Отзыв

официального оппонента
на диссертационную работу Олега Павловича Орлова
“Комбинаторно-вероятностные задачи для случайных выборок”,
представленную на соискание ученой степени кандидата
физико-математических наук по специальности 1.1.4 – теория
вероятностей и математическая статистика.

В работе решается три типа задач комбинаторно-вероятностного типа. Их объединяет идея описания асимптотических свойств последовательности независимых экспериментов. В первой главе диссертации изучаются межрекордные наполнения — новая характеристика в теории рекордов, вторая глава посвящена экстремальным расстояниям до ближайшего соседа, в третьей главе рассмотрена задача о числе пустых ячеек, связанная с классической схемой случайных размещений частиц по ячейкам.

Диссертация имеет теоретический характер. Ее результаты представляют значительный интерес для математической статистики и ее приложений. В частности: математическая теория рекордов используется для моделирования в спорте, страховании, гидрологии; расстояния до ближайших соседей используются при построении статистических критериев равномерности распределения, при оценке энтропии, в алгоритмах классификации; к задаче о размещении частиц по ячейкам сводятся многие статистические критерии.

Математическая теория рекордов, по существу, берет свое начало со статьи Чендлера (1952). Она стимулировала появление огромного количества работ по данной тематике. В настоящее время тематика продолжает развиваться.

В первой главе доказан ряд новых теорем о свойствах совместных распределений межрекордных наполнений, получены явные формулы для одномерных распределений двупараметрического семейства межрекордных наполнений и описаны их асимптотические представления. Определим рекордные значения последовательности независимых одинаково распределенных положительных случайных величин X_i, $i \in \mathbb{N}$, с общей непрерывной функцией распределения $F(x)$, $x \in \mathbb{R}$. Рекордные моменты $L(n)$, $n \in \mathbb{N}$, определяются соотношениями

$$L(1) = 1; L(n) = \min\{t > L(n-1) : X_t > X_{L(n-1)}\}, n > 1,$$

а $X_{L(n)}$, $n \in \mathbb{N}$, будет последовательностью рекордов. Автор исследует межрекордные наполнения, равные количеству случайных величин в последовательности X_n с индексами между $L(i)$ и $L(i+1)$ и попавших в промежуток между рекордами $X_{L(j-1)}$ и $X_{L(j)}$, $1 \leq j \leq i$. Особенностью данных объектов является независимость характеристик относительно непрерывной функции распределения $F(x)$. При доказательстве предельных теорем (ЦПТ, УЗБЧ, закон повторного логарифма) автор существенно использует последний факт и выбирает в качестве $F(x)$ удобное для вычислений экспоненциальное распределение. Следует отметить, что задача может быть сформулирована в терминах неравновозможного распределения частиц по ячейкам с добавлением новой ячейки после появления нового рекорда.

Во второй главе диссертации изучаются предельные распределения экстремальных расстояний до ближайшего соседа. Изучение случайные величин, связанных с расстояниями до ближайших соседей, имеют богатую историю. В частности, в монографии Barbour, Holst, Janson *Poisson approximation* (1992) рассмотрен случай многомерного компактного однородного риманова многообразия. В диссертации доказана предельная теорема для максимального расстояния до ближайшего соседа в произвольном метрическом пространстве с определенной на нем вероятностной мерой. Главное отличие от имевшихся ранее работ заключается в том, что метрическое пространство может быть дискретным (например, двоичным кубом). Кроме этого, для доказательства предельных теорем для распределения минимального расстояния между точками выборки и распределении максимального расстояния до ближайшего соседа используется метод Чена-Стейна. Довольно интересен результат теоремы 16, где доказано, что расстояние Хемминга для максимально отстоящих друг от друга ближайших соседей для независимой равновероятной выборки из $\{0,1\}^T$ при выполнении ряда условий может принимать асимптотически ровно два значения с найденными в явном виде вероятностями.

В третьей главе рассматривается равновозможное и независимое друг от друга размещения n частиц по N ячейкам. Для среднего числа пустых ячеек $\mu_0 = \mu_0(n, N)$ верна формула $\mathbb{E}\mu_0 = N(1 - N^{-1})^n$. При $n, N \to \infty$ выделяют пять областей, связывающих n и N:

- *левая область:* $n^2(2N)^{-1} \to \lambda < \infty$, где λ – константа,
- *левая промежуточная область:* $nN^{-1} \to 0$ и $n^2(2N)^{-1} \to \infty$,
- *центральная область:* $c_1 < nN^{-1} \leq c_2 < \infty$, где c_1, c_2 — константы,
- *правая промежуточная область:* $nN^{-1} \to \infty$ и $\mathbb{E}\mu_0 \to \infty$,
- *правая область:* $nN^{-1} \to \infty$ и $\mathbb{E}\mu_0 \to \lambda$, где λ – константа.

В центральной, левой промежуточной и правой промежуточной областях число пустых ячеек асимптотически нормально. В правой области распределение числа пустых ячеек слабо сходится к распределению Пуассона с параметром λ. В левой области распределение случайной величины $\mu_0 + n - N$ слабо сходится к распределению Пуассона с параметром λ. Эти результаты дают полное описание всех невырожденных предельных распределений для μ_0 и собраны в монографии Колчина, Севастьянова, Чистякова "Случайные размещения".

При переходе от интегральных предельных теорем для μ_0 к локальным остается много открытых проблем. Автор устраняет ряд этих пробелов. Им описана асимптотика $\mathbb{P}(\mu_0 = k)$ при $k \geq 0$ и $n, N \to \infty$, причем $1 < c_1 < n(N-k)^{-1} \leq c_2 < \infty$, где c_1, c_2 — константы. Этот результат существенно обобщает имевшиеся ранее и применим в центральной, левой, левой промежуточной областях.

Замечаний по тексту диссертации и автореферата нет.

Автор продемонстрировал владение широким спектром теоретико-вероятностных методов в комбинаторных задачах дискретной математики. Им решен ряд задач, имеющих важное значение как для развития комбинаторно-вероятностных методов в дискретной математике, так и для приложений в математической статистике.

Все результаты, выносимые на защиту, получены лично автором, обладают внутренним единством, снабжены подробными доказательствами, доложены на ряде международных и всероссийских конференций, на научных семинарах МГУ и МИАН, опубликованы в четырех статьях в журнале "Дискретная математика", рекомендованном ВАК и индексируемым в Web of Science Core Collection. Две из них подготовлены лично автором, а две в соавторстве. Часть результатов из работ с соавторами приводится в тексте, как полученная другими авторами, что отделяет вклад соавторов. Для всех решаемых задач приводится очень подробная история их постановки, решения в процессе ослабления ограничений и усложнения моделей и суть обобщений, полученных автором. В диссертационной работе используется 74 литературных источника.

Диссертация соответствует паспорту научной специальности 1.1.4 – теория вероятностей и математическая статистика, удовлетворяет всем требованиям, предъявляемым к диссертациям на соискание ученой степени кандидата наук, в Положении о присуждении ученых степеней, а ее автор, Орлов Олег Павлович, безусловно, заслуживает присуждения ученой степени кандидата физико-математических наук.

Официальный оппонент В.А. Топчий

03.11.2022

Топчий Валентин Алексеевич
доктор физико-математических наук (специальность 01.01.05 – теория вероятностей и математическая статистика), профессор, ведущий научный сотрудник Лаборатории комбинаторных и вычислительных методов алгебры и логики ОФ ИМ СО РАН,
e-mail: topchij@gmail.com , тел. +7(903) 981-21-00.

Омский филиал Федерального государственного бюджетного учреждения науки Институт математики им. С.Л. Соболева Сибирского отделения Российской академии наук, 644043, г. Омск, ул. Певцова, 13,
E-mail: admin@ofim.oscsbras.ru , тел. +7(3812) 23-65-67.

Подпись В.А. Топчия заверяю:

Ученый секретарь ОФ ИМ СО РАН,
кандидат физ.-мат. наук С.В. Тиховская

07.11.2022