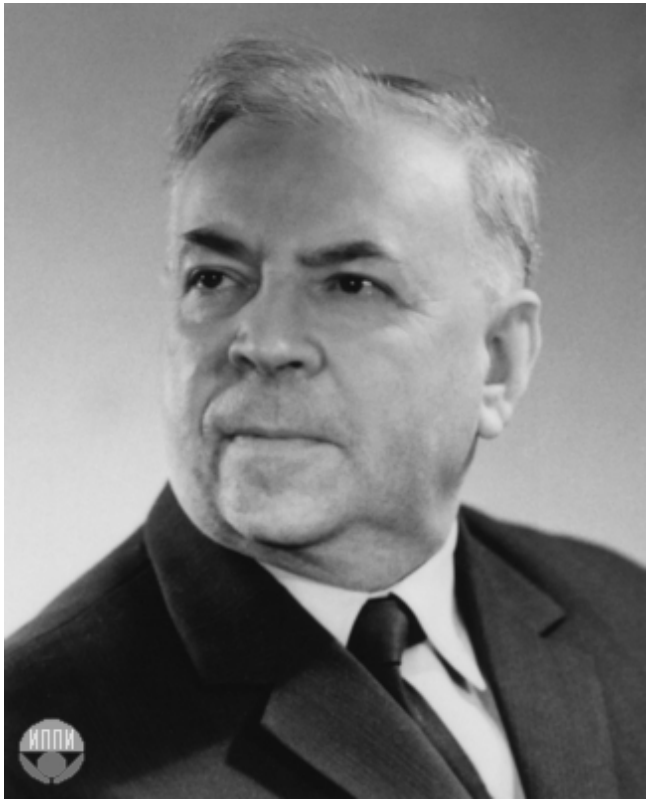
Отдел дискретной математики в Математическом институте им.В.А.Стеклова РАН

90 лет МИАН

Москва
15 мая 2024 г.

1953 г. – 1994 г. : Отдел прикладных расчетов
1994 г. по настоящее время : Отдел дискретной математики

После окончания Второй мировой войны американским криптографам в ходе операции «Венона» удалось расшифровать переписку советских разведчиков. Руководством СССР были приняты решения об укреплении криптографической службы.
К решению криптографических задач был привлечен ряд известных математиков, а в 1953 году в Математическом институте им.В.А.Стеклова АН СССР (МИАН) был образован Отдел прикладных расчетов (ОПР).



Константин Константинович Марджанишвили (1903–1981)
Сотрудник МИАН с 1934 г.
заведующий ОПР в 1953–1981 гг.
Доктор физ.-матем.наук с 1949 г.
Член-корр. АН СССР с 1964 г., с 1974 г. — академик

Владимир Константинович Захаров (1925–1988)
Сотрудник МИАН с 1956 г.
заведующий ОПР в 1981–1988 гг.
Доктор физ.-матем.наук с 1972 г.

Борис Александрович Севастьянов (1923–2013)
Сотрудник МИАН с 1948 г.
заведующий ОПР в 1988–1994 гг.
Доктор физ.-матем.наук с 1968 г.
Член-корреспондент АН СССР с 1984 г.

В Математическом институте с самого начала существовали отделы, проводившие прикладные исследования, и в нескольких случаях они становились основой новых научно-исследовательских институтов.

Например, в предвоенные и военные годы А.Н.Колмогоров и Н.В.Смирнов, С.Л.Соболев, Л.В.Келдыш и А.А.Марков разрабатывали теорию стрельбы и движения снарядов, И.М.Гельфанд и Л.А.Люстерник участвовали в составлении таблиц для определения параметров движения самолета, С.Н.Бернштейн разрабатывал метод радиопеленгации кораблей, А.О.Гельфонд участвовал в работе Главного штаба ВМФ СССР, К.К.Марджанишвили проводил расчеты прочности различных конструкций.

В 1967 г. институт был награжден орденом Ленина за большой вклад в решение задач, имеющих прикладное значение, а в 1984 году — орденом Октябрьской Революции.

В 1994 году была проведена реорганизация некоторых отделов Института, в частности, ликвидирован Отдел прикладных расчетов и создан Отдел дискретной математики, сотрудники которого должны вести исследования по общему плану Института и могут выполнять договорные работы. Заведующим ОДМ стал А.М.Зубков.

Как при основании Отдела прикладных расчетов, эта реорганизация была связана со сложным периодом в жизни нашей страны.

Из интервью В.В.Путина (2021 г.): «В России 90-х в качестве советников работали сотни иностранцев. Одним таким советником работал кадровый сотрудник ЦРУ. Наверное, он был не один.» «Американцы старались попасть на оборонные предприятия РФ, они лезли везде. Даже если не были кадровыми сотрудниками ЦРУ, наверняка отчитывались и в Госдеп, и в ЦРУ о результатах проделанной ими работы. У нас же было межправсоглашение, в рамках которого на некоторых предприятиях ядерного комплекса сидели американские специалисты. Сидели на наших предприятиях и ходили туда, как на работу. Каждый день.»

В 90-е годы было ликвидировано большое число предприятий и прикладных институтов. В легкой промышленности объем производства упал на 86%, в стройматериалах и лесобумажной — около 50%, в оборонзаказе общее падение составило около 80%.

В работе Отдела прикладных расчетов принимали участие члены-корреспонденты АН СССР А.А.Марков, А.О.Гельфонд.
С момента основания Отдела в нем работал Б.А.Севастьянов. Многие сотрудники Отдела в разные годы защищали кандидатские и докторские диссертации.

Основным направлением деятельности Отдела прикладных расчетов было решение сложных математических задач, связанных с теоретической криптографией.

