Краткий очерк жизни и творчества академика А. А. Маркова

А. В. Булинский

Прошло 100 лет со дня смерти А. А. Маркова. Его жизнь — яркий пример служения отечественной науке. В нашей статье даётся краткая биография учёного и рассматриваются некоторые полученные им результаты, относящиеся к теории вероятностей. Акцент делается на последующем развитии идей выдающегося математика. Изложение ориентировано на широкий круг читателей, однако приведены и сведения, представляющие интерес для специалистов.

§ 1. Биография

1.1. Детство. Обучение в гимназии

Выдающийся русский математик Андрей Андреевич Марков родился 2 (14) июня 1856 года в Рязани в семье чиновника. В детстве Андрей тяжело болел костным туберкулёзом, ходил на костылях, поскольку одна нога не разгибалась в колене. Семья переехала в Петербург, и там после удачной операции нога у мальчика стала разгибаться. Далее он передвигался без костылей, хотя и немного прихрамывал всю жизнь. В 1866 году А. Маркова определили в Пятую петербургскую гимназию. От гимназии у Андрея Маркова остались тяжёлые воспоминания. Его отец, А. Г. Марков, писал: «Опять меня вызывали к директору. Ничем

Автор статьи — д. ф.-м. н., профессор, лауреат премии имени М. В. Ломоносова за научные работы (МГУ, 2010), лауреат премии имени А. Н. Колмогорова РАН (2021), лауреат Общенациональной премии «Профессор года (2022)» в номинации «Физико-математические науки».

А. А. Марков, 1886 год

Андрей не хочет заниматься, кроме математики». Однако это не совсем точное наблюдение. А. Марков проявлял острый интерес к явлениям общественной жизни. Когда Андрей был в выпускном классе, директор вызвал его отца и заявил ему, что не потерпит у себя в гимназии «атеистов и нигилистов». Андрею Григорьевичу удалось убедить директора принять сына вновь в выпускной класс. В аттестате А. Маркова было указано, что «любознательность весьма в удовлетворительной степени развита к физико-математическим наукам». Педагогический совет выставил А. Маркову по математике 5, и эта же оценка была получена им на выпускных испытаниях. По латинскому, греческому и немецкому языкам стояли тройки. Сохранились записные книжки юного А. Маркова, в которых он давал решения сложных математических задач. Ему показалось, что он изобрёл новый метод решения дифференциальных уравнений с постоянными коэффициентами. Об этом он написал вице-президенту С.-Петербургской академии наук В. Я. Буняковскому, который в письме (от 13 ноября 1873 года) гимназисту А. Маркову указал, что предлагаемый метод «уступает со стороны простоты общеизвестному», а также отметил ряд пробелов в тексте автора.

1.2. Обучение и работа в университете

В 1874 году Андрей Марков, получив аттестат, поступил на физико-математический факультет С.-Петербургского университета. Поражает сильный состав педагогов факультета того периода. Достаточно сказать, что теорию чисел и теорию вероятностей читал П. Л. Чебышёв, органическую химию — А. М. Бутлеров, неорганическую химию — Д. И. Менделеев. В отличие от гимназии, А. Марков был одним из лучших студентов. За весь период обучения в С.-Петербургском университете он имел лишь две четвёрки, а остальные — пятёрки (одну четвёрку ему поставил на экзамене Д. И. Менделеев, а другую он получил по богословию). В 1877 году студентам математикам была предложена конкурсная тема по дифференциальным уравнениям для самостоятельного исследования. Студент А. Марков был награждён за выполненное исследование золотой медалью. В 1878 году А. Марков окончил университет со степенью кандидата по разряду математических наук (так тогда назывался диплом о высшем образовании) и был оставлен при университете «для приготовления к профессорскому званию». Так началась работа А. Маркова в С.-Петербургском университете, продолжавшаяся всю его жизнь с небольшими перерывами.

Вспоминая годы обучения в университете, А. Марков писал, что среди его профессоров «следует упомянуть П. Л. Чебышёва, Е. И. Золотарёва и в особенности А. Н. Коркина». В 1880 году А. Марков защитил магистерскую диссертацию «О бинарных квадратичных формах положительного определителя» и был удостоен степени магистра математики (аналог нашей кандидатской диссертации), об этой работе можно прочитать на с. 146–193 книги [1]. После этого он подаёт прошение ректору Императорского Санкт-Петербургского университета профессору А. Н. Бекетову об утверждении его в должности приват-доцента (в современном понимании ассистента). Интересно, что для вступления в должность полагалось удачно прочитать две лекции на заданные факультетом темы. А. А. Маркову это удалось. Далее Андрей Андреевич подготовил докторскую диссертацию, и после успешной защиты был утверждён в степени доктора математики (аналог нынешней докторской диссертации) в 1885 году. В 1882 году П. Л. Чебышёв прекратил читать лекции в университете, и А. А. Марков начал читать лекции по теории вероятностей. До этого он преподавал дифференциальное исчисление и высшую алгебру, интегральное исчисление, аналитическую геометрию, теорию чисел. В 1886 году А. А. Марков был утверждён экстраординарным профессором по кафедре чистой математики (соответствует современной должности доцента). В 1893 году он назначен ординарным профессором (соответствует нынешней должности профессора) и читал лекции вплоть до 1921 года, когда серьёзно заболел.

