{t+1} = q_1, \dots, x_{t+m} = q_m\} \equiv \frac{1}{2^m}, \quad \forall t \in \mathbb{Z}, m > 0, q_1, \dots, q_m \in \{0, 1\},$$

которую обозначим $\mathbf{U} \in \text{MC}(0)$. Семейства $\text{MC}(s)$ вложены друг в друга:

$$\text{MC}(0) \subset \text{MC}(1) \subset \text{MC}(2) \subset \dots$$

Их объединение – семейство распределений вероятностей цепей Маркова всех конечных порядков на бесконечных двоичных последовательностях – обозначим $\text{MC} ::= \cup_{s \in \mathbb{N}_0} \text{MC}(s)$. Равномерная мера \mathbf{U} лежит во всех семействах $\text{MC}(s)$, $s \in \mathbb{N}_0$. Обозначим через $\text{TMC}(s)$ касательное пространство точки \mathbf{U} в многообразии $\text{MC}(s)$, $s \in \mathbb{N}_0$. Пространство $\text{TMC}(s)$ имеет размерность $\dim(\text{TMC}(s)) = 2^s$. Как и семейства $\text{MC}(s)$, касательные пространства $\text{TMC}(s)$ вложены друг в друга:

$$\text{TMC}(0) \subset \text{TMC}(1) \subset \text{TMC}(2) \subset \dots$$

Их объединение – бесконечномерное касательное пространство точки \mathbf{U} в многообразии MC – обозначим $\text{TMC} ::= \cup_{s \in \mathbb{N}_0} \text{TMC}(s)$. На пространстве TMC определен тензор Фишера-Римана – скалярное произведение:

$$\langle \tau_1, \tau_2 \rangle_{\text{TMC}} \in \mathbb{R}, \quad \tau_1, \tau_2 \in \text{TMC}.$$

Вероятностную меру на множестве двоичных слов $(x_1, \dots, x_s) \in \{0, 1\}^s$ длины $s \in \mathbb{N} = \{1, 2, \dots\}$ будем называть инвариантной к сдвигам, если для нее одинаково распределены подслова (x_1, \dots, x_{s-1}) и (x_2, \dots, x_s) длины $s-1$ (при $s=1$ все распределения инвариантны к сдвигам). Отметим, что из этого условия следует одинаковое распределение всех подслов равных длин m , $1 < m < s$. Обозначим через $\text{SI}(s)$ семейство невырожденных (приписывающих всем двоичным словам строго положительные вероятности) инвариантных к сдвигам вероятностных мер на множестве двоичных слов длины $s \in \mathbb{N}$. Через $\mathbf{U}(s)$ обозначим равномерную вероятностную меру на множестве $\{0, 1\}^s$ двоичных слов длины $s \in \mathbb{N}$. Очевидно равномерная вероятностная мера инвариантна к сдвигам: $\mathbf{U}(s) \in \text{SI}(s)$. Касательное пространство точки $\mathbf{U}(s)$ в многообразии $\text{SI}(s)$ обозначим $\text{TSI}(s)$. На этом касательном пространстве также определен тензор Фишера-Римана:

$$\langle \sigma_1, \sigma_2 \rangle_{\text{TSI}(s)} \in \mathbb{R}, \quad \sigma_1, \sigma_2 \in \text{TSI}(s), \quad s \in \mathbb{N}.$$

Существует тесная связь между введенными семействами Марковских вероятностных мер $\text{MC}(s)$, $s \in \mathbb{N}_0$, на бесконечных двоичных последовательностях и сдвигово инвариантных вероятностных мер $\text{SI}(s)$, $s \in \mathbb{N}$, на конечных двоичных словах. Например, сдвигово инвариантно маргинальное распределение вероятностей любого конечного m -подслова $(x_{t+1}, \dots, x_{t+m}) \in \{0, 1\}^m$, $t \in \mathbb{Z}$, $m > 0$, стационарной цепи Маркова любого конечного порядка $s \in \mathbb{N}_0$. При этом для $m = s+1$ распределение вероятностей $\mu_{\text{SI}(s+1)} \in \text{SI}(s+1)$ $(s+1)$ -подслова $(x_{t+1}, \dots, x_{t+s+1})$ (ввиду стационарности, это распределение не зависит от t) однозначно задает распределение вероятностей $\mu_{\text{MC}(s)} \in \text{MC}(s)$ всей цепи Маркова $\{x_t\}_{t \in \mathbb{Z}}$. Другими словами, между многообразиями $\text{SI}(s+1)$ и $\text{MC}(s)$, $s \in \mathbb{N}_0$, существует биективное соответствие, в направлении $\phi_s : \text{MC}(s) \rightarrow \text{SI}(s+1)$ представляющее собой взятие маргинального распределения вероятностей $(s+1)$ -подслов. При $s \geq m > 0$ определена естественная операция проекции $\psi_{s \rightarrow m} : \text{SI}(s) \rightarrow \text{SI}(m)$, представляющая собой взятие маргинального распределения вероятностей m -подслов в s -слове (в случае $s = m$ это тождественное преобразование $\psi_{s \rightarrow s} = \text{id}$). Сопряженное к данной проекции преобразование:

$$\tilde{\psi}_{s \rightarrow m} ::= \phi_m^{-1} \circ \psi_{s+1 \rightarrow m+1} \circ \phi_s : \text{MC}(s) \rightarrow \text{MC}(m), \quad s \geq m \geq 0,$$

определяет естественную “аппроксимирующую” проекцию на множестве Марковских распределений, когда распределение цепи Маркова порядка s аппроксимируется распределением цепи Маркова меньшего порядка m таким образом, что у исходной и аппроксимирующей цепей Маркова одинаковые маргинальные распределения вероятностей $(m + 1)$ -подслов. На уровне касательных пространств проекция $\tilde{\psi}_{s \rightarrow m}$ превращается в ортогональный проектор $\text{TMC}(s)$ на $\text{TMC}(m)$ относительно тензора Фишера – Римана $\langle \cdot, \cdot \rangle_{\text{TMC}}$, а отображение ϕ_s – в линейное биективное отображение $\text{TMC}(s)$ на $\text{TSI}(s + 1)$.

В докладе описывается конструкция касательных пространств TMC , $\text{SI}(s)$, $s \in \mathbb{N}$, и тензоров Фишера – Римана на них. Даются явные формулы связи между тензорами $\langle \cdot, \cdot \rangle_{\text{TMC}}$ и $\langle \cdot, \cdot \rangle_{\text{SI}(s)}$ и биективного преобразования между $\text{TMC}(s)$ и $\text{TSI}(s + 1)$. Приводятся примеры решаемых с помощью описанной теории задач [1–3], связанных с асимптотическими свойствами статистических тестов “чистой случайности” (равномерного распределения) двоичной последовательности, используемых в известных багареях тестов, таких как NIST. В частности, рассматриваются задачи вычисления асимптотической мощности тестов и выявления асимптотической зависимости между тестами. Приводится критерий асимптотической независимости статистик двух тестов типа хи-квадрат в случае истинной нулевой гипотезы (равномерного распределения) и в случае асимптотического сближения альтернативы с нулевой гипотезой.

Список литературы

- [1] *Волошко В. А., Трубей А. И.* О мощности тестов многомерной дискретной равномерности, используемых для статистического анализа генераторов случайных последовательностей // Журнал Белорусского государственного университета. Математика. Информатика. 2022. 1:26–37. doi: 10.33581/2520-6508-2022-1-26-37
- [2] *Voloshko V. A., Kharin Yu. S., Trubey A. I.* On Power Comparison for Some Tests on Pure Randomness under Markov High-Order Dependencies // Proc. XIII Int. Conf. “Computer Data Analysis and Modeling”. 2022. P. 211–217.
- [3] *Волошко В. А.* Об асимптотических свойствах семейства χ^2 -тестов чистой случайности двоичной последовательности // Сборник материалов II международной научной конференции “Теоретическая и прикладная криптография”. 2023. С. 15–43.

Расчет стационарного распределения времени обслуживания и длительности периода занятости в односерверной системе

¹Александр С. Головин, ²Александр С. Румянцев

¹Институт прикладных математических исследований КарНЦ РАН,
Петрозаводск, Россия;

²Новгородский государственный университет им. Ярослава Мудрого,
Великий Новгород, Россия

E-mail: golovin@krc.karelia.ru; ar0@krc.karelia.ru

Односерверная система обслуживания является важным базовым объектом в теории систем массового обслуживания, для которого доступен ряд явных результатов (см., напр. [1]). В то же время, такая система является (стохастической) верхней границей (с точки зрения нагрузки) для ряда многосерверных систем [2]. Кроме того, система может использоваться для исследования многосерверных систем методом точной выборки [3]. В этой связи, возникает задача расчета стационарных характеристик односерверной системы, а именно, распределения нагрузки и длительности цикла занятости (в терминах числа обслуженных клиентов) в стационарном режиме.

В общем случае в системе типа M/G/1 (на вход которой поступает пуассоновский поток заявок с интенсивностью λ , а времена обслуживания имеют произвольное распределение F_S), при условии $\rho = \lambda ES < 1$, длительность периода занятости T может быть получена по следующей формуле [1, с. 250]

$$P\{T = n\} = \int_0^\infty \frac{(\lambda u)^{n-1}}{n!} e^{-\lambda u} dF_S^{(n)}(u),$$

где $F_S^{(n)}$ есть n -кратная свертка распределения времени обслуживания S с собой. Для расчета распределения времени ожидания F_W можно воспользоваться соотношением

$$F_W(x) = P(W \leq x) = (1 - \rho) \sum_{k=0}^{\infty} \rho^k F_{S_e}^{(k)}(x),$$

где $F_{S_e}(x) = \frac{1}{ES} \int_0^x (1 - F_S(u)) du$ есть распределение стационарного незавершенного времени обслуживания.

© А. С. Головин, А. С. Румянцев, 2024

Для ряда случаев указанные вычисления удается провести в явном виде, если доступны выражения для соответствующих сверток. Таким случаем, в частности, является система, в которой времена обслуживания имеют распределение Парето второго типа,

$$F_S(x) = 1 - \left(\frac{x_0}{x_0 + x} \right)^\alpha, \quad x_0, \alpha > 0.$$

В то же время, вычисление соответствующих выражений представляет определенные сложности.

В данной работе предложен подход к вычислению выражений для распределения времени обслуживания и длительности периода занятости, основанный на работах [4–6].

Работа поддержана РФФ, проект 21-71-10135.

Список литературы

- [1] *Cohen J. W.* The single server queue. Elsevier North-Holland, 1982. 694 p.
- [2] *Морозов Е. В., Румянцев А. С.* Вероятностные модели многопроцессорных систем: стационарность и моментные свойства // Информатика и ее применения. 2012. 3(6):99–106.
- [3] *Sigman K.* Using the M/G/1 queue under processor sharing for exact simulation of queues // Annals of Operations Research. 2016. 1–2(241):23–34.
- [4] *Ramsay C. M.* Exact waiting time and queue size distributions for equilibrium M/G/1 queues with Pareto service // Queueing Systems. 2007. 4(57):147–155.
- [5] *Ramsay C. M.* The Distribution of Sums of I.I.D. Pareto Random Variables with Arbitrary Shape Parameter // Communications in Statistics – Theory and Methods. 2008. 14(37):2177–2184.
- [6] *Pearson J. W., Olver S., Porter M. A.* Numerical methods for the computation of the confluent and Gauss hypergeometric functions // Numerical Algorithms. 2017. 3(74):821–866.

Сложность конечных случайных структур

¹Данила А. Демин, ²Максим Е. Жуковский

¹Московский физико-технический институт, Долгопрудный, Россия;

²University of Sheffield, Sheffield, UK

E-mail: demin.da@phystech.edu; m.zhukovskii@sheffield.ac.uk

Случайная n -элементная структура сигнатуры σ – это случайный элемент некоторого подмножества всех n -элементных моделей сигнатуры σ . Например, случайный граф $G(n, p)$ – это случайная модель сигнатуры $\sigma = \{\sim\}$, для которой выполнено $\forall x \forall y (x \sim y) \leftrightarrow (y \sim x) \wedge \forall x \neg(x \sim x)$, такая, что все события $x \sim y$ для различных x и y независимы и имеют вероятность p . Случайная структура без условий $D^\sigma(n, p_1, \dots, p_s)$ – это случайная модель сигнатуры $\sigma = \{P_1^{(d_1)}, P_2^{(d_2)}, \dots, P_s^{(d_s)}\}$, где события $P_i(x_1, \dots, x_{d_i})$ имеют вероятность p_i и независимы по всем i и наборам (x_1, \dots, x_{d_i}) , в частности для сигнатуры, состоящей из одного предиката P арности d , мы получаем случайную n -элементную интерпретацию $D^{(d)}(n, p)$.

Для формул первого порядка сигнатуры σ рассматриваются предельные свойства вероятностей их выполнения в случайных n -элементных интерпретациях σ . Например, в работах Глебского Ю. В., Когана Д. И., Лиогонького М. И., Таланова В. А. [1] и Фейгина Р. [2] доказан 0–1 закон для $D^\sigma(n, p_1, \dots, p_s)$ (и для $G(n, p)$ тоже), где $p_i = \text{const}$, то есть то, что для любой формулы в сигнатуре σ вероятность её выполнения в $D^\sigma(n, p_1, \dots, p_s)$ стремится к нулю или к единице при $n \rightarrow \infty$.

Мы вводим понятие сложности первого порядка случайной структуры как проекции множества последовательностей $(\Pr(D_n \models \varphi))_{n \in \mathbb{N}}$ по всем замкнутым формулам первого порядка φ на пространство ℓ_∞/c_0 . Выполнение 0–1 закона и закона сходимости переформулируется как свойство сложности первого порядка для случайной структуры. Мы приводим примеры случайных структур вида $G(n, p(n))$ и $G(n \mid \psi)$ (то есть равномерных случайных графов, для которых выполняется некоторая замкнутая формула первого порядка ψ), для которых сложность первого порядка удовлетворяет одному из свойств:

- содержится в d -мерном пространстве, но не в $(d - 1)$ -мерном пространстве;
- не содержится ни в каком конечномерном пространстве, но вполне ограничена;

© Д. А. Демин, М. Е. Жуковский, 2024

- не вполне ограничена.

Также мы показали, что сложность первого порядка структуры $G(n, n^{-\alpha})$ для рационального $\alpha \in (0, 1)$ не вполне ограничена.

Далее мы вводим понятие стохастической сводимости первого порядка и доказываем, что она позволяет переносить сложность первого порядка между случайными структурами, а значит и 0–1 закон, и закон сходимости. Далее мы приводим примеры результатов о выполнении 0–1 закона, полученных сводимостью к D^σ . Также, с помощью сводимостей, мы доказываем то, что для любых целых неотрицательных r и k структуры $D^{(r+2)}(n, \frac{(r+1)\ln n}{n})$ и $D^{(r+3)}(n, \frac{(r+1+\frac{1}{k})\ln n}{n})$ не подчиняются 0–1 закону, а для рациональных $\beta \in (\frac{2}{3}, 1)$ сложность первого порядка структуры $D^{(r+3)}(n, \frac{(r+\beta)\ln n}{n})$ не вполне ограничена.

Затем, мы обобщаем результат Ларраури, Мюллера и Ноя [3] о том, что существует такая константа c_0 , что множество пределов вероятностей выполнения формул первого порядка для неориентированных d -однородных гиперграфов $G^{(d)}(n, \frac{c}{n^{d-1}})$ плотно в $[0, 1]$ тогда и только тогда, когда $c \geq c_0$, на H -ориентированные гиперграфы (то есть d -однородные гиперграфы, для которых гиперребром считается H -орбита последовательности из d различных вершин под действием группы $H \leq S_d$ на вершинах гиперребра перестановками).

В конце мы показываем, что задача проверки того, что $G(n | \varphi)$ подчиняется 0–1 закону по данной формуле φ не является перечислимой.

Список литературы

- [1] *Glebskii Y. V., Kogan D. I., Liogon'kii M. I., Talanov V. A.* Range and degree of realizability of formulas in the restricted predicate calculus // *Cybernetics*. 1969. 5(2):142–154.
- [2] *Fagin R.* Probabilities on finite models // *The Journal of Symbolic Logic*. 1976. 41(1):50–58.
- [3] *Larrauri A., Müller T., Noy M.* Limiting probabilities of first order properties of random sparse graphs and hypergraphs // *Random Structures & Algorithms*. 2022. 60(3):506–526.

Majority dynamics на конечных деревьях

Георгий Е. Захаров

Московский физико-технический институт, Долгопрудный, Россия

E-mail: georgiizakharov0@gmail.com

Majority dynamics это процесс на графах который описывается следующим образом. Вначале каждая вершина $i \in V$ графа G имеет *изначальное мнение* $\xi_0(i) \in \{-1, 1\}$. Затем на каждом шаге $t \in \mathbb{N}$ она принимают мнение большинства своих соседей, т. е.

$$\xi_t(i) = \text{sign} \sum_{j \in N_G(i)} \xi_{t-1}(j),$$

где $N_G(i)$ обозначает множество соседей i в G . Для удобства мы рассматриваем графы только с нечетными степенями, поэтому $\sum_{j \in N_G(i)} \xi_t(j)$ никогда не равно 0.

Голес и Оливос [1] доказали, что процесс стабилизируется на конечных деревьях G . Более формально, выполняется следующее свойство *периодичности длины два*: для каждой вершины i и всех достаточно больших t , $\xi_{t+2}(i) = \xi_t(i)$. Тем не менее, вопрос о наименьшем времени до стабилизации

$$\tau(T; \xi_0) = \min\{t : \xi_{t+2}(i) = \xi_t(i) \text{ для всех } i \in V(T)\}$$

все еще открыт. Мы изучаем время до стабилизации majority dynamics на *конечных деревьях*.

Во-первых, для любого конечного дерева T мы нашли точное значение наилучшего времени до стабилизации

$$\tau(T) := \max_{\xi_0 \in \{-1, 1\}^{V(T)}} \tau(T; \xi_0).$$

Назовем укорененное дерево T *совершенным k -арным*, если все вершины, кроме листьев, имеют $k+1$ соседей и все листья находятся на одинаковом расстоянии от корня. Для случайной величины $\xi_0 \in \{-1, 1\}^{V(T)}$ будем говорить, что некое событие выполняется *с высокой вероятностью*, если вероятность этого события, стремится к единице, при стремлении размера графа к бесконечности. Мы доказали, что время стабилизации на совершенном 2-арном дереве с равномерно распределенным случайным вектором начальных мнений ξ_0 на его вершинах с высокой вероятностью

линейно зависит от диаметра дерева и отличается нетривиальной мультипликативной константой от наихудшего времени. Кроме того, мы доказали, что для фиксированного четного $k > 2$ и равномерно распределенного случайного вектора начальных мнений ξ_0 на вершинах совершенного k -арного дерева T диаметра D верно $\tau(T; \xi_0) = \Omega(\sqrt{D})$ с высокой вероятностью.

Это совместная работа с Итаем Беньямини и Максимом Жуковским.

Список литературы

- [1] *Goles E., Olivos J.* Periodic behaviour of generalised threshold functions // *Discrete Mathematics*. 1980. 30:187–189.

Теоремы об итерациях случайных отображений конечных множеств

Андрей М. Зубков

Математический институт им. В. А. Стеклова РАН, Москва, Россия

E-mail: zubkov@mi-ras.ru

Отображения конечных множеств возникают как математические модели в разных областях дискретной математики: теории графов, теории алгоритмов, теории вероятностей и т. п. В докладе планируется сделать небольшой обзор результатов, связанных со случайными отображениями конечных множеств и дополняющих результаты, описанные в книгах В. Ф. Колчина «Случайные отображения» и «Случайные графы» (М.: Наука, 1984, 2000).

Будут рассмотрены результаты о распределении структурных характеристик графов, соответствующих типичным случайным отображениям: числа связанных компонент, их размеров, длин циклов, а также отрезков аперiodичности, начинающихся в случайно выбранной вершине. Представляет практический интерес исследование вероятности принадлежности нескольких вершин одной связной компоненте случайного отображения.

Значительное число публикаций посвящено изучению различных характеристик итераций случайных отображений конечных множеств. Получен ряд результатов об изменении объема образа конечного множества с ростом числа итераций, например, о числе итераций, при котором образ становится одноточечным.

Распределения характеристик случайных отображений изучались не только для равновероятных распределений на множестве всех отображений конечного множества в себя, но и для неравновероятных распределений (в частности, условных распределений, когда какие-то характеристики отображения фиксированы).

Для итераций случайных неравновероятных взаимно однозначных отображений конечных множеств изучалась скорость сходимости распределения образа одной точки к равновероятному распределению, а также скорость сходимости распределения итераций случайных отображений к равновероятному распределению на множестве подстановок на заданном множестве.

О промежуточных значениях размерностей квантования вероятностных мер

Александр В. Иванов

Институт прикладных математических исследований КарНЦ РАН,
Петрозаводск, Россия

E-mail: alvlivanov@krc.karelia.ru

Квантованием вероятностной меры называется приближение данной меры мерами с конечными носителями. В рамках теории квантования определены размерности квантования (верхняя $\overline{D}(\mu)$ и нижняя $\underline{D}(\mu)$) вероятностной меры μ . Пусть (X, ρ) – метрический компакт и $P(X)$ – пространство вероятностных мер на X с метрикой ρ_P Канторовича – Рубинштейна, определяемой по формуле

$$\rho_P(\mu, \nu) = \sup\{|\mu(f) - \nu(f)| : f \in \text{Lip}_1(X)\},$$

где $\text{Lip}_1(X)$ – множество вещественных функций на X , удовлетворяющих условию Липшица с константой 1, и $\mu(f) = \int f d\mu$. Носителем $\text{supp}(\mu)$ меры $\mu \in P(X)$ называется наименьшее замкнутое подмножество X полной меры. Для $\mu \in P(X)$ и $\varepsilon > 0$ через $N(\mu, \varepsilon)$ обозначается наименьшее число точек в носителе ε -приближения μ . Если носитель μ бесконечен, то $N(\mu, \varepsilon)$ неограниченно возрастает при $\varepsilon \rightarrow 0$. Скорость этого возрастания характеризует размерность квантования $D(\mu)$ меры μ :

$$D(\mu) = \lim_{\varepsilon \rightarrow 0} \frac{\log N(\mu, \varepsilon)}{-\log \varepsilon}. \quad (1)$$

Если указанный предел не существует, то рассматривают верхний и нижний пределы и получают (соответственно) верхнюю $\overline{D}(\mu)$ и нижнюю $\underline{D}(\mu)$ размерности квантования меры μ .

Известно, что размерности квантования вероятностной меры μ , определенной на метрическом компакте (X, ρ) , не превосходят соответствующих емкостных размерностей \dim_B ее носителя:

$$\overline{D}(\mu) \leq \overline{\dim}_B(\text{supp}(\mu)), \quad \underline{D}(\mu) \leq \underline{\dim}_B(\text{supp}(\mu)). \quad (2)$$

Емкостная размерность является аналогом размерности квантования для пространства замкнутых подмножеств компакта X . Для замкнутого

$F \subset X$ и $\varepsilon > 0$ через $N(F, \varepsilon)$ обозначается наименьшее число точек в ε -приближении F по метрике Хаусдорфа (фактически $N(F, \varepsilon)$ – наименьшее число точек в ε -сети F), и $\dim_B F$ определяется по формуле (1) с заменой $N(\mu, \varepsilon)$ на $N(F, \varepsilon)$.

Неравенства (2) определяют верхнюю границу размерностей квантования мер с заданным носителем и порождают следующие естественные вопросы о промежуточных значениях размерностей квантования:

Верно ли, что для любого числа $b \in [0, \overline{\dim}_B X]$ ($b \in [0, \underline{\dim}_B X]$) существует вероятностная мера μ_b на компакте X , для которой $\overline{D}(\mu_b) = b$ ($\underline{D}(\mu_b) = b$)?

Для верхних размерностей вопрос о промежуточных значениях решен положительно в [1]. Его решение основано на существовании в X замкнутого подмножества F наперед заданной верхней емкостной размерности $\overline{\dim}_B F = b$ для любого $b \in [0, \overline{\dim}_B X]$. Для нижней размерности квантования аналогичный подход невозможен, поскольку в работе [2] построен пример метрического компакта X размерности $\dim_B X = 1$, все непустые собственные замкнутые подмножества которого нульмерны в смысле $\underline{\dim}_B$. Тем не менее, для нижних размерностей поставленный выше вопрос также решается положительно. Таким образом, значения нижней размерности квантования вероятностных мер на X заполняют весь диапазон $[0, \underline{\dim}_B X]$ независимо от наличия «провалов» в множестве значений нижней емкостной размерности замкнутых подмножеств компакта X . Более того, доказана следующая

Теорема. Пусть X – метрический компакт и $\dim_B X = a \leq \infty$. Тогда для любых чисел b, c таких, что $0 \leq b \leq c \leq a$ существует мера $\mu \in P(X)$, для которой $\underline{D}(\mu) = b$, $\overline{D}(\mu) = c$.

Финансовое обеспечение исследования осуществлялось из средств федерального бюджета на выполнение государственного задания КарНЦ РАН (Институт прикладных математических исследований КарНЦ РАН).

Список литературы

- [1] Иванов А. В. О множестве значений размерности квантования вероятностных мер на метрическом компакте // Сибирский матем. журнал. 2022. 63(5):1074–1081. doi: 10.33048/smzh.2022.63.509
- [2] Иванов А. В. О промежуточных значениях емкостных размерностей // Сибирский матем. журнал. 2023. 64(3):540–545. doi: 10.33048/smzh.2023.64.307

Стратегии оптимальной остановки в игре на разорение

Анна А. Ивашко

Институт прикладных математических исследований КарНЦ РАН,
Петрозаводск, Россия

E-mail: aivashko@krc.karelia.ru

Рассматривается многошаговая игра в дискретном времени и конечным горизонтом n , связанная с задачей о разорении ([1]). Пусть у каждого из двух игроков (I и II) имеется некоторое количество капитала. Игроки на каждом шаге игры разыгрывают единицу капитала. Предполагается, что у игрока I вероятность выиграть равна p , а у игрока II $q = 1 - p$. Соответственно, на каждом шаге капитал игрока либо увеличивается на единицу, либо уменьшается. На каждом шаге игры каждый из участников платит величину c , $0 \leq c < 1$. Игра заканчивается, если у одного из игроков закончится капитал, то есть он разорится. В этом случае этот игрок получит в качестве выигрыша 0, а его противник получит выигрыш, равный 1. Если до момента времени n игра не закончилась, то оба игрока ничего не получают. В данной задаче необходимо найти оптимальные стратегии остановки для максимизации ожидаемых выигрышей игроков.

Задача о разорении, если игрок останавливается как только количество его капитала уменьшится на 1, была рассмотрена в работе [2]. В данной работе исследованы две постановки задачи: а) у одного из игроков начальный капитал конечный, а у другого бесконечный; б) у обоих игроков капитал бесконечный. Были найдены границы оптимальной остановки и выигрыши игроков для обоих случаев. Проведено численное моделирование оптимальных стратегий и выигрышей игроков для различных значений p и n .

Список литературы

- [1] *Феллер В.* Введение в теорию вероятностей и ее приложения. М.: Мир, 1984.
- [2] *Mazalov V. V., Ivashko A. A.* Harmonic Numbers in Gambler's Ruin Problem // Lecture Notes in Computer Science. Springer, Cham. 2023. 13930:278–287. doi: 10.1007/978-3-031-35305-5_19

An explicit form of the Kolmogorov constant in the theory of continuous-time Markov Stochastic Branching Systems

Azam A. Imomov

Karshi State University, Karshi city, Uzbekistan

E-mail: imomov_azam@mail.ru

In the recent paper [1], the famous constant of the theory of subcritical stochastic branching Galton-Watson systems, announced by Kolmogorov [2] in 1938, was calculated in explicit dependence on the structural parameters of the system. Our purpose in this talk is to find an explicit form of an analogical constant for the case of noncritical continuous-time Markov branching systems. In this issue, we will essentially use the method from [1].

Consider the continuous-time Markov branching system with branching rates $\{a_k, k \in \mathbb{N}_0\}$, where

$$a_k \geq 0 \quad \text{and} \quad 0 < a_0 < -a_1 = \sum_k a_k < \infty \quad \text{for} \quad k \in \mathbb{N}_0 \setminus \{1\},$$

where $\mathbb{N}_0 = \{0\} \cup \mathbb{N}$ and \mathbb{N} is the set of natural numbers. This system forms a reducible, homogeneous and continuous-time Markov chain with a state space consisting of two classes: $\mathcal{S}_0 = \{0\} \cup \mathcal{S}$, where $\mathcal{S} \subset \mathbb{N}$ is the class of possible essential communicating states, and $\{0\}$ is an absorbing state. Denoting by $Z(t)$ the population size at time $t \in \mathcal{T}$, we have a *homogeneous-continuous-time Markov branching (MB) system* with branching rates $\{a_k, k \in \mathcal{S}_0\}$; see [3].