Например, в 1956 г. А.А.Марков доказал, что все биективные отображения множества слов над конечным алфавитом, не увеличивающие расстояние Хемминга между словами, сводятся к перестановкам и заменам отдельных букв. Эта работа опубликована в 2003 г. во втором томе его Избранных трудов.

Одной из важных для криптографии математических задач является так называемая задача дискретного логарифмирования, которая в общем виде состоит в решении уравнений $a^x = b$ или $ax = b$, где a и b — элементы циклической группы G по умножению или сложению соответственно.

В настоящее время на гипотезе о сложности решения этой задачи в случаях, когда G — подгруппа большого простого порядка группы точек эллиптической кривой над конечным полем, основывается представление о стойкости некоторых систем шифрования, позволяющих абонентам организовывать обмен секретной информацией, пользуясь только незащищенным каналом связи.

В общем случае наилучшие известные алгоритмы вычисления дискретного логарифма имеют сложность порядка $\sqrt{|G|}$ операций.

Алгоритм такой сложности предложил А.О.Гельфонд в 60-е годы, за рубежом первая аналогичная публикация появилась в 1971 г. (Baby Step/Giant Step алгоритм).

В.И.Нечаев (Матем. заметки, 1994 г.) доказал, что если алгоритм вычисления дискретного логарифма использует только групповые операции, то его сложность по порядку не меньше $\sqrt{|G|}$ операций; за рубежом такой результат был получен позже.

Для групп специального вида, допускающих использование не групповых операций (например, разложение на множители элементов $\mathbb{Z}_p$ как целых чисел), существуют методы со сложностью порядка $\exp(c(\ln|G|)^{1/3}(\ln\ln|G|)^{2/3})$.

В ОПР изучались уравнения относительно подстановок (А.И.Павлов, М.П.Минеев, В.Ф.Колчин, В.Е.Тараканов). Пример: А.И.Павлов, Дискретная математика, 1989.
Если Q_n — число решений $(X_1, \ldots, X_k)$ системы уравнений относительно подстановок из S_n

$$X_1^{m_1} = X_2^{m_2} = \ldots = X_k^{m_k} = e,\ X_i X_j = X_j X_i\ (1 \leqslant i, j \leqslant k),$$

где e — тождественная подстановка, то при $m = m_1 \ldots m_k$ для некоторых целых чисел h_d

$$\frac{Q_n}{n!} \sim \frac{C n^{-n/m}}{\sqrt{2\pi m n}} \exp\left\{\sum\nolimits_{d|m} \frac{h_d}{d}\, n^{d/m}\right\}, n \to \infty,$$

и для случайно выбираемого решения системы вектор $(\eta_1, \ldots, \eta_k)$ чисел циклов асимптотически нормален.

В.Е.Тараканов получил оценки для максимальной глубины прямоугольных (0,1)-матриц с заданными суммами по строкам и столбцам, т.е. минимального числа строк, образующих подматрицу без нулевых столбцов (Матем.сборник, 1973).

Максимальная глубина (0,1)-матрицы как матрицы инцидентности гиперграфа совпадает с минимальным числом ребер, покрывающих все вершины.

Для широкого класса регулярных гиперграфов В.Е.Тараканов доказал справедливость гипотезы Райзера о неравенстве, связывающем трансверсальное число и число независимости (Матем.заметки, 1997).

В.Е.Тараканов и В.Н.Сачков — соавторы монографии «Комбинаторика неотрицательных матриц», изд-во ТВП, 2000, и Amer. Math. Soc., 2002.

Ф.М.Малышев и В.Е.Тараканов, обобщая известные в комбинаторике (v, k, λ)-конфигурации, ввели понятие (v, k)-конфигураций как таких совокупностей k-элементных подмножеств $X_0, X_1, \ldots, X_{v-1}$ множества $X = \{1, \ldots, v\}$, что соответствующие им (0,1)-матрицы инцидентности L невырождены, имеют ровно k единиц в каждой строке и в каждом столбце, и обратные (над $\mathbb{R}$) матрицы L^{-1} тоже имеют ровно k единиц в каждой строке и в каждом столбце. Доказан ряд теорем о (v, k)-конфигурациях и построены их бесконечные семейства (Матем.сборник, 2001).

(v, k)-конфигурации существуют только при нечетных k.

В последние годы Ф.М.Малышев вместе с учениками разработали ряд новых семейств построения (v, k)-конфигураций при конкретных небольших значениях k. Такие матрицы используются при построении обратимых преобразований двоичных строк.

Если X — конечное множество, то *граф де Брейна* — это ориентированный граф с вершинами $(x_1, \ldots, x_r) \in X^r$ и ребрами $(x_1, x_2, \ldots, x_r) \to (x_2, \ldots, x_r, x_{r+1})$. Это граф переходов конечного автомата — регистра сдвига. В графе де Брейна между любыми двумя вершинами существует единственный путь длины r.

Графы, обладающие последним свойством, называются *обобщенными графами де Брейна*. Их (0,1)-матрицы инцидентности Y являются решениями матричных уравнений вида $Y^r = J$, где J — матрица, состоящая из единиц.

Ф.М.Малышев доказал, что в случае $X = \{0, 1\}$ существует не менее $12r - 33$ попарно не изоморфных обобщенных графов де Брейна (Дискретная математика, 2020).

Предельные теоремы для характеристик случайных графов и отображений (чисел и размеров связных компонент, длин циклов, расстояний до циклов, высоты и ширины случайных деревьев и т.п.) изучали многие сотрудники отдела.

Монографии В.Ф.Колчина «Случайные отображения» и «Случайные графы» переведены на английский язык.

Для криптографии особую важность представляют последовательности независимых случайных величин, принимающих с равными вероятностями значения из конечного алфавита (отождествляемого с $\mathbb{Z}_n$).