Преподавание А. А. Маркова не отличалось внешними эффектами. По воспоминаниям профессора физики Б. П. Вейнберга, «некоторые из нас увлекались способом изложения А. А. Маркова, каждым словом как бы заколачивавшего гвоздь за гвоздём по одной прямой линии, с которой не давал сходить истине». Это же можно сказать и об учебных пособиях А. А. Маркова, в частности, о знаменитой книге «Исчисление вероятностей» [2]. Обычно на экзаменах А. А. Марков ставил много пятёрок и четвёрок. Однако экзамен мог продолжаться несколько дней.

Уже в молодости А. А. Марков не только был крупным учёным, но и проявил твёрдость, отстаивая свои убеждения. Очень актуально звучит высказывание А. А. Маркова: «Я не отрицаю пользы составления учебных планов и программ, но не для контроля над отдельными преподавателями, а для установления надлежащей полноты и последовательности преподавания, причём программы могут быть только самого общего характера, не стесняющие свободы преподавания». В декабре 1905 года на заседании Совета он высказал мнение, что кафедра богословия должна быть отделена от всех факультетов и преподавание богословия не должно быть обязательным. Требование знать латинский язык А. А. Марков считал вредным. Он неоднократно писал заявления в различные инстанции вплоть до министра просвещения, отстаивая свои взгляды по различным вопросам общественной жизни, выступал в защиту студентов.

1.3. Деятельность в Академии наук

В 1896 году Общее собрание Академии наук избрало А. А. Маркова ординарным академиком (аналог современного звания академика). В этом качестве Андрей Андреевич активно участвовал в научно-организационной работе, входил в различные комиссии. Например, был членом комиссии по реформе календаря в России. В 1905 году эта комиссия приняла решение о целесообразности перехода от юлианского календаря к григорианскому. Однако данное решение было реализовано только в 1918 году. Выдающийся математик В. А. Стеклов (имя которого теперь носит Математический институт РАН) подчёркивал, что Андрей Андреевич был «неизменно-пречестным человеком», хотя часто и вступал в споры, не обращая внимания на принятую дипломатическую форму дискуссий. При этом А. А. Марков излагал свои взгляды, невзирая на то, какое положение занимал человек, с которым шла полемика. А. А. Марков подверг резкой критике научные труды С. В. Ковалевской, К. Пирсона, П. А. Некрасова, Б. Б. Голицына и др. (в некоторых случаях он признавал, что судил слишком строго). Принято считать, что Андрей Андреевич был абстрактным теоретиком, не интересовавшимся практическими задачами. Однако это не так. Например, он принял участие в работе по совершенствованию расчётов эмеритальных касс, созданных в России в середине XIX века и осуществлявших обязательное страхование государственных служащих (название происходит от латинского слова «emeritus», что означает «заслуженный»). С 1909 года до конца жизни А. А. Марков был членом комиссии по изучению научного наследия Л. Эйлера и изданию его сочинений. А. А. Марков был одним из инициаторов документа «Записка о нуждах просвещения в России», опубликованного в газетах в январе 1905 года за подписью 342 научных работников, в том числе 16 академиков. В этом документе говорилось о тяжёлом положении школы начальной, средней и высшей. Авторы указывали, что политика правительства в области народного просвещения задерживает духовный рост народа, что необходимо коренное преобразование общественного строя. Там же отмечалось тяжёлое положение преподавателей высшей школы, указывалось на их падающий научный и нравственный авторитет. После свержения самодержавия А. А. Марков продолжает активную работу в Академии. В ноябре 1918 года он избирается представителем Академии в Совет по делам статистики Центрального статистического управления.

1.4. Отношение к самодержавию и религии

А. А. Марков не только без колебаний откликался на события университетской и академической жизни, но и принимал близко к сердцу события государственного масштаба. Например, когда в петербургской газете «Новое время» в 1904 году по поводу неудач в русско-японской войне было напечатано, что «все мы виноваты, и теперь лишь пожинаем плоды посеянного нами», А. А. Марков на следующий день направил письмо издателю газеты А. В. Суворину, в котором написал: «В № 10334 (от 6 дек.) Вашей газеты неудачная война с Японией поставлена в вину всем. Такое обвинение всем представляет, с одной стороны, клевету, а с другой — попытку скрыть настоящих виновных». А. А. Марков отказался войти в комиссию по подготовке празднования в 1913 году 300-летия Дома Романовых. Параллельно этому он развернул кампанию по празднованию 200-летия закона больших чисел Бернулли. 1 декабря 1913 года на торжественном заседании Академии наук он выступил с докладом «Очерк развития закона больших чисел как совокупности математических теорем». По воспоминаниям В. А. Стеклова, мест в зале не хватило на всех желающих слушать этот доклад. Андрей Андреевич был возмущён отменой (по указанию Николая II) избрания писателя М. Горького (А. М. Пешкова) почётным академиком и выступал за пересмотр этого решения. В 1912 году

А. А. Марков обратился в Священный синод с просьбой отлучить его от Церкви. В своём прошении он написал: «Надеюсь, что достаточным основанием для отлучения может служить ссылка на мою книгу „Исчисление вероятностей“, где ясно выражено моё отрицательное отношение к сказаниям, лежащим в основании еврейской и христианской религии». В итоге было зафиксировано его отпадение от Церкви.