Let $P_{ij}(t)$ be the transition probabilities from $i \in \mathcal{S}$ to $j \in \mathcal{S}_0$ in time t . To study the evolution of the system, it is sufficient to determine probabilities $\mathbf{p}_j(t) := P_{1j}(t)$, since $P_{ij}(t)$ are an i -fold convolution of $\mathbf{p}_j(t)$ which is called the branching property.

Consider the the generating functions (GFs)

$$f(s) = \sum_{j \in \mathcal{S}_0} a_j s^j \quad \text{and} \quad F(t; s) := \sum_{j \in \mathcal{S}_0} \mathbf{p}_j(t) s^j$$

The branching property implies that $\sum_{j \in \mathcal{S}_0} P_{ij}(t) s^j = [F(t; s)]^i$ as well.

We consider the noncritical system, i. e. $m := f'(1-) \neq 0$.

Let q be an extinction probability of the system initiated by single founder and denote $R(t; s) := q - F(t; s)$. Letting $\mathcal{H} := \min\{t : Z(t) = 0\}$ be an extinction time, we observe that $Q(t) := R(t; 0) = \mathbb{P}\{t < \mathcal{H} < \infty\}$ is the bridled-survival probability of the system at time t . In particular, in subcritical case $\mathbb{P}\{\mathcal{H} < \infty\} = 1$ and hence $Q(t) = \mathbb{P}\{Z(t) > 0\}$.

In subcritical case, Sevastyanov [4] has proved that if $f''(1) < \infty$ then

$$\frac{Q(t)}{m^t} \longrightarrow \mathcal{K} \quad \text{as } t \rightarrow \infty, \quad (1)$$

where \mathcal{K} is a certain constant depending on the form of $f(s)$. Earlier, in 1938, the discrete-time analogue of Sevastyanov's result was established by Kolmogorov [2]. In this regard, we refer to \mathcal{K} as the Kolmogorov constant. Indeed, it follows that

$$\frac{m^t}{Q(t)} = \frac{\mathbb{E}Z(t)}{\mathbb{P}\{Z(t) > 0\}} = \mathbb{E} \left[Z(t) \mid Z(t) > 0 \right] \longrightarrow 1/\mathcal{K}$$

as $t \rightarrow \infty$. So the constant \mathcal{K} can be interpreted as an equivalence ratio of the expectation $\mathbb{E}Z(t) = m^t$ of the population size to the survival probability $\mathbb{P}\{Z(t) > 0\}$ of the subcritical system. Thus, it is of principal interest to obtain an explicit expression for this coefficient depending on the numerical parameter of the system under consideration.

Let $\beta := \exp\{f'(q)\} < 0$ and $U_q[0, 1) := \{[0, q) \cup (q, 1)\}$ be a unit interval with a punctured point q . In our purpose, it is fundamentally important to find the limit of the function

$$\mathcal{A}_q(t; s) := \frac{R(t; s)}{\beta^t}$$

for $s \in U_q[0, 1)$. Indeed, after finding the function $\mathcal{A}_q(s) = \lim_{t \rightarrow \infty} \mathcal{A}_q(t; s)$, we obtain the desired Kolmogorov constant as $\mathcal{K}_q = \mathcal{A}_q(0)$. Further we will be accompanied by the condition

$$\left[m \neq 0 \quad \text{and} \quad f''(1-) < \infty \quad \text{for } m < 0 \right] \quad [\mathbf{K}]$$

that we call the Kolmogorov condition.

Lemma. *Let Kolmogorov condition [K] holds. Then the following relation is valid:*

$$\frac{1}{\mathcal{A}_q(t; s)} = \frac{1}{q - s} + \Delta(t; s),$$

where

$$\Delta(t; s) = \int_0^t b_q(u; s) \beta^u du,$$

therein $2b_q(t; s) \rightarrow f''(q)$ as $t \rightarrow \infty$ uniformly in $s \in U_q[0, 1)$.

This lemma entails our main result, which we state in the following theorem.

Theorem. *Let Kolmogorov condition [K] be satisfied. Then*

$$\mathcal{K}_q = \frac{q}{1 + q\gamma},$$

where $\gamma = b_q/|\ln \beta|$ and $2b_q := f''(q)$. In particular, the Kolmogorov constant announced in the asymptote (1) for the subcritical system is

$$\mathcal{K} = \frac{1}{1 + \gamma},$$

where $\gamma = b/|m|$ and $2b := f''(1-)$.

References

- [1] *Imomov A. A., Murtazaev M.* On the Kolmogorov constant explicit form in the theory of Discrete-time Stochastic Branching Systems // J. Appl. Probab. 2024:1–15. doi://10.1017/jpr.2023.85
- [2] *Kolmogorov A. N.* K resheniyu odnoy biologicheskoy zadachi // Reports of SRI Math. and Mech. at Tomsk Univ. 1938. 2:7–12. (in Russian)
- [3] *Sevastyanov B. A.* Branching Progresses. M: Nauka, 1971. (in Russian)
- [4] *Sevastyanov B. A.* The theory of Branching stochastic process // Uspekhi Mat. Nauk. 1951. 6(46):47–99.

Нетривиальные и предписанные раскраски кнезеровского графа

Эдуард Л. Иноземцев

Московский физико-технический институт, Долгопрудный, Россия

E-mail: eduard_inozemtsev@bk.ru

Первая задача, рассмотренная в данной работе, это следующее обобщение теоремы Эрдеша-Ко-Радо, которое впервые было предложено Эрдемем: для заданных n, k, s найти максимальный размер семейства $\mathcal{F} \subset \binom{[n]}{k}$, являющегося объединением не более чем s пересекающихся семейств. Естественная гипотеза (сделанная Эрдемем) звучит так

Гипотеза. Пусть даны s пересекающихся семейств k -множеств $\mathcal{F}_1, \dots, \mathcal{F}_s$. Тогда $|\bigcup \mathcal{F}_i| \leq \binom{n}{k} - \binom{n-s}{k}$.

Франкл и Фюреди рассмотрели эту задачу для двух пересекающихся семейств. Они доказали следующую теорему.

Теорема 1 (Франкл–Фюреди, [1]). Пусть $\mathcal{F} \subset \binom{[n]}{k}$ – это объединение двух пересекающихся семейств и $n > \frac{1}{2}(3 + \sqrt{5})k$, тогда $|\mathcal{F}| \leq \binom{n}{k} - \binom{n-2}{k}$. Равенство выполнено только в том случае, когда \mathcal{F} – это объединение двух звёзд.

Эллис и Лифшиц [2] доказали обобщение этой теоремы:

Теорема 2 (Эллис–Лифшиц, [2]). Для любого целого $s \geq 2$ существует положительная целая константа $C = C(s)$ такая, что выполнено следующее утверждение. Пусть $n \geq 2k + Ck^{3/2}$ и \mathcal{F} это объединение не более чем s пересекающихся семейств, тогда $|\mathcal{F}| \leq \binom{n}{k} - \binom{n-s}{k}$. И равенство выполнено только в том случае, когда \mathcal{F} – это объединение s звёзд.

Заметим, что константа $C(s)$ здесь зависит от s , поэтому результат интересен для постоянного s . Мы получаем тот же вывод в другом режиме, когда $n \geq 2k^2 - 2k + s$.

Теорема 3. 1. Пусть $n > 2k^2 - 2k + s$, и пусть \mathcal{F} является объединением s пересекающихся семейств. Тогда $|\mathcal{F}| \leq \binom{n}{k} - \binom{n-s}{k}$ для достаточно больших k . Равенство имеет место, только если \mathcal{F} является объединением s звезд.

2. Пусть $n \geq 2.6k^2$, и пусть \mathcal{F} является объединением s нетривиальных пересекающихся семейств, тогда $|\mathcal{F}| < \binom{n}{k} - \binom{n-s}{k}$ для достаточно больших k .

© Э. Л. Иноземцев, 2024

Обратим внимание на то, что результат теоремы 3 можно переформулировать в терминах раскрасок кнезеровского графа:

Теорема 4. Пусть $n > 2k^2 - 2k + s$. Тогда наибольшее подмножество вершин кнезеровского графа $KG_{n,k}$, которое можно правильно покрасить в s цветов, имеет размер не более чем $\binom{n}{k} - \binom{n-s}{k}$ для достаточно больших k .

Вторая задача, которую мы рассматриваем в этой работе, это вопрос о предписанных раскрасках кнезеровского графа. Предписанное хроматическое число $\text{ch}(G)$ графа G это такое минимальное s , что для любой задачи списков допустимых цветов $L(v)$ размера s вершинам $v \in G$ существует правильная раскраска вершин G , которая для каждой $v \in G$ использует цвет из $L(v)$.

В [3] Буланкина и Купавский доказали, что достаточно больших n и $3 \leq k \leq n^{\frac{1}{2} - \frac{1}{s}}$ верно, что

$$\text{ch}(KG_{n,k}) \geq \frac{1}{2s^2} n \log n.$$

В этой работе мы немного улучшаем эту оценку и показываем, что

Теорема 5. Пусть $4 \leq k \leq \left(\frac{n}{1+\varepsilon}\right)^{1/2}$ для любого $\varepsilon > 0$. Тогда

$$\text{ch}(KG_{n,k}) > \left(\frac{1}{8} - \varepsilon\right) \frac{n \log n}{2}$$

для достаточно больших n . В случае $k = 3$ выполнена оценка

$$\text{ch}(KG_{n,k}) > \left(\frac{1}{12} - \varepsilon\right) \frac{n \log n}{2}.$$

Список литературы

- [1] Frankl P., Füredi Z. Extremal problems concerning Kneser graphs // Journal of Combinatorial Theory, Series B. 1986. 40(3):270–284. doi: 10.1016/0095-8956(86)90084-5
- [2] Ellis D., Lifshitz N. On the union of intersecting families // Combinatorics, Probability and Computing. 2019. 28(6):826–839. doi: 10.1017/S096354831900004X
- [3] Bulankina V., Kupavskii A. Choice number of Kneser graphs // Discrete Mathematics. 2022. 345(11):113097. doi: 10.1016/j.disc.2022.113097

Обобщенная схема размещения: некоторые аспекты развития и применения

Андрей В. Колчин

ФГБОУ ВО «Московский автомобильно-дорожный государственный
технический университет (МАДИ)», Москва, Россия

E-mail: akolchin@madi.ru

Комбинаторика сыграла важную роль в начале развития теории вероятностей, и эти два раздела математики продолжают развиваться в тесном взаимодействии. В настоящее время теория вероятностей, предлагая новые подходы к решению задач дискретной математики, как бы отдает долги комбинаторике. Среди этих новых подходов отметим хорошо развитые в теории вероятностей методы асимптотического анализа, которые успешно используются при решении сложных комбинаторных задач. Комбинаторные задачи и методы занимают значительное место в исследованиях по теории вероятностей. Среди многочисленных работ в этой области можно выделить несколько направлений: комбинаторные задачи в теории случайных процессов, задачи, связанные со случайными отображениями и случайными графами, задачи размещения частиц по ячейкам.

Для решения широкого круга подобных комбинаторных задач весьма плодотворным оказывается *вероятностный подход* [1, 2]. Если распределение вероятностей задано на множестве рассматриваемых комбинаторных структур, то числовые характеристики этих структур можно рассматривать как случайные величины и анализировать их вероятностными методами. При таком вероятностном подходе мы автоматически ограничиваемся рассмотрением типичных структур, которые составляют основную массу рассматриваемого множества, и исключаем из рассмотрения небольшую долю структур с нестандартными свойствами.

Вероятностный подход впервые был использован в почти современном виде В. Л. Гончаровым, применившим его к изучению множества S_n всех подстановок степени n и серий в случайных $(0, 1)$ -последовательностях [3]. Среди тех, трудами которых развивалась вероятностная комбинаторика в России, были С. Н. Бернштейн, Н. В. Смирнов, В. Е. Степанов, ее успехи тесно связаны с блестящей Российской вероятностной школой, школой А. А. Маркова, П. Л. Чебышёва, А. М. Ляпунова, А. Я. Хинчина, А. Н. Колмогорова, Ю. В. Прохорова.

При анализе случайных структур успешно применялись разнообразные вероятностные методы; в частности, в вероятностной комбинаторике находит успешное применение *обобщенная схема размещения*, позволяющая сводить ряд комбинаторных задач к задачам о суммах независимых случайных величин, классическому объекту изучения в теории вероятностей. Обобщенная схема размещения была введена в 1968 г. В. Ф. Колчиным [4] и заняла заметное место в асимптотических исследованиях в вероятностной комбинаторике. Свое название эта схема получила в связи с тем, что она является обобщением классической задачи о случайном размещении частиц по ячейкам. Она оказалась удобным средством исследования таких интереснейших процессов, как эволюция случайных графов, случайных лесов, систем линейных уравнений со случайными коэффициентами, случайных подстановок, в том числе в связи с построением и анализом вычислительных алгоритмов [5]. В настоящее время активные исследования асимптотического поведения различных комбинаторных объектов с использованием обобщенной схемы размещения ведутся, в частности, Ю. Л. Павловым в Карельском научном центре РАН, А. Н. Чупруновым в Чувашском государственном университете [6] и И. Фазекашем в Дебреценском университете [7].

в обобщенной схеме размещения частиц распределение заполнений ячеек представимо как условное распределение независимых случайных величин при условии, что их сумма принимает фиксированное значение. Пусть η_1, \dots, η_N – неотрицательные целочисленные случайные величины, рассматриваемые как некоторые числовые характеристики комбинаторной структуры из N компонент, состоящей из n элементов, такие, что $\eta_1 + \dots + \eta_N = n$. Если существуют независимые случайные величины ξ_1, \dots, ξ_N такие, что совместное распределение η_1, \dots, η_N допускает представление

$$P\{\eta_1 = k_1, \dots, \eta_N = k_N\} = P\{\xi_1 = k_1, \dots, \xi_N = k_N \mid \xi_1 + \dots + \xi_N = n\},$$

где k_1, \dots, k_N – произвольные целые числа, то говорят, что η_1, \dots, η_N образуют обобщенную схему размещения с параметрами n и N и независимыми случайными величинами ξ_1, \dots, ξ_N . Случайные величины η_1, \dots, η_N интерпретируются как заполнения ячеек.

А. Н. Чупруновым активно изучаются обобщенные схемы размещения случайного числа K частиц по N ячейкам [6]. И. Фазекашем анализируются расширения обобщенных схем размещения, где в N ячеек размещаются либо по крайней мере n частиц, либо не более n частиц [7].

Изучение характеристик обобщенной схемы размещения по сути сводится к задачам о суммах независимых случайных величин. Однако в большинстве применений обобщенной схемы возникает необходимость в

локальных предельных теоремах в *схеме серий* при различном поведении параметров схемы [8–10].

Список литературы

- [1] *Erdős P., Spencer J. H.* Probabilistic Methods in Combinatorics. New York: Academic Press, 1974. 106 p.
- [2] *Alon N., Spencer J. H.* The Probabilistic Method. New York: Wiley–Interscience, 2000. 305 p. doi: 10.1002/9780470277331
- [3] *Гончаров В. Л.* Из области комбинаторики // Известия Академии наук СССР. Сер. математическая. 1944. 8(1):3–48.
- [4] *Kolčinas V.* Viena ribinių teoremų klasė sąlyginiams pasiskirstymams // Lietuvos Matematikos Rinkinys. 1968. 8(1):53–63.
- [5] *Колчин В. Ф., Чистяков В. П.* Комбинаторные задачи теории вероятностей // Итоги науки и техники. Сер. Теория вероятностей, математическая статистика, теоретическая кибернетика. 1974. 11:5–45.
- [6] *Чупрунов А. Н., Фазекаш И.* Об обобщенной схеме размещения со случайным числом частиц // Дискретная математика. 2012. 24(2):149–153. doi: 10.4213/dm1190
- [7] *Fazekás I., Porvásznyik B.* Some limit theorems for generalized allocation schemes // Miskolc Math. Notes. 2015. 16(2):817–832. doi: 10.18514/MMN.2015.1461
- [8] *Колчин А. В.* Предельные теоремы для обобщенной схемы размещения // Дискретная математика. 2003. 15(4):148–157. doi: 10.4213/dm224
- [9] *Колчин А. В., Безродный Б. Ф.* О некоторых аспектах развития обобщенной схемы размещения // Труды Карельского научного центра РАН. 2019. 7:21–29. doi: 10.17076/mat1050
- [10] *Kolchin A. V., Bezrodnyy B. F., Leeva M. A.* On some aspects of evolution of generalized allocation schemes // J. Math. Sci. 2022. 267(2):246–254. doi: 10.1007/s10958-022-06131-0

Рамсеевские множества в пространстве Минковского при радиальных раскрасках

Екатерина А. Костина, Андрей Б. Купавский

Московский физико-технический институт, Долгопрудный, Россия

E-mail: ekostina02@gmail.com

Пусть \mathcal{M} – метрическое пространство, содержащее хотя бы 2 точки. Хроматическим числом $\chi(\mathbb{R}_p^n, \mathcal{M})$ назовём наименьшее число цветов, необходимое для покраски всех элементов \mathbb{R}_p^n таким образом, чтобы не было одноцветных подпространств, изометричных \mathcal{M} . Метрическое пространство \mathcal{M} назовём ℓ_p -рамсеевским, если $\chi(\mathbb{R}_p^n, \mathcal{M}) \rightarrow \infty$ при $n \rightarrow \infty$. Случай $p = 2$ приводит к рассмотрению евклидовой теории Рамсея [1], в то время как $p = \infty$ порождает теорию Рамсея относительно шах-нормы [2].

В данной работе рассмотрены $\chi(\mathbb{R}_p^n, \mathcal{B}_k)$ при $p \in \mathbb{N}$, где \mathcal{B}_k - арифметическая прогрессия длины k , для специального вида раскрасок. Нормальными назовём раскраски \mathbb{R}_p^n , полученные из раскраски \mathbb{R}_+ в соответствующее количество цветов сопоставлением элементу $x \in \mathbb{R}_p^n$ цвета числа $\|x\|_p^p$. Хроматическое число относительно таких раскрасок будем обозначать $\chi_{norm}(\mathbb{R}_p^n, \mathcal{B}_k)$.

Для четных p были получены следующие результаты:

Теорема 1. Для $p \in 2\mathbb{N}$ и некоторой константы c выполнено $\chi(\mathbb{R}_p^n, \mathcal{B}_{p+1}) \leq c$ для любого $n \in \mathbb{N}$.

Теорема 2. Для $p \in 2\mathbb{N}$ и $k \leq p$ $\chi_{norm}(\mathbb{R}_p^n, \mathcal{B}_k) \rightarrow \infty$ при $n \rightarrow \infty$.

Обозначим элементы прогрессии через v_1, v_2, \dots, v_k . Будем считать, что разность прогрессии равна единичному вектору u . Пусть $x_i = \|v_i\|_p^p \forall i = 1..k$, также обозначим $s_q = \sum (v_1^j)^q (u^j)^{p-q}$. Тогда x_i можно воспринимать как линейные комбинации $s_q, q = 0..p$.

Для $k > p$ эти линейные комбинации будут зависимы, то есть существуют коэффициенты $c_i \in \mathbb{N}$ и константы $c \neq 0$ таких, что $\sum_{i=1}^{p+1} c_i x_i = c$.

По этим коэффициентам можно построить нормальные раскраски \mathbb{R}_p^n в константное число цветов, исключаяющие одноцветные изометрические копии \mathcal{B}_k .

Для прогрессий меньшей длины такой линейной зависимости x_i нет. Более того, при больших размерностях n для многих наборов x_1, \dots, x_k

существуют соответствующие арифметические прогрессии v_1, \dots, v_k . Используя это наблюдение, доказана теорема 2.

Аналогичным образом переходя к рассмотрению s_q для нечётных p задачу поиска одноцветных арифметических прогрессий в R_p^n можно свести к теореме ван дер Вардена. С помощью этого получен следующий результат:

Теорема 3. *Для нечётного $p \in \mathbb{N}$ и произвольного $k \in \mathbb{N}$ при $n \geq p + 2$ выполнено $\chi_{\text{norm}}(R_p^n, \mathcal{B}_k) = \infty$.*

Список литературы

- [1] Erdős P., Graham R.L., Montgomery P., Rothschild B. L., Spencer J., Straus E. G. Ramsey theorems I, II, III.
- [2] Frankl N., Kupavskii A., Sagdeev A. Max-norm Ramsey Theory // European Journal of Combinatorics. 2024. 118:103918. doi: 10.1016/j.ejc.2024.103918

Модулярность в модели посаженного разбиения

Михаил М. Кошелев

Московский физико-технический институт, Москва, Россия

E-mail: mkoshelev99@gmail.com

В данной работе мы получим верхние и нижние оценки модулярности в модели посаженного разбиения для случая равных частей. Начнем с определения данной модели.

Определение 1. *Зафиксируем натуральные числа k, n_1, \dots, n_k и действительные числа $p, q \in [0, 1]$. Тогда моделью посаженного разбиения называется случайный граф, в котором множество вершин есть множество пар $(i, j), i \in \{1, \dots, k\}, j \in \{1, \dots, n_i\}$, а ребра проводятся независимо с вероятностью p для вершин с совпадающей первой координатой и вероятностью q для вершин с различной первой координатой. При этом вершины $(i, 1), \dots, (i, n_i)$ называются i -м кластером.*

Отметим, что в некоторых работах под моделью посаженного разбиения понимается несколько видоизмененная модель, в которой распределение вершин по кластерам также является случайным (см., например, [1]).

Одной из наиболее популярных задач, связанных с моделью посаженного разбиения, является задача о восстановлении исходного разбиения на кластеры по графу со стертými метками (см., например, [2, 3]). В случае $p > q$ представляется разумным использование для этого алгоритма кластеризации, который бы разбил наш граф на подграфы с большой плотностью, которые будут являться хорошими приближениями исходных кластеров графа. В настоящее время разработано большое количество эвристических алгоритмов, решающих данную задачу, большинство из которых основано на таком свойстве графа, как его модулярность.

Впервые модулярность появилась в работе [4], и, несмотря на некоторые недочеты (см. [5]), быстро стала одной из самых популярных эвристик, используемых в алгоритмах кластеризации (см. [6, 7]). Дадим формальное определение модулярности. Начнем с модулярности разбиения.

Определение 2. *Пусть $\mathcal{A} = \{A_1, A_2, \dots, A_k\}$ – разбиение множества вершин (взвешенного) графа G . Модулярностью разбиения \mathcal{A} назовем величину*

$$q(\mathcal{A}) = \sum_{A \in \mathcal{A}} \frac{e(A)}{e(G)} - \sum_{A \in \mathcal{A}} \frac{(\sum_{v \in A} \deg(v))^2}{4e^2(G)},$$

где $e(A)$ – сумма весов ребер в индуцированном подграфе G с множеством вершин A , а $\deg(v)$ – сумма весов ребер, один из концов которых совпадает с v .

Теперь осталось определить модулярность графа G как максимум из модулярностей его разбиений.

Определение 3. Модулярностью графа G называется величина

$$q^*(G) = \max_{\mathcal{A} \text{ – разбиение } G} q(\mathcal{A}).$$

В данной работе мы получим верхние и нижние оценки модулярности графов в модели посаженного разбиения в случае равных размеров кластеров. Сформулируем основной результат работы.

Теорема 1. Пусть n_i, k_i – последовательность натуральных чисел, а $p_i, q_i \in [0, 1]$. Пусть также

$$n_i k_i \rightarrow \infty, p_i(n_i - 1) + q_i(n_i k_i - k_i) = \omega \left(\sqrt{n_i k_i} \right).$$

Тогда с вероятностью 1

$$\begin{aligned} \liminf_{i \rightarrow \infty} \frac{(n_i - 1)p_i}{p_i(n_i - 1) + q_i(n_i k_i - k_i)} - \frac{1}{k_i} &\leq \liminf_{i \rightarrow \infty} q^*(G(n_i, k_i, p_i, q_i)) \\ &\leq \liminf_{i \rightarrow \infty} \frac{\max\{p_i, |(n_i - 1)p_i - n_i q_i|\}}{p_i(n_i - 1) + q_i(n_i k_i - k_i)} \end{aligned}$$

и

$$\begin{aligned} \limsup_{i \rightarrow \infty} \frac{(n_i - 1)p_i}{p_i(n_i - 1) + q_i(n_i k_i - k_i)} - \frac{1}{k_i} &\leq \limsup_{i \rightarrow \infty} q^*(G(n_i, k_i, p_i, q_i)) \\ &\leq \limsup_{i \rightarrow \infty} \frac{\max\{p_i, |(n_i - 1)p_i - n_i q_i|\}}{p_i(n_i - 1) + q_i(n_i k_i - k_i)}. \end{aligned}$$

Для доказательства теоремы нам потребуется перенести некоторые утверждения из спектральной теории графов на случай взвешенных графов. Кроме того, мы докажем общее утверждение о связи модулярности в случайных графах с модулярностью взвешенных графов. Для его формулировки нам понадобится дать вспомогательное определение.

Определение 4. Пусть G взвешенный граф, сумма весов ребер, выходящих из каждой его вершины равна d , а матрица W , составленная из весов его ребер имеет собственные значения $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. Тогда будем говорить, что $G - (n, d, \lambda)$ -взвешенный граф, где $\lambda = \max\{\lambda_2, -\lambda_n\}$.

Теорема 2. Пусть H_n – последовательность (n, d_n, λ_n) -взвешенных графов с весами в $[0, 1]$, а H'_n – случайные графы на n вершинах, в которых ребра проводятся независимо и с вероятностями, совпадающими с их весом в H_n . Пусть также

$$d_n = \omega(n^{1/2}), \quad \limsup_{n \rightarrow \infty} \frac{\lambda_n}{d_n} < 1,$$

тогда

$$|q^*(H_n) - q^*(H'_n)| \xrightarrow{n, n} 0, \quad n \rightarrow \infty.$$

Список литературы

- [1] *Bickel P. J., Chen A.* A nonparametric view of network models and Newman–Girvan and other modularities // Proc. Nat. Acad. Sci. India Sect. A. 2009. 106:21068.
- [2] *Abbe E., Sandon C.* Community Detection in General Stochastic Block models: Fundamental Limits and Efficient Algorithms for Recovery // IEEE 56th Annual Symposium on Foundations of Computer Science. 2015. P. 670–688.
- [3] *Lei J., Rinaldo A.* Consistency of spectral clustering in stochastic block models // Ann. Statist. 2015. 43(1):215–237.
- [4] *Newman M. E. J., Girvan M.* Finding and evaluating community structure in networks // Phys. Rev. E. 2004. 69:26–113.
- [5] *Fortunato S., Barthelemy M.* Resolution limit in community detection // Proceedings of the National Academy of Sciences. 2007. 104(1):36–41.
- [6] *Clauset A., Newman M. E. J., Moore C.* Finding community structure in very large networks // Phys. Rev. E. 2004. 70:066111.
- [7] *Newman M. E. J.* Fast algorithm for detecting community structure in networks // Phys. Rev. E. 2004. 69:066133.

Построение нормального распределения по статистическим данным о COVID-19

Галина А. Курина, Мария Ю. Зиновьева

Воронежский государственный университет, Воронеж, Россия

E-mail: kurina@math.vsu.ru; shkondamari@mail.ru

В докладе приводятся результаты проверки выдвинутой в [1] гипотезы о нормальном распределении количества пострадавших во время пандемии COVID-19 на примере второй волны в Канаде, в качестве которой берется промежуток времени с 03.11.2020 по 13.02.2021. Количество заразившихся в день во вторую волну изменяется в пределах промежутка [1; 11383], количество умерших за день – в пределах промежутка [0; 257], а количество выздоровевших за день – в пределах промежутка [2; 14030].

Используются данные по дням для числа заразившихся, умерших и выздоровевших в Канаде, представленные ранее на сайте <https://horoshotam.ru/kanada/coronavirus>. К сожалению, в настоящее время ежедневные данные на этом сайте заменили данными по неделям. Схема исследования та же, что и в [1], где детально изучалась первая волна заболеваемости в Канаде, пришедшая на март – июнь 2020 года.

Как указано в [2], один из основоположников медицинской статистики Уильям Фарр, по-видимому, был первым, кто установил, что во время эпидемии оспы в Англии и Уэльсе в 1837–1839 годах график числа смертей за квартал следовал примерно нормальной кривой.