Согласно теореме Шеннона идеальный шифр — это сложение знаков сообщения $m_1, m_2, \ldots, m_T \in \mathbb{Z}_n$ с элементами $\gamma_1, \gamma_2, \ldots, \gamma_T$ реализации такой последовательности: $m_1+\gamma_1, m_2+\gamma_2, \ldots, m_T+\gamma_T$. При этом последовательность $\gamma_1, \gamma_2, \ldots, \gamma_T$ должна быть известна только отправителю и получателю, а у остальных о ней не должно быть никакой информации.

Однако на практике последовательности $\gamma_1, \gamma_2, \ldots, \gamma_T$ получают с помощью технических устройств, и нельзя гарантировать, что порождаемые ими последовательности неотличимы от реализаций «идеальных» случайных последовательностей, не говоря уже о сбоях в их работе.

Поэтому для криптографии важны исследования:
а) способов проверки гипотезы H_0 о том, что конкретная последовательность является реализацией независимых случайных величин, имеющих равновероятное распределение,
б) способов обнаружения особенностей конкретной последовательности, не позволяющих считать ее соответствующей гипотезе H_0.

Для последовательности $\gamma_1, \gamma_2, \ldots, \gamma_T$ элементов $\mathbb{Z}_n$ можно вычислять частоты ν_j, $j \in \mathbb{Z}_n$ появлений элементов, частоты $\nu_{j_1,\ldots,j_s}$ появлений s-цепочек $(j_1, \ldots, j_s) \in \mathbb{Z}_n^s$ и сравнивать эти частоты с множествами их наиболее вероятных значений в случае, когда гипотеза H_0 о независимости и равновероятности выполняется.

Известные теоремы теории вероятностей об асимптотической нормальности сумм независимых случайных величин применимы к частотам s-цепочек только в случаях, когда объем выборки по порядку превосходит число всех возможных s-цепочек, а оно экспоненциально зависит от s.

В 60-е годы в СССР и за рубежом стала разрабатываться теория «случайных размещений частиц по ячейкам» — распределений частот знаков в последовательностях дискретных случайных величин. С точки зрения математической статистики эта теория относится к задачам о «малых выборках», когда объем выборки T по порядку либо меньше числа N возможных исходов, либо незначительно отличается от него.

Если число N исходов велико, то вместо частот рассматриваются случайные величины $\mu_r(N,T)$ — количества исходов, которые появились в выборке ровно $r = 0, 1, \ldots$ раз.

Типичные случаи: «левая область», когда $T \sim CN^{1-1/r}$ и распределение $\mu_r(N,T)$ сходится к распределению Пуассона, «центральная область», когда $T \sim CN$ и распределения $\mu_r(N,T)$ асимптотически нормальны, и «правая область», когда $T = N\ln N + rN\ln\ln N - N\ln(\lambda r!) + o(N)$ и распределение $\mu_r(N,T)$ сходится к распределению Пуассона.

Рассматривались различные задачи, связанные со случайными величинами $\mu_r(N,T)$, в том числе построение и оптимизация статистических критериев. Монография В.Ф.Колчина, Б.А.Севастьянова, В.П.Чистякова «Случайные размещения» опубликована в 1976 г. в издательстве «Наука», в 1978 году — в издательстве J.Wiley & Sons.

Частоты цепочек отражают глобальные свойства последовательностей. Локальные особенности могут обнаруживаться числами повторений s-цепочек или максимальными длинами повторяющихся цепочек.

В равновероятной двоичной последовательности длины 2^s при $s \to \infty$ каждая s-цепочка в среднем появляется один раз, но длина максимальной цепочки, появляющейся в этой последовательности хотя бы два раза, с вероятностью, близкой к 1, отличается от $2s$ на небольшую величину.

Цикл работ, начавшийся в 1974 г. с совместной работы с В.Г.Михайловым, интенсивно развивался им и другими отечественными и зарубежными авторами. Кроме точных повторений цепочек рассматривались как повторения с малым числом ошибок, так и обобщенные повторения (с точностью до переименования элементов), а также повторения в последовательностях зависимых случайных величин.

Задачи о случайных размещениях, повторениях цепочек и т.п. связаны с исследованием сумм зависимых случайных индикаторов, принимающих только значения 0 и 1.

Б.А.Севастьянов в 1972 г. построил удобный вариант метода моментов для доказательства сходимости распределений сумм зависимых индикаторов к распределению Пуассона.

В это же время за рубежом был разработан метод Чена–Стейна доказательства сходимости сумм зависимых индикаторов к распределению Пуассона с оценками скорости сходимости. Другие оценки точности пуассоновской аппроксимации для U-статистик вида $\sum_{i_1,\dots,i_d=1}^{n} f_{i_1,\dots,i_d}(X_{i_1},\dots,X_{i_d})$ были получены А.М.Зубковым (Матем. заметки, 1977).

Разрабатывались методы доказательства асимптотической нормальности сумм зависимых случайных индикаторов, в частности, U-статистик (В.Г.Михайлов, В.П.Чистяков).

Большое место среди публикаций сотрудников ОДМ занимают работы по теории ветвящихся процессов.

Термин «ветвяшиеся процессы» был введен А.Н.Колмогоровым и стал общепринятым; до этого в западной литературе использовались термины «cascade processes», «multiplicative processes» и т.п.

Классический ветвящийся процесс Гальтона–Ватсона описывает эволюцию численностей поколений в популяции частиц, каждая из которых независимо от остальных порождает случайное число потомков с одним и тем же распределением. Если $\xi_{t,j}$ — число потомков j-й частицы t-го поколения, а Z_t — число частиц t-го поколения, то процесс определяется рекуррентным соотношением $Z_{t+1} = \sum_{j=1}^{Z_t} \xi_{t,j}, t \geqslant 0$.