1.5. Семья

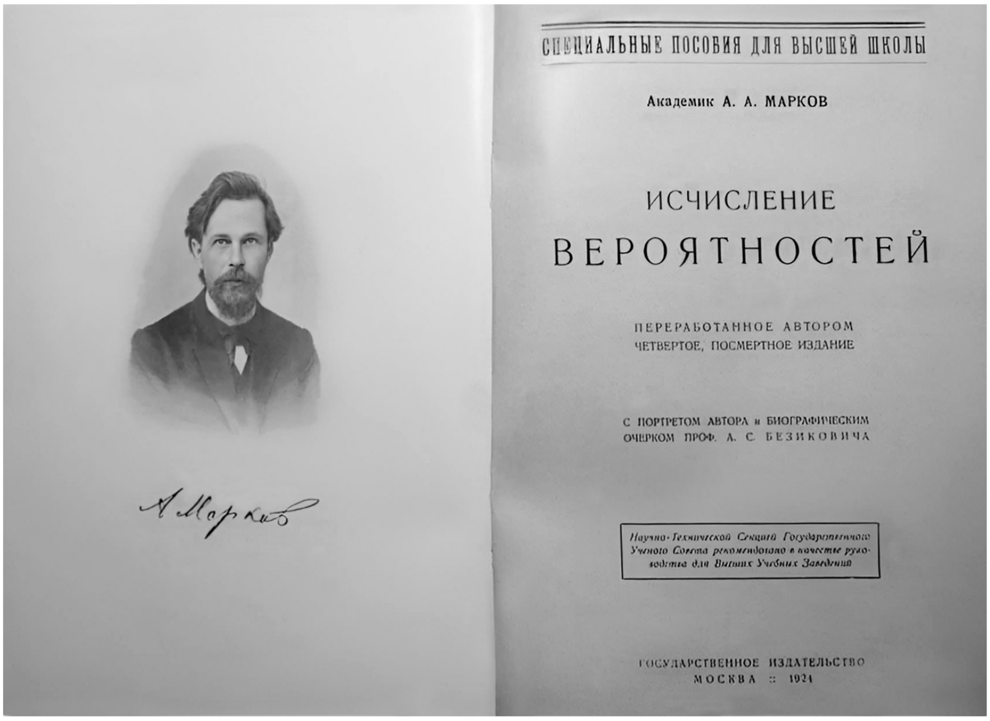
В 1883 году А. А. Марков женился на Марии Ивановне Вальватьевой. Они были знакомы с детства, поскольку одно время отец Андрея был управляющим в имении помещиков Вальватьевых. Сильные взаимные чувства у молодых людей возникли, когда студент Андрей Марков стал заниматься математикой с Марией, которой в гимназии не давались точные науки. Её мать, Екатерина Александровна, желала для дочери мужа с более определённым положением, чем студент. Свадьба состоялась лишь тогда, когда Андрей Марков был уже приват-доцентом, готовился защищать докторскую диссертацию и становиться профессором. В конце 90-х годов А. А. Марков с женой переехали в жилой дом Академии наук, поскольку Андрей Андреевич уже был избран академиком. Этот дом № 2 на 7-й линии Васильевского острова украшен множеством мемориальных досок в память о выдающихся учёных, живших в нём.

Своих детей у Марковых не было 20 лет. Поэтому они взяли на воспитание трёх осиротевших детей своих дальних родственников. Только в 1903 году у Марковых родился сын Андрей. Мальчик рос болезненным, родители его обожали и проявляли трогательную заботу. Сохранились письма А. А. Маркова, в которых непримиримый публицист предстаёт любящим отцом, заботящимся не только о здоровье сына, но и о его воспитании.

А. А. Марков был разносторонне одарённым человеком. Он проявил себя знатоком культуры древнего Ирана (Персии), был замечательным шахматистом, с большим успехом участвовал в шахматных турнирах по переписке, дружил с основателем отечественной шахматной школы М. И. Чигориным. Сын А. А. Маркова стал известным советским математиком, был избран членом-корреспондентом АН СССР.

1.6. Последние годы жизни

Октябрьскую революцию А. А. Марков (в отличие, например, от К. А. Тимирязева) не понял. А. А. Марков в письмах жалуется на неустроенность быта, хотя он и был включён в список лиц, которым необходим усиленный паёк питания. 5 марта 1921 года А. А. Марков сооб-



Посмертное издание классического
труда А. А. Маркова «Исчисление вероятностей»

щил Общему собранию Академии наук, что из-за отсутствия обуви лишён возможности посещать заседания Академии. Комиссия под председательством А. М. Горького выделила обувь выдающемуся учёному. Однако 25 мая 1921 года А. А. Марков заявил следующее: «Наконец, я получил обувь; но она не только дурно сшита, но и совершенно не подходит мне по своим размерам. Полученную мною обувь я предлагаю поместить в Этнографическом музее, как образец материальной культуры текущего момента, ради чего я готов ею пожертвовать».