Для построения кривой нормального распределения область значений рассматриваемой случайной величины (соответственно число заболевших, умерших или выздоровевших) разбивается на «оптимальное» число k равных интервалов, определяемое по эвристической формуле Стёрджесса [3]

$$k = 1 + [\log_2 n],$$

где $[x]$ означает целую часть числа x , n – объем выборки. Если число значений случайной величины, попавших в некоторый интервал, мало, то интервалы объединяются.

Для вычисления выравнивающих (теоретических) частот n'_i используется известная формула

$$n'_i = n(\Phi(t_{2i}) - \Phi(t_{1i})),$$

где

$$t_{ji} = (x_{ji} - \bar{x}_B) / S,$$

\bar{x}_B – выборочная средняя, S – исправленное выборочное среднее квадратическое отклонение, $x_{ji}, j = 1, 2$ – концы i -го промежутка разбиения, и Φ – функция Лапласа.

При проведении вычислений использовались встроенные функции приложения Microsoft Office Excel.

Обозначим через x_i середину i -го промежутка разбиения. По точкам с координатами (x_i, n'_i) строится кривая нормального распределения. Точки с координатами (x_i, n_i) , где n_i – эмпирические частоты, изображаются на графиках синим цветом.

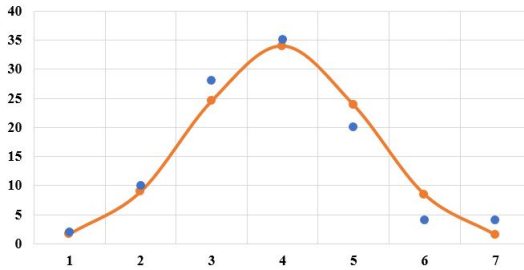


Рис. 1: Кривая нормального распределения, построенная по данным о числе заразившихся в Канаде во вторую волну

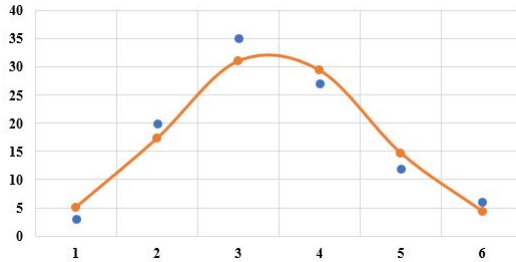


Рис. 2: Кривая нормального распределения, построенная по данным о числе умерших в Канаде во вторую волну

Для проверки гипотезы о предполагаемом нормальном законе распределения используется критерий согласия Пирсона. Согласно этому критерию при уровне значимости $\alpha = 0,05$ для второй волны в Канаде нет

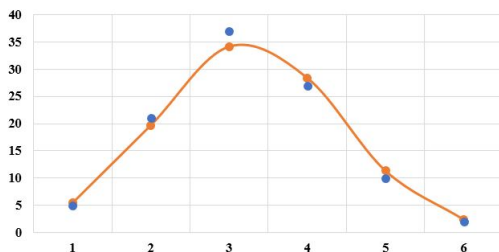


Рис. 3: Кривая нормального распределения, построенная по данным о числе выздоровевших в Канаде во вторую волну

оснований отклонить гипотезу о нормальном распределении количества заразившихся, умерших и выздоровевших.

Кроме Канады, были исследованы данные по трём рассматриваемым показателям для нескольких волн в Болгарии, Бразилии, Великобритании, Германии, Доминиканской республике, Индии, Италии, Кении, Китае, Норвегии, России, США, Турции, Франции, Чили, Швейцарии. В подавляющем большинстве случаев при уровне значимости $\alpha = 0,05$ в силу критерия согласия Пирсона нет оснований отвергнуть гипотезу о нормальном распределении отдельного показателя.

Список литературы

- [1] *Курина Г. А., Зиновьева М. Ю., Золотарева Е. А.* Построение нормального распределения по данным СМИ о COVID-19 // Вестник ВГУ. Серия: Системный анализ и информационные технологии. 2023. 3:134–142. doi: 10.17308/sait/1995-5499/2023/3/134-142
- [2] *Martin W. J.* The epidemic curve of smallpox // The Journal of Hygiene. 1934. 34(1):10–29.
- [3] *Sturges H. A.* The choice of a class interval // Journal of the American Statistical Association. 1926. 21(153):65–66.

Методы Монте-Карло по схеме марковских цепей для обучения факторизационных машин

Арина О. Маслова¹, Олег В. Лукашенко²

¹ Петрозаводский государственный университет, Петрозаводск, Россия

² Институт прикладных математических исследований КарНЦ РАН,
Петрозаводск, Россия

E-mail: ari.maslova03@yandex.ru; lukashenko@krc.karelia.ru

Пусть дана обучающая выборка $\mathcal{D} = \{\mathbf{x}_i, y_i\}_{i=1}^n$, где $\mathbf{x}_i = (x_1, \dots, x_p)^T \in \mathbb{R}^p$ – вектор признаков, $y_i \in \{0, 1\}$ – целевая переменная. Рассматривается следующая модель для решения задачи бинарной классификации:

$$\begin{aligned} y_i &\sim \text{Ber}(p_i), \\ p_i &= F(y(\mathbf{x}_i)), \quad i = 1, \dots, n, \end{aligned}$$

где $F(x) = \frac{1}{1+e^{-x}}$; функция $y(\mathbf{x})$ – модель факторизационных машин [1], которая представляет собой специальный вариант линейной модели со взаимодействиями факторов, где веса взаимодействия определяются скалярными произведениями низкоразмерных векторов $\mathbf{v}_i, \mathbf{v}_j \in \mathbb{R}^k$:

$$y(\mathbf{x}) = w_0 + \sum_{j=1}^p w_j x_j + \sum_{i=1}^p \sum_{j=i+1}^p (\mathbf{v}_i, \mathbf{v}_j) x_i x_j.$$

Модель факторизационных машин впервые была предложена в работе [1]. В статье [2] для обучения данной модели был предложен байесовский подход, суть которого состоит в задании априорных распределений на параметры модели и последующем анализе апостериорного распределения. В этой же статье были разработаны ряд алгоритмов для вычисления оценок, в том числе на основе методов Монте-Карло по схеме марковских цепей (МСМС). В настоящей работе рассматривается применение специального варианта схемы Гиббса (Gibbs sampling) с дополнительными переменными, имеющими распределение Пуа–Гамма (Poisson–Gamma) [3]. Проведены численные эксперименты на синтетических и реальных данных, иллюстрирующие качество предложенного подхода.

Список литературы

- [1] *Rendle S.* Factorization machines // In Proceedings of the 10th IEEE International Conference on Data Mining. IEEE Computer Society. 2010. P. 995–1000.
- [2] *Freudenthaler C., Schmidt-Thieme L., Rendle S.* Bayesian factorization machines // In Proceedings of the NIPS Workshop on Sparse Representation and Low-rank Approximation. 2011. P. 995–1000.
- [3] *Polson N. G., Scott J. G., Windle J.* Bayesian Inference for Logistic Models Using Pólya–Gamma Latent Variables // Journal of the American Statistical Association. 2013. 108(504):1339–1349.

Теоретико-игровые подходы к анализу загруженности городских дорог на примере г. Петрозаводск

Наталия Н. Никитина

Институт прикладных математических исследований КарНЦ РАН,
Петрозаводск, Россия

E-mail: nikitina@krc.karelia.ru

Одной из актуальных прикладных задач является задача моделирования и анализа транспортной системы города или региона. Как правило, транспортная система моделируется при помощи графа, в котором ребра – транспортные пути – соединяют коммуникационные узлы – перекрестки, остановки маршрутного транспорта, склады, железнодорожные станции, аэропорты и т. п.

Теоретико-игровые методы позволяют моделировать дорожные ситуации с участием нескольких лиц (например, автомобилей, пассажиров общественного транспорта, фирм-грузоперевозчиков и др.), которые совместно используют дорожную сеть и имеют те или иные личные предпочтения в выборе маршрутов. На основе смоделированного поведения игроков можно анализировать загруженность транспортных магистралей, прогнозировать возникновение пробок при авариях, появлении новых районов города или введении новых маршрутов, находить оптимальные маршруты общественного транспорта в случае как кооперативного, так и некооперативного (эгоистичного) поведения игроков.

В данной работе исследуется загруженность городских дорог в равновесиях по Нэшу – дорожных ситуациях, устойчивых относительно единичных отклонений игроков, и в равновесиях по Вардропу – распределениях участников движения между маршрутами с равными задержками на всех используемых маршрутах. Представленные математические модели апробированы на дорожном графе г. Петрозаводск [1].

Рассмотрим дискретную игру заполнения между игроками n типов $N = \{N_1, \dots, N_n\}$, число которых составляет $|N_t| = n_t, t = 1, \dots, n$. Игроки выбирают ресурсы из множества $M = \{1, \dots, m\}$. С каждым ресурсом j связаны функции задержек $c_j^{(1)}(k), \dots, c_j^{(n)}(k)$, где верхний индекс (t) – тип игрока.

Стратегией игрока $i \in N$ является выбор подмножества ресурсов. Пусть профиль выбранных стратегий $s = \{s_i, i \in N\}$. Каждому профилю стратегий s соответствует вектор заполнения $k(s) = (k_1(s), \dots, k_m(s))$, $k_j(s) = k_j^{(1)}(s) + \dots + k_j^{(n)}(s)$, $j \in M$. Здесь $k_j^{(t)}(s)$ – число игроков типа t , выбравших ресурс j . Имеет место соотношение $\sum_{j=1}^m k_j^{(t)}(s) = n_t$ для $t = 1, \dots, n$.

Выигрыш игрока зависит от типа данного игрока и стратегий, выбранных остальными игроками, и строится следующим образом. При выборе ресурса $j \in M$ игроки первого типа получают выигрыш $-c_j^{(1)}(k)$, где k – общее число игроков, выбравших данный ресурс j . Игроки второго типа получают выигрыш $-c_j^{(1)}(k) - c_j^{(2)}(k_j^{(2)} + \dots + k_j^{(n)})$, где k – общее число игроков, выбравших данный ресурс j , и $k_j^{(t)}$ – число игроков типа t , выбравших данный ресурс j ($t = 2, \dots, n$). Аналогичным образом строятся функции выигрыша игроков всех остальных типов. Наконец, игроки n -го типа получают выигрыш $-c_j^{(1)}(k) - c_j^{(2)}(k_j^{(2)} + \dots + k_j^{(n)}) - \dots - c_j^{(n)}(k_j^{(n)})$. Выигрыш здесь берется со знаком минус, поскольку c_j^t представляют собой затраты игрока типа t .

Таким образом, для профиля стратегий $s = \{s_i, i \in N\}$ выигрыш i -го игрока, имеющего тип t , имеет вид

$$H_i^{(t)} = - \sum_{j \in s_i} \sum_{p=1}^t c_j^{(p)} \left(\sum_{l=p}^n k_j^{(l)}(s) \right).$$

Для краткости обозначим общее число игроков, использующих ресурс j в профиле стратегий s и имеющих тип не младше p , как $K_j^{(p)}(s) = \sum_{l=p}^n k_j^{(l)}(s)$. Доказано, что данная игра является потенциальной, потенциал имеет вид

$$P(s) = - \sum_{j \in M} \sum_{p=1}^n \sum_{k=1}^{K_j^{(p)}(s)} c_j^{(p)}(k),$$

выполняется свойство конечности улучшений (FIP), и алгоритм последовательных улучшений сходится к равновесию по Нэшу в чистых стратегиях. Для нахождения равновесия по Нэшу возможно также применение различных аналитических и численных методов.

Для фрагмента дорожной сети г. Петрозаводск найдены равновесия по Нэшу для смешанных потоков транспорта (легковые автомобили, автобусы и грузовики) при различных характеристиках транспортного потока.

Предложены способы регулирования загруженности отдельных участков дороги без существенных ухудшений выигрышей игроков.

Рассмотрим граф городской сети, в котором установилась ситуация равновесия по Вардропу. В такой ситуации ни один участник дорожного движения не может сократить свое время в пути, единолично отклонившись от выбранного маршрута. Известно, что при модификациях транспортной сети равновесие по Вардропу может нарушаться, и для достижения нового равновесия потребуется длительное время для перераспределения игроков по другим маршрутам. Однако на практике не менее важно оценивать последствия тех модификаций транспортной сети, при которых не происходит полного изменения равновесия. Это может быть, например, кратковременное перекрытие участка дороги из-за аварии. В таком случае исходное распределение игроков по маршрутам можно использовать для вычисления различных характеристик ребер транспортного графа, например, центральности. Центральность представляет собой вещественнозначную функцию на множестве ребер и выражает потенциальную важность ребра с точки зрения структурных свойств графа, позволяя ранжировать участки дорог.

В данной работе представлена мера центральности, при вычислении которой используется исходное равновесие по Вардропу, что позволяет учесть не только структурные свойства графа, но и фактическое распределение участников дорожного движения. Показано, что полученные характеристики ребер согласуются с данными о загруженности дорог, и центральности ребер могут также быть использованы в качестве начальных данных для вычисления равновесия по Вардропу.

Работа выполнена при поддержке гранта Российского научного фонда № 22-11-20015, проводимого совместно с органами власти Республики Карелия с финансированием из Фонда венчурных инвестиций Республики Карелия (ФВИ РК).

Список литературы

- [1] *Ermolin N. A., Khitraya V. A., Khitryi A. V., Mazalov V. V., Nikitina N. N.* Modeling of the City's Transport Network Using Game-Theoretic Methods on the Example of Petrozavodsk // Contributions to Game Theory and Management. 2022. 15:18–31. doi: 0.21638/11701/spbu31.2022.02

Sizes of trees in a Galton-Watson forest in the critical case with infinite variance

Yury L. Pavlov

Institute of Applied Mathematical Research of KRC RAS,
Petrozavodsk, Russia

E-mail: pavlov@krc.karelia.ru

We consider Galton-Watson forests with N rooted trees and n nonroot vertices. Assume that the forest-generating Galton-Watson branching process G starts with N particles and the random variable ξ equal to the number of offspring of each particle has the distribution

$$\mathbf{P}\{\xi = k\} = \frac{h(k+1)}{(k+1)^\tau}, \quad k = 0, 1, 2, \dots, \quad \tau \in (2, 3), \quad (1)$$

where $h(x)$ is a slowly varying function at infinity. Let the process G is critical, i. e. $\mathbf{E}\xi = 1$. Relation (1) imply that that the random variable ξ has infinite variance. Denote by η the size of maximal tree in the forest and by μ_r the number of trees with r vertices. In [1–3] limit theorems were proved for the η and μ_r in three zones in which N and n tend to infinity: $n/N \rightarrow 0$, $0 < C_1 \leq n/N \leq C_2 < \infty$, $n/N \rightarrow \infty$. But in the last zone the limit distributions of μ_r were obtained only in $n/N \rightarrow \infty$ with a small convergence rate. Below we give the new limit theorems for μ_r . Let $F(z)$ be the generating function of the distribution (1). We are concerned with those function $F(z)$ of the form

$$F(z) = z + (1-z)^{\tau-1}L(1-z),$$

where $L(x)$ is a slowly varying function at zero and

$$\lim_{x \rightarrow 0} L(x) = C,$$

where C is a positive constant.

It is easy to verify that the distribution (1) belongs to the zone of attraction of a stable law with exponent $\tau - 1$. We assume also that it is zone of normal attraction. Denote ν the random variable equal to the total number of particles in the process G until its extinction if $N = 1$ and let $q_k = \mathbf{P}\{\nu = k\}$, $k = 1, 2, \dots$

Theorem 1. Let $N, n \rightarrow \infty$ in such a way that $n/N^{\tau-1} \rightarrow \infty$. Then, for positive integers k

$$\mathbf{P}\{\mu_r = k\} = \frac{1 + o(1)}{\sqrt{2\pi N q_r (1 - q_r)}} e^{-u_r^2/2}$$

uniformly with respect to $u_r = (k - Nq_r)/\sqrt{Nq_r(1 - q_r)}$ in any fixed finite interval.

Theorem 2. Let $N, n, r \rightarrow \infty$ in such a way that $n/N^{\tau-1} \rightarrow \infty$. Then, for positive integers k

$$\mathbf{P}\{\mu_r = k\} = \frac{1}{k!} (Nq_r)^k e^{-Nq_r} (1 + o(1))$$

uniformly with respect to $u_r = (k - Nq_r)/\sqrt{Nq_r}$ in any fixed finite interval.

The idea for proving Theorems 1 and 2 is based on generalized particle allocation scheme proposed and studied by V. F. Kolchin [4].

This work was supported by the federal budget of the Russian Federation within the state assignment of the Rarelian Research Centre of the Russian Academy of Sciences.

References

- [1] *Pavlov Yu. L.* The maximum tree of a random forest in the configuration graph // Sbornik: Mathematics. 2021. 212(9):1329–1346. doi: 10.1070/SM9481
- [2] *Pavlov Yu. L., Cheplyukova I. A.* Sizes of trees in a random forest and configuration graphs // Branching processes and related topics. Trudy Math. Inst. Steklova. 2022. 326:298–315. doi: 10.4213/tm4216
- [3] *Pavlov Yu. L.* On the maximal Galton-Watson forest tree with infinite variance of the offspring // Discrete mathematics. 2023. 35(2):78–92. (in Russian). doi: 10.4213/dm1765
- [4] *Kolchin V. F.* Random graphs. Cambridge: Cambridge Univ. Press, 1999. 252 p. doi: 10.1017/CBO9780511721342

Статистическое моделирование динамики популяций в некоторых задачах иммунологии

Николай В. Перцев

Институт математики им. С. Л. Соболева СО РАН, Новосибирск, Россия
Институт вычислительной математики им. Г. И. Марчука РАН,
Москва, Россия
E-mail: homlab@ya.ru

Представлен подход к построению и численному исследованию простейшей стохастической модели динамики популяции в непрерывно-дискретной постановке с немарковскими ограничениями для индивидуумов. Приведенный подход используется для разработки стохастических моделей, возникающих в задачах иммунологии.

Формализация простейшей модели такова. Изучается популяция индивидуумов A . Символ A означает некоторого или произвольно выбранного индивидуума популяции A . Полагаем, что D и B означают популяции, в которые поступают индивидуумы A после завершения своего пребывания в популяции A . Популяция D может отражать погибших индивидуумов, популяция B – следующую стадию развития индивидуумов популяции A . Пусть вещественная переменная t означает время, $[0; T_{mod}]$ – промежуток моделирования, $(t; t + h) \subset [0; T_{mod}]$ – бесконечно малый промежуток времени, $h \rightarrow +0$. Под $\rho_A(z)$ будем понимать функцию от целочисленной неотрицательной переменной $z \in Z_+$, удовлетворяющую следующему условию: $0 < \rho_A(z) \leq \rho_A^*$ для всех $z \in Z_+$, где $\rho_A^* > 0$ – некоторая константа. Обозначим через $A(t)$ численность популяции A в момент времени t , $t \in [0; T_{mod}]$. В начальный момент времени $t = 0$ популяция A либо не содержит, либо содержит некоторое количество первоначально существующих индивидуумов, $A(0) = A_0 = const \geq 0$. Для каждого фиксированного $t > 0$ под $A(t)$ понимается неотрицательная целочисленная случайная величина.

Зафиксируем t и положим, что $A(t) = x$ – неотрицательная целочисленная константа. Введем набор постулатов модели.

Н1. Независимо от предшествующих моменту времени t событий в течение промежутка $(t; t + h)$ с вероятностью $\rho_A(x)h + o(h)$ популяция A пополняется одним индивидуумом, поступающим из S ; вероятность поступления из S в популяцию A более одного индивидуума за $(t; t + h)$

равна $o(h)$; популяция A не пополняется из S за $(t; t+h)$ с вероятностью $1 - \rho_A(x)h + o(h)$.

Н2. Пусть индивидуум \mathcal{A} появился в популяции A в некоторый момент времени $t_{\mathcal{A}} \leq t$. Положим, что $\xi_{\mathcal{A}}$ означает экспоненциально распределенную случайную величину с параметром $\mu_A > 0$, описывающую продолжительность времени до перехода \mathcal{A} в популяцию D . Через $\omega_{\mathcal{A}}$ обозначим длительность пребывания \mathcal{A} в популяции A до перехода этого индивидуума в популяцию B . Величина $\omega_{\mathcal{A}}$ задается одним из трех способов: 1) $\omega_{\mathcal{A}} = \omega = \text{const} > 0$, 2) $\omega_{\mathcal{A}} = \omega(t)$, где $\omega(t)$ – положительная, ограниченная сверху функция, такая, что функция $t + \omega(t)$ является монотонно возрастающей, 3) $\omega_{\mathcal{A}}$ – случайная величина, распределенная на конечном промежутке $[0; \omega_*]$, $\omega_* > 0$, с функцией распределения $F_{\omega_A}(a)$, $F_{\omega_A}(+0) = 0$. Случайные величины $\xi_{\mathcal{A}}$, $\omega_{\mathcal{A}}$ являются независимыми между собой, не зависят от $t_{\mathcal{A}}$ и не зависят от поведения остальных индивидуумов популяции. Индивидуум \mathcal{A} покидает популяцию A в момент времени $t_{\mathcal{A}} + \min\{\xi_{\mathcal{A}}, \omega_{\mathcal{A}}\}$, переходя в D или B .

Пусть $N_A(t)$ – количество индивидуумов, поступивших в популяцию A за промежуток времени $(-\omega_*; t]$, $t \in [0; T_{mod}]$; $j = 1, 2, \dots, N_A(t)$ – порядковый номер поступления очередного индивидуума \mathcal{A} в популяцию A , $\mathcal{A}(j)$ – индивидуум популяции A с порядковым номером j . Зафиксируем $t \in [0; T_{mod}]$ и введем $\Omega_A(t)$ – семейство уникальных типов индивидуумов популяции A . Если $N_A(t) = 0$, то $\Omega_A(t) = \emptyset$. Пусть $N_A(t) \geq 1$. Тогда $\Omega_A(t) = \left\{ (t_{\mathcal{A}(j)}, \varphi_{\mathcal{A}(j)}, \eta_{\mathcal{A}(j)}) : t_{\mathcal{A}(j)} \leq t, 1 \leq j \leq N_A(t) \right\}$, где $t_{\mathcal{A}(j)}$ – момент поступления $\mathcal{A}(j)$ в популяцию A , $\varphi_{\mathcal{A}(j)} = \min\{\xi_{\mathcal{A}(j)}, \omega_{\mathcal{A}(j)}\}$ – длительность пребывания $\mathcal{A}(j)$ в популяции A , $\eta_{\mathcal{A}(j)}$ – индикатор «судьбы» $\mathcal{A}(j)$:

если $\xi_{\mathcal{A}(j)} \leq \omega_{\mathcal{A}(j)}$, то $\eta_{\mathcal{A}(j)} = 0$ (переход в D),

если $\xi_{\mathcal{A}(j)} > \omega_{\mathcal{A}(j)}$, то $\eta_{\mathcal{A}(j)} = 1$ (переход в B).

Динамику популяции A опишем с помощью случайного процесса

$$H(t) = (A(t), N_A(t), \Omega_A(t)), \quad t \in [0; T_{mod}],$$

полагая, что $A(0) = N_A(0) = A_0$ – заданная константа,

$$\Omega_A(0) = \Omega_A^{(0)}, \quad \text{если } N_A(0) \geq 1,$$

$$\Omega_A(0) = \emptyset, \quad \text{если } N_A(0) = 0,$$

семейство $\Omega_A^{(0)}$ содержит фиксированный набор «троек», отражающих первоначально существующих индивидуумов.

Для численного исследования динамики популяции A построен алгоритм моделирования на основе метода Монте-Карло. Алгоритм моделирования предусматривает изменения пар $(t_m, H(t_m))$, $m = 0, 1, 2, \dots$, $t_m < T_{mod}$, где $t_0 = 0$ – начальный момент времени, $H(t_0)$ – начальное состояние популяции, $t_m > 0$ – моменты очередного скачкообразного изменения компонент процесса $H(t)$. При реализации алгоритма на ЭВМ использованы методы и датчики псевдослучайных чисел, приведенные в [1].

Одно из приложений описанного подхода связано с разработкой непрерывно-дискретных стохастических моделей в задачах иммунологии. В качестве примера приведены результаты вычислительных экспериментов в рамках двух моделей. Первая модель описывает перемещение клеток и вирусных частиц по лимфатической системе человека с учетом нестационарных длительностей перемещения между лимфоузлами [2]. Вторая модель демонстрирует начальный этап развития ВИЧ-1 инфекции в лимфоузле при заражении CD4+ Т-лимфоцитов [3]. Каждая из двух моделей допускает вспомогательное аналитическое исследование. Результаты аналитических исследований используются при планировании вычислительных экспериментов со стохастическими моделями.

Работа выполнена при поддержке Российского научного фонда, проект 23-11-00116.

Список литературы

- [1] Михайлов Г. А., Войтишек А. В. Численное статистическое моделирование. Методы Монте-Карло. М.: Академия, 2006. 368 с.
- [2] Логинов К. К., Перцев Н. В., Тончий В. А. Стохастическое моделирование компартментных систем с трубками // Математическая биология и биоинформатика. 2019. 14(1):188–203. doi: 10.17537/2019.14.188
- [3] Pertsev N. V., Loginov K. K. Stochastic Modeling in Immunology Based On a Stage-Dependent Framework with Non-Markov Constraints for Individual Cell and Pathogen Dynamics // Mathematical Biology and Bioinformatics. 2023. 18(2):543–567. doi: 10.17537/2023.18.543

О некоторых результатах исследования графа цитирования статей российских математических журналов

Андрей А. Печников

Институт прикладных математических исследований КарНЦ РАН,
Петрозаводск, Россия

E-mail: pechnikov@krc.karelia.ru

В докладе рассматривается ряд вопросов, связанных с анализом библиографического цитирования научных публикаций. Одной из формальных моделей цитирования является граф цитирования, в котором вершинами служат статьи, а дуги отражают факт цитирования. В нашем графе дуга $(p1, p2)$, связывающая вершины, соответствующие статьям $p1$ и $p2$, существует тогда, когда в статье $p1$ имеется хотя бы одна библиографическая ссылка на статью $p2$.

Сама идея такого подхода не нова, но даже трудно представить, что ей почти 60 лет [1]. При этом накопление библиографической информации и появление новых баз данных способствует не только применению уже разработанных методов к новым объектам исследования, но и появлению новых вопросов, а значит и новых результатов.

Исследование базируется на данных о 56 российских журналах и некоторых журналах стран ближнего зарубежья, которые индексируются в информационной системе портала Math-Net.Ru с 2000 по 2020 годы полным или значительным объемом выпусков (это почти 62 тысячи научных статей, авторами которых являются около 33 тысяч ученых) [2]. Граф цитирования статей, построенный по этим данным, содержит 40,9 тысяч вершин, имеющих ненулевую инцидентность, и 79 тысяч дуг. Сразу можно сделать вывод, что около 21 тысячи статей не имеют входящих и исходящих ссылок. Максимальная компонента связности содержит почти 28 тысяч вершин и 64 тысяч дуг (вторая по размерности компонента содержит всего 78 вершин). Обозначим граф, являющийся максимальной компонентой, $G(V, E)$.

Анализ графа $G(V, E)$ стандартными средствами показывает достаточно предсказуемые результаты. Например, распределение входящих степеней вершин не соответствует дискретному степенному закону: из 28 тысяч вершин 8,5 тысяч не имеют входящих ссылок, 6666 вершин имеют

по одной ссылке, 4 тысячи – по 2. Впрочем, этому есть и частичное содержательное объяснение: вершины в графе имеют различный «возраст» от 0 до 20 лет.

В работе [3] показано, что для журналов, проиндексированных в MathNet.Ru, время полужизни библиографических ссылок равно 8 годам. «Время полужизни» цитируемости (или более точно «медиана хронологического распределения ссылок») определяется следующим образом: это такой момент времени, в который половина рассматриваемых ссылок относится к статьям, опубликованным позднее медианы, а другая половина – к статьям, более ранним, чем медиана.

С учетом этого, для того, чтобы частично «сгладить» возрастную разницу вершин было выполнено следующее преобразование графа цитирования статей:

1. сформировано множество вершин V_{12} удалением из множества вершин V всех вершин, соответствующих статьям, опубликованным после 2012 года;
2. сформировано множество дуг E_{12} , для чего из множества E были выбраны все дуги, являющиеся входящими по отношению к вершинам множества V_{12} ,
3. во множество V_{12} добавлены все вершины, инцидентные дугам из E_{12} и не вошедшие в V_{12} на шаге 1.

Построенный граф обозначим $G_{12}(V_{12}, E_{12})$. Он содержит 24 тысячи вершин и более 45 тысяч дуг.

В отличие от графа $G(V, E)$ граф $G_{12}(V_{12}, E_{12})$ распадается на более чем 400 компонент связности, максимальная из которых по-прежнему большая (20,7 тыс. вершин), а вторая по размерности компонента содержит 273 вершины. Распределение входящих степеней вершин по-прежнему не соответствует дискретному степенному закону.

Один из вопросов, рассмотренных на этом графе, сформулируем так: какова доля статей в журнале, имеющих хотя бы одну входящую ссылку, по отношению ко всем статьям журнала, опубликованным за 2000-2012 годы? Обозначим эту долю *fract*.

Для 19 журналов, полностью проиндексированных за исследуемый временной период, наибольшее значение *fract*=0,61 имеет журнал «Математический сборник», а наименьшее *fract*=0,18 – «Вестник Московского университета. Серия 1: Математика. Механика».

Поскольку оба этих журнала являются известными высокорейтинговыми российскими журналами, входят в перечень ВАК, RSCI и различные международные базы, возникает вопрос об адекватности полученных результатов.

Для анализа был взят так называемый «белый список» Межведом-

ственной рабочей группы Минобрнауки России [4]. Каково же было удивление автора, когда оказалось, что «Математический сборник» в «белом списке» имеет уровень (*level*) 1 (наивысший из четырех), а «Вестник Московского университета. Серия 1: Математика. Механика» - уровень 3, как бы полностью подтверждая соотношение *fract*.

Из 56 рассматриваемых журналов 44 входят в "белый список". Корреляция между значениями *fract* и *level* для этих журналов достаточно заметная.

Однако большее значение полученные результаты имеют для содержательного анализа «выбросов». К примеру, журнал «Regular and Chaotic Dynamics» имеет в «белом списке» первый уровень, а его значение *fract* всего 0,29. Связано ли это с тем, что журнал издается только на английском языке? Или он посвящен достаточно узким вопросам, и поэтому имеет относительно мало ссылок с других российских математических журналов?

Представляется, что ответы на такие вопросы позволяют более точно принимать экспертные решения по официальным оценкам научной деятельности.

Автор приносит благодарность разработчикам портала Math-Net.Ru за любезно предоставленные данные.

Список литературы

- [1] Price D. J. D. S. Networks of scientific papers: The pattern of bibliographic references indicates the nature of the scientific research front // Science. 1965. 149(3683):510–515.
- [2] Знаменская Е. А., Печников А. А., Чебуков Д. Е. Анализ соавторства в математических журналах Math-Net.Ru // Научный сервис в сети Интернет: труды XXIV Всероссийской научной конференции. М.: ИПМ им. М. В. Келдыша. 2022. С. 190–202. doi: 10.20948/abrau-2022-5
- [3] Печников А. А., Чебуков Д. Е. Исследование библиографических ссылок в Math-Net.Ru с использованием графа цитирования журналов // Электронные библиотеки. 2021. 24(5):923–943. doi: 10.26907/1562-5419-2021-24-5-923-943
- [4] «Белый список» научных журналов. URL: <https://journalrank.rcsi.science/ru/> (дата обращения 05.02.2024).

О моделировании односерверных систем обслуживания с устареванием информации

¹Ирина В. Пешкова, ²Андрей С. Алексеев

¹Петрозаводский государственный университет, Петрозаводск, Россия;

²Институт прикладных математических исследований КарНЦ РАН,
Петрозаводск, Россия

E-mail: iaminova@petrsu.ru; endray2@mail.ru

Понятие возраста информации было введено в 2011 году в [1] для количественной оценки актуальности имеющихся знаний о состоянии системы [2]. При поступлении новой информации возникает необходимость принять решение: считать ее или нет. Выбор имеет очень важное значение и зависит от различных параметров, таких как время задержки, длина пакета и частота поступления информационных пакетов [3].

Мы рассматриваем односерверные системы обслуживания без буфера, в которых при поступлении с некоторой вероятностью $0 < p < 1$ принимается решение о считывании информации, а также два типа блокировки старых сообщений. Входной поток τ имеет гиперэкспоненциальное распределение с функцией распределения (ф. р.) вида:

$$A(t) = 1 - qe^{-\lambda_1 t} - (1 - q)e^{-\lambda_2 t}, \quad t \geq 0,$$

где $\lambda_1, \lambda_2 \geq 0$ и $0 \leq q \leq 1$. Обслуживание (чтение сообщения) распределено в соответствии с законом Вейбулла

$$B(t) = 1 - e^{-t^\alpha}, \quad \alpha > 0, \quad t \geq 0.$$

Обозначим через T_n – момент прихода n -го сообщения в систему, $n \geq 1$, $T_0 = 0$. Пусть $\tau_n := T_{n+1} - T_n$ – интервалы между поступлениями сообщений на сервер. Введем для каждого $n \geq 1$ индекс допуска сообщения χ_n следующим образом:

$$\chi_n = \begin{cases} 1, & \text{если сообщение, прибывшее в момент } T_n, \text{ допущено} \\ 0, & \text{иначе.} \end{cases}$$

Пусть $P(\chi_n = 1) = p$. Обозначим через T'_n время ухода сообщения из системы, s_n – время чтения сообщения, поступившего в момент T_n . Будем называть сообщение n *успешным*, если оно покидает систему после

© И. В. Пешкова, А. С. Алексеев, 2024

полного прочтения. Обозначим N_t – номер ближайшего к моменту времени t сообщения, допущенного к прочтению ($\chi_{N_t} = 1$). Тогда случайный процесс

$$\Delta(t) = t - T_{N_t}, \quad t > 0$$

называется *возрастом информации* [1].

Рассмотрим два типа систем. В *системе с обновлением старых заявок* при поступлении нового сообщения находящееся на прочтении сообщение немедленно снимается и покидает систему. В *системе с блокированием новых заявок* каждое новое сообщение после принятия решения о допуске блокируется, если в момент его поступления сервер занят.

Для рассмотренных систем были получены следующие эмпирические оценки: среднее время простоя системы, средний возраст сообщения и количество “успешных” заявок в системе при $\lambda_1 = 6.4$, $\lambda_2 = 2.1$ и различных $\alpha = 0,5; \dots, 3$.

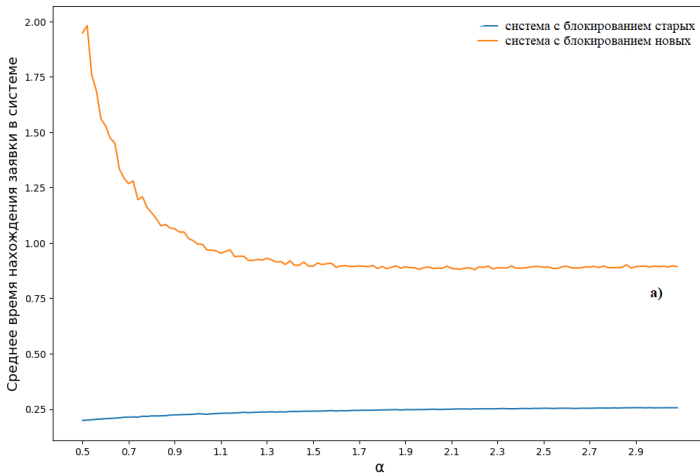


Рис. 1: Зависимость среднего возраста сообщения от α .

Из рис. 1 видно, что в системе с обновлением старых заявок параметр формы в распределении Вейбулла слабо влияет на средний возраст заявок в системе: оно монотонно растет от минимального значения 0.1991 до своего максимума 0.25539. Для второй системы средний возраст сообщения убывает с ростом α до значения $\alpha = 1$, после чего начинает медленно увеличиваться.

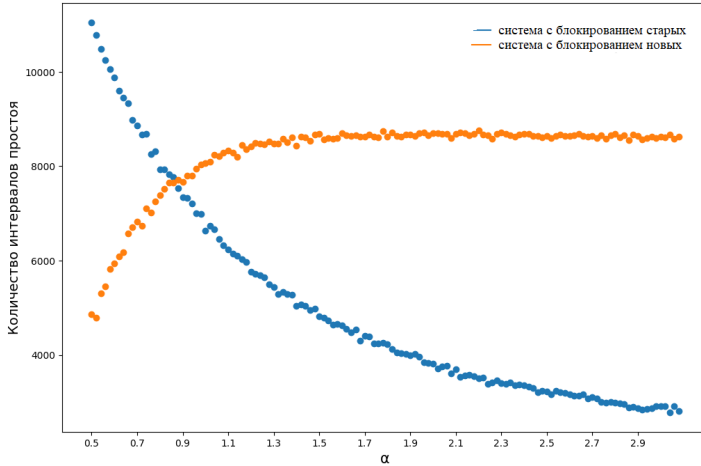


Рис. 2: Зависимость числа “успешных” сообщений от α .

На рис. 2 изображены графики зависимости количества интервалов “успешных” сообщений от параметра α распределения Вейбулла. В системе с обновлением старых заявок зависимость – монотонно убывающая, а в системе с блокированием новых сообщений – монотонно возрастающая, при $\alpha \approx 1$ графики пересекаются.

Работа выполнена при финансовой поддержке РФФ (проект 21-71-10135).

Список литературы

- [1] *Kaul S., Gruteser M., Rai V., and Kenney J.* Minimizing age of information in vehicular networks // 8th Annual IEEE Communications Society Conference on Sensor, Mesh and Ad Hoc Communications and Networks (SECON). 2011. P. 350–358.
- [2] *Gruteser M., Kaul S., Yates R.* Real-time status: How often should one update? // Proceedings – IEEE INFOCOM. 2012. P. 2731–2735.
- [3] *Zazanis M. A., Konstantopoulos T., Kesidis G.* The distribution of age-of-information performance measures for message processing systems // Queueing Systems. 2020, 95(3):1–48.

Регенеративное моделирование систем обслуживания с суперпозицией входных процессов восстановления

¹Ирина В. Пешкова, ²Евсей В. Морозов, ³Микеле Пагано

¹Петрозаводский государственный университет, Петрозаводск, Россия;

²Институт прикладных математических исследований КарНЦ РАН,
Петрозаводск, Россия,

²Новгородский государственный университет им. Ярослава Мудрого,
Великий Новгород, Россия;

³Университет Пизы, Пиза, Италия

E-mail: iaminova@petrsu.ru; emorozov@krc.karelia.ru

Рассматривается стохастическая система обслуживания $\sum_{i=1}^n GI_i/G/1$, у которой входной процесс является суперпозицией n независимых стационарных процессов восстановления. Для моделирования современных телекоммуникаций важность таких процессов обусловлена сложной структурой входных потоков, содержащих самые разные компоненты, включая, например, ‘тяжелые’ видеопотоки и короткие (‘легкие’) сообщения, например, показания датчиков.

Хотя такие системы изучаются достаточно давно (см., например, [1–6]), однако в данной работе предлагается метод, основанный на построении искусственных моментов регенерации входного процесса в случае, когда распределение интервалов между заявками обладает так называемым тяжелым хвостом. Насколько известно авторам такой подход к данной модели применен впервые.

Хорошо известно, что, если входные процессы в системе не являются пуассоновскими, то описывающие систему случайные процессы (число заявок, нагрузка), вообще говоря, не являются регенерирующими [2, 7]. С другой стороны, свойство регенерации значительно упрощает анализ стационарности и производительности системы обслуживания, а также позволяет осуществить надежное доверительное оценивание ее характеристик на основе регенеративной версии ЦПТ.

В настоящем исследовании предлагается исходную систему, вообще говоря, не обладающую свойством (классической) регенерации, аппроксимировать регенеративной системой $GI/G/1$ с помощью построения стационарного входного процесса восстановления, задающего интервалы между приходами заявок в систему [6].

Пусть i -й входной поток задан независимыми одинаково распределенными (н. о. р.) интервалами $\{\tau_k^i, k \geq 0\}$ между поступлениями заявок с функцией распределения (ф. р.) A_i и средним $E\tau^i = 1/\lambda_i$, а интервал до 1-го события имеет ф. р. стационарного перескока, т. е.,

$$\lambda_i \int_0^x (1 - A(u)) du, \quad i = 1, \dots, n.$$

Обозначим через $\{S_k^i, k \geq 1\}$ н. о. р. времена обслуживания заявок i -го потока с ф. р. B_i , $i = 1, \dots, n$. Известно, что хвост ф. р. интервала входного потока, аппроксимирующего суперпозицию входных потоков, имеет вид [6]:

$$1 - A(x) = \sum_{i=1}^n \frac{\lambda_i}{\lambda} (1 - A_i(x)) \prod_{j, j \neq i} \lambda_j \int_x^{\infty} (1 - A_j(x)) dx, \quad (1)$$

где $\lambda = \lambda_1 + \dots + \lambda_n$. Распределение (1) есть n -компонентная смесь распределений с коэффициентами пропорции смеси (вероятностями) λ_i/λ . Ф. р. времени обслуживания в данной системе (в предположении дисциплины обслуживания FIFO) можно представить в виде смеси

$$B(x) := \sum_{i=1}^n (\lambda_i/\lambda) B_i(x), \quad x \geq 0. \quad (2)$$

Таким образом, вместо исходной (не регенерирующей) системы $\sum_{i=1}^n GI_i/G/1$ с суперпозицией n входных процессов восстановления (и n классами заявок) исследуется аппроксимирующая система $GI/G/1$, в которой н. о. р. интервалы входного потока имеют распределение (1), а (н. о. р.) времена обслуживания – распределение (2). Условие стационарности (обеих систем) имеет вид [7]:

$$\rho = \sum_{i=1}^n \lambda_i ES^i < 1.$$

В аппроксимирующей системе можно построить моменты регенерации (моменты прихода заявок в пустую систему). Численный анализ показал хорошую точность аппроксимации, когда интервалы компонент входного процесса имеют (вообще говоря, различные) распределения Вейбулла с тяжелым хвостом.

Работа выполнена при финансовой поддержке РФФИ (проект 21-71-10135).

Список литературы

- [1] *Albin S. L.* On Poisson Approximations for Superposition Arrival Processes in Queues // *Management Science*. 1982. 28(2):126–137.
- [2] *Asmussen S., Schmidli H., Schmidt V.* Tail probabilities for non-standard risk and queueing processes with subexponential jumps // *Advances in Applied Probability*. 1999. 31(2):422–447.
- [3] *Lam T. and Lehoczky J. P.* Superposition of Renewal Processes // *Advances in Applied Probability*. 1991. 23(1):64–85.
- [4] *Lawrence A. J.* Dependency of Intervals Between Events in Superposition Processes // *Journal of the Royal Statistical Society. Series B (Methodological)*. 1973. 35(2):306–315.
- [5] *Newell G. F.* Approximations for Superposition Arrival Processes in Queues // *Management Science*. 1984. 30(5):623–632.
- [6] *Whitt W.* Approximating a point process by a renewal process, I: Two basic methods // *Operat. Res.* 1982. 30:125–147.
- [7] *Morozov E., Steyaert B.* *Stability Analysis of Regenerative Queueing Models*. Springer, Cham, 2021.

Метод пробных функций для нахождения оценок скорости сходимости однородных цепей Маркова к инвариантному распределению

Александр В. Резлер

Новосибирский государственный университет, Новосибирск, Россия

E-mail: rezlers123@gmail.com

Мы исследуем однородные цепи Маркова в произвольном польском пространстве состояний \mathcal{X} . Как известно, для доказательства стабильности (положительной возвратности) цепи широко распространен так называемый критерий Фостера. Его суть заключается в определении некоторой пробной функции $L(x)$ такой, что математическое ожидание ее приращения за один шаг от элементов цепи отрицательно вне некоторого множества и ограничено внутри него. В работах С. Г. Фосса (см., например, [1, теорема 1]) было доказано обобщение теоремы на многошаговый случай.

Впрочем, при использовании многошагового критерия Фостера для исследования эргодичности марковских цепей, остается открытым вопрос об оценивании скорости их сходимости к инвариантному распределению. Как известно, она напрямую зависит от свойств распределения момента возвращения цепи в положительно возвратное множество.

В данной работе мы обобщаем одну из фундаментальных теорем метода пробных функций (см. [2, теорема 5.2.1]) на случай растущих и, более того, случайных шагов, а также приводим примеры марковских цепей, где использование полученных результатов уместно.

Список литературы

- [1] *Foss S., Konstantopoulos T.* An overview of some stochastic stability methods // Journal of Operation Research Society Japan. 2004. 47(4):275–303. Zbl 1134.93412
- [2] *Kalashnikov V. V.* Mathematical methods in queuing theory. – Dordrecht etc.: Kluwer acad. publ., – IX, 377 p. – (Mathematics and its applications; Vol. 003271). 1994. ISBN 0-7923-2568-0.

Multicriteria dynamic games with asymmetric horizons

Anna N. Rettieva

Institute of Applied Mathematical Research of KRC RAS,
Petrozavodsk, Russia

E-mail: annaret@krc.karelia.ru

The approaches to obtain optimal behavior in dynamic multicriteria games [1] with different planning horizons are constructed. The multicriteria Nash equilibrium is obtained via the Nash bargaining scheme, and the cooperative equilibrium is determined by the modified bargaining procedure. To illustrate presented approaches a multicriteria bioresource management problem with different planning horizons is investigated.

Consider a bicriteria dynamic game with n participants in discrete time. The players exploit a common resource and wish to optimize k different criteria. The state dynamics is in the form

$$x_{t+1} = f(x_t, u_{1t}, \dots, u_{nt}), \quad x_1 = x, \quad (1)$$

where $x_t \geq 0$ is the resource size at time $t \geq 0$, $f(x_t, u_{1t}, \dots, u_{nt})$ denotes the growth function, and $u_{it} \geq 0$ gives the exploitation rate of player i at time t , $i \in N = \{1, \dots, n\}$.

We explore a model in which the players possess heterogeneous planning horizons. By assumption, players harvest the fish stock during m_1, \dots, m_n steps, respectively, $m_1 < \dots < m_n$ and $m_1 > m_0 = 1$. Therefore, during the time period $[m_i, m_{i+1}]$ $n - i$ players exploit the same resource stock, and the problem consists in evaluating their optimal strategies.

The payoff functions of the players are defined by

$$J_i = \begin{pmatrix} J_i^1 = \sum_{l=1}^i \sum_{t=m_{l-1}}^{m_l} \delta^t g_i^1(x_{lt}, u_{lt}, \dots, u_{nt}) \\ \dots \\ J_i^k = \sum_{l=1}^i \sum_{t=m_{l-1}}^{m_l} \delta^t g_i^k(x_{lt}, u_{lt}, \dots, u_{nt}) \end{pmatrix}, \quad i \in N, \quad (2)$$

where $g_i^j(\cdot) \geq 0$ gives the instantaneous utility, $j = 1, \dots, k$, $i \in N$, $\delta \in (0, 1)$ denotes a common discount factor, and x_t possess the dynamics

$$x_{t+1} = f(x_t, u_{it}, \dots, u_{nt}), \quad t \in [m_{i-1} + 1, m_i], \quad i \in N, \quad x_1 = x.$$

We design the noncooperative equilibrium in dynamic multicriteria game (1), (2) applying the Nash bargaining products [2].

Then, we find a multicriteria cooperative equilibrium as a solution of modified bargaining scheme with the multicriteria Nash equilibrium playing the role of status quo points [3, 4].

A discrete-time game-theoretic bioresource management model (harvesting problem) with different planning horizons is treated to show how the presented concepts work.

This research was supported by the Russian Science Foundation: grant No. 22-11-20015, <https://rscf.ru/project/22-11-20015/>, jointly with support of the authorities of the Republic of Karelia with funding from the Venture Investment Foundation of the Republic of Karelia.

References

- [1] *Shapley L. S.* Equilibrium points in games with vector payoffs // Naval Research Logistic Quarterly. 1959. 6:57–61.
- [2] *Rettieva A. N.* Equilibria in dynamic multicriteria games // Int. Game Theory Review. 2017. 19(1):1750002.
- [3] *Mazalov V. V., Rettieva A. N.* Asymmetry in a cooperative bioresource management problem // In: Game-Theoretic Models in Mathematical Ecology, Nova Science Publishers, 2015. P. 113–152.
- [4] *Rettieva A. N.* Dynamic multicriteria games with asymmetric players // Journal of Global Optimization. 2022. 83:521–537.

Анализ статистики расстояний на основе максимального общего подграфа для деревьев с упорядоченными вершинами

Александр А. Рогов, Николай Д. Москин, Роман В. Воронов

Петрозаводский государственный университет, Петрозаводск, Россия

E-mail: rogov@petsu.ru; moskin@petsu.ru; rvoronov@petsu.ru

В работе исследуется распределение расстояний между деревьями с упорядоченными вершинами на основе максимального общего подграфа (см. [1]). Знать распределение расстояний необходимо для сравнения графовых моделей, которые возникают, например, при анализе синтаксических структур текстов [2].

Рассмотрим расстояние $d(G, F) = 1 - \min_{i=1, \dots, k} \left(\frac{|mcs'(g_{\min(i,m)}, f_i)|}{i} \right)$ (здесь $|G|$ – число вершин графа G). Отметим, что для двух помеченных графов $G = (V_1, E_1)$, $F = (V_2, E_2)$ максимальным общим индуцированным помеченным подграфом $mcs'(G, F)$ назовем граф $mcs'(G, F) = (V, E)$ с максимальным числом вершин, такой, что $V \subset V_1$, $V \subset V_2$ и $mcs'(G, F)$ равен подграфам графов G и F , индуцированным подмножеством V . Для анализа статистики расстояний были реализованы два алгоритма: эвристический и полный перебор с заданными условиями (при этом деревья генерировались при помощи кода Прюфера [3]).

Таблица 1: Распределение значений метрики для пар деревьев с 6 вершинами

Значение метрики d	Доля пар графов s
0	0,007
0,167	0,018
0,2	0,064
0,25	0,211
0,33	0,517
0,4	0,532
0,5	1

Обозначим число упорядоченных вершин в деревьях с помощью n . Рассмотрим долю пар графов (s) со значением метрики не больше, чем заданное расстояние d при малых значениях n . Например, в таблице 1 представлены расчеты для $n = 6$. Если при меньших n ($n = 3, 4, 5$) аналогично значение метрики не превосходит 0,5, то при больших n наблюдается другая картина.

Среднее значение расстояния d при увеличении числа вершин n постепенно растет и стабилизируется в точке 0,587 (таблица 2). Здесь x_1 – количество вершин, x_2 – среднее значение, x_3 – коэффициент асимметрии, x_4 – длина малых расстояний с общим количеством меньшим 5% меньше, x_5 – длина малых расстояний с общим количеством меньшим 10% меньше, x_6 – длина больших расстояний с общим количеством меньшим 10% больше, x_7 – длина больших расстояний с общим количеством меньшим 5% больше. 90% диапазон значений метрики сужается с ростом n и является более островершинным по сравнению с нормальным законом. Распределение значений расстояний является мультимодальным (см. рис. 1), при этом с ростом n наблюдается рост числа мод. Гипотеза о соответствии распределения нормальному закону при $n = 100$ была отвергнута на уровне значимости 0,1.

Таблица 2: Распределение значений метрики для пар деревьев при разных значениях числа вершин n

x_1	x_2	x_3	x_4	x_5	x_6	x_7
10	0,562	-0,317	0,440	0,453	0,625	0,652
15	0,575	-0,255	0,469	0,495	0,638	0,649
20	0,579	-0,230	0,486	0,509	0,630	0,648
30	0,583	-0,187	0,530	0,533	0,632	0,633
40	0,585	-0,139	0,531	0,549	0,623	0,627
50	0,586	-0,133	0,540	0,553	0,619	0,624
75	0,587	-0,103	0,548	0,560	0,613	0,621
100	0,587	-0,085	0,556	0,561	0,610	0,616

Четвертая-седьмая колонки дают возможность проанализировать расстояния, которые встречаются реже. Если близких объектов не более 10% среди всех, то при $n = 10$ маленьким расстоянием определим такое расстояние, которое меньше 0,453. При этом критическое значение длины маленького расстояния растет с ростом n от 0,453 до 0,561. Похожим образом можно проанализировать большие расстояния. Расстояния между критическими значениями (при $n = 10$ от 0,453 до 0,625) можно определить, как часто встречающиеся.

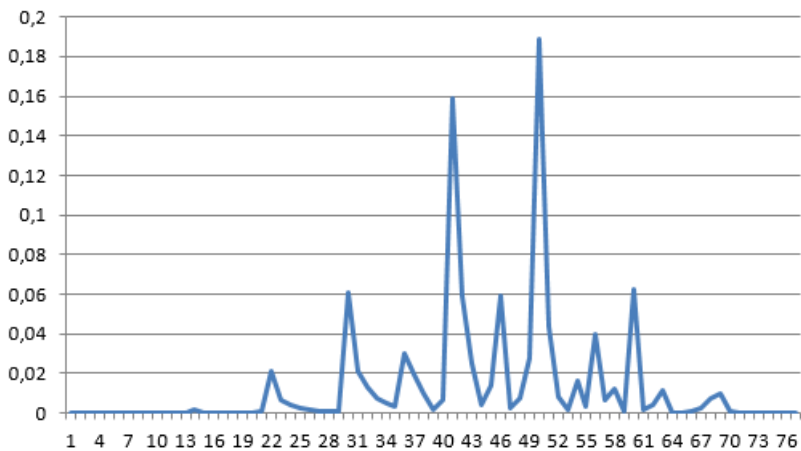


Рис. 1: Распределение частот расстояний для $n = 30$

Список литературы

- [1] *Москин Н. Д.* Метрика для сравнения графов с упорядоченными вершинами на основе максимального общего подграфа // Прикладная дискретная математика. 2021. 52:105–113. doi: 10.17223/20710410/52/7
- [2] *Севбо И. П.* Графическое представление синтаксических структур и стилистическая диагностика. Киев: Наукова Думка, 1981. 192 с.
- [3] *Prüfer H.* Neuer Beweis eines Satzes über Permutationen // Archiv für Mathematik und Physik. 1918. 27:742–744.

Предельные совместные распределения статистик критериев пакета NIST и их обобщений

Максим П. Савелов

Московский государственный университет им. М. В. Ломоносова,
Москва, Россия

E-mail: savelovmp@gmail.com

Одним из наиболее популярных инструментов, используемых для проверки качества генераторов случайных двоичных последовательностей, является пакет NIST [1]. В докладе будут обсуждаться результаты работ [2, 3], в которых получены предельные совместные распределения статистик критериев пакета NIST и их обобщений в задаче о проверке гипотезы H_0 , в соответствии с которой тестируемая последовательность является последовательностью независимых случайных величин с распределением Бернулли с параметром $\frac{1}{2}$. Альтернативная гипотеза H_1 соответствует схеме серий, в которой распределение исследуемой последовательности сближается с ее распределением при H_0 . Частным случаем гипотезы H_1 является схема серий из однородных строго стационарных цепей Маркова фиксированного порядка, переходные вероятности которых «близки» к $\frac{1}{2}$.

Также в докладе будет обсуждаться более общий случай, когда при H_0 тестируемая последовательность является последовательностью независимых одинаково распределенных случайных величин с полиномиальным распределением и гипотеза H_1 «сближается» с H_0 с ростом размера выборки.

Список литературы

- [1] *Rukhin A., Soto J., Nechvatal J., Smid M., Barker E., Leigh S., Levenson M., Vangel M., Banks D., Heckert A., Dray J., Vo S.* A statistical test suite for random and pseudorandom number generators for cryptographic applications, NIST Special Publication 800-22 Revision 1a, ed. L. E. Bassham III, NIST, April 2010.

- [2] *Савелов М. П.* Мощность критерия, основанного на одновременном применении «Monobit Test», «Frequency Test within a Block» и «Serial Test» // Дискрет. матем. 2023. 35(4):79–114. doi: 10.4213/dm1802
- [3] *Савелов М. П.* Мощность критерия, основанного на одновременном применении «Monobit Test», «Frequency Test within a Block» и обобщения критерия «Approximate Entropy Test» // Дискрет. матем. 2024. 36(1):67–102. doi: 10.4213/dm1814

Оценка остаточного члена в разложениях строго-устойчивого закона

Вячеслав В. Саенко

НИТИ им. С.П. Капицы, Ульяновского государственного университета,
Ульяновск, Россия

E-mail: vvsenko@inbox.ru

Получены оценки остаточных членов для разложений в ряд плотности вероятности и функции распределения строго-устойчивого закона с характеристической функцией

$$\hat{g}(t, \alpha, \theta, \lambda) = \exp \left\{ -\lambda |t|^\alpha \exp \left\{ -i \frac{\pi}{2} \alpha \theta \operatorname{sign} t \right\} \right\}, \quad t \in \mathbf{R},$$

где $\alpha \in (0, 2]$, $|\theta| \leq \min(1, 2/\alpha - 1)$, $\lambda > 0$ [1].

В случае $x \rightarrow 0$ для разложений плотности вероятности и функции распределения в степенной ряд справедлива теорема

Теорема 1. *В случае $x \rightarrow 0$ для любого допустимого набора параметров (α, θ) за исключением значений $\theta = \pm 1$ для плотности вероятности $g(x, \alpha, \theta)$ и функции распределения $G(x, \alpha, \theta)$ справедливы следующие представления в виде ряда*

$$\begin{aligned} g(x, \alpha, \theta) &= g_N^0(x, \alpha, \theta) + R_N^0(x, \alpha), \\ G(x, \alpha, \theta) &= \frac{1 - \theta}{2} + G_N^0(x, \alpha, \theta) + \mathcal{R}_N^0(x, \alpha), \end{aligned}$$

где

$$\begin{aligned} g_N^0(x, \alpha, \theta) &= \frac{1}{\alpha\pi} \sum_{n=0}^{N-1} \frac{x^n}{n!} \Gamma\left(\frac{n+1}{\alpha}\right) \sin\left(\frac{\pi}{2}(n+1)(1-\theta)\right), \\ |R_N^0(x, \alpha)| &\leq \frac{1}{\alpha\pi} \frac{|x|^N}{N!} \Gamma\left(\frac{N+1}{\alpha}\right) \\ G_N^0(x, \alpha, \theta) &= \frac{1}{\pi\alpha} \sum_{n=0}^{N-1} \frac{x^{n+1}}{(n+1)!} \Gamma\left(\frac{n+1}{\alpha}\right) \sin\left(\frac{\pi}{2}(n+1)(1-\theta)\right), \\ |\mathcal{R}_n^0(x, \alpha, \theta)| &\leq \frac{|x|^{N+1}}{\alpha\pi(N+1)!} \Gamma\left(\frac{N+1}{\alpha}\right). \end{aligned}$$

© В. В. Саенко, 2024

В случае $x \rightarrow \infty$ оказывается справедлива следующая теорема.

Теорема 2. В случае $x \rightarrow \pm\infty$ для любого допустимого набора параметров (α, θ) , за исключением значений $\theta = \pm 1$, для плотности вероятности $g(x, \alpha, \theta)$ и функции распределения $G(x, \alpha, \theta)$ справедливы представления в виде степенного ряда

$$g(x, \alpha, \theta) = g_N^\infty(|x|, \alpha, \theta^*) + R_N^\infty(|x|, \alpha),$$

$$G(x, \alpha, \theta) = \frac{1}{2}(1 + \text{sign}(x)) - \text{sign}(x) (G_N^\infty(|x|, \alpha, \theta^*) + \mathcal{R}_N^\infty(|x|, \alpha)),$$

где $\theta^* = \theta \text{sign}(x)$ и

$$g_N^\infty(x, \alpha, \theta) = \frac{1}{\pi} \sum_{n=1}^{N-1} \frac{(-1)^{n+1}}{n!} \Gamma(\alpha n + 1) \sin\left(\frac{\pi}{2} \alpha n (1 + \theta)\right) x^{-\alpha n - 1}, \quad x > 0,$$

$$|R_N^\infty(x, \alpha)| \leq \frac{x^{-\alpha N - 1}}{\pi N!} (\Gamma(\alpha N + 1) + x^{-\alpha} \Gamma(\alpha(N + 1) + 1)), \quad x > 0.$$

$$G_N^\infty(x, \alpha, \theta) = \frac{1}{\pi} \sum_{n=1}^{N-1} \frac{(-1)^{n+1}}{n!} \Gamma(\alpha n) \sin\left(\frac{\pi}{2} \alpha n (1 + \theta)\right) x^{-\alpha n}, \quad x > 0,$$

$$|\mathcal{R}_N^\infty(x, \alpha)| \leq \frac{x^{-\alpha N}}{\pi N!} (\Gamma(\alpha N) + x^{-\alpha} \Gamma(\alpha(N + 1))), \quad x > 0.$$

Полученные оценки остаточных членов оказываются полезными в задачах, связанных с вычислением плотностей строго-устойчивых законов в области больших или малых значениях координаты x . Действительно, основным методом вычисления плотности строго-устойчивого закона является использование интегрального представления. Однако, в области больших или малых значений координаты x численные методы интегрирования оказываются неспособны корректно вычислить интеграл, входящий в это интегральное представление, и выдают неверный результат. Представленные оценки остаточных членов позволяют решить эту проблему. Используя эти оценки, могут быть получены выражения для пороговых координат x_ε , при которых абсолютная погрешность вычисления плотности вероятности и функции распределения с помощью рядов из теорем 1 и 2 не будет превышать требуемого уровня точности ε , при заданных α, N . В случае $x \rightarrow 0$ выражение для пороговой координаты может быть получено в явном виде. В случае $x \rightarrow \pm\infty$, в явном виде получить выражение для пороговой координаты не удастся и необходимо численно решать уравнение. Однако, проблем с численным решением этого уравнения не возникает.

Полученные оценки остаточных членов позволяют исследовать сходимость рядов. Степенные ряды из теоремы 1 в случае $\alpha < 1$ являются

асимптотическими, в случае $\alpha > 1$ являются сходящимися, а в случае $\alpha = 1$ ряды сходятся в области $|x| < 1$. Ряды из теоремы 2 в случае $\alpha < 1$ являются сходящимися, в случае $\alpha > 1$ являются асимптотическими, а в случае $\alpha = 1$ ряды сходятся в области $|x| > 1$. Поведение этих рядов в случаях $\alpha < 1$ и $\alpha > 1$ хорошо известны [1, 2]. Случай $\alpha = 1$ ранее не был исследован. Провести исследование этого случая удалось благодаря полученным оценкам остаточных членов.

В докладе будут представлены доказательства приведенных теорем и выражения для пороговых координат. Также будут представлены результаты тестовых расчетов, подтверждающие справедливость полученных оценок.

Список литературы

- [1] *Золотарев В. М.* Одномерные устойчивые распределения. М.: Наука, 1983. 304 с.
- [2] *Uchaikin V. V., Zolotarev V. M.* Chance and stability Stable Distributions and their Applications. VSP, Utrecht, 1999.

Нахождение остовных подграфов в случайных графах

Ольга И. Серкова

Московский физико-технический институт, Долгопрудный, Россия
E-mail: kalinichenko.oi@phystech.edu

Тема поиска остовных подграфов в случайных графах занимает умы математиков с момента формулировки задачи Эрдешем и Реньи в 1960 году [1]. В основном интересует поиск пороговой вероятности присутствия того или иного подграфа в случайном графе.

Наиболее общий результат в области поиска остовных подграфов был получен Риорданом [2].

Здесь и далее $v(H)$ – количество вершин в графе H , $e(H)$ – количество ребер в графе H , $N = C_n^2$.

Теорема 1 (Риордан). Пусть H – граф на n вершинах с максимальной степенью Δ , $n \rightarrow \infty$, $\gamma = \sup_{H' \subset H, v(H') \geq 3} \frac{e(H')}{v(H')-2}$, $\alpha = e(H)/N$, pN целое и выполняются условия

1. $\alpha N \geq n$
2. $pN, (1-p)\sqrt{n} \rightarrow \infty$
3. $np^\gamma/\Delta^4 \rightarrow \infty$.

Тогда граф $G(n, pN)$ а. н. н. содержит подграф, изоморфный H .

В силу условий теоремы результат аналогичен в случае $G(n, p)$ вместо $G(n, pN)$.

Одной из интересных задач в области остовных подграфов случайного графа является поиск степеней гамильтонова цикла. Напомним, что степенью гамильтонова цикла является гамильтонов цикл, в котором каждая вершина соединена не только с предыдущей, но с k предыдущими вершинами.

Пороговая вероятность для гамильтоновых циклов была получена Позой в [3]. Из результата Риордана следует, что при $k \geq 3$ и $p \gg n^{-\frac{1}{k}}$ граф будет содержать k -ю степень гамильтонова цикла. При этом, нетрудно получить, что $n^{-\frac{1}{k}}$ является и верхней оценкой пороговой вероятности при

$k \geq 3$, то есть для $k \geq 3$ это значение является пороговой вероятностью. Для $k = 2$ задача была решена в статье 2021 года Кана, Нараянана и Парк [4].

Однако никто пока не задавался «обратным» вопросом: какова максимальная степень k такая, что в случайном графе а. п. н. существует k -я степень гамильтонова пути при константном p , например при $p = 1/2$? То есть, каково число k такое, что при $k' < k$ а. п. н. в случайном графе есть k' -я степень гамильтонова цикла, а при $k' > k$ а. п. н. её нет, и чему может быть равно это k ? Нетрудно показать, что при $k > \log_2 n$ в графе $G(n, 1/2)$ а. п. н. нет k -й степени гамильтонова цикла.

В рамках работы доказана теорема, улучшающая результаты Риордана для некоторых классов графов, в том числе помогающая улучшить нижнюю оценку на максимальную степень гамильтонова цикла в случайном графе:

Теорема 2. Пусть H – граф на n вершинах с максимальной степенью Δ , где $n \rightarrow \infty$, при этом существуют такие $x > 0, k > 0, \beta > 0$, что каждый подграф $F \subset H$, имеющий $v(F) \leq n/x$ вершин и с компонент связности, имеет не более $kx - \beta$ ребер, а все остальные подграфы имеют не более kx ребер. Пусть также $\alpha = e(H)/N, u = p^{-1} - 1$ и выполняются условия

1. $\alpha N \geq n$
2. $pN, (1 - p)\sqrt{n} \rightarrow \infty$
3. $\frac{np^2}{\Delta^4} \rightarrow \infty$
4. $\alpha^3 N p^{-2} \rightarrow 0$
5. $\frac{4e^2 \Delta^2 (2u)^k}{n} < q < 1$,
6. Выполняется одно из:
 - (a) $ne(2u)^{k-\beta} = o(1)$
 - (b) $ne(2u)^{k-\beta} \leq 1$ и $q = o(1)$
 - (c) $ne(2u)^{k-\beta} > 1$ и $\Delta^4 (2u)^{3k-\beta} = o(n)$
7. $n \gg x^2 \max(\frac{\beta}{k}, 1) \ln n$

Тогда граф $G(n, p)$ а. п. н. содержит подграф, изоморфный H .

Заметим, что условия 3 и 4 теоремы 2 следуют из условия 3 теоремы Риордана, так что теорема отличается от теоремы Риордана условиями 5-7 вместо условия 3.

Также отметим, что k -я степень гамильтонова цикла удовлетворяет условиям теоремы для $k = k, \Delta = 2k, x = k, \beta = C_{k+1}^2$. При этом результат Риордана дает нижнюю оценку $k' = \log n - 4 \log \log n - \omega(n)$, $\omega(n) \rightarrow \infty$ на максимальную степень гамильтонова цикла в $G(n, 1/2)$. Теорема 2 дает оценку $k' = \log n - 2 \log \log n - O(1)$.

Также полученный результат применим для нахождения пороговой вероятности вхождения k -вырожденных графов. Напомним, что k -вырожденный граф – это неориентированный граф, в котором каждый подграф имеет вершины со степенью, не превосходящей k . Теорема 2 позволяет, в частности, найти пороговую вероятность в случае константного $k \geq 3, \Delta = O(1)$ и оценку на точную пороговую вероятность в случае $k \geq 5, \Delta = O(1)$, а теорема Риордана дает возможность лишь найти пороговую вероятность.

Существуют и другие примеры применения теоремы. Таким образом, она вносит вклад в развитие области поиска остовных подграфов в случайных графах.

Работа выполнена при поддержке Российского фонда фундаментальных исследований, грант 20-51-56017.

Список литературы

- [1] *Erdős P., Rényi A.* On the evolution of random graphs // Magyar Tud. Akad. Mat. Kutató Int. Közl. 1960. 5:17–61.
- [2] *Riordan O.* Spanning subgraphs of random graphs // Combinatorics, Probability & Computing. 2000. 9:125–148.
- [3] *Pósa L.* Hamiltonian circuits in random graphs // Discrete Mathematics. 1976. 14:359–364.
- [4] *Kahn J., Narayanan B., Park J.* The threshold for the square of a Hamilton cycle // Proc. Amer. Math. Soc. 2021. 149:3201–3208. doi: 10.1090/proc/15419

Двусторонние оценки квантилей биномиального распределения

Александр А. Серов

Математический институт им. В. А. Стеклова РАН, Москва, Россия

E-mail: serov@mi-ras.ru

Эффективная реализация статистических тестов с заданными вероятностями ошибок требует информации о значениях квантилей различных уровней распределения тестовой статистики. Если распределение тестовой статистики является биномиальным или аппроксимируется таким распределением, то необходимы оценки квантилей биномиального распределения. Существует множество работ, посвященных эмпирическим и теоретическим оценкам квантилей различных распределений и, в частности, квантилей биномиального распределения (см., например, [2–12]). В тезисах представлены явные формулы для верхних и нижних оценок квантилей биномиального распределения для произвольных значений параметров распределения. Применение полученных оценок на практике показывает, что в большинстве случаев для произвольных значений параметров n , p и квантили уровня R биномиального распределения разность верхней и нижней оценок не превосходит 1.

Пусть $X_{n,p}$ – случайная величина, имеющая биномиальное распределение с параметрами (n, p) :

$$\mathbf{P}\{X_{n,p} \leq k\} = \sum_{0 \leq m \leq k} C_n^m p^m (1-p)^{n-m}.$$

Обозначим через $Q(n, p, R)$ квантиль уровня R биномиального распределения с параметрами (n, p) , которая определяется следующим образом

$$Q(n, p, R) = \min_{0 \leq k \leq n} \{k : \mathbf{P}\{X_{n,p} \leq k\} \geq R\} \in \{0, 1, \dots, n\}.$$

В частности,

$$\begin{aligned} Q(n, p, 0) &= 0 \text{ и } Q(n, p, 1) = n, \\ Q(n, 1/2, 1/2) &= (n-1)/2, \text{ если } n - \text{нечетно,} \\ Q(n, 1/2, 1/2) &= n/2, \text{ если } n - \text{четно.} \end{aligned}$$

Далее будем использовать стандартные обозначения

$$\lfloor x \rfloor = [x] = \max\{m \in \mathbb{Z} \mid m \leq x\},$$

$$\lceil x \rceil = -\lfloor -x \rfloor = \min\{m \in \mathbb{Z} \mid m \geq x\}.$$

Теорема 1. *Справедливы следующие оценки при $0 < R < 1/2$*

$$r_1 - 1 \leq Q(n, p, R) \leq r_2,$$

$$\text{где } q = 1 - p, \quad r_1 = \left\lceil np + \frac{q-p}{6} \Phi^{-1}(R)^2 + \Phi^{-1}(R) \sqrt{npq + \frac{(q-p)^2 \Phi^{-1}(R)^2}{6^2}} \right\rceil,$$

$$r_2 = \left\lceil \frac{p(6-5p)n + \frac{3}{2} \Phi^{-1}(1-R)^2 q^2 - |\Phi^{-1}(1-R)| \frac{q}{2} \sqrt{24n(6-5p)p + \Phi^{-1}(1-R)^2 (3+p)^2}}{(6-5p) + \Phi^{-1}(1-R)^2 (1-p)/n} \right\rceil.$$

при $1/2 < R \leq 1$

$$r_3 \leq Q(n, p, R) \leq r_1,$$

$$\text{где } r_3 = \left\lceil \frac{p(1+5p)n + \frac{p}{2}(4-3p)\Phi^{-1}(R)^2 + \frac{p}{2} |\Phi^{-1}(R)| \sqrt{24n(1+(4-5p)p) + \Phi^{-1}(R)^2 (p-4)^2}}{(1+5p) + 2p\Phi^{-1}(R)^2/n} \right\rceil,$$

при $R = 1/2$

$$\lfloor np \rfloor \leq Q(n, p, R) \leq \lceil np \rceil.$$

Доказательство Теоремы 1 по существу основано на следующей теореме 2 из [13], где было представлено полное доказательство явных двусторонних оценок функции распределения биномиального закона.

Теорема 2 ([13]). *Пусть $\text{sing}(x) = \frac{x}{|x|}$ при $x \neq 0$, $\text{sing}(0) = 0$,*

$$H(x, p) = x \ln \left(\frac{x}{p} \right) + (1-x) \ln \left(\frac{1-x}{1-p} \right),$$

и пусть $\{C_{n,p}(k)\}_{k=0}^n$ возрастающая последовательность:

$$C_{n,p}(0) = (1-p)^n, \quad C_{n,p}(n) = 1 - p^n,$$

$$C_{n,p}(k) = \Phi \left(\text{sign} \left(\frac{k}{n} - p \right) \sqrt{2nH \left(\frac{k}{n}, p \right)} \right), \quad 1 \leq k < n. \quad (1)$$

Тогда для любых $k = 0, 1, \dots, n-1$ и $p \in (0, 1)$

$$C_{n,p}(k) \leq \mathbf{P}\{X_{n,p} \leq k\} \leq C_{n,p}(k+1) \quad (2)$$

и равенства справедливы только при $k = 0$ и $k = n-1$.

Список литературы

- [1] *Kaas R., Buhrman J. M.* Mean, median and mode in binomial distribution // *Statistica Neerlandica*. 1980. 34(1):13–18.
- [2] *Bahadur R. R.* A note on quantiles in large samples // *Ann. Math. Stat.* 1966. 37(3):577–580.
- [3] *Falk M.* On the estimation of the quantile density function // *Stat. Probab. Lett.* 1986. 4:69–73.
- [4] *Hsu J. C., Nelson B. L.* Control variates for quantile estimation // *Manage. Sci.* 1990. 36(7):835–851.
- [5] *Avramidis A. N., Wilson J. R.* Correlation-induction techniques for estimating quantiles in simulation // *Oper. Res.* 1998. 46:574–591.
- [6] *Jin X., Fu M. C., Xiong X.* Probabilistic error bounds for simulation quantile estimation // *Manag. Sci.* 2003. 49:230–246.
- [7] *Wu W. B.* On the Bahadur representation of sample quantiles for dependent sequences // *Ann. Stat.* 2005. 33(4):1934–1963.
- [8] *Ghosh S., Pasupathy R.* Low-storage online estimators for quantiles and densities // In: *Proc. of the 2013 Winter Simulation Conf. (WSC'13): Simulation: Making Decisions in a Complex World*. Washington, DC, USA. IEEE Press. 2013. P. 778–789. doi: 10.1109/WSC.2013.6721470
- [9] *Short M.* Improved inequalities for the Poisson and Binomial distribution and upper tail quantile functions // *Hindawi*. 2013. Art. 412958. doi: 10.1155/2013/412958
- [10] *Briggs K., Ying F.* How to estimate quantiles easily and reliably // *Mathematics Today*. 2018. P. 26–29.
- [11] *Short M.* On binomial quantile and proportion bounds: With applications in engineering and informatics // *Communication in Statistics – Theory and Methods*. 2023. 52(12):4183–4199. doi: 10.1080/03610926.2021.1986540
- [12] *Dong H., Nakayama M. K.* A Tutorial on Quantile Estimation via Monte Carlo // *Monte Carlo and Quasi-Monte Carlo Methods. MCQMC 2018*. Springer Proceedings in Mathematics & Statistics. 2020. 324:3–30. doi: 10.1007/978-3-030-43465-6_1
- [13] *Zubkov A. M., Serov A. A.* A complete proof of universal inequalities for distribution function of binomial law // *Theory Probab. Appl.* 2013. 57(3):539–544. doi: 10.1137/S0040585X97986138

Хроматическое число плоскости в условиях раскраски жордановыми множествами и многоугольниками

Георгий М. Соколов, В. А. Воронов

Московский физико-технический институт, Москва, Россия

E-mail: SokG2000@yandex.ru

Одной из классических задач комбинаторной геометрии является задача определения хроматического числа плоскости, то есть наименьшего числа цветов $\chi(\mathbb{R}^2)$, в которое можно раскрасить плоскость так, чтобы любые две точки, находящиеся на единичном расстоянии, имели разный цвет. Известно, что $5 \leq \chi(\mathbb{R}^2) \leq 7$ [1, 2], однако точное значение хроматического числа плоскости пока не определено.

Можно рассматривать различные модификации этой задачи. Например, можно потребовать, чтобы плоскость разбивалась на одноцветные области определённого вида. В работах [3, 4] требуется, чтобы границы между областями были непрерывными кривыми и в любой ограниченной области число вершин карты было конечным. В этих условиях получается доказать, что потребуется не менее 6 цветов. Усиливая ограничения, можно потребовать, чтобы границы были ломаными, а одноцветные области – многоугольниками.

В 2003 Экзо предложил постановку задачи о хроматическом числе плоскости с запрещенным интервалом расстояний [5]. Требуется, чтобы отсутствовала пара одноцветных точек, расстояние между которыми принадлежит интервалу $[1 - \varepsilon, 1 + \varepsilon]$. Обозначим соответствующее хроматическое число $\chi_{[1-\varepsilon, 1+\varepsilon]}(\mathbb{R}^2)$. В работе [6] было показано, что при $0 < \varepsilon < 1.15$ выполнено равенство $\chi_{[1-\varepsilon, 1+\varepsilon]}(\mathbb{R}^2) = 7$.

В настоящей работе это равенство распространено на случай разбиения на многоугольники и, при некоторых дополнительных условиях, на случай разбиения на ограниченные непрерывными кривыми области. Для разбиения непрерывными кривыми мы требуем дополнительно, чтобы любая единичная окружность пересекалась с границами областей по конечному множеству точек и чтобы не было вершин, в которых сходится четыре или более областей трех цветов. Доказательство активно использует идеи из работы [6] об интервальном хроматическом числе.

Список литературы

- [1] *de Grey A. D. N. J.* The Chromatic Number of the Plane Is at Least 5 // Geombinatorics. 2018. 28(1):18–31.
- [2] *Hadwiger H., Debrunner H., Klee V.* Combinatorial geometry in the plane // Courier Corporation, 2015.
- [3] *Woodall D. R.* Distances realized by sets covering the plane // Journal of Combinatorial Theory, Series A. 1973. 14(2):187–200.
- [4] *Townsend S. P.* Colouring the plane with no monochrome unit // Geombinatorics. 2005. 14(4):184–193.
- [5] *Exoo G.* ε -unit distance graphs // Discrete & Computational Geometry. 2005. 33:117–123.
- [6] *Voronov V.* The chromatic number of the plane with an interval of forbidden distances is at least 7. 2024. doi: 10.48550/arXiv.2304.10163

Blotto game with testing (The Locks, Bombs and Testing model)

Isaac M. Sonin

University of NC at Charlotte, Charlotte, USA

E-mail: imsonin@uncc.edu

Allocating scarce resources across targets that feature different probabilities of success is a classic problem in studies of conflict (Borel (1921), Shubik and Weber (1981), Powell (2007), Kovenock and Roberson (2012)), electoral politics (Snyder (1989), Myerson (1993), Laslier and Picard (2002)), auctions (Szentes and Rosenthal (2003), Kvasov (2007), project portfolio selection (Clark and Konrad (2008)), marketing (Friedman (1958), Hauser and Shugan (2008)), property rights (Chowdhury, S. M., et al (2018), etc.

One of the most popular and well-studied models is the classic Colonel Blotto game in which two players distribute limited resources between different sites (battlefields) with the goal of winning more sites Borel (1921), Gross and Wagner (1950), Blackett (1958), Roberson (2006), Powell (2007), S. Hart(2007, 2015), and many others.

Blotto Game with Testing, or equivalently General Locks, Bombs and Testing (G-LBT) model can be described as follows. As in most attack-defense games, there are two players with different roles. We call one of them the Defender (DF) and the other, the Attacker (AT). They are “fighting” over n “boxes” (sites, battlefields, targets, time slots, etc.) with possibly distinct cost values c_i , $i = 1, \dots, n$. AT is trying to destroy these boxes by placing “bombs” that can result in explosions and destruction. One or more bombs can be placed into the same box. AT has m , $m = 1, 2, \dots$ identical bombs to allocate among n boxes. A box is destroyed if at least one explosion occurs, and the explosions of different bombs in the same site or in different sites are independent. DF is trying to protect the boxes by distributing among them identical “locks”. A lock is a protection device which, when placed in a box, prevents its destruction with any number of bombs in it. Obviously, “locks” and “bombs” are just the names for discrete units of resources of protection and destruction. The number of locks k , $1 \leq k < n$, can be fixed, or it can be a random variable obtained when a lock appears in site i with probability λ_i independently of other boxes. The crucial feature of both models, in contrast to the classic Colonel Blotto game, is that AT can and will *test* every box, trying to find boxes without locks. The result of testing is

signal $s = (s_1, \dots, s_n)$, where each $s_i = 0, 1$. For every signal AT will select an allocation of m available bombs along n boxes. The goal of AT is to maximize the *total expected value* of destroyed boxes, and the goal of DF is to minimize it. Thus, we obtain a zero-sum game though there are variations of the basic models leading to a non zero-sum game.

We formally present this game in the *strategic form*, and thus, we obtain classical zero-sum game with finite number of pure strategies so the classical von Neuman's Minimax Theorem (1928) about the existence of von Neumann-Nash (NE) equilibrium can be applied. Less formal description has a dynamic structure. At the first time DF selects her strategy – possibly randomized allocation of k locks among n boxes. Then a random experiment is performed, signal s , based on prior π and parameters of testing is obtained, and only AT receives this signal. At the second time, AT assuming that she knows π , can calculate the *posterior distribution* of the positions of locks, and using this distribution, can obtain vector of the probabilities $\alpha_i(s|\pi)$ of *absence* of a lock in site i . Then AT, using a vector of values of boxes, can find her optimal response to each signal, allocation of her m bombs among n boxes, maximizing total expected value of destroyed sites. The DF, without knowing signal s , can repeat all of these calculations and hence both AT and DF know the total expected loss $v(\pi, \varphi_{opt}(\pi))$. Then, at the third time, DF selects π_* minimizing this loss. The solution of the LBT game is a triplet $(\pi_*, \varphi_{opt}(\pi_*), v_*)$, where v_* is the value of the game. In our paper we present three results.

Theorem 1. Short version. *With general values vector c , k locks and $m = 1$ all NE points in a problem without testing have the following structure: The optimal strategy of DF is to protect the first k_* boxes, where $k < k_* \leq n$, selecting such π_* to make AT indifferent to which of k_* boxes to attack, but more attractive than remaining boxes. The optimal strategy of AT is to allocate a bomb at random between sites $1, 2, \dots, k_*$. If $1 < m \leq k_*$, then all bombs will be allocated in a similar way, one bomb in a box.*

We call such a strategy of DF, where some boxes are left unprotected, a “green zone” strategy, where the green zone consists of boxes having some randomized protection. There is this kind of the “green zone” in Baghdad. The main importance of the Theorem 1 lies in the fact that an optimal strategy obtained there using a polynomial algorithm, gives an “estimate” of strategy π_* for a general case with testing. The full version of Theorem 1 contains a nontrivial recursive polynomial algorithm to obtain critical threshold k_* .

Our second theorem covers the “mixed” case, when for boxes in so called “transparent” set A , testing is perfect, and for all other boxes testing is non-informative. The main statement of Theorem 2 is that the optimal strategy of DF is to create two green zones, G_1 and G_2 . Each of the boxes in set $G_1 \subseteq A$,

receives one lock; the remaining locks are used to create green zone G_2 with the boxes, that are not in A . The optimal strategy of AT is to allocate a bomb at random between sites in the second green zone G_2 . Our third result is Theorem 3, which gives a full explicit solution to G-LBT for the case $n = 2$, $k = 1$.

The direct precursor of the current paper is K. Sonin and Wright (2024) [1], which offered a model of intelligence gathering in combat and used highly detailed data about Afghan rebel attacks, insurgent-led spy networks, and counterinsurgency operations to analyze its implications. A more general and abstract mathematical model called the Locks, Bombs and Testing (LBT) model was described in I. Sonin (2019) [2]. The solution for a related model is given in Li and I. Sonin (2021) [3].

References

- [1] *Sonin K., Wright A. L.* Rebel Capacity, Intelligence Gathering, and Combat Tactics // American Journal of Political Science. 2023.
- [2] *Sonin I.* Bayesian Game of Locks, Bombs and Testing. 2019. arXiv:1906.01163.
- [3] *Liu L., Sonin I.* Locks, Bombs and Testing: The Case of Independent Locks // In: Piunovskiy A., Zhang Y. (eds.) Modern Trends in Controlled Stochastic Processes. ECC, Springer. 2021. 41:248–265.

Об оценке параметров асимметричного TSP распределения и его применении

Александр В. Степанов, А. Г. Чуновкина

ВНИИ метрологии им. Д. И. Менделеева, Санкт-Петербург, Россия

E-mail: stepanov17@yandex.ru

В работе рассматривается стандартное двустороннее степенное распределение (Standard Two-Sided Power Distribution, STSP) [1,2], плотность которого определена на отрезке $[0, 1]$ и имеет вид:

$$f(x) = \begin{cases} p \left(\frac{x}{\theta}\right)^{p-1}, & 0 < x \leq \theta, \\ p \left(\frac{1-x}{1-\theta}\right)^{p-1}, & \theta \leq x < 1; \end{cases}$$

здесь $p > 0$, $0 \leq \theta \leq 1$ (случай $p = 1$ отвечает равномерному распределению, $p = 2$ — треугольному). Данное семейство распределений достаточно широко и разнообразно, но при этом имеет простое математическое описание, что сильно облегчает как математические выкладки, так и моделирование с его использованием (например, для вывода статистических критериев [3]).

Задача оценивания параметров асимметричных распределений (а также производных характеристик, таких, как интервал охвата для заданного уровня доверия), имеющих ограниченное множество значений, мода которых значительно смещена к границе этого множества, нередко возникает при рассмотрении практических задач, где существуют естественные ограничения на обрабатываемые (измеренные) значения. В качестве примера такой задачи можно привести измерение уровня содержания (концентрации) вещества в смеси [4], множество значений которого как раз совпадает с отрезком $[0, 1]$ (от 0% до 100%), и плотность распределения часто заметно выше вблизи того или иного конца данного отрезка. Не умаляя общности, можно ограничиться рассмотрением верхней границы множества допустимых значений (равной единице) и считать, что $\theta \geq 0.5$.

В работах [1, 2] предложены оценки по методу максимального правдоподобия параметров p , θ (обозначим их \hat{p} , $\hat{\theta}$, соответственно). В первой части данной работы изучаются характеристики точности указанных оценок. Оценка $\hat{\theta}$ имеет многомодальное асимметричное распределение. Установлено, что его математическое ожидание смещено в сторону 0 при

$\theta > 0.5$ тем сильнее, чем ближе θ к единице. Аналогично, оценка \hat{p} является смещенной в сторону больших значений. Методом Монте-Карло получены оценки указанных смещений, их дисперсии и соответствующие доверительные интервалы, в зависимости от значений исходных параметров p, θ , а также длины выборки n .

Получены явные выражения для границ односторонних и двусторонних интервалов охвата, показано, что их оценки, полученные на основе оценок $\hat{p}, \hat{\theta}$, также являются смещенными, оценено влияние их смещения на величину уровня доверия. Также рассмотрен вопрос построения кратчайших интервалов охвата.

Вторая часть посвящена задаче аппроксимации произвольных непрерывных распределений, определенных на отрезке $[0, 1]$, распределениями из семейства STPS, и подбору данных распределений по экспериментальным данным. Критерием качества в данном случае считаем близость характеристик аппроксимирующего распределения (в частности, математического ожидания, дисперсии, границ интервала охвата), к аналогичным характеристикам аппроксимируемого распределения. Другим критерием, имеющим практический интерес, представляется сохранение уровня доверия при переходе к аппроксимирующей плотности распределения. Иными словами, ожидается, что доверительная вероятность, полученная путем интегрирования аппроксимируемой плотности распределения, по интервалу охвата, оцененному с использованием ее STSP-аппроксимации, не будет заметно отличаться в меньшую сторону от заданного уровня доверия. Проверка последнего критерия проводилась методом Монте-Карло (для различных n), при этом в качестве аппроксимируемых были, в частности, рассмотрены равномерное и усеченное нормальное распределения, плотность которых сосредоточена вблизи правого конца данного отрезка.

Помимо непосредственной оценки параметров аппроксимирующего распределения, исходя из метода максимального правдоподобия, также рассмотрен метод выбора наиболее подходящего распределения из рассматриваемого STSP-семейства на основе обратного отображения [5]. Заметим, что в данном случае используемая функция, обратная к функции распределения, имеет особенно простой вид:

$$F^{-1}(y) = \begin{cases} \sqrt[p]{\theta^{p-1}y}, & 0 \leq y \leq \theta, \\ 1 - \sqrt[p]{(1-\theta)^{p-1}(1-y)}, & \theta < y \leq 1 \end{cases}$$

(напомним, что выбор осуществляется путем поиска пары параметров p, θ из некоторого конечного множества, доставляющей минимум функционалу $\sum_{i=1}^n (F^{-1}(\frac{i}{n}) - x_{(i)})^2$, где $x_{(i)}$ – вариационный ряд для исходной выборки). Применимость данного метода была ранее исследована авторами [6] для случая TSP-распределения более общего вида.

В третьей части работы рассмотрена задача построения (коррекции) асимметричного интервала охвата с использованием Байесовского подхода вблизи верхней границы области допустимых значений (т.е. вблизи точки $x = 1$) экспериментальных данных. Рассмотрены варианты выбора априорной плотности распределения из асимметричного семейства TSP распределений, даны рекомендации по выбору априорной плотности, в зависимости от параметров эксперимента (таких, как близость экспериментальных данных к верхней границе допустимых значений, величина приписываемой им неопределенности, а также формализованная априорная информация об экспериментальных данных).

Список литературы

- [1] *Van Dorp J. R., Kotz S.* The Standard Two-Sided Power Distribution and its Properties // *The American Statistician*. 2002. 56(2):90–99.
- [2] *Kotz S., Van Dorp J. R.* Beyond Beta: Other Continuous Families of Distributions with Bounded Support and Applications. World Scientific Publishing, 2004. 289 p.
- [3] *Stepanov A. V., Chunovkina A. G.* On testing of the homogeneity of variances for two-sided power distribution family // *Accreditation and Quality Assurance*. 2023. 28:129–137.
- [4] Руководство ЕВРАХИМ/СИТАК CG 4: Количественное описание неопределенности в аналитических измерениях. 3-е изд., перевод (УНИИМ), 2018. 158 с.
- [5] *Тырсин А.Н.* Метод подбора наилучшего закона распределения непрерывной случайной величины на основе обратного отображения // *Вестник ЮУрГУ, Серия Математика. Механика. Физика*. 2017. 9(1):31–38.
- [6] *Степанов А. В., Чуновкина А. Г.* Об одном методе подбора закона распределения из семейства TSP и его свойствах // *Материалы международной Воронежской зимней математической школы*. Воронеж, ВГУ. 2024. С. 238–241.

Метод Стейна и характеристические функции

Александр Н. Тихомиров

Физико-математический институт Коми Научного Центра УрО РАН,
Сыктывкар, Россия

E-mail: tikhomirov@ipm.komisc.ru

В 1972 году была опубликована работа Ч. Стейна "A bound for the error in the normal approximation to the distribution of sum of dependent random variables," с которой берет свое начало один из самых популярных методов оценки точности аппроксимации распределений в теории вероятностей и математической статистике. Ч. Стейн в своей статье высказывался о преимуществе прямых методов без использования аппарата характеристических функций. Однако, как показало время, во многих задачах аппарат характеристических функций весьма эффективен.

Метод Стейна – один из основных методов в теории вероятностей и математической статистике оценки точности аппроксимации распределения некоторой случайной величины W распределением другой случайной величины Z . Две составляющих метода Стейна:

- Структура, позволяющая преобразовать проблему оценки погрешности аппроксимации в задачу вычисления некоторого функционала от W .
- Набор методов для вычисления математических ожиданий, представленных в предыдущем пункте. Детали методов сильно зависят от свойств W и функционала.

Для аппроксимирующего распределения $F(x) = \Pr\{Z < x\}$ первую часть метода можно описать следующим образом.

Сначала находим подходящий оператор $A := A(F) = AZ$ (называемый оператором Стейна) и широкий класс функций $\mathcal{F}(A)$ (класс Стейна), такие, что

$\Pr\{Z < x\} = F(x)$ тогда и только тогда, когда $E[Af(Z)] = 0$ для всех $f \in \mathcal{F}(A)$. Далее, пусть H – класс, определяющий метрику на множестве распределений:

$$d_H(W, Z) := \sup_{h \in H} |Eh(W) - Eh(Z)|.$$

Предположим, для любой функции $h \in H$ существует функция $f_h \in \mathcal{F}$, такая, что

$$h(x) - Eh(Z) = Af(x).$$

Обозначим $\mathcal{F}(H) = \{f \in \mathcal{F}(A) : \text{существует } h \in H \text{ и } f = f_h\}$. Положим $x = W$ и возьмем математическое ожидание от обеих частей равенства. Получим

$$Eh(W) - Eh(Z) = E[Af(W)].$$

Имеет место неравенство

$$d_H(Z, W) \leq \sup_{f \in \mathcal{F}(H)} |E[Af(W)]|.$$

В работе 1976 года автор (см. [1]) предложил новую версию метода Стейна, сузив класс $\mathcal{F}(A)$ до класса функций $\Phi(A) = \{\varphi_t(x) = \exp\{itx\}, t \in \mathbb{R}\}$. Равенство $E[A\varphi(Z)] = 0$ для всех функций $\varphi_t(x) \in \Phi(A)$ необходимо и достаточно для идентификации распределения случайной величины Z . В частности, для аппроксимации нормальным законом оператор Стейна приводит к равенству $E[A\varphi_t(W)] = g'_W(t) + tg_W(t) =: \varepsilon_W(t)$, где $g_W(t) = E\varphi_t(W)$ — характеристическая функция случайной величины W . Оценки величины $\varepsilon_W(t)$ приводят к оценкам разности характеристических функций $g_W(t) - g_Z(t)$. Для оценки расстояния между распределениями случайных величин можно воспользоваться неравенствами с характеристическими функциями. Например неравенством Берри — Эссеена в случае центральной предельной теоремы. Более подробно об оценках расстояний между распределениями через их характеристические функции можно найти в обзоре С. Г. Бобкова [2].

Структура случайной величины W может быть различна. От этого во многом зависит метод оценки математического ожидания функционала $E[A\varphi_t(W)]$. В случае аппроксимации нормальным законом (оценки в центральной предельной теореме) чаще всего $W = \sum_{j=1}^n \xi_j$ — сумма случайных величин ξ_1, \dots, ξ_n , с $E\xi_j = 0$, зависимых, слабо зависимых или независимых. Метод, предложенный Стейном для оценки выражений вида $\sum_{j=1}^n E\xi_j f(W)$, возникающих в данном случае, можно условно назвать "локальным секционированием". Для любого фиксированного $j = 1, \dots, n$ выбирается множество "соседей" \mathbb{I}_j ($j \in \mathbb{I}_j$) так, чтобы величина $W^{(j)} := \sum_{k \notin \mathbb{I}_j} \xi_k$ была близка по распределению к W , и величина $\sum_{j=1}^n \sum_{k \in \mathbb{I}_j} E\xi_j \xi_k$ была близка к EW^2 .

В случае аппроксимации распределением Пуассона с параметром $\lambda > 0$ целочисленных случайных величин, можно также использовать аппарат характеристических функций. Оператор Стейна для распределения Пуассона имеет вид $Af(x) = \lambda f(x+1) - xf(x)$. Впервые он был предложен

Ченом (L. Chen) в [3]. Применяя оператор Стейна к характеристическим функциям, мы приходим к равенству $E[A\varphi_t(W)] = i\lambda e^{it}g_W(t) - g'_W(t)$.

С помощью характеристических функций можно оценивать точность в аппроксимации биномиальным распределением $Bi(n, p)$ с параметрами n и p . В этом случае оператор Стейна имеет вид $Af(x) = \frac{p}{1-p}(n-x)f(x+1) - xf(x)$. Применяя его к характеристическим функциям, получим $E[A\varphi_t(W)] = \frac{1}{1-p}(npe^{it}g_W(t) - (1-p+pe^{it})g'_W(t))$.

Более детальное описание применения характеристических функций в методе Стейна можно найти в работах [4–7].

Список литературы

- [1] *Тихомиров А. Н.* О скорости сходимости в центральной предельной теореме для слабо зависимых величин // Вестник Ленинградского университета. Серия "Математика. Механика. Астрономия". 1976. С. 158–159.
- [2] *Бобков С. Г.* Близость вероятностных распределений в терминах преобразований Фурье–Стилтьеса // Успехи математических наук. 2016. 71(6):37–98.
- [3] *Chen L. H. Y.* Poisson approximation for dependent trials // Annals of Probability. 1975. 3:534–545.
- [4] *Arras B., Mijoule G., Poly G., Swan Y.* A new approach to the Stein–Tikhomirov method: with applications to the second Wiener chaos and Dickman convergence. 2016. doi: 10.48550/arXiv.1605.06819
- [5] *Formanov Sh.* On the Stein–Tikhomirov method and its applications in nonclassical limit theorems // Discrete Mathematics and Applications. 2007. 17(1):23–36. doi: 10.1515/DMA.2007.004
- [6] *Eichelbacher P., Redno B.* Kolmogorov bounds for decomposable random variables and subgraph counting by the Stein–Tikhomirov method. 2021. doi: 10.48550/arXiv.2107.03775
- [7] *Arras B., Mijoule G., Poly G., Swan Y.* A new approach to the Stein–Tikhomirov method. 2017. doi: 10.48550/arXiv.1605.06819v2

Functional limit theorems for purtical numbers of cells in a general allocation scheme

István Fazekas¹, Alexey N. Chuprunov²

¹Debrecen University, Debrecen, Hungary; ²ChuGY, Cheboksary, Russia

E-mail: fazekasi@inf.unideb.hu; achuprunov@mail.ru

In 1968 V. F. Kolchin introduced a notion of a generalized allocation scheme. Let $\xi_1, \xi_2, \dots, \xi_N$ be independent integer valued nonnegative identically distributed random variables. We say that the random variables η_1, \dots, η_N fit to the generalized allocation scheme of n particles by N cells if

$$\mathbf{P}\{\eta_1 = k_1, \dots, \eta_N = k_N\} = \mathbf{P}\left\{\xi_1 = k_1, \dots, \xi_N = k_N \left| \sum_{i=1}^N \xi_i = n \right.\right\},$$

where k_1, k_2, \dots, k_N are nonnegative integer numbers such that $k_1 + k_2 + \dots + k_N = n$.

Various models of discret probability theory such as random forests, random permutations, random allocations, urn schemes are particular cases of a general allocation scheme (see [1, 2]).

Let $0 < K < N$. The aim of the talk is to study the convergence in distribution in Skhorohod space of the random processes

$$X_{n,K,N}(t) = \sum_{i=1}^{[tK]} \eta_i, \quad 0 \leq t \leq 1.$$

Observe that $X_{n,K,N}(t)$ is a number of particles in the first $[tK]$ cells.

We will use $\xi_i = \xi_i(\beta)$, $1 \leq i \leq N$, with the distribution

$$p_k = p_k(\beta) = \mathbf{P}\{\xi_i = k\} = \frac{b_k \beta^k}{k! B(\beta)}, \quad k = 0, 1, 2, \dots,$$

where $B(\beta) = \sum_{k=0}^{\infty} \frac{b_k}{k!} \beta^k$, $b_k \geq 0$, and the radius of convergence of this series $R > 0$.

Denote:

$$m(\beta) = \mathbf{E}\xi_i(\beta), \quad \sigma^2(\beta) = \mathbf{E}(\xi_i(\beta) - m(\beta))^2.$$

We will use the property

$$Al(l_1) : \quad a_0 = 0, \dots, a_{l-1} = 0, \quad a_l > 0, \quad a_{l+1} = 0, \dots, a_{l-1} = 0, \quad a_{l+1} > 0,$$

© I. Fazekas, A. N. Chuprunov, 2024

for some $0 \leq l < l_1$. For $l = 0$ the property $Al(l_1)$ was introduced by V. F. Kolchin and A. V. Kolchin.

Observe that $\xi_i(\beta) = l$ if and only if $B(\beta) = a_l\beta^l$. So $\xi_i(\beta)$ is not a constant if and only if $Al(l_1)$ is valid for some $0 \leq l < l_1$.

Really a general allocation scheme with the random variables $\xi_i = \xi_i(\beta)$, $1 \leq i \leq N$, does not depend from β .

We will denote: \xrightarrow{D} is a convergence in distribution in Skhorohod space, \mathbf{N}_0 is a set of nonnegative integer numbers, X_λ is a homogenous Poisson process with the intensity λ , W is Wiener process, W_0 is Brownian bridge, F_n is an empirecal process.

Denote:

$$X'_{n,K,N}(t) = X_{n,K,N}(t) - l[tK], \quad 0 \leq t \leq 1.$$

Theorem 1. Consider a general allocation scheme of n particles by N cells defined by the random variables $\xi_i = \xi_i(\beta)$, $1 \leq i \leq N$, having the property $Al(l_1)$. Let $N \rightarrow \infty$ and $n - lN = (l_1 - l)m$, where $m \in \mathbf{N}$ is a fixed number. Then we have

$$\frac{X'_{n,N,N}}{l_1 - l} \xrightarrow{D} mF_m.$$

Theorem 2. Consider a general allocation scheme of n particles by N cells defined by the random variables $\xi_i = \xi_i(\beta)$, $1 \leq i \leq N$, having the property $Al(l_1)$. Suppose that $n, K, N \rightarrow \infty$ such that $n - lN \rightarrow \infty$, $\frac{n-lN}{N} \rightarrow 0$, $\frac{(n-lN)K}{N} \rightarrow \lambda$, where $0 < \lambda < \infty$. Then we have

$$\frac{X'_{n,K,N}}{l_1 - l} \xrightarrow{D} X_\lambda.$$

We will use the conditions:

$$m(\beta) = l + \frac{n}{N}. \quad (1)$$

$$\sqrt{2\pi N}\sigma(\beta)\mathbf{P}\{S_N(\beta) = k\} = \exp\left(-\frac{(k - Nm(\beta))^2}{2N\sigma^2(\beta)}\right) + o(1) \quad (2)$$

uniformly by $k \in \mathbf{N}_0$,

$$K\sigma^2(\beta) \rightarrow \infty, \quad (3)$$

$$\frac{K}{N} \rightarrow s, \quad (4)$$

where $0 \leq s \leq 1$. Let the random proseece $W_0^{\{s\}}$, where $0 < s \leq 1$ is a parameter, defined by the formular

$$W_0^{\{s\}}(t) = \frac{1}{\sqrt{s}}W_0(st), \quad 0 \leq t \leq 1.$$

Denote:

$$Y_{n,K,N}(t) = X_{n,K,N}(t) - [tK] \left(l + \frac{n}{N} \right), \quad 0 \leq t \leq 1.$$

Theorem 3. Consider a general allocation scheme of n particles by N cells defined by the random variables $\xi_i = \xi_i(\beta)$, $1 \leq i \leq N$, having the property $Al(l_1)$. If $N, K, n \rightarrow \infty$ such that (1), (2), (3), (4), are valid then we have

$$\frac{Y_{n,K,N}}{\sqrt{K}\sigma(\beta)} \xrightarrow{D} W_0^{\{s\}} \quad \text{for } 0 < s \leq 1, \quad \frac{Y_{n,K,N}}{\sqrt{K}\sigma(\beta)} \xrightarrow{D} W \quad \text{for } s = 0.$$

Using Kolchin local theorems from Theorem 3 we obtain

Theorem 4. Consider a general allocation scheme of n particles by N cells defined by the random variables $\xi_i = \xi_i(\beta)$, $1 \leq i \leq N$, having the property $Al(l_1)$. If $N, K, n \rightarrow \infty$ such that $\frac{n}{N} \rightarrow l + 0$, $\frac{nK}{N} \rightarrow \infty$ and (4) is valid then we have

$$\frac{Y_{n,K,N}}{\sqrt{\frac{Kn}{N}}} \xrightarrow{D} W_0^{\{s\}} \quad \text{for } 0 < s \leq 1, \quad \frac{Y_{n,K,N}}{\sqrt{\frac{Kn}{N}}} \xrightarrow{D} W \quad \text{for } s = 0.$$

Theorem 5. Consider a general allocation scheme of n particles by N cells defined by the random variables $\xi_i = \xi_i(\beta)$, $1 \leq i \leq N$, having the property $Al(l_1)$. Let the numbers C_1, C_2 be such that $l < C_1 < C_2 < \infty$ and it exists $\beta_2 \in (0, R)$ with the property: $m(\beta_2) = C_2$. Suppose that $n, K, N \rightarrow \infty$ such that $C_1 < l + \frac{n}{N} < C_2$ and (4) is valid. Then we have

$$\frac{Y_{n,K,N}}{\sqrt{K}\sigma(\beta)} \xrightarrow{D} W_0^{\{s\}} \quad \text{for } 0 < s \leq 1, \quad \frac{Y_{n,K,N}}{\sqrt{K}\sigma(\beta)} \xrightarrow{D} W \quad \text{for } s = 0,$$

there β is such that (1) is valid.

References

- [1] Kolchin V. F. Random graphs. Cambridge: Cambridge University Press, 1999. 252 p.
- [2] Pavlov Yu. L. Random forests. Utrecht: VSP, 2000. 122 p.

Предельный процесс для последовательных количеств слов, встретившихся в тексте ровно один раз

Шахзод Шухрат угли Файзуллаев

Новосибирский государственный университет, Новосибирск, Россия

E-mail: s.faizullaev@g.nsu.ru

Независимые одинаково распределенные случайные величины X_1, X_2, \dots, X_n имеют целочисленное степенное распределение:

$$p_k := \mathbf{P}(X_1 = k) = l(\alpha, k)k^{-\alpha}, \quad k \geq 1, \quad (1)$$

где $\alpha > 1$ – неизвестный параметр, $l(\alpha, k) > 0$ – функция такая, что выполнено условие нормировки

$$\sum_{k=1}^{\infty} l(\alpha, k)k^{-\alpha} = 1,$$

и $l(\alpha, x)$ – медленно меняющаяся функция при $x \rightarrow \infty$ для любого $\alpha > 1$, то есть $l(\alpha, cx)/l(\alpha, x) \rightarrow 1$ для любого $c > 0$.

Итак, наша идея состоит в том, чтобы сравнить количество разных слов и слов, которые встречаются во всем тексте только один раз, R_n и $R_{n,1}$, с теми же количествами в первой половине текста, $R_{\lfloor n/2 \rfloor}$ и $R_{\lfloor n/2 \rfloor,1}$. Обозначим $\theta = 1/\alpha$, $0 < \theta < 1$.

Бахадур [1] был первым, кто объяснил закон увеличения числа различных слов на основе закона Ципфа и элементарной вероятностной модели схемы бесконечной урны. Кей [5] изучил асимптотику числа уникальных слов. В частности, показано, что если $\lim_{i \rightarrow \infty} p_{i+1}/p_i = 1$, то $R_{n,1} \rightarrow_p \infty$ при $n \rightarrow \infty$. Если $\limsup_{i \rightarrow \infty} p_{i+1}/p_i < 1$, то $\mathbf{E}R_{n,1}$ равномерно ограничено.

Карлин [4] доказал сходимость $R_n/\mathbf{E}R_n \rightarrow 1$, $R_{n,1}/\mathbf{E}R_{n,1} \rightarrow 1$ п. н. и нашел асимптотику

$$\mathbf{E}R_n \sim \Gamma(1 - \theta)L(\theta, n)n^\theta, \quad \mathbf{E}R_{n,1} \sim \theta\Gamma(1 - \theta)L(\theta, n)n^\theta \quad (2)$$

при $n \rightarrow \infty$.

Чебунин и Ковалевский [3] предложили оценку

$$\hat{\theta} = R_{n,1}/R_n. \quad (3)$$

Из СЗБЧ и (2) эта оценка сильно состоятельная, $\hat{\theta} \rightarrow \theta$ п. н. при $n \rightarrow \infty$.

Мы используем функциональную центральную предельную теорему для векторного процесса $(R_{[nt]}, R_{[nt],1})$, $0 \leq t \leq 1$.

Пусть для $t \in [0, 1], k \geq 1$

$$Y_n(t) = \frac{R_{[nt]} - \mathbf{E}R_{[nt]}}{\sqrt{\mathbf{E}R_n}}, \quad Y_{n,k}(t) = \frac{R_{[nt],k} - \mathbf{E}R_{[nt],k}}{\sqrt{\mathbf{E}R_n}}.$$

Следующая лемма является прямым следствием теоремы Чебунина и Ковалевского [2]. Эта лемма является основой наших теорем, которые дают новый метод статистического анализа текстов.

Лемма. Если (1) выполнено, $\theta = 1/\alpha$, то случайный вектор $(Y_n(1), Y_n(1/2), Y_{n,1}(1), Y_{n,1}(1/2))$ слабо сходится к нормальному вектору $(Y(1), Y(1/2), Y_1(1), Y_1(1/2))$ с нулевым математическим ожиданием и ковариационной матрицей $G(\theta) = (g_{ij}(\theta))_{i,j=1}^4$, $g_{ji}(\theta) = g_{ij}(\theta)$,

$$\begin{aligned} g_{11}(\theta) &= c_{00}(1, 1) = 2^\theta - 1, \\ g_{12}(\theta) &= c_{00}(1/2, 1) = (3/2)^\theta - 1, \\ g_{13}(\theta) &= c_{01}(1, 1) = -\theta 2^{\theta-1}, \\ g_{14}(\theta) &= c_{10}(1/2, 1) = -\theta (1/2 - (3/2)^{\theta-1}), \\ g_{22}(\theta) &= c_{00}(1/2, 1/2) = 1 - (1/2)^\theta, \\ g_{23}(\theta) &= c_{01}(1/2, 1) = \theta (1/2 - (3/2)^{\theta-1}), \\ g_{24}(\theta) &= c_{01}(1/2, 1/2) = -\theta/2, \\ g_{33}(\theta) &= c_{11}(1, 1) = \theta (1 - 2^{\theta-2}(1 - \theta)), \\ g_{34}(\theta) &= c_{11}(1/2, 1) = \theta (1/2)^\theta (1 - (3/2)^{\theta-2}(1 - \theta)), \\ g_{44}(\theta) &= c_{11}(1/2, 1/2) = \theta ((1/2)^\theta - (1 - \theta)/4). \end{aligned}$$

Нам нужно более ограничительное условие по вероятностям:

$$p_k = ck^{-1/\theta}(1 + o(k^{-1/2})) \text{ as } k \rightarrow \infty, \quad c > 0, \quad 0 < \theta < 1. \quad (4)$$

Теорема 1. Если (4) выполняется, то статистика

$$V_n^{(1)} = \sqrt{R_n} \left(\frac{R_{n,1}}{R_n} - \frac{R_{\lfloor n/2 \rfloor,1}}{R_{\lfloor n/2 \rfloor}} \right)$$

слабо сходится к нормальной случайной величине $V^{(1)} = Y_1(1) - \theta Y(1) - 2^\theta Y_1(1/2) + \theta 2^\theta Y(1/2)$ с нулевым математическим ожиданием и дисперсией $b_1^2(\theta) = m_1(\theta)G(\theta)m_1^T(\theta)$, где $m_1(\theta) = (-\theta, \theta 2^\theta, 1, -2^\theta)$.

Теорема 2. Если (4) выполняется, то статистика

$$V_n^{(2)} = \sqrt{R_n} \left(\frac{R_{\lfloor n/2 \rfloor}}{R_n} - \frac{R_{\lfloor n/2 \rfloor,1}}{R_{n,1}} \right)$$

слабо сходится к нормальной случайной величине $V^{(2)} = Y(1/2) - 2^{-\theta} Y(1) - \theta^{-1} Y_1(1/2) + 2^{-\theta} Y_1(1)$ с нулевым математическим ожиданием и дисперсией $b_2^2(\theta) = m_2(\theta)G(\theta)m_2^T(\theta)$, где $m_2(\theta) = (-2^{-\theta}, 1, 2^{-\theta}, -\theta^{-1})$.

Список литературы

- [1] *Bahadur R. R.* On the number of distinct values in a large sample from an infinite discrete distribution // Proceedings of the National Institute of Sciences of India, Supp. II. 1960. 26A:67–75.
- [2] *Chebunin M., Kovalevskii A.* Functional central limit theorems for certain statistics in an infinite urn scheme // Statistics and Probability Letters. 2016. 119:344–348.
- [3] *Chebunin M., Kovalevskii A.* Asymptotically normal estimators for Zipf's law // Sankhya A. 2019. 81:482–492.
- [4] *Karlin S.* Central Limit Theorems for Certain Infinite Urn Schemes // Journal of Mathematics and Mechanics. 1967. 17(4):373–401.
- [5] *Key E. S.* Rare Numbers // Journal of Theoretical Probability. 1992. 5(2):375–389.

Вероятностная оценка ширины дерева поиска в задаче изоморфного вхождения подграфа

Игорь А. Федотов

Нижегородский государственный технический университет им. П. Е.
Алексеева, Дзержинск, Россия

E-mail: optimal.control@gmail.com

Мотивацией начала работы изначально были сугубо утилитарные соображения по оценке требуемой алгоритму [1] памяти при работе в режиме обхода дерева поиска в ширину. Проблемы с аллокацией памяти были препятствием в развитии tree-search методов, миграции dfs \rightarrow bfs, и sru \rightarrow gru. Переход к вероятностной оценке движения по дереву поиска сделаем на основе замены основных булевых операций, последовательность которых строго определена структурой паттерн графа [2] (булевой матрицей смежности A) их вероятностной оценкой. Понятия плотности для графов далее будем отождествлять с вероятностью. Для этого рассмотрим два 8-битных регистра (в little-endian соглашении)

mov AL, 00110101b; (53_{dec})

mov BL, 00111000b; (56_{dec})

and AL, BL; AL = 00110000b (48_{dec}) общие вершины 3,4

popcnt AX, AX; (AX = 2)

вершины 7 и 2 соответственно. Примем I_A – число единиц принадлежащих только множеству AL (2), I_B – число единиц принадлежащих только множеству BL (1), I_{AB} – число единиц принадлежащих обоим множествам AL, BL (2), n – количество элементов в множестве (8). Тогда оценки для основных побитовых логических операторов будут:

not(\sim) : $P_{not}^A = 1 - P_A$

and(&) : $P_{and}^{AB} = \frac{I_{AB}}{n} = P(A) \cdot P(B/A) = P(B) \cdot P(A/B) =$

$\frac{I_A + I_{AB}}{n} \cdot \frac{I_{AB}}{I_A + I_{AB}} = \frac{I_{AB}}{n} \approx P_A \cdot P_B$

or(|) : $P_{or}^{AB} = \frac{I_A + I_B + I_{AB}}{n} = P(A) + P(B) - P(AB) =$

$\frac{I_A + I_{AB} + I_B + I_{AB}}{n} - \frac{I_{AB}}{n} = \frac{I_A + I_B + I_{AB}}{n} \approx P_A + P_B - P_A \cdot P_B$

xor(^) : $P_{xor}^{AB} = \frac{I_A + I_B}{n} = P(A) + P(B) - 2 \cdot P(AB) =$

$\frac{I_A + I_{AB} + I_B + I_{AB}}{n} - 2 \frac{I_{AB}}{n} = \frac{I_A + I_B}{n} \approx P_A + P_B - 2 \cdot P_A \cdot P_B$

И оценка популяции результата будет равна например **popcnt**(**and** AL, BL) $\approx P_A \cdot P_B \cdot n$, где n размер регистра, или

$\text{popcnt}(\text{and } AL, BL) \approx \frac{\text{popcnt}(AL) \cdot \text{popcnt}(BL)}{n} = \frac{4 \cdot 3}{8} = 1.5$. Для ориентированных графов ситуация немного сложнее, будет четыре типа ребер *none*, *in*, *out*, *both*:

none (несуществующие) : $P_{\text{none}} = (1 - P_{\text{in}}) \cdot (1 - P_{\text{out}}) = \overline{P_{\text{in}}} \cdot \overline{P_{\text{out}}}$

in (строго входящие) : $P_{\text{in}} = P_{\text{in}} \cdot \overline{P_{\text{out}}} = P_{\text{in}} \cdot (1 - P_{\text{out}}) = P_{\text{in}} - P_{\text{both}}$

out (строго исходящие) : $P_{\text{out}} = P_{\text{out}} \cdot \overline{P_{\text{in}}} = P_{\text{out}} \cdot (1 - P_{\text{in}}) = P_{\text{out}} - P_{\text{both}}$

both (двойные) : $P_{\text{both}} = P_{\text{in}} \cdot P_{\text{out}}$

Сделанные выкладки позволяют априорно оценить результат последовательности векторных операций над регистрами с известной заселённостью, осталось предварительно оценить заселённость в регистрах по уровню дерева и сформулировать правило вычисления вероятностей и оценки размера. Плотность популяций графа без петель выражается как

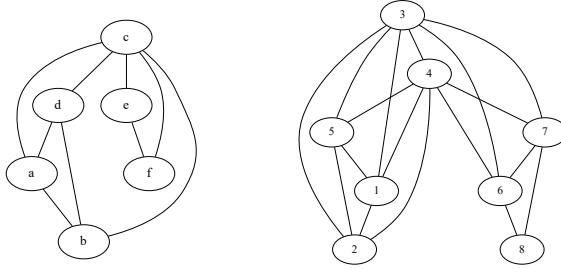


Рис. 1: Модельная задача subgraph изоморфизма

$$P_e = \frac{1}{(n-1)} \cdot \text{diag}(\mathbf{d}) = \frac{1}{(n-1)} D,$$

где \mathbf{d} – вектор степеней, n = кол-во вершин графа, а само решение модельной задачи приведено в табл. 1 и 2.

Таблица 1: Вероятности появления биекции у вершины Pattern графа

v_i	deg	RI: P_{v_i}	VF3 : P_{v_i}	Possible Matchings	Ave: P_e	Ave: P_e
a	3	$\frac{4}{6}$	$\frac{7}{8}$	1,2,3,4,5,6,7	$\frac{32}{49}$	$\frac{17}{49}$
b	3	$\frac{4}{6}$	$\frac{7}{8}$	1,2,3,4,5,6,7	$\frac{32}{49}$	$\frac{17}{49}$
c	5	$\frac{1}{6}$	$\frac{2}{8}$	3,4	$\frac{6}{7}$	$\frac{1}{7}$
d	3	$\frac{4}{6}$	$\frac{7}{8}$	1,2,3,4,5,6,7	$\frac{32}{49}$	$\frac{17}{49}$
e	2	$\frac{2}{6}$	$\frac{8}{8}$	1,2,3,4,5,6,7,8	$\frac{34}{56}$	$\frac{22}{56}$
f	2	$\frac{2}{6}$	$\frac{8}{8}$	1,2,3,4,5,6,7,8	$\frac{34}{56}$	$\frac{22}{56}$

Таблица 2: Оценка ширины дерева поиска

core level	v_i	consist.cond.	$P_{v,c}$	$P_{v,c}$ val	P_v	Breeding factor (b)	eval Width w (real)
0	c	$\bar{\forall}$	1	1	$2/8$	2	2 (2)
1	b	$r[l_0]$	P_c^d	$6/8$	$7/8$	6	12 (12)
2	a	$r[l_0] \wedge r[l_1]$	$P_c^e \cdot P_b^f$	$3 \cdot 2^6 / 7^8$	$7/8$	$3 \cdot 2^6 / 7^8 \approx 3.9184$	47.02 (36)
3	d	$r[l_0] \wedge r[l_1] \wedge r[l_2]$	$P_c^e \cdot P_b^f \cdot P_a^g$	$3 \cdot 2^{11} / 7^8$	$7/8$	$3 \cdot 2^{11} / 7^8 \approx 2.5589$	120.3221 (60)
4	e	$r[l_0] \wedge r[l_1] \wedge r[l_2] \wedge r[l_3]$	$P_c^e \cdot P_b^f \cdot P_a^g \cdot P_d^h$	$3 \cdot 2^{17} / 7^8$	$7/8$	$3 \cdot 2^{17} / 7^8 \approx 0.2864$	34.4546 (24)
5	f	$r[l_0] \wedge r[l_1] \wedge r[l_2] \wedge r[l_3] \wedge r[l_4]$	$P_c^e \cdot P_b^f \cdot P_a^g \cdot P_d^h \cdot P_e^i$	$3 \cdot 17^4 / 2 \cdot 7^8$	$7/8$	$3 \cdot 2^2 \cdot 17^4 / 7^8 \approx 0.1739$	5.9902 (24)
Expr:		$\bigcup_{i \in N} r[l_i] \wedge \bigcap_{j \in P} r[l_j]$				$P_{c,e_i} \cdot P_{v_i} \cdot n$	$b_i \cdot w_{i-1} = \prod_{j=0}^i b_j = \prod_{j=0}^i P_{c,e_j} \cdot \prod_{j=0}^i P_{v_j} \cdot n^{i+1}$
Σ :							$W = \sum_{i=0}^{m-1} w_i = \sum_{i=0}^{m-1} \prod_{j=0}^i b_j = \mathbf{221.79}$ (158)

Сформируем вероятностную матрицу P которая позволяет рассчитать коэффициент размножения (breeding factor по аналогии с атомными реакторами) на слое дерева поиска и оценить его ширину. Она будет записана относительно ядра (внутреннего понятия tree-search алгоритмов – неполная логически непротиворечивая биекция паттерн вершин на их таргет отображения).

$$P = \text{diag}(P_v) + (\text{diag}(P_{out}) \cdot U(A) + (I - \text{diag}(P_{out})) \cdot U(\bar{A} - I)) + (L(A) \cdot \text{diag}(P_{in}) + L(\bar{A} - I) \cdot (I - \text{diag}(P_{in})))$$

Данный вид матрицы обладает некоторыми свойствами, так на диагонали стоят вероятности матчнга вершин паттерн графа без учета ребер, которые до этой работы уже по-разному рассчитывались в [3–5], верхний надтреугольник суть вероятности извлечения исходящих ребер для соответствующего уровня, нижний – входящих ребер. Уровень дерева поиска начинается с элемента [0,0] и движется по диагонали до [m,m] (m - кол-во вершин паттерн графа). Атрибуты при вершинах и петли учитываются в диагональных элементах, а для учета атрибутов на ребрах необходимо ввести доп. матричный множитель, который формируется также на основе структуры матрицы A . При расчетах, впрочем удобнее пользоваться верхнетреугольной формой этой матрицы:

$$P^U = \left(D_v^{1/2} + D_{out} \cdot U(A) + (I - D_{out}) \cdot U(\bar{A} - I) \right) \circ \left(D_v^{1/2} + D_{in} \cdot U(A^T) + (I - D_{in}) \cdot U(\bar{A}^T - I) \right)$$

где \circ – Адамарово покомпонентное произведение [6], $L(A), U(A)$ – нижний и верхний треугольники матрицы A , \bar{A} – логическое отрицание (дополнение графа), A^T – транспонирование матрицы. Верхнетреугольная форма, помимо упрощения вычислений удобна еще и тем, что она как минимум положительно полуопределённая $P^U \succcurlyeq 0$. Но на самом деле если она полуопределённая то решения задачи subgraph изоморфизма не будет, для разрешимости задачи она должна быть положительно определенной $P^U \succ 0$. Поскольку на диагонали стоят вероятности матчнга

вершин (они же и собственные значения), и если хотя бы одна из них равна 0, то полной биекции не будет существовать.

Тогда к. размножения (b_i) и ширину слоя (w_i) можно записать как:

$$b_i = n \cdot \left(p_{ii} \cdot \prod_{j=0}^{i-1} p_{ij} \cdot \prod_{k=0}^{i-1} p_{ki} \right) = n \cdot \left(p_{ii} \cdot \prod_{j=0}^{i-1} p_{ji}^U \right)$$

$$w_i = b_i \cdot w_{i-1} = n^{i+1} \cdot \prod_{j,k=0}^i p_{j,k}$$

и общий размер дерева:

$$\mathbf{W} = \sum_{i=0}^{m-1} w_i$$

Список литературы

- [1] Intel Subgraph isomorphism algorithm (2019-20). github.com/oneapi-src/oneDAL/tree/master/cpp/oneapi/dal/algo/subgraph_isomorphism
- [2] Федотов И. А. Особенности алгоритмической и архитектурной оптимизации алгоритмов сопоставления графов на примере задачи изоморфного вхождения подграфа // Графы, игры и модели: материалы междунар. науч. конф., Майкоп: АГУ. 2022. С. 42–46. URL: <https://ggm.cmcagu.ru/>
- [3] Carletti V., Foggia P., Saggese A., Vento M. Challenging the time complexity of exact subgraph isomorphism for huge and dense graphs with VF3 // IEEE transactions on pattern analysis and machine intelligence. 2018. 40(4):804–818. doi: 10.1109/TPAMI.2017.2696940
- [4] Carletti V., Foggia P., Ritrovato P., Vento M., Vigilante V. A Parallel Algorithm for Subgraph Isomorphism // Lecture Notes in Computer Science. 2019. 11510:141–151. doi: 10.1007/978-3-030-20081-7_14
- [5] Bonnici V., Giugno R., Pulvirenti A., Shasha D., Ferro A. A subgraph isomorphism algorithm and its application to biochemical data // BMC Bioinformatics. 2013, 14(7):S13. doi: 10.1186/1471-2105-14-S7-S13
- [6] Хорн Р., Джонсон Ч. Матричный анализ. М.: Мир, 1989. 656 с.

Критерий отношения правдоподобия проверки сложных гипотез s -мерной равномерности двоичных последовательностей

Юрий С. Харин

НИИ прикладных проблем математики и информатики,
Белорусский государственный университет, Минск, Беларусь

E-mail: kharin@bsu.by

В криптографии генерация случайности сводится к генерации равномерно распределенной случайной последовательности (РПСЦ). При этом неизбежно возникает сопутствующая проблема статистического тестирования (проверки гипотезы $H_0 = \{\text{наблюдаемая последовательность есть РПСЦ}\}$ против альтернативы $H_1 = \overline{H_0}$), для решения которой существуют десятки «батарей тестов»: NIST, STS, DieHarder, gJrand, и другие.

Проведенный в [1] обзор существующих тестов выявляет следующие недостатки: 1) многие тесты ориентированы на проверку лишь частных случаев альтернативы H_1 ; 2) многие тесты не фиксируют семейство альтернатив и не имеют оценок мощности w_T ; 3) при включении нескольких тестов в «батарею» не удается учесть их вероятностную зависимость. Значимой критике подвергнута «батарея» NIST STS [2], что привело к решению о начале «ревизии» этой «батареи тестов» [3].

Существующие статистические тесты и построенные на их основе батареи обладают еще одним существенным недостатком, приводящим к следующему парадоксу: при увеличении длины $T \rightarrow +\infty$ наблюдаемых последовательностей, порождаемых реально существующими источниками случайности, вероятность принятия альтернативы H_1 стремится к единице: $w_T \rightarrow 1$, поэтому «увеличением длины проверяемой последовательности T можно забраковать любой реальный источник случайности». Объяснение этого «парадокса» состоит в следующем. Реальные источники случайности – не идеальны, они отклоняются от гипотезы H_0 по своему вероятностному распределению на некоторую величину $\varepsilon \neq 0$, а тесты, включаемые в «батарею» обладают свойством состоятельности проверки простой гипотезы H_0 против сложной альтернативы H_1 : $w_T \rightarrow 1$ при $T \rightarrow +\infty$. Таким образом, этот «парадокс» порожден некорректностью

математической модели нулевой гипотезы H_0 : нулевая гипотеза должна быть сложной и задавать допустимую величину ε отклонений от РРСП. В связи с этим актуальна рассматриваемая в докладе задача построения сложных нулевых гипотез H_0^ε , критерия правдоподобия заданного уровня значимости α_0 для проверки $H_0^\varepsilon, \overline{H_0^\varepsilon}$.

Пусть на вероятностном пространстве $(\Omega, \mathcal{F}, \mathbf{P})$ наблюдается двоичная случайная последовательность длины $T=n \cdot s$, разбитая на n последовательных s -грамм ($s \in \mathbb{N}$): $X=X_1^T=(x_1, \dots, x_T) \equiv (X_1^s, X_{s+1}^{2s}, \dots, X_{(n-1)s+1}^T) \in V^T, V=\{0, 1\}$, независимых в совокупности и одинаково распределенных с распределением вероятностей $p=(p_{J_1^s}) \in \mathcal{P}, p_{J_1^s}=\mathbf{P}\{X_{(i-1)s+1}^{is}=J_1^s\}$,

$$\mathcal{P}=\left\{p=(p_{J_1^s}) : p_{J_1^s} \geq 0, J_1^s \in V^s, \sum_{J_1^s \in V^s} p_{J_1^s}=1\right\}$$

– семейство (симплекс) всевозможных s -мерных распределений на V^s .

В дальнейшем для упрощения обозначений биективно перейдем от мультииндекса $J_1^s=(j_1, \dots, j_s) \in V^s$ к одномерному индексу $k \in \{0, 1, \dots, K-1\}$, $K=2^s$.

В «батареях тестов» используются тесты статистической проверки простой нулевой гипотезы $H_0=\{p=p_0\}$, $p_0=(p_{0k}), p_{0k} \equiv 1/K$, против сложной альтернативы $\overline{H_0}$. Во избежание парадокса «неадекватности модели нулевой гипотезы» введем в рассмотрение сложную гипотезу об s -мерной равномерности: $H_0^\varepsilon=\{p=(p_k) \in \mathcal{P}_0^\varepsilon\}$,

$$\mathcal{P}_0^\varepsilon=\left\{p \in \mathcal{P} : \|p-p_0\|=\sqrt{\sum_{k=0}^{K-1} (p_k-1/K)^2} \leq \varepsilon\right\}$$

– пересечение симплекса \mathcal{P} с гипершаром в R^K заданного радиуса ε с центром в точке p_0 s -мерного равномерного распределения; $\varepsilon \in (0, \sqrt{1-1/K})$ – достаточно малый параметр нулевой гипотезы, определяющий допустимые отклонения от точечной гипотезы H_0 . Гипотеза H_0 является предельной по отношению к H_0^ε : $H_0^\varepsilon \rightarrow H_0$ при $\varepsilon \rightarrow 0$.

Примем обозначения: $S_\varepsilon=\{p \in \mathcal{P} : \|p-p_0\|=\varepsilon\} \subset \mathcal{P}_0^\varepsilon$ – пересечение гиперсферы радиуса ε с центром в точке p_0 с симплексом \mathcal{P} ; ν_k – частота встречаемости s -граммы k среди n s -грамм последовательности X ; $\hat{p}=(\hat{p}_k), \hat{p}_k=\nu_k/n, k=0, 1, \dots, K-1$, – несмещенная, состоятельная, эффективная оценка максимального правдоподобия для p_k ; $l(X;p)=n \sum_{k=0}^{K-1} \hat{p}_k \ln p_k$ – логарифмическая функция правдоподобия (ЛФП); $l_{S_\varepsilon}^*(X)=\max_{p \in S_\varepsilon} l(X;p)=l(X;p^*)$ – максимальное значение ЛФП на гиперсфере S_ε , $\lambda(X)=\ln\left(\left(\sup_{p \in \mathcal{P}_0^\varepsilon} \mathbf{P}(X;p)\right) / \left(\sup_{p \in \mathcal{P}} \mathbf{P}(X;p)\right)\right) \in [0, 1]$.

Лемма. Логарифмическая статистика отношения правдоподобия для проверки сложных гипотез $H_0^\varepsilon, H_1^\varepsilon$ об s -мерной равномерности по наблюдаемой двоичной последовательности $X \in V^T$ имеет вид:

$$\lambda(X) = \begin{cases} l_{S_\varepsilon}^*(X) - n \sum_{k=0}^{K-1} \hat{p}_k \ln \hat{p}_k, & \text{если } \|\hat{p} - p_0\| > \varepsilon, \\ 0, & \text{если } \|\hat{p} - p_0\| \leq \varepsilon. \end{cases}$$

Отметим, что задача нахождения точки $p^* \in S_\varepsilon$, в которой ЛФП $l(X; p)$ достигает максимума, оказывается нетривиальной оптимизационной задачей, решение которой получено в [4]. В условиях Леммы ближайшей к точке $\hat{p} \in \mathcal{P}$ точкой $p^{**} \in \mathcal{P}_0^\varepsilon$ является точка гиперболы $S_\varepsilon : p^{**} = p_0 + (\varepsilon / \|\hat{p} - p_0\|) \times (\hat{p} - p_0)$, причем справедливо асимптотическое разложение: $p_k^* = p_k^{**} + O(\varepsilon^2)$, $k=0, 1, \dots, K-1$, и более точное разложение $p_k^* = p_k^{**} + g(\hat{p}_k, \|\hat{p} - p_0\|) \varepsilon^2 + O(\varepsilon^3)$, где для $g(\cdot)$ имеется явное выражение.

Обозначим: $c = \varepsilon^2 + 1/K$, $p_+ = \max_k \hat{p}_k$, $\Phi^{-1}(\cdot)$ – квантильная функция для $\mathcal{N}_1(0, 1)$, $H(p)$ – энтропия Шеннона.

Теорема. Если $p_+ > c$, $1 + K\varepsilon(e_K^+ - \varepsilon) > 0$, то при $T \rightarrow +\infty$, асимптотический размер теста

$$d = d(X) = \begin{cases} 0, & \text{sign}(\|\hat{p} - p_0\| - \varepsilon) \times (-2\tilde{\lambda}(\hat{p})/K)^{1/2} \leq \gamma_n, \\ 1, & \text{в противном случае,} \end{cases}$$

$$\tilde{\lambda}(p) = H(p) - \ln K + \varepsilon K \|p - p_0\|,$$

$$\gamma_n = \frac{\Phi^{-1}(1 - \alpha_0)}{\sqrt{n}} \sqrt{1/K + \varepsilon(e_K^+ - \varepsilon)}, \quad e_K^+ = \frac{1 - 2/K}{\sqrt{1 - 1/K}} \leq 1.$$

Получены асимптотические оценки мощности для этого теста, и доказана его состоятельность. Выполнены компьютерные эксперименты, иллюстрирующие работоспособность построенного теста [4].

Список литературы

- [1] Харин Ю. С. и др. Криптология. Минск: БГУ, 2013. 512 с.
- [2] Zubkov A. M., Serov A. A. Testing the NIST Statistical Test Suite on artificial pseudorandom sequences // Математические вопросы криптографии. 2019. 10(2):89–96.
- [3] Decision to Revise NIST SP 800-22 Rev. 1a [Electronic resource].
- [4] Харин Ю. С., Зубков А. М. О статистической проверке сложных гипотез об s -мерном равномерном распределении вероятностей двоичных последовательностей // Дискр. матем. 2024. 36(1):116–135.

Переходные явления в ветвящихся процессах в случайной среде: докритический случай

Виктор В. Харламов

Московский государственный университет им. М. В. Ломоносова,
Москва, Россия

E-mail: vi.v.kharlamov@gmail.com

Рассмотрим два семейства производящих функций

$$\{f_y, y \in Y\}, \quad \{f_{y,i,n}, y \in Y, 0 \leq i < n\},$$

где (Y, \mathcal{G}) – измеримое пространство. Последовательность независимых и одинаково распределённых случайных элементов $\Xi = \{\xi_i, i \in \mathbb{N}\}$ со значениями в (Y, \mathcal{G}) будем называть *случайной средой*.

1. Положим $F_{k-1} := f_{\xi_k}$ при каждом $k \in \mathbb{N}$.
2. Разыграем независимые случайные величины (с. в.) $Y_{i,k}$ с п. ф. F_{k-1} , $i, k \in \mathbb{N}$.
3. Положим $Z_0 = 1$, $Z_k := \sum_{i=1}^{Z_{k-1}} Y_{i,k}$ при $k \in \mathbb{N}$.

Последовательность $\mathcal{Z} = \{Z_k, k \geq 0\}$ будем называть *ветвящимся процессом в случайной среде* (ВПСС).

Положим

$$X_i := \log F'_{i-1}(1), \quad S_0 := 0, \quad S_k := X_1 + \dots + X_k, \quad i, k \in \mathbb{N}.$$

Последовательность $\{S_k, k \geq 0\}$ будем называть *сопровождающим случайным блужданием* (ССБ) для ВПСС \mathcal{Z} .

Если $EX_1 = 0$, то ВПСС \mathcal{Z} будем называть *критическим*. Козлов М. В. в работе [1] получил асимптотическое поведение $P(Z_n > 0)$ для критического ВПСС с дробно-линейной производящей функцией. Общий случай был рассмотрен в работах [2, 3].

В настоящей работе мы изучаем переходные явления для вероятности невырождения ВПСС. Для исследования переходных явлений мы будем использовать модель, введённую Харламовым В. В. в работе [4].

1. Положим $F_{k-1,n} := f_{\xi_{k,k-1,n}}$ при всех натуральных $k \leq n$.

© В. В. Харламов, 2024

2. Разыграем независимые с. в. $Y_{i,k,n}$ с п. ф. $F_{k-1,n}$ при всех натуральных i и $k \leq n$.
3. При каждом натуральном n определим набор $\{Z_{k,n}, k \leq n\}$. Положим $Z_{0,n} = 1$, $Z_{k,n} := \sum_{i=1}^{Z_{k-1,n}} Y_{i,k,n}$ при $k \in \mathbb{N}$.

Набор с. в. $\{Z_{k,n}, 0 \leq k \leq n, Z_{0,n} = 1\}$ назовём *возмущённым ветвящимся процессом в случайной среде* Ξ (ВВПСС).

Введём обозначения

$$a_{i,n} := \log F'_{i-1,n}(1) - \log F'_{i-1}(1), \quad b_{k,n} := \sum_{i=1}^k a_{i,n}, \quad i, k \leq n.$$

В работе [4] Харламов В. В. получил условия, при которых асимптотики вероятностей невырождения ВПСС и ВВПСС эквивалентны. Главным из этих условий являлось ограничение на степень отклонения ССБ: $b_{k,n} = o(\sqrt{k})$. В этой работе мы рассматриваем случай перехода из критической в докритическую область: $b_{k,n}$ отрицательно и достигает уровня $-\sqrt{n}$ при величинах k порядка n .

Предположение 1. *Найдётся функция $g: [0, 1] \rightarrow [0, \infty)$, удовлетворяющая следующим условиям.*

1. *При всех $k \leq n$ почти наверное $b_{k,n} = -g(k/n)\sqrt{n}$.*
2. *$g(x)$, $x \in [0, 1]$, является непрерывной функцией.*
3. *$g(x) > 0$ при $x > 0$ и $g(0) = 0$.*
4. *$g(x) = o(\sqrt{x})$ при $x \rightarrow 0$.*

Основной результат этой работы состоит в следующем утверждении.

Теорема. *Пусть $EX_1 = 0$ и $EX_1^{2+\delta} < \infty$ при некотором $\delta > 0$. При выполнении предположения 1 и некоторых технических предположениях справедлива эквивалентность*

$$\mathbb{P}(Z_{n,n} > 0) \sim \gamma \mathbb{P}(Z_n > 0), \quad n \rightarrow \infty,$$

где

$$\gamma = \mathbb{P}(W_t^+ \geq g(t), t \in [0, 1]).$$

Список литературы

- [1] *Козлов М. В.* Об асимптотике вероятности невырождения критических ветвящихся процессов в случайной среде // Теория вероятностей и ее применения. 1976. 21(4):813–825.
- [2] *Geiger J., Kersting G.* The survival probability of a critical branching process in a random environment // Theory of Probability & Its Applications. 2001. 45(3):517–525.
- [3] *Afanasyev V., Geiger J., Kersting G., Vatutin V.* Criticality for branching processes in random environment // Ann. Probab. 2005. 33(1):645–673.
- [4] *Харламов В. В.* Асимптотика вероятности невырождения почти критических ветвящихся процессов в случайной среде // Математический сборник. 2024. 215(1):131–152.

О числе деревьев заданного объема леса Гальтона – Ватсона с ограниченным числом вершин

Елена В. Хворостянская

Институт прикладных математических исследований КарНЦ РАН,
Петрозаводск, Россия

E-mail: cher@krc.karelia.ru

Пусть G_N – однородный ветвящийся процесс Гальтона – Ватсона с N начальными частицами, занумерованными числами $1, \dots, N$. Распределение числа прямых потомков каждой частицы процесса G_N задается равенствами

$$p_k = \mathbf{P} \{ \xi = k \} = \frac{1}{(k+1)^\tau} - \frac{1}{(k+2)^\tau}, \quad k = 0, 1, 2, \dots \quad (1)$$

Обозначим через τ_0 значение параметра τ , для которого $\mathbf{E}\xi = \zeta(\tau, 2) = 1$, где $\zeta(s, v) = \sum_{k=0}^{\infty} (k+v)^{-s}$ – обобщенная дзета-функция, $\tau_0 \approx 1,7$. Будем считать, что значения параметра τ распределения (1) находятся в интервале $\tau_0 \leq \tau < \infty$ и процесс G_N является критическим или докритическим. Нетрудно видеть, что распределение (1) не имеет конечной дисперсии, если $\tau_0 \leq \tau \leq 2$.

Рассматривается лес Гальтона – Ватсона $\mathcal{F}'_{N,n}$, состоящий из N деревьев и содержащий не более n вершин. Получены предельные распределения случайной величины $\mu_r(\mathcal{F}')$, равной числу деревьев объема r леса $\mathcal{F}'_{N,n}$, при различном характере стремления N, n к бесконечности, при этом значения r могут быть фиксированными или стремящимися к бесконечности. Кроме того, рассматриваются как фиксированные значения параметра τ , так и изменяющиеся вместе N, n : $\tau = \tau(N, n)$.

Для доказательства предельных теорем используется аналог обобщенной схемы размещения не более чем n частиц по N ячейкам [1]. Обобщенная схема размещения n частиц по N ячейкам была введена и подробно изучена В. Ф. Колчиным (см., например, [2]).

Финансовое обеспечение исследований осуществлялось из федерального бюджета на выполнение государственного задания Карельского научного центра РАН.

Список литературы

- [1] *Чупрунов А. Н., Фазекаш И.* Аналог обобщенной схемы размещения. Предельные теоремы для числа ячеек заданного объема // *Дискретная математика*. 2012. 24(1):140–158.
- [2] *Колчин В. Ф.* Случайные отображения. М.: Наука, 1984. 208 с.

Большие уклонения для ветвящихся процессов в случайной среде с замораживанием

Мария А. Ходякова

Московский государственный университет им. М. В. Ломоносова,
Москва, Россия

E-mail: khodyakova.mari@mail.ru

Пусть $\eta = (\eta_1, \eta_2, \dots)$ – последовательность независимых одинаково распределённых случайных величин, $\{f_y(s), y \in \mathbb{R}\}$ – семейство производящих функций распределений. Определим ветвящийся процесс в случайной среде (ВПСС) $\mathcal{Z} = (Z_n, n \geq 0)$. При фиксации среды η предположим, что $(X_{i,j}, i, j > 0)$ – неотрицательные целочисленные независимые случайные величины, причем при каждом i случайные величины $X_{i,j}, j > 0$, одинаково распределены в соответствии с производящей функцией f_{η_i} . Тогда ветвящийся процесс в случайной среде есть случайная последовательность

$$Z_0 = 1, \quad Z_n = \begin{cases} X_{n,1} + \dots + X_{n,Z_{n-1}}, & Z_n > 0, \\ 0, & Z_n = 0. \end{cases} \quad (1)$$

Положим $\xi_i = \ln f'_{\eta_i}(1)$, тогда сопровождающее случайное блуждание для ВПСС \mathcal{Z} определяется как

$$S_0 = 0, \quad S_n = \sum_{i=1}^n \xi_i, \quad n > 0.$$

В работе [1] А. В. Шкляевым получена точная асимптотика вероятностей для $\ln Z_n$ в интегро-локальном и интегральном виде, а также показана связь с распределением сопровождающего блуждания S_n .

Рассмотрим ветвящийся процесс в случайной среде с замораживанием (ВПССЗ), предложенный В. А. Вагутиным и введенный И. Д. Коршуновым в работе [2]. ВПССЗ отличается от обычного ВПСС тем, что каждое значение среды определяется на несколько поколений. Более конкретно, пусть задана некоторая детерминированная целочисленная положительная последовательность $\{\tau_k, k \in \mathbb{N}\}$, где τ_k будем называть длительностью

k -го замораживания. Рассмотрим при каждом фиксированном n величину $k(n)$, определяющуюся соотношением

$$k(n) = \max\{j : \tau_1 + \dots + \tau_{j-1} < n\}.$$

При фиксированной среде $\boldsymbol{\eta}$ при каждом i рассмотрим неотрицательные целочисленные независимые случайные величины $X_{i,j}$, $j > 0$, одинаково распределенные в соответствии с производящей функцией $f_{\eta_{k(n)}}$. Тогда ветвящийся процесс в случайной среде с замораживанием есть случайная последовательность $\mathcal{Z} = (Z_n, n \geq 0)$, удовлетворяющая формуле (1). Сопровождающее случайное блуждание для ВПССЗ естественно определять формулой

$$S_0 = 0, \quad S_n = \sum_{i=1}^{k(n)} \tau_i \xi_i,$$

где $\xi_i = \ln f'_{\eta_i}(1)$, $\tau_{k(n)} := n - \tau_1 - \dots - \tau_{k(n)-1}$.

Применим подход, описанный Шкляевым в работе [3], для исследования больших уклонений для ВПССЗ. Пусть (A_n, B_n) – случайные векторы, Y_0 – некоторая случайная величина, тогда случайной рекуррентной последовательностью называется последовательность $\{Y_n, n \geq 0\}$, заданная уравнением

$$Y_n = A_n Y_{n-1} + B_n, \quad n > 0.$$

Пусть $\{\zeta_i, i \in \mathbb{N}\}$ – независимые одинаково распределённые нерешётчатые случайные величины, $\{\tau_i, i \in \mathbb{N}\}$ – положительная целочисленная равномерно ограниченная детерминированная последовательность. Положим $A_n = e^{\zeta_{k(n)}}$, что соответствует замораживанию последовательности $\{A_n, n \geq 1\}$ на промежутки длительностей τ_k , $k \geq 1$. Мы рассматриваем такие векторы (A_n, B_n) , что $(A_{\tau_1+\dots+\tau_j+1}, A_{\tau_1+\dots+\tau_j+2}, \dots)$ не зависит от $(A_i, B_i, i \leq \tau_1 + \dots + \tau_j)$ и B_i удовлетворяют некоторым моментным ограничениям. Заметим, что B_i могут быть зависимыми между собой.

Рассмотрим преобразование Крамера для ζ_i с параметром h , где $h \in [0, h^+]$, $h^+ > 0$:

$$R(h) = \mathbf{E}e^{h\zeta_1} < \infty, \quad m(h) = (\ln R(h))'.$$

Верна следующая теорема.

Теорема. *При некоторых моментных условиях на $\{B_i\}_{i=1}^\infty$ и $\{\zeta_i\}_{i=1}^\infty$ при $n \rightarrow +\infty$ для всех положительных последовательностей Δ_n , стремящихся к нулю достаточно медленно, верно соотношение*

$$\begin{aligned} \mathbf{P}(\ln Y_n \in [x, x + \Delta_n)) &= (1 + o(1))c_n \mathbf{P}(S_n \in [x, x + \Delta_n)) = \\ &= \frac{(1 + o(1))c_n \Delta_n}{\sqrt{2\pi}\sigma_n(h_n)} \exp\left(-n\Lambda_n\left(\frac{x}{n}\right)\right), \end{aligned}$$

где для $h_n = h_n(x/n)$, $\Lambda_n(x/n)$, $\sigma_n(h_n)$ и $c_n = c_n(x/n)$ получены явные представления. При этом $o(1)$ равномерно мало по таким x , что $x/n \in [\theta_1, \theta_2] \subset (m^*, m^+]$ для некоторых m^* и m^+ .

Из полученного интегро-локального результата прямо вытекает аналогичное интегральное представление для вероятностей $\mathbf{P}(\ln Y_n \geq x)$. Из данного более общего факта нами получена теорема об асимптотике вероятностей больших уклонений для ВПССЗ.

Список литературы

- [1] Шкляев А. В. Большие уклонения ветвящегося процесса в случайной среде. II. // Дискретная математика. 2020. 32(1):135–156.
- [2] Коршунов И. Д. Ветвящиеся процессы в случайной среде с замораживаниями // Дискретная математика. 2023. 35(3):20–36.
- [3] Шкляев А. В. Большие уклонения ветвящегося процесса в случайной среде. I. // Дискретная математика. 2019. 31(4):102–115.

Предельные распределения числа вершин заданной степени условного конфигурационного графа

Ирина А. Чеплюкова

Институт прикладных математических исследований КарНЦ РАН,
Петрозаводск, Россия

E-mail: chia@krc.karelia.ru

В последние годы изучение случайных графов, предназначенных для моделирования различных сложных сетей коммуникаций, становится все более актуальным направлением для специалистов по вероятностным методам дискретной математике (см, например, [1, 2]). Одним из наиболее известных видов таких моделей является конфигурационный граф.

Рассмотрим конфигурационный граф, содержащий N вершин. Степени вершин этого графа являются независимыми одинаково распределенными случайными величинами, распределение которых удовлетворяет следующему условию: при $k \rightarrow \infty$

$$\mathbf{P}\{\eta = k\} \sim \frac{d}{k^g \ln^h k},$$

$d > 0$, $h \geq 0$, $2 < g < 3$, где случайная величина η равна степени любой вершины графа. Исследуются распределения числа вершин заданной степени в таком графе при условии, что сумма степеней всех вершин равна n . Получены предельные теоремы для данной характеристики при $N, n \rightarrow \infty$ так, что $n, N \rightarrow \infty$, $n/N^{(3g-4)/(2g-2)} \rightarrow \infty$.

Через μ_r обозначим случайную величину, равную числу вершин степени r в рассматриваемом графе. Справедливы следующие утверждения.

Теорема 1. Пусть $n, N \rightarrow \infty$, $n/N^{(3g-4)/(2g-2)} \rightarrow \infty$, $N/(r^g \ln^h r) \rightarrow \infty$. Тогда равномерно относительно целых неотрицательных k таких, что $u_r = \frac{k - Np_r}{\sqrt{Np_r(1-p_r)}}$ лежит в любом фиксированном конечном интервале

$$\mathbf{P}\{\mu_r = k\} = \frac{1 + o(1)}{\sqrt{2\pi Np_r(1-p_r)}} e^{-u_r^2/2}.$$

Теорема 2. Пусть $n, N, r \rightarrow \infty$, $n/N^{(3g-4)/(2g-2)} \rightarrow \infty$. Тогда равномерно относительно целых неотрицательных k таких, что $(k - Nr)/\sqrt{Nr}$ лежит в любом фиксированном конечном интервале

$$\mathbf{P} \{ \mu_r = k \} = \frac{(Nr)^k}{k!} e^{-Nr} (1 + o(1)).$$

В основе доказательства этих результатов лежит обобщенная схема размещения частиц по ячейкам, введенная и изученная В. Ф. Колчиным (см., например, [3]).

Список литературы

- [1] *Hofstad R.* Random graphs and complex networks. Cambridge Univ. Press. Cambridge, 2017. 321 p.
- [2] *Durrett R.* Random graph dynamics. Cambridge Univ. Press. Cambridge, 2007. 212 p.
- [3] *Колчин В. Ф.* Случайные графы. М.: Физматлит, 2000. 256 с.

Об асимптотике вероятности связности графов Эрдёша-Реньи и неоднородных случайных блужданиях

Борис Б. Чиняев

Московский государственный университет им. М. В. Ломоносова,
Москва, Россия

E-mail: bchinyayev.msu@gmail.com

В докладе мы исследуем случайные графы в модели Эрдёша – Реньи. В данной модели рассматривается граф $G(n, p)$ с множеством вершин $V = V_n = \{1, \dots, n\}$ и матрицей смежности C с элементами $c_{i,j}$, $i < j$, полученными реализацией набора независимых одинаково распределенных величин с распределением Бернулли с параметром $p = p(n)$.

Нас будет интересовать асимптотика вероятности $P_n(p)$ связности рассматриваемого графа при $n \rightarrow \infty$, $p(n) \rightarrow 0$. Известно множество результатов [1, 2] для данной задачи:

1. Пусть $p(n) = (\ln n + \alpha + o(1))/n$, $\alpha > 0$.

$$P_n(p) = e^{-e^{-\alpha}}(1 + o(1)), \quad n \rightarrow \infty.$$

2. Пусть $p(n) = c/n$, $c > 0$. Тогда

$$P_n(p) = \left(1 - \frac{c}{e^c - 1}\right) (1 - (1 - c/n)^n)^n (1 + o(1)), \quad n \rightarrow \infty.$$

3. Пусть $p(n) = o(1/n^2)$, $n \rightarrow \infty$. Тогда

$$P_n(p) = n^{n-2} p^{n-1} (1 + o(1)), \quad n \rightarrow \infty.$$

Методы, использованные в указанных выше работах базируются на комбинаторных оценках. В докладе будет предложен новый подход к исследованию вероятности связности случайного графа Эрдёша – Реньи, основанный на исследовании неоднородных случайных блужданий.

Удастся показать, что вероятность связности сводится к вероятности того, что некоторый мост, построенный по неоднородному случайному блужданию, является экскурсией.

Теорема 1 (Вероятность связности). Пусть $G(n, p)$ – граф в модели Эрдёша – Реньи. Тогда вероятность, что граф связан

$$P_n(p) = (1 - (1 - p)^{n-1})^{n-1} \mathbf{P}(S_k \geq 0, 0 < k < n - 1 | S_{n-1} = 0),$$

$S_k = \sum_{i=1}^k X_i$, при этом X_i – независимые величины и $X_{i+1} \sim \text{Poiss}(\lambda_i)$, где

$$\begin{aligned} \lambda_i &= q(1 - p)^{(i-1)}, \\ q &= (n - 1)p / (1 - (1 - p)^{n-1}). \end{aligned}$$

Стоит отметить, что предложенный метод единообразно сводит задачу связности к указанной задаче из теории неоднородных случайных блужданий вне зависимости от зависимости $p(n)$ от n . Помимо доказательства приведенных выше результатов, мы дополнительно получим следующую теорему.

Теорема 2. Пусть $G(n, p)$ – граф в модели Эрдёша – Реньи с вероятностью ребра $p = \lambda_n / (n - 1)$. Пусть $\lambda_n = o(1)$, причем $\lambda_n n^{1/2} / \ln n \rightarrow +\infty$, $n \rightarrow \infty$. Тогда

$$\mathbf{P}(C_n) \sim \frac{e}{2} \lambda_n^2 \left(1 - \left(1 - \frac{\lambda_n}{n-1} \right)^{n-1} \right)^{n-1}, \quad n \rightarrow \infty.$$

Список литературы

- [1] Erdős P., Rényi A. On random graphs I. // Publ. Math. Debrecen. 1959. 6:290–297.
- [2] Степанов В. Е. О вероятности связности случайного графа $\mathcal{G}_m(t)$ // Теория вероятн. и ее примен. 1970. 15(1):56–68.
- [3] Райгородский А. М. Модели случайных графов и их применения // Труды Московского физико-технического института 2. 2010. 4:130–140.

Дробные раскраски случайных гиперграфов

Дмитрий А. Шабанов

Московский физико-технический институт, Долгопрудный, Россия

E-mail: shabanov.da@mipt.ru

В докладе рассматривается проблема поиска пороговых вероятностей для свойств раскрасок случайных гиперграфов в биномиальной модели $H(n, k, p)$. Напомним, что если $n > k \geq 2$ – натуральные числа и $p \in (0, 1)$, то модель случайного k -однородного гиперграфа $H(n, k, p)$ формируется как схема Бернулли на k -подмножествах некоторого n -элементного множества: каждое включается в $H(n, k, p)$ в качестве ребра независимо с вероятностью p . Случайный гиперграф $H(n, k, p)$ является естественным обобщением классической модели $G(n, p)$ – модели Эрдеша–Реньи случайного графа.

Мы изучаем так называемые дробные раскраски случайных гиперграфов. Пусть $H = (V, E)$ – гиперграф и заданы натуральные числа $a > b \geq 1$. Обозначим $[a] = \{1, \dots, a\}$. Отображение $f : V \rightarrow \binom{[a]}{b}$ называется *дробной $(a : b)$ -раскраской*, т. е. мы каждой вершине сопоставляем набор из b цветов из заданного общества множества a цветов. Дробная $(a : b)$ -раскраска f называется *правильной*, если в ней каждое ребро не имеет общего цвета, сопоставленного всем вершинам. Формально, для любого $A \in E$ должно быть выполнено

$$\bigcap_{v \in A} f(v) = \emptyset.$$

Отметим, что при $b = 1$ мы получаем классическое определение правильной раскраски в a цветов, когда все ребра не должны быть одноцветными. Если гиперграф допускает правильную дробную $(a : b)$ -раскраску, то он называется $(a : b)$ -раскрашиваемым.

В модели $H(n, k, p)$ при фиксированных $k \geq 3$, $a > b \geq 1$, свойство $(a : b)$ -раскрашиваемости имеет так называемую *точную пороговую вероятность*, т. е. существует такая функция $\hat{p}_{a,b} = \hat{p}_{a,b}(n)$, что для любого $\varepsilon > 0$ выполнено

$$\lim_{n \rightarrow \infty} \mathbb{P}(H(n, k, p) \text{ } (a : b)\text{-раскрашиваем}) = \begin{cases} 1, & \text{если } p \leq (1 - \varepsilon)\hat{p}_{a,b}, \\ 0, & \text{если } p \geq (1 + \varepsilon)\hat{p}_{a,b}. \end{cases}$$

Существование $\widehat{p}_{a,b}$ вытекает из общих результатов [1] о точных пороговых вероятностях в случайных гиперграфах.

Наиболее точные оценки $\widehat{p}_{a,b}$ в настоящее время известны для случая $a = 2$ и $b = 1$, т. е. для свойства 2-раскрашиваемости. Для удобства обозначим $p = cn/\binom{n}{k}$. В работе [2] было доказано, что существует такая функция $\varepsilon(k) = 2^{-k(1+o_k(1))}$, что

$$\lim_{n \rightarrow \infty} \mathbb{P}(H(n, k, p) \text{ 2-раскрашиваем}) \rightarrow \begin{cases} 1, & c < 2^{k-1} \ln 2 - \frac{\ln 2}{2} - \frac{1}{4} - \varepsilon(k), \\ 0, & c > 2^{k-1} \ln 2 - \frac{\ln 2}{2} - \frac{1}{4} + \varepsilon(k). \end{cases}$$

Целью работы было изучение пороговой вероятности для свойства дробной (4:2)-раскрашиваемости. Основным результатом, который удалось получить, сформулирован в следующей теореме.

Теорема. Пусть $p = cn/\binom{n}{k}$. Существуют такие экспоненциально быстро стремящиеся к нулю положительные функции $g_1(k)$ и $g_2(k)$, что для всех достаточно больших k выполнено

$$\mathbb{P}(H(n, k, p) \text{ (4:2)-раскрашиваем}) \rightarrow \begin{cases} 1, & c < 2^{k-2} \ln 6 - \frac{\ln 6}{2} - g_1(k), \\ 0, & c > 2^{k-2} \ln 6 - \frac{\ln 6}{2} + g_2(k). \end{cases}$$

Доказательство опирается на метод второго момента, а также на решение нескольких оптимизационных задач на множестве матриц с неотрицательными элементами.

Доклад основан на совместной работе с П. А. Захаровым.

Список литературы

- [1] *Hatami H., Molloy M.* Sharp thresholds for constraint satisfaction problems and homomorphisms // *Random Structures Algorithms*. 2000. 33(3):310–332. doi: 10.1002/rsa.20225
- [2] *Coja-Oghlan A., Panagiotou K.* Catching the k-NAESAT Threshold // *Proc. 44th STOC*. 2012. P. 899–908.

Асимптотика вероятности невырождения критической рекуррентной последовательности

Александр В. Шкляев, П. И. Сенько

Московский государственный университет им. М. В. Ломоносова,
Москва, Россия

E-mail: alexander.shklyaev@math.msu.ru

Будем называть случайной рекуррентной последовательностью последовательность $\{Y_n\}$, представленную в виде

$$Y_{n+1} = A_n Y_n + B_n, \quad n \geq 0, \quad (1)$$

где (A_i, B_i) – некоторые величины. В случае независимых одинаково распределенных (н. о. р.) (A_i, B_i) данная последовательность хорошо известна и исследована рядом авторов. Одной из наиболее известной здесь является работа [1]. Однако, для многих целей рассмотрение н. о. р. A_i , но зависимых и разнораспределенных B_i сохраняет содержательность такого рода представления. Впервые такая модель вводилась в работе [2] в целях применения к теории больших уклонений ветвящихся процессов в случайной среде и ветвящимся процессам в случайной среде с иммиграцией (см. [3]), однако, те же результаты удалось применить к двуполюмым ветвящимся процессам в случайной среде (см. [4]) и другим моделям.

Ключевая идея использования такого рода представления в том, что последовательность $\{\ln Y_n\}$ оказывается близкой к случайному блужданию $\{S_n\}$ (более конкретно, см. [5]) с шагами $\xi_i := \ln A_{i-1}$, которое мы назовем сопровождающим к последовательности $\{Y_n\}$, если рассматриваются траектории блуждания с “большими” значениями S_n . При этом на B_i достаточно наложить достаточно слабые моментные условия. Соответственно, можно применять это представление для анализа надкритических последовательностей ($\mathbf{E}\xi > 0$) в зоне нормальных, умеренных и больших уклонений или критических и докритических последовательностей ($\mathbf{E}\xi = 0$ и $\mathbf{E}\xi < 0$ соответственно) в зоне больших уклонений, как это проделано в работах [2–5]. В настоящей работе впервые рассматривается тот же подход к критической последовательности с $\mathbf{E}\xi = 0$ в случае нормальных уклонений.

Более конкретно, предположим, что последовательность случайных величин Y_n с целыми неотрицательными значениями образует марковскую последовательность в случайной среде, а именно существует такая последовательность н.о.р. величин $\boldsymbol{\eta} = (\eta_0, \eta_1, \dots)$, при фиксации которой Y_{n+1} образует неоднородную марковскую цепь. Предположим, что справедливо представление (1), причем A_i измерима относительно $\sigma(\eta_i)$, $i \geq 0$. Также предположим, что вектор $(B_0, \dots, B_{i-1}, \eta_0, \dots, \eta_{i-1})$ не зависит от (η_i, \dots) при всех i . Будем считать, что цепь обладает поглощающим состоянием 0. Исследуется асимптотическое поведение вероятности $\mathbf{P}(Y_n > 0)$ при $n \rightarrow \infty$.

При дополнительных условиях для критической последовательности удается получить асимптотическое соотношение

$$\mathbf{P}(Y_n > 0) \sim \frac{c}{\sqrt{n}}, \quad n \rightarrow \infty,$$

где c – некоторая константа. При этом в частном случае ветвящегося процесса в случайной среде накладываемые условия близки к условиям, накладываемым в работе [6] в случае сопровождающих блужданий с конечной дисперсией. При близких условиях отсюда также удается получить асимптотику вероятности невырождения двуполого ветвящегося процесса в случайной среде.

В докладе будет сформулирована общая теорема и ее приложения к конкретным моделям, будут изложены идеи доказательства и их сравнения с ветвящимися процессами в случайной среде.

Все указанные результаты получены совместно со студентом 5 курса механико-математического факультета МГУ имени М. В. Ломоносова П. И. Сенько.

Список литературы

- [1] *Kesten H.* Random difference equations and renewal theory for products of random matrices // Acta Math. 1973. 131:207–248. doi: 10.1007/BF02392040
- [2] *Шкляев А. В.* Большие отклонения ветвящегося процесса в случайной среде. I // Дискрет. матем. 2019. 31(4):102–115. doi: 10.1515/dma-2021-0025
- [3] *Шкляев А. В.* Большие отклонения ветвящегося процесса в случайной среде. II // Дискрет. матем. 2020. 32(1):135–156. doi: 10.1515/dma-2021-0039

- [4] *Шкляев А. В.* Большие уклонения ветвящегося процесса с частицами двух полов в случайной среде // Дискрет. матем. 2023. 35(3):125–142. doi: 10.1515/dma-2021-0025
- [5] *Шкляев А. В.* Условная функциональная предельная теорема для случайной рекуррентной последовательности при условии совершения ею большого уклонения // Теория вероятн. и ее примен. 2024. 69(1):125–147.
- [6] *Afanasyev V. I., Geiger J., Kersting G., Vatutin V. A.* Criticality for branching processes in random environment // The Annals of Probability. 2005. 33(2):645–673. doi: 10.1214/009117904000000928

Асимптотика числа A -отображений с остаточным членом

Арсен Л. Якимив

Математический институт им. В. А. Стеклова, Москва, Россия

E-mail: arsen@mi.ras.ru

Памяти В. Ф. Колчина посвящается

Доклад состоит из двух частей. В первой части приводится краткий обзор результатов В. Ф. Колчина в области случайных отображений, которые являются наиболее близкими к работам автора. Эти результаты в основном представлены в широко известных монографиях Валентина Фёдоровича [1, 2]. Во второй части приведём следующий конкретный результат в указанной области, недавно полученный автором.

Пусть \mathfrak{S}_n – полугруппа отображений множества X из n элементов в себя и $V_n(A)$ – совокупность отображений из \mathfrak{S}_n , размеры контуров которых принадлежат множеству A . Такие объекты были рассмотрены В. Н. Сачковым в 1972 году [3] и названы A -отображениями. В предыдущей работе автора [4] была найдена асимптотика для $|V_n(A)|$ в предположении, что у множества A существует положительная асимптотическая плотность ϱ во множестве натуральных чисел, а именно:

$$\lim_{n \rightarrow \infty} \frac{|k : k \in A, k \leq n|}{n} = \varrho > 0. \quad (1)$$

В настоящем докладе для некоторого более узкого класса множеств A дано уточнение полученного в [4] утверждения со степенным понижением остатка. Положим при $k \in N$

$$p(k) = [s^k] \exp \left(\sum_{m \in A} \frac{s^m}{m} \right), \quad B(k) = \sum_{l=1}^k lp(l).$$

Здесь и далее $[s^k]g(s)$ есть коэффициент при s^k аналитической в окрестности нуля функции $g(s)$. Мы будем предполагать, что при $k \rightarrow \infty$

$$B(k) = ck^\alpha(1 + O(k^{-\beta})) \quad (2)$$

для некоторых положительных постоянных α, β и c (поскольку это не улучшит наших оценок, считаем далее, что $\beta < 1$). Как несложно проверить, из (1) следует, что постоянная α в (2) равна $\varrho + 1$ [4, соотношение (20)].

Через $\lambda_n(A)$ обозначим число циклических элементов случайного отображения $\sigma_n(A)$, имеющего равномерное распределение на множестве $V_n(A)$. Положим

$$I_\varrho = \int_0^\infty x^\varrho \exp\left(-\frac{x^2}{2}\right) dx.$$

Теорема. Пусть выполнены соотношения (1) и (2). Тогда

$$|V_n(A)| = c(1 + \varrho)n^{n-(1-\varrho)/2}(I_\varrho + O(n^{-\beta/2})),$$

при $n \rightarrow \infty$. Кроме этого, для произвольного фиксированного $z > 0$

$$\mathbb{P}\left\{\frac{\lambda_n}{\sqrt{n}} \leq z\right\} = \frac{1}{I_\varrho} \int_0^z x^\varrho \exp\left(-\frac{x^2}{2}\right) dx + O(n^{-\beta/2}).$$

Приведём некоторые примеры множеств A , для которых выполнено соотношение (2).

Будем говорить, что множество A принадлежит классу F_1 , если $A = \bigcup_{i=1}^M A_i$, где $M \in \mathbb{N}$, $A_i = \{m \in \mathbb{N} : m = a_i k + b_i, k = 0, 1, 2, \dots\}$ и целые $a_i > 1$, $1 \leq b_i \leq a_i - 1$, $(a_i, b_i) = 1$, причём прогрессии A_i и A_j при $i \neq j$ не пересекаются.

Будем говорить, что множество A принадлежит классу F_2 , если $A = \{m \in \mathbb{N} : m/k_i \notin \mathbb{N}, i = 1, \dots, s\}$ для некоторых $s \in \mathbb{N}$ и $k_1, \dots, k_s \in \mathbb{N}$ таких, что $k_i \geq 2$, $i = 1, \dots, s$ и $(k_i, k_j) = 1$ при $i \neq j$.

Пусть множество $A \in F_1 \cup F_2$. В работе А. И. Павлова [5] доказано, что в этом случае для некоторых положительных постоянных C и β (зависящих от A) при $n \rightarrow \infty$

$$p(n) = Cn^{\varrho-1} (1 + O(n^{-\beta})). \quad (3)$$

Из (3) несложно вывести, что при $A \in F_1 \cup F_2$ имеет место соотношение (2) с $c = C/(\varrho + 1)$. Следующий пример показывает, что возможны ситуации, когда выполнено предположение (2), но (3) не имеет места. Пусть $A = a\mathbb{N}$, где натуральное число $a > 1$ – фиксировано. Тогда для множества A выполнено (2). Тем не менее, в этом случае соотношение (3) не имеет места, поскольку $p(n) = 0$ при $n \in \mathbb{N} \setminus A$.

Список литературы

- [1] *Колчин В. Ф.* Случайные отображения. М.: Наука, 1984, 208 с.
- [2] *Колчин В. Ф.* Случайные графы. М.: Физматлит, 2000, 255 с.
- [3] *Сачков В. Н.* Отображения конечного множества с ограничениями на контур и высоту // Теория вероятн. и ее примен. 1972. 17(4):679–694.
- [4] *Якымив А. Л.* О числе циклических точек случайного A -отображения // Дискрет. матем. 2013. 25(3):116–127.
- [5] *Павлов А. И.* О двух классах подстановок с теоретико-числовыми ограничениями на длины циклов // Мат. заметки. 1997. 62(6):881–891.