Ветвящиеся процессы делятся на три класса по значению математического ожидания $A = \mathbf{E}\xi_{t,j}$ числа потомков одной частицы:
при $A < 1$ процесс называется докритическим и с вероятностью 1 вырождается: $\mathbf{P}\{Z_t > 0\}$ экспоненциально убывает при $t \to \infty$,
при $1 < A < \infty$ процесс называется надкритическим, последовательность Z_t с положительной вероятностью растет как геометрическая прогрессия A^t,
при $A = 1$ процесс называется критическим, он вырождается с вероятностью 1, но значительно медленнее, чем в докритическом случае, и предельные условные распределения таких процессов при условии невырождения к моменту наблюдения, как правило, имеют явно описываемый вид, не зависящий от конкретного распределения числа потомков.

Например, если $A = \mathbf{E}\xi_{t,j} = 1$ и $B = \mathbf{D}\xi_{t,j} < \infty$, то $\mathbf{P}\{Z_t > 0\} \sim \frac{2}{Bt}$ и $\lim_{t\to\infty} \mathbf{P}\left\{\frac{2Z_t}{Bt} \leqslant x\right\} = 1 - e^{-x}$ (теорема Яглома).

Основы этого направления были заложены А.Н.Колмогоровым в 40-х годах XX века; важная роль в его развитии принадлежит Б.А.Севастьянову. Его первая работа (совместная с А.Н.Колмогоровым) о финальных вероятностях для ветвящихся процессов появилась в 1947 г. После этого вплоть до 2007 года Б.А. регулярно публиковал результаты для различных модификаций ветвящихся процессов: ветвящихся процессов с непрерывным временем, ветвящихся процессов с несколькими типами частиц, ветвящихся процессов с перемещением частиц в пространстве, ветвящихся процессов с превращениями, зависящими от возраста частиц, регулируемых ветвящихся процессов, ограниченных снизу ветвящихся процессов, ветвящихся процессов с взаимодействием частиц.

Линейность по Z_t рекуррентного соотношения $Z_{t+1} = \sum_{j=1}^{Z_t} \xi_{t,j}$ позволяет сводить исследование свойств ветвящихся процессов к исследованию итераций производящих функций; в случае регулируемых ветвящихся процессов такой линейности обычно нет, и тогда приходится использовать другие методы.

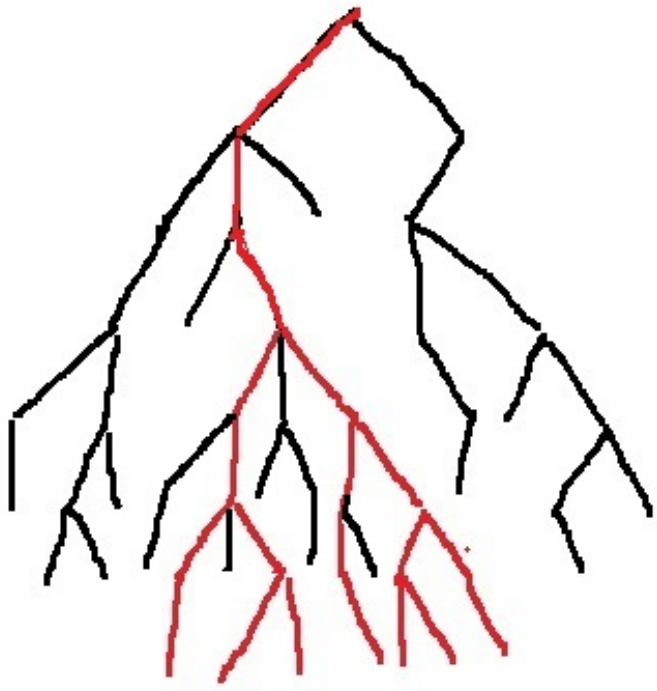
Гальтон и Ватсон ввели ветвящийся процесс, чтобы объяснить эффект исчезновения фамилий английских пэров; фамилии передавались только потомкам мужского пола, и при их отсутствии линия прерывалась. Имея в виду этот эмпирический опыт, Гальтон и Ватсон (некорректно) пришли к выводу о том, что любые ветвящиеся процессы вырождаются.

В 1971 г. Б.А.Севастьянов опубликовал монографию «Ветвящиеся процессы», которая была переведена на немецкий язык.

Б.А.Севастьянов был научным руководителем не менее 12 кандидатских диссертаций по теории ветвящихся процессов, в том числе В.И.Афанасьева, В.А.Ватутина, Е.Е.Дьяконовой, А.М.Зубкова, В.И.Хохлова, В.П.Чистякова, В.И.Шерстнева, А.Л.Якымива (МИАН).

Теорией ветвящихся процессов стали заниматься в Новосибирске, Ташкенте, Киеве и других городах, и возникла отечественная школа теории ветвящихся процессов, которая интенсивно работает и сейчас.

В настоящее время в отделе дискретной математики исследования по теории ветвящихся процессов проводят В.А.Ватутин, В.И.Афанасьев, Е.Е.Дьяконова. Основными направлениями их исследований являются более сложные модели:
условные распределения траекторий при условии невырождения,
ветвящиеся процессы с несколькими типами частиц,
ветвящиеся процессы в случайной среде,
сопровождающие случайные блуждания.



Генеалогическое дерево ветвящегося процесса
Редуцированный процесс
Расстояние до ближайшего общего предка (1975, 1977)

Ветвящиеся процессы с несколькими сообщающимися типами частиц $T_1, T_2, \ldots, T_n$ аналогичны процессам с одним типом частиц. Качественно другими свойствами обладают разложимые ветвящиеся процессы; например, если превращения типов имеют вид $T_1 \to T_2 \to \ldots \to T_n$. Для критических ветвящихся процессов с одним типом частиц $(n = 1)$ вероятность невырождения за время $t \to \infty$ имеет асимптотику $\frac{2}{Bt}$, а для $n > 1$ асимптотика вероятности невырождения имеет вид $Ct^{-2^{1-n}}$ (Савин 1959, Полин 1976, Матем.сборник).

Для разложимых критических ветвящихся процессов с несколькими типами частиц В.А.Ватутин описал предельную структуру редуцированных процессов при условии невырождения (Теория вероятн. и ее примен., 2014, 2015).
$\left[0, \frac{1}{2^{N-1}}\right) \sqcup \left[\frac{1}{2^{N-1}}, \frac{1}{2^{N-2}}\right) \sqcup \left[\frac{1}{2^{N-2}}, \frac{1}{2^{N-3}}\right) \sqcup \ldots \sqcup \left[1 - \frac{1}{2}, 1\right)$

Если законы размножения частиц зависят от номера поколения, то процесс становится неоднородным по времени и его свойства изменяются.

Чтобы сохранить в какой-то мере однородность, в 1961 году были введены ветвящиеся процессы со случайной средой: законы размножения частиц в каждом поколении одинаковы, но для разных поколений они определяются значениями среды, т.е. выбираются случайно.

Фактически такие процессы получаются как объединение (смесь) траекторий разных неоднородных процессов. Свойства ветвящихся процессов со случайной средой определяются как законами размножения, так и распределениями значений случайной среды.

При этом возникает огромное число разных постановок задач и результатов.

Пусть A_t — математическое ожидание числа потомков частиц t-го поколения. Для процессов со случайной средой это независимые случайные величины, они могут быть как больше, так и меньше 1, поэтому уже вопрос о классификации процессов на докритические, критические и надкритические нетривиален.

Полная классификация построена в 2005 году в статье В.И.Афанасьева, В.А.Ватутина, Д.Гейгера, Г.Керстинга (Ann. Inst. H.Poincare, Probab., Statist.). Она основывается на предельном поведении сопровождающих случайных блужданий $S_n = \ln A_1 + \ln A_2 + \ldots + \ln A_n$:
$\mathbf{P}\{\lim_{n\to\infty} S_n = +\infty\} = 1 \Rightarrow$ надкритический,
$\mathbf{P}\{\lim_{n\to\infty} S_n = -\infty\} = 1 \Rightarrow$ докритический,
$\mathbf{P}\{\limsup_{n\to\infty} S_n = +\infty\} = \mathbf{P}\{\liminf_{n\to\infty} S_n = +\infty\} = 1$
$\Rightarrow$ критический.

В множестве докритических процессов со случайной средой выделяются еще три подкласса.

Для ветвящихся процессов со случайной средой сотрудниками отдела опубликованы десятки статей. В них:

найдены асимптотики вероятностей невырождения, предельные условные распределения чисел частиц и траекторий при условии невырождения,

исследованы связи распределений чисел частиц со свойствами траекторий сопровождающих случайных блужданий (в частности, внезапное вырождение, «бутылочные горлышки», неблагоприятная среда),

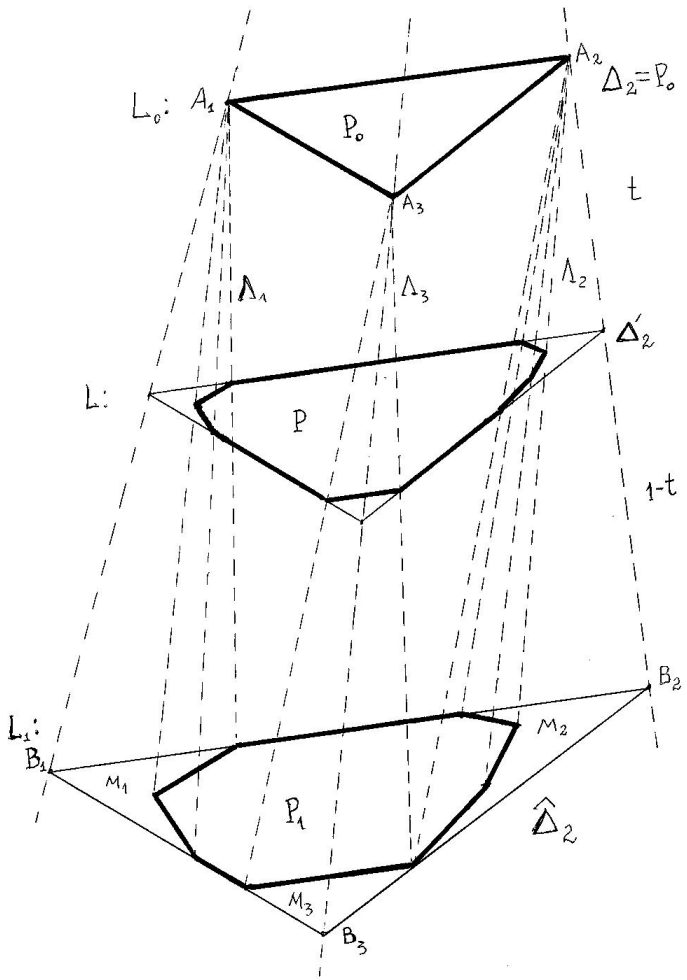
изучались ветвящиеся процессы с несколькими типами частиц в случайной среде.

Точные формулировки довольно громоздки.

В заключение приведу несколько примеров разнородных достаточно просто описываемых результатов, полученных в последние годы сотрудниками отдела.

В геометрии известно неравенство Брунна–Минковского с богатой историей.