20 июля 1922 года А. А. Марков скончался от заражения крови при хирургической операции. Он был похоронен на Митрофаниевском кладбище Петрограда. В 1954 году прах А. А. Маркова перенесён на Литераторские мостки Волкова кладбища Ленинграда.

Андрей Андреевич Марков прожил очень яркую жизнь. Для нас он является не только выдающимся учёным, обогатившим мировую науку классическими результатами, но и примером гражданского мужества при отстаивании своих взглядов. Много интересного материала о жизни А. А. Маркова содержит книга [3].

§ 2. Некоторые результаты А. А. Маркова в области теории вероятностей и их развитие

С именем Маркова связаны крупные достижения в теории чисел, приближении функций, исчислении конечных разностей, проблеме моментов. Однако центральное место в его научном творчестве занимают работы по теории вероятностей и математической статистике. Комментарии к ряду трудов А. А. Маркова можно прочитать, например, в приложениях к [3, 4] и в [5]. Объём данной статьи позволяет рассмотреть лишь часть деятельности Андрея Андреевича в области теории вероятностей, отметив последующее развитие некоторых идей выдающегося учёного. Желательно знакомство читателя с базовыми понятиями случайной величины и независимости таких величин. Ряд используемых далее терминов поясняется, чтобы сделать изложение доступным широкой аудитории.

2.1. Закон больших чисел

Очень давно было подмечено, что при большом числе опытов, производимых в одинаковых условиях так, что результаты этих опытов не влияют друг на друга, частота наступления изучаемого явления (события) приобретает устойчивое поведение и как бы совершает небольшие колебания возле некоторого числа. Например, К. Пирсон (1857–1936) подбросил обычную монету 24 000 раз, герб выпал 12 012 раз. Таким образом, частота выпадения герба в этой серии опытов составила 12 012/24 000, что очень близко к 1/2. Однако не следует думать, что при увеличении числа бросаний монеты частота будет сходиться к пределу 1/2 в обычном понимании предела числовой последовательности. Оказалось, что можно построить математическую модель, описывающую подобные эксперименты, и в рамках этой модели дать строгое объяснение упомянутой устойчивости частот. Для удобства читателей напомним некоторые базовые понятия.

Случайные эксперименты описываются с помощью пространства элементарных исходов Ω, состоящего из точек ω. Каждая точка ω отвечает одному из возможных исходов. Можно сказать, что в случайном эксперименте результат неизвестен до его проведения (в отличие, например, от эксперимента по нагреванию воды до 100 °C при нормальном давлении, когда она превращается в пар). Например, при высоком подбрасывании монеты не удаётся предсказать, какой стороной она выпадет на поверхность. При n бросаниях монеты исход ω удобно описывать вектором в $\mathbb{R}^n$ с компонентами, равными нулю или единице (единица соответствует выпадению «герба», а нуль — «решки»). Если представить, что монета подбрасывается неограниченное количество раз, то элементарный исход — это бесконечная последовательность, имеющая на каждом месте нуль или единицу. Заметим, что бесконечные последовательности из нулей и единиц дают пример несчётного пространства Ω.

После построения Ω выделяется класс $\mathscr{F}$ его подмножеств, называемых событиями. Если Ω дискретно, т. е. состоит из конечного или счётного числа элементарных исходов, то в качестве $\mathscr{F}$ рассматриваются все подмножества из Ω. В других случаях, как правило, выделяется непустой класс подмножеств, обладающих следующими свойствами. Если $A \in \mathscr{F}$, то *дополнение* $\overline{A} := \Omega \setminus A \in \mathscr{F}$. Если $A_1, A_2, \ldots \in \mathscr{F}$, то $\bigcup_{n=1}^{\infty} A_n \in \mathscr{F}$. Тогда автоматически и $\bigcap_{n=1}^{\infty} A_n \in \mathscr{F}$ (сказанное относится и к объединению конечного числа множеств из $\mathscr{F}$, а также и к их пересечению). Указанные два свойства задают систему подмножеств Ω, называемую *сигма-алгеброй* (σ-алгеброй). Очевидно, $\Omega \in \mathscr{F}$ и пустое множество $\varnothing \in \mathscr{F}$. Сигма-алгебры подмножеств можно рассматривать не только в Ω, но в других множествах S, например, для $S = \mathbb{R}$. Разумеется, совокупность всех подмножеств произвольного (непустого) множества является σ-алгеброй.

Затем на *измеримом пространстве* $(\Omega, \mathscr{F})$ вводится мера. Неотрицательная счётно-аддитивная функция $\mathbb{Q}$, заданная на $\mathscr{F}$, называется *мерой*. Счётная аддитивность означает, что для последовательности $A_1, A_2, \ldots$ множеств из $\mathscr{F}$, которые попарно не пересекаются (т. е. $A_i \cap A_j = \varnothing$ при $i \neq j$, $i, j \in \mathbb{N}$), справедливо равенство:

$$\mathbb{Q}\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} \mathbb{Q}(A_n). \tag{1}$$