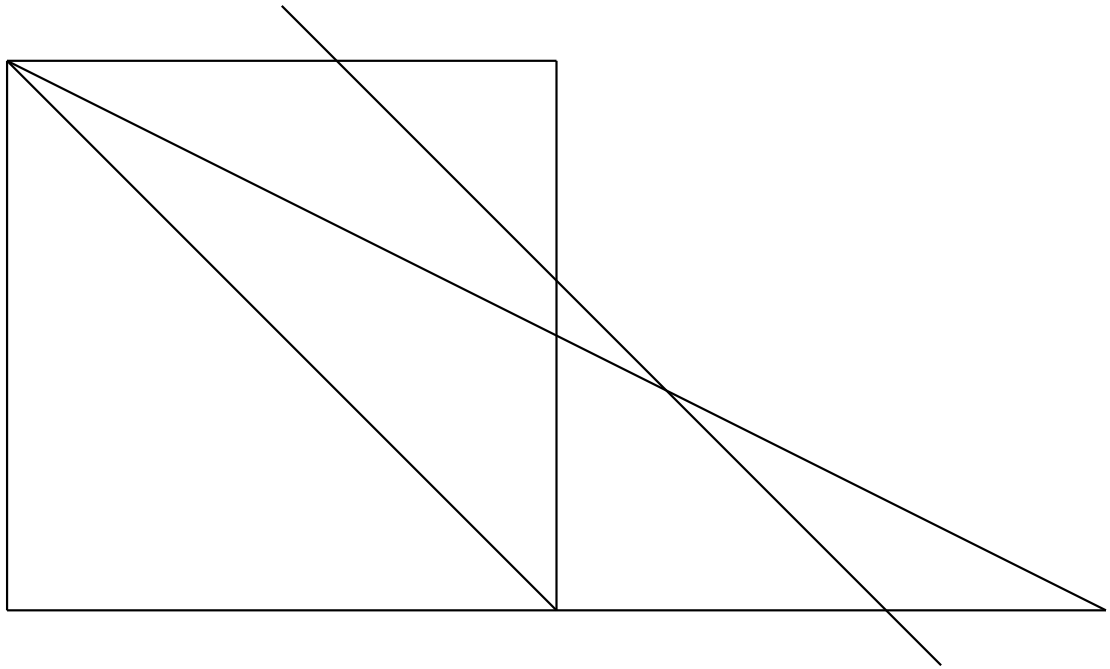
Теорема. *Пусть в двух параллельных гиперплоскостях* $L_0, L_1 \in \mathbb{R}^{n+1}, n \geqslant 1$, *содержатся выпуклые тела* P_0, P_1 *одинакового* n*-мерного объёма* $v > 0$ *и пусть* P *— сечение выпуклой оболочки их объединения гиперплоскостью* L, *параллельной* L_0, L_1 *и находящейся строго между ними. Тогда* n*-мерный объём тела* P *не меньше* v, *причём если он равен* v, *то* P_1 *получается из* P_0 *параллельным переносом.*



Ф.М.Малышев, Завершение доказательства теоремы Брунна элементарными средствами. — Чебышевский сборник, 2021, т. 22, вып. 2, с. 160–182.

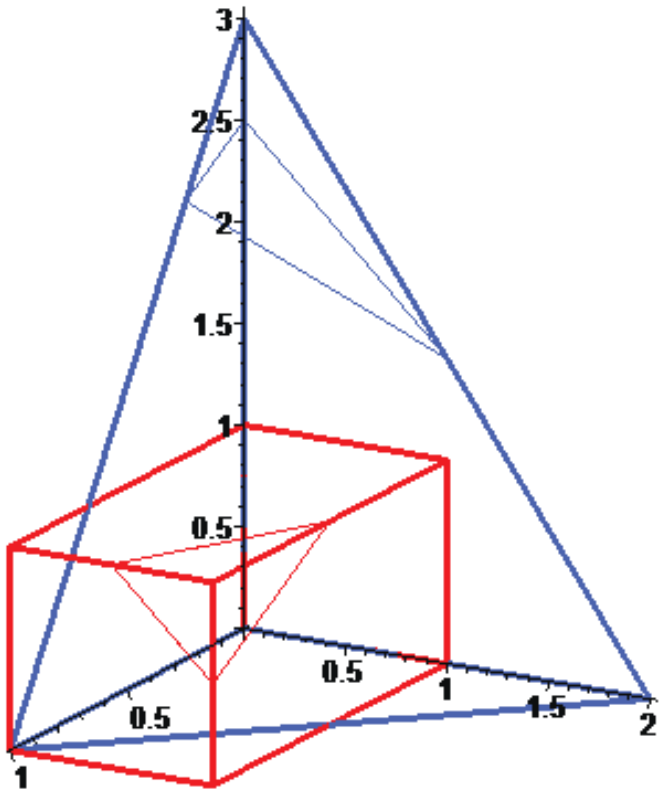
$$\mathbf{P}\{\max\{\xi_1,\ldots,\xi_n\}\leqslant 1\}=\int_0^n C(t)\lambda^n e^{-\lambda t}dt=$$

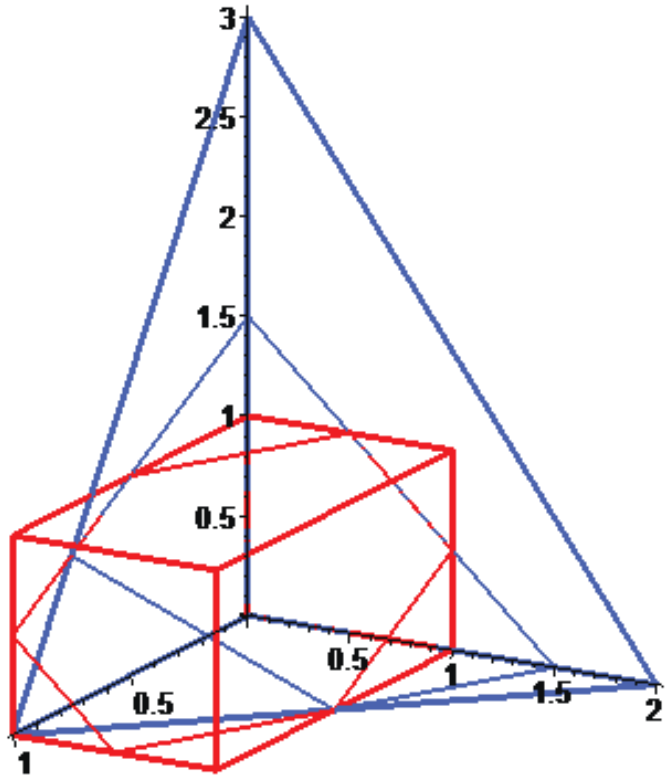
$$=\mathbf{P}\left\{\xi_1+\tfrac{\xi_2}{2}+\ldots+\tfrac{\xi_n}{n}\leqslant 1\right\}=\int_0^n S(t)\lambda^n e^{-\lambda t}dt\quad\forall\lambda>0$$



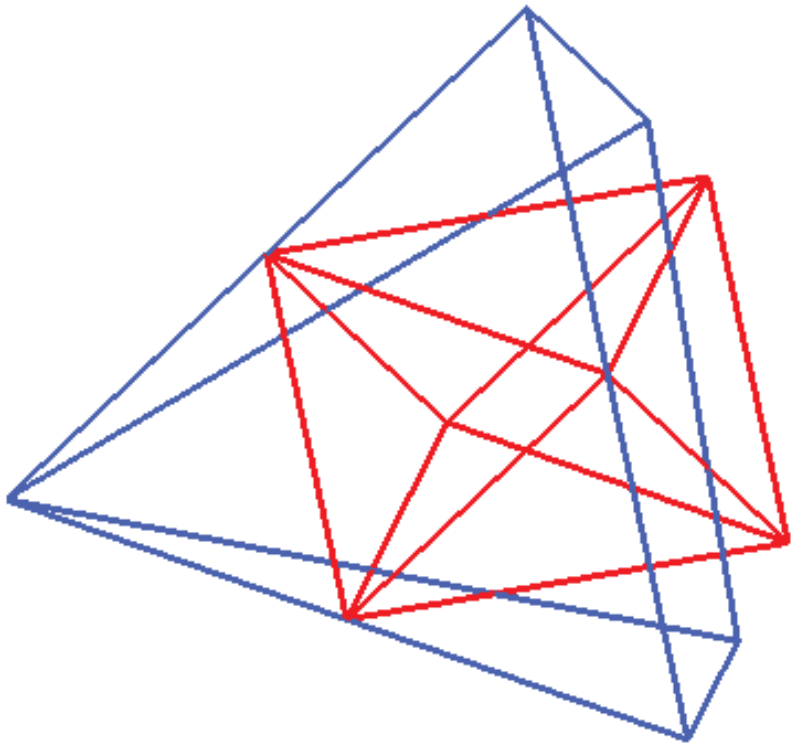
А.М.Зубков Вероятностное доказательство геометрической теоремы. Матем. заметки, 1979, т. 26, вып. 6, с. 957–959

Пересечения куба и тетраэдра двумя плоскостями:
$L_{2.5} = \{(x_1, x_2, x_3) : x_1 + x_2 + x_3 = 2.5\}$ и
$L_{1.5} = \{(x_1, x_2, x_3) : x_1 + x_2 + x_3 = 1.5\}$

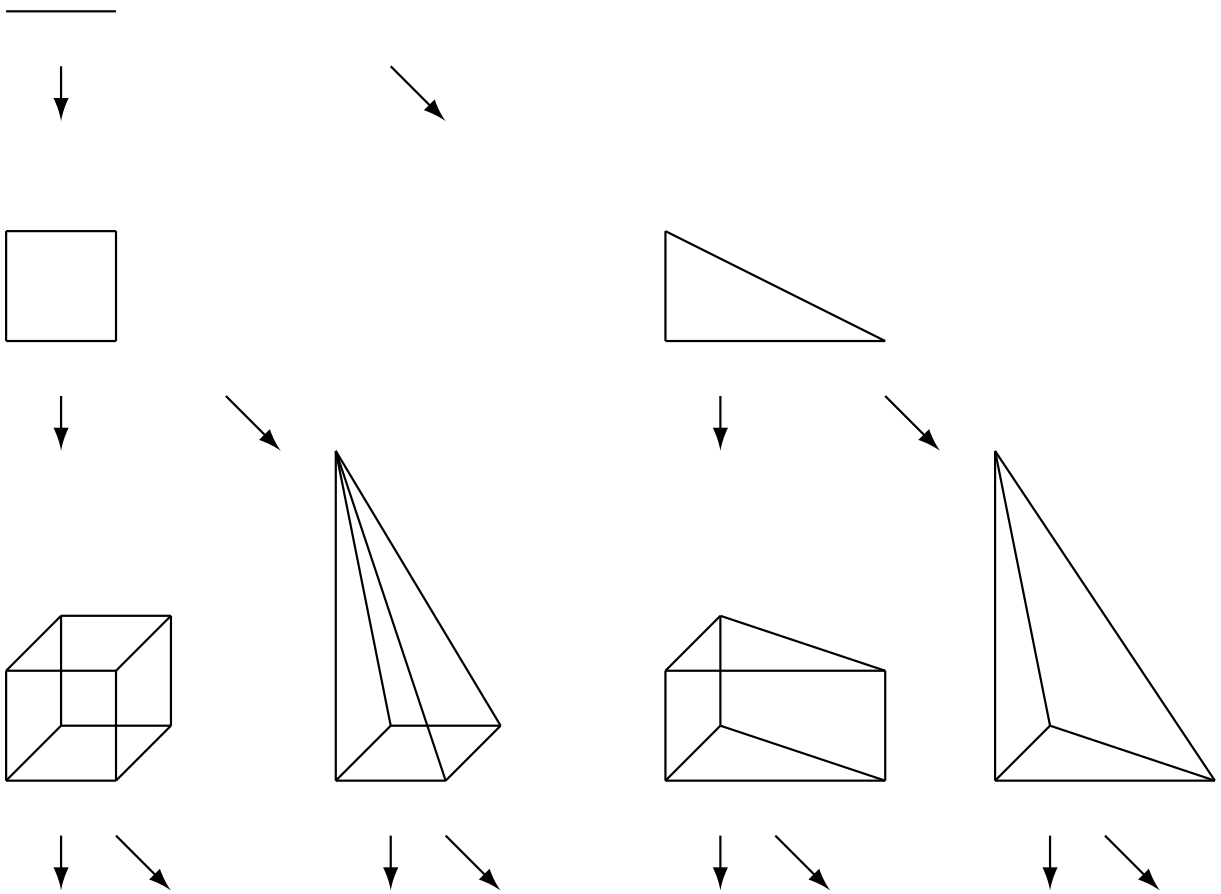




Трехмерные пересечения 4-мерных куба и симплекса гиперплоскостью $L_2 = \{(x_1, x_2, x_3) : x_1 + x_2 + x_3 = 2\}$



Продолжение:



Малышев Ф.М. Семейство равновеликих n-мерных многогранников, удовлетворяющих принципу Кавальери. — Матем. заметки, 2015, т. 97, вып. 2, 231–248

Двусторонние неравенства для сумм $\sum_{m=0}^{k} C_n^m p^m (1-p)^{n-m}$

$S_{n,p}$ — число успехов в n независимых испытаниях с вероятностью успеха $p \in (0,1)$,

$$\mathbf{P}\{S_{n,p} = k\} = C_n^k p^k (1-p)^{n-k},$$

$$\mathbf{P}\{S_n \leqslant k\} = \sum_{m=0}^{k} C_n^m p^m (1-p)^{n-m} = (k+1) C_n^{k+1} \int_p^1 t^k (1-t)^{n-k-1} dt.$$

Теорема Муавра–Лапласа, 1730–1812. *Если* $p = \mathrm{const}$, *то при любом фиксированном* $x \in (-\infty, \infty)$

$$\mathbf{P}\{S_{n,p} = k\} = \frac{1}{\sqrt{2\pi np(1-p)}} \exp\left\{-\frac{(k-np)^2}{np(1-p)}\right\} + O\left(\frac{1}{\sqrt{n}}\right), \quad n \to \infty,$$

$$\lim_{n\to\infty} \mathbf{P}\{S_{n,p} \leqslant np + x\sqrt{np(1-p)}\} = \Phi(x) \stackrel{\text{def}}{=} \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-u^2/2} du.$$

$$H(x,p) = x\ln\tfrac{x}{p} + (1-x)\ln\tfrac{1-x}{1-p},$$
$$\operatorname{sgn}(x) = \tfrac{x}{|x|}, \text{ если } x \neq 0,\ \operatorname{sgn}(0) = 0,$$

$\{C_{n,p}(k)\}_{k=0}^{n+1}$ — монотонно возрастающие по k последовательности:

$$C_{n,p}(0) = 0,\ C_{n,p}(n+1) = 1,$$
$$C_{n,p}(k) = \Phi\left(\operatorname{sgn}\left(\tfrac{k}{n} - p\right)\sqrt{2nH\left(\tfrac{k}{n}, p\right)}\right),\ 1 \leqslant k \leqslant n.$$

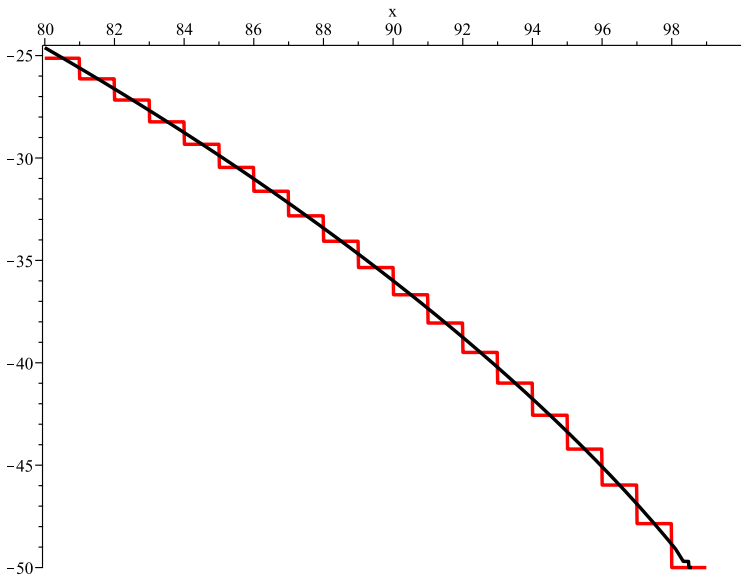
Теорема. *Если* $k = 0, 1, \ldots, n$ *и* $0 < p < 1$, *то*

$$C_{n,p}(k) < \mathbf{P}\{S_{n,p} \leqslant k\} < C_{n,p}(k+1).$$

Иначе говоря:

$$\ldots < \mathbf{P}\{S_{n,p} \leqslant k-1\} < C_{n,p}(k) < \mathbf{P}\{S_{n,p} \leqslant k\} <$$
$$< C_{n,p}(k+1) < \mathbf{P}\{S_{n,p} \leqslant k+1\} < C_{n,p}(k+2) < \ldots$$

Зубков А. М., Серов А. А. Полное доказательство универсальных неравенств для функции распределения биномиального закона. — Теория вероятностей и ее применения, 2012, т. 57, № 3, с. 597–602.



Графики $\log_{10}(1 - \mathrm{Bin}(100, 0.3)(x))$ и $\log_{10}(1 - C_{100,0.3}(x))$.