Понятие меры является естественным обобщением понятий длины, площади и объёма, используемых в геометрии. Обращаясь к (1), уместно вспомнить, что ещё Аристотеля в философском плане интересовало соотношение между целым и частями. Если $\mathbb{Q}(\Omega) = 1$, то мера называется вероятностной (или просто *вероятностью*). Вероятностную меру обычно обозначают $\mathbb{P}$. В дискретном пространстве мера задаётся последовательностью (или набором) $(p_i)_{i \in I}$, где $I = \mathbb{N}$ (или $I = \{1, \ldots, r\}$, r — число элементов конечного Ω), $p_i \geqslant 0$ для $i \in I$ и $\sum_{i \in I} p_i = 1$. А именно, $\mathbb{P}(\{\omega_i\}) := p_i$, $i \in I$, тогда для любого $A \subset \Omega$ имеем $\mathbb{P}(A) = \sum_{i:\ \omega_i \in A} p_i$. Показывается, что так введённая функция $\mathbb{P}$ является мерой на σ-алгебре всех подмножеств дискретного пространства Ω. Когда Ω конечно и имеет мощность r, выбор $p_i = 1/r$, где $i = 1, \ldots, r$, приводит к «классическому определению вероятности» (все элементарные исходы равновероятны). Тройка объектов $(\Omega, \mathscr{F}, \mathbb{P})$ называется *вероятностным пространством*.

Основной интерес представляют функции, заданные на пространстве Ω. Например, важно знать не только то, что исход ω привёл к выигрышу в лотерее, но желательно представлять размер выигрыша, т. е. величину $X(\omega)$. Для построения общей теории рассматриваются функции, обладающие «хорошими свойствами.» *Случайная величина* X определяется как «измеримое отображение» Ω в $\mathbb{R}$, т. е. $\{\omega\colon X(\omega) \leqslant x\} \in \mathscr{F}$ для $x \in \mathbb{R}$. Существуют эквивалентные определения, например, $\{\omega\colon X(\omega) \in B\} \in \mathscr{F}$ для каждого «борелевского» подмножества прямой $\mathbb{B}$ (*борелевские множества* составляют минимальную σ-алгебру, содержащую все интервалы вида (a, b), где $-\infty < a \leqslant b < \infty$). Важно, что случайная величина индуцирует *распределение вероятностей* на σ-алгебре $\mathscr{B}(\mathbb{R})$ борелевских подмножеств прямой по формуле

$$\mathbb{P}_X(B) := \mathbb{P}(\omega\colon X(\omega) \in B), \quad B \in \mathscr{B}(\mathbb{R}).$$

Обычно пишут $\mathbb{P}(X \in B)$. Если Ω дискретно, то случайной величиной является любая функция $X : \Omega \to \mathbb{R}$. В таком случае распределение вероятностей $\mathbb{P}_X$ на подмножествах множества $\{x_i, i \in I\}$, состоящего из различных значений функции X, задаётся вектором $(p_i)_{i \in I}$, где $p_i := \mathbb{P}(X = x_i)$, $i \in I$. Случайные величины называются *одинаково распределёнными*, если их распределения совпадают на $\mathscr{B}(\mathbb{R})$. Для дискретных случайных величин X, Y это означает, что они принимают значения в одном и том же множестве $\{u_i, i \in I\}$, и при этом $\mathbb{P}(X = u_i) = \mathbb{P}(Y = u_i)$ для всех $i \in I$, где I конечно или счётно.

Говорят, что последовательность $X_1, X_2, \ldots$ состоит из *независимых* случайных величин, заданных на некотором вероятностном пространстве $(\Omega, \mathscr{F}, \mathbb{P})$, если для каждого $n \in \mathbb{N}$ и любых борелевских множеств $B_1, B_2, \ldots$ верно равенство

$$\mathbb{P}\Bigl(\bigcap_{j=1}^{n} \{X_j \in B_j\}\Bigr) = \prod_{j=1}^{n} \mathbb{P}(\{X_j \in B_j\}). \tag{2}$$

Определение, эквивалентное (2), получается, если в качестве B_j рассматривать произвольные полупрямые вида $B_j = (\infty, x_j]$, $x_j \in \mathbb{R}$, $j = 1, \ldots, n$. Понятие независимости отражает обычное предположение, что ни одна из рассматриваемых величин «не влияет на поведение других». Понятие независимости систем случайных величин сводится к независимости порождённых ими σ-алгебр. Независимые системы событий рассматриваются, например, в [6, § 8, гл. 1].

Важную роль играют числовые характеристики случайных величин. Если X принимает значение x_j с вероятностью p_j, где $j = 1, \dots, r$, то *математическое ожидание* X определяется формулой

$$\mathbb{E}X = \sum_{j=1}^{r} x_j p_j.$$

Представим, что в точке x_j сосредоточена масса p_j, $j = 1, \dots, r$, тогда $\mathbb{E}X$ — это координата центра тяжести данной системы. Поэтому $\mathbb{E}X$ также называют *средним значением*. Для величины X с бесконечным числом значений $\mathbb{E}X$ может принимать конечные или бесконечные значения, а может и не существовать. Когда число различных значений величины X счётно (составляет множество $\{x_1, x_2, \dots\}$), $\mathbb{E}X$ равно сумме ряда $\sum_{j=1}^{\infty} x_j p_j$, если этот ряд сходится абсолютно. Аналогично

$$\mathbb{E}h(X) = \sum_{j=1}^{\infty} h(x_j) p_j$$

при условии абсолютной сходимости ряда для измеримой функции $h\colon \mathbb{R} \to \mathbb{R}$ (в пространстве $\mathbb{R}$ берётся σ-алгебра $\mathscr{B}(\mathbb{R})$). Если величины X_k имеют распределение Бернулли с параметром p, т. е. $\mathbb{P}(X_k = 1) = p$ и $\mathbb{P}(X_k = 0) = 1 - p$, где $k \in \mathbb{N}$ и $p \in (0, 1)$, то $\mathbb{E}X_k = 1 \cdot p + 0 \cdot (1 - p) = p$. Для случайных величин, множество значений которых необязательно дискретно, математическое ожидание вводится как интеграл Лебега функции X по мере $\mathbb{P}$. Эта конструкция приводится практически в любом университетском учебнике по теории вероятностей.

Многие годы, начиная с Я. Бернулли (1665–1705), изучался «слабый» вариант закона больших чисел для последовательности независимых случайных величин $X_1, X_2, \dots$, имеющих конечное математическое ожидание. Речь шла об условиях, обеспечивающих для любого $\varepsilon > 0$ справедливость соотношения

$$\Delta_n(\varepsilon) := \mathbb{P}\left(\omega\colon \left|\frac{1}{n}\sum_{k=1}^{n} X_k(\omega) - \frac{1}{n}\sum_{k=1}^{n} \mathbb{E}X_k\right| > \varepsilon\right) \to 0, \quad n \to \infty. \tag{3}$$

Соотношение (3) показывает, что при больших n (и при выполнении некоторых условий) средние арифметические случайных величин $X_1, \dots, X_n$ приобретают как бы неслучайное поведение, близкое к поведению числовой последовательности $\frac{1}{n}\sum_{k=1}^{n} \mathbb{E}X_k$. Если в каждом опыте только фиксируется, произошло или не произошло некоторое событие, часто именуемое «успехом», то для описания полученных результатов используются величины $X_1, X_2, \ldots$, принимающие значения 1 или 0. Обычно значение 1 величины X_k показывает, что в k-м опыте наступил «успех». Если величины $X_1, X_2, \ldots$ имеют распределение Бернулли, то выполнение (3) в такой схеме означало бы должную близость к p частоты появления единицы в n опытах (частота «успеха» в n опытах будет $\frac{1}{n}\sum_{k=1}^{n} X_k$, при этом $\frac{1}{n}\sum_{k=1}^{n} \mathbb{E}X_k = p$). Последовательность случайных величин $Y_1, Y_2, \ldots$ называется сходящейся *по вероятности* к случайной величине Y, если $\mathbb{P}(|Y_n - Y| \geqslant \varepsilon) \to 0$ для каждого $\varepsilon > 0$ при $n \to \infty$. Соотношение (3) означает, что последовательность случайных величин

$$\frac{1}{n}\sum_{k=1}^{n} X_k - \frac{1}{n}\sum_{k=1}^{n} \mathbb{E}X_k$$

сходится к нулю по вероятности, когда $n \to \infty$. Подчеркнём, что (3) не обязано выполняться для каждой последовательности случайных величин $X_1, X_2, \ldots$.

Интересно, что при определённых условиях можно даже выяснить, сколь большие n следует брать для того, чтобы при заданных положительных ε и δ выполнялось неравенство $\Delta_n(\varepsilon) < \delta$. Это легко объяснить с помощью *неравенства Чебышёва*, которое называют также неравенством Бьенэме — Чебышёва. А именно, для любой случайной величины X, имеющей конечную дисперсию $\operatorname{var} X := \mathbb{E}(X - \mathbb{E}X)^2$, и каждого $\varepsilon > 0$

$$\mathbb{P}(|X - \mathbb{E}X| \geqslant \varepsilon) \leqslant \frac{\operatorname{var} X}{\varepsilon^2}. \tag{4}$$

Отметим, что (4) — следствие установленного позднее *неравенства Маркова*, которое утверждает, что для любой неубывающей неотрицательной функции f, заданной на $[0, \infty)$, произвольной случайной величины X и каждого $\varepsilon > 0$

$$f(\varepsilon)\mathbb{P}(|X| \geqslant \varepsilon) \leqslant \mathbb{E}f(|X|). \tag{5}$$

Это неравенство нетривиально, когда $\mathbb{E}f(|X|) < \infty$ и $f(\varepsilon) > 0$, и тогда можно написать, что $\mathbb{P}(|X| \geqslant \varepsilon) \leqslant \mathbb{E}f(|X|)/f(\varepsilon)$. Очевидно, если в (5) вместо X подставить $X - \mathbb{E}X$ и выбрать $f(x) = x^2$ для $x \geqslant 0$, то получится (4).

Если $X = \frac{1}{n}\sum_{k=1}^{n} X_k$, то из (4) вытекает (3), когда $X_1, X_2, \ldots$ независимы и $\operatorname{var} X_k \leqslant C$ при всех $k \in \mathbb{N}$, где C — положительная константа. Действительно, $\Delta_n(\varepsilon) \leqslant C/(n\varepsilon^2)$, так как дисперсия суммы независимых величин равна сумме их дисперсий и $\operatorname{var}(aX) = a^2 \operatorname{var} X$ для $a \in \mathbb{R}$. Если $|X_k(\omega)| \leqslant b$ для всех $\omega \in \Omega$ и некоторой константы $b > 0$, то $\operatorname{var} X_k \leqslant b^2$, где $k \in \mathbb{N}$. Следовательно, если модули независимых величин ограничены некоторой константой, то закон больших чисел (3) справедлив. Если независимые X_k имеют распределение Бернулли с $p = 1/2$, то $\operatorname{var} X_k = 1/4$ и $\Delta_n(0{,}01) \leqslant 0{,}01$ согласно (4) при $n \geqslant 250\,000$. Оказывается, эту оценку выбора n можно заметно улучшить. Например, в силу неравенства Хёфдинга (установленного в 1963 году) $\Delta_n(\varepsilon) \leqslant \leqslant 2e^{-2\varepsilon^2 n}$, откуда $\Delta_n(0{,}01) \leqslant 0{,}01$ при $n \geqslant 26\,500$.

Глубокий анализ соотношения (3) показывает, что величины под знаком модуля в (3), при определённых условиях, сходятся к нулю на множестве точек ω, составляющих событие вероятности единица. Тогда говорят, что выполняется «усиленный закон больших чисел». Сходимость случайных величин для «почти всех» $\omega \in \Omega$ называется *сходимостью почти наверное*, из неё вытекает сходимость по вероятности (обратное утверждение, вообще говоря, неверно). А. А. Марков не обращался к этой теме, его интересовал закон больших чисел в слабой форме (внесённый им вклад рассматривается в разделе 2.2), хотя первый результат в этом направлении получил Э. Борель (1871–1956) в 1909 году. Установленный А. Н. Колмогоровым (1903–1987) классический усиленный закон больших чисел утверждает, что

$$\frac{1}{n}\sum_{k=1}^{n} X_k \to a \in \mathbb{R}$$

с вероятностью единица при $n \to \infty$ для независимых одинаково распределённых величин $X_1, X_2, \ldots$ тогда и только тогда, когда существует $\mathbb{E}X_1$, и при этом $a = \mathbb{E}X_1$. Напомним, что случайные величины X и Y имеют одинаковое распределение, если для любого «хорошего» (т. е. *борелевского*) подмножества B прямой $\mathbb{P}(\omega: X(\omega) \in B) = \mathbb{P}(\omega: Y(\omega) \in B)$. Эквивалентным образом можно сказать, что $\mathbb{P}(X \leqslant x) = \mathbb{P}(Y \leqslant x)$ при любом $x \in \mathbb{R}$, т. е. совпадают *функции распределения* этих величин. А. Я. Хинчин (1894–1959) для последовательности независимых одинаково распределённых величин $X_1, X_2, \ldots$ получил критерий выполнения для каждого $\varepsilon > 0$ соотношения $\mathbb{P}(|S_n/n - c_n| > \varepsilon) \to 0$, когда $n \to \infty$, где $S_n := X_1 + \ldots + X_n$, $c_n \in \mathbb{R}$, $n \in \mathbb{N}$. Удивительно, что наличие математического ожидания у слагаемых не является необходимым условием справедливости этого соотношения для некоторой последовательности $(c_n)_{n \in \mathbb{N}}$, а необходимо и достаточно, чтобы $n\mathbb{P}(|X_1| \geqslant n) \to 0$, $n \to \infty$. При этом можно описать способ выбора c_n, $n \in \mathbb{N}$.

Исследование поведения средних арифметических последовательности случайных величин дало импульс созданию современной эргодической теории.

2.2. Исследование зависимых случайных величин

С начала XX века стали интенсивно разрабатываться модели, описываемые зависимыми величинами, т. е. не являющимися независимыми. Интерес Андрея Андреевича к последовательностям зависимых величин был связан с тем, что профессор П. А. Некрасов (1853–1924), ректор Московского университета с 1893 по 1898, президент Московского математического общества с 1903 по 1905, утверждал, что для справедливости закона больших чисел необходимо условие их независимости. На этой основе П. А. Некрасов пытался обосновать некоторые религиозно-философские взгляды о свободе воли, а А. А. Марков был решительным противником таких взглядов. А. А. Марков понял, что если конечны дисперсии величин $X_1, X_2, \ldots$, то в силу (4) слабый закон больших чисел верен без предположения независимости членов рассматриваемой последовательности, когда $\operatorname{var}(\sum_{k=1}^n X_k)/n^2 \to 0$ при $n \to \infty$.

Огромная заслуга А. А. Маркова заключается в том, что он ввёл и начал изучать в 1906 году (см. [7]) последовательности случайных величин $X_0, X_1, \ldots$, принимающих значения в некотором конечном (или счётном) множестве I, механизм зависимости которых допускает очень наглядное описание. Если величины $X_0, X_1, \ldots$ рассматривать как случайный процесс с «дискретным временем» $n = 0, 1, 2, \ldots$, можно сказать, что поведение X_n, где $n \in \mathbb{N}$, определяется не «всей предысторией» $X_0, \ldots, X_{n-1}$, а лишь величиной X_{n-1}. Однако ошибочно думать, что это выливается в рекуррентную формулу вида $X_n = f(X_{n-1})$, где f — некоторая неслучайная функция, $n \in \mathbb{N}$ (подробности ниже в этом разделе). В качестве аналогии можно напомнить, что в рамках механики Ньютона движение материальной точки (на которую не действуют силы трения) в каждый момент времени t определяется её мгновенными координатами $\vec{x}(t)$ и скоростью $\vec{v}(t)$, а не этими характеристиками во все предшествующие моменты времени. В современных обозначениях упомянутая зависимость последовательности случайных величин означает, что для каждого натурального n и любых $i, j, i_0, \ldots, i_{n-1} \in I$

$$\mathbb{P}(X_{n+1} = j \mid X_n = i,\ X_{n-1} = i_{n-1},\ \ldots,\ X_0 = i_0) = \mathbb{P}(X_{n+1} = j \mid X_n = i) \quad (6)$$

при $\mathbb{P}(X_n = i,\ X_{n-1} = i_{n-1},\ \ldots,\ X_0 = i_0) \neq 0$ (по определению условная вероятность события A при условии события B задаётся формулой $\mathbb{P}(A|B) := \mathbb{P}(A \cap B)/\mathbb{P}(B)$, когда $\mathbb{P}(B) \neq 0$). Формула (6) показывает, что значения, принимаемые случайными величинами X_k при $k = 0, \ldots, n-1$, могут быть исключены из рассмотрения. А. А. Марков назвал этот вид зависимости цепным. Термин «цепь Маркова» для такой последовательности $X_0, X_1, \ldots$ стал классическим, его ввёл в 1926 году С. Н. Бернштейн (1880–1968). Цепь называется однородной, если правая часть формулы (6) не меняется для всех n при фиксированных i, j. Для однородной цепи Маркова матрица $P=(p_{i,j})_{i,j\in I}$, составленная из вероятностей перехода $p_{i,j}:=\mathbb{P}(X_1=j\,|\,X_0=i)$ за единицу времени из состояния i в состояние j, является «стохастической», т. е.

$$p_{i,j}\geqslant 0 \text{ при всех } i, j\in I, \quad \sum_{j\in I} p_{i,j}=1 \text{ для каждого } i\in I. \tag{7}$$

Теперь в курсах теории случайных процессов доказывается следующее утверждение (см., например, [6, гл. VI, § 7, 8], где рассматриваются и неоднородные цепи). На некотором $(\Omega, \mathscr{F}, \mathbb{P})$ можно построить случайные величины $X_0, X_1, \ldots$, образующие однородную цепь Маркова такую, что данная стохастическая матрица P будет матрицей переходных вероятностей, а любой набор неотрицательных чисел $p_i(0)$, для которого $\sum_{i\in I} p_i(0)=1$, определяет начальное распределение, т. е. $\mathbb{P}(X_0=i)=p_i(0)$, $i\in I$. Важно подчеркнуть, что до появления фундаментальной монографии А. Н. Колмогорова [8], первое издание которой опубликовано в 1933 году, у исследователей не было возможности утверждать, что на некотором вероятностном пространстве существует семейство случайных величин с заданной структурой зависимости, в частности, с зависимостью марковского типа. Парадоксальность ситуации заключалась в том, что математиками до появления этой книги изучались свойства объектов, существование которых не было доказано. Точнее говоря, исследователи не задумывались над тем, удастся ли определить на некотором вероятностном пространстве $(\Omega, \mathscr{F}, \mathbb{P})$ семейство изучаемых случайных величин так, чтобы оно обладало некоторыми предполагаемыми свойствами (например, чтобы величины оказывались независимыми и при этом имели требуемые распределения вероятностей). Заметим, что теорема Ломницкого — Улама (см., например, теорему 2 гл. I в [6]), опубликованная в 1934 году, дала возможность вводить любые семейства независимых случайных величин (даже со значениями в произвольных измеримых пространствах), имеющих заданные распределения вероятностей.

По-видимому, название «цепь» связано с тем, что при сборке реальной металлической цепи осуществляется присоединение каждого последующего звена к предыдущему. А. А. Марков рассматривал и «сложные цепи», для которых поведение X_n определялось величинами $X_{n-m}, \ldots, X_{n-1}$, где натуральное число $m>1$.

А. А. Марков описал последовательность зависимых величин, образующих цепь с двумя состояниями, для которой выполняется слабый закон больших чисел.

Поразительно, что идея цепной зависимости возникла у А. А. Маркова при анализе чередования гласных и согласных звуков в романе А. С. Пушкина «Евгений Онегин». Андрей Андреевич вручную проанализировал первые 20 000 букв этого знаменитого романа (исключив из рассмотрения пропуски между словами, знаки препинания и т. п.) и подсчитал частоту гласных в данной выборке, частоты появления гласной после согласной, а также согласной после гласной. Он интерпретировал эти результаты как наблюдения значений набора случайных величин с цепной зависимостью. Андрей Андреевич также заметил, что обработка первых 100 000 знаков повести С. Т. Аксакова «Детские годы Багрова внука» даёт иные частоты упомянутых выше комбинаций гласных и согласных, чем в романе А. С. Пушкина. Известный революционер и исследователь Н. А. Морозов (1854–1946) увидел в этом замечательную возможность математическими методами идентифицировать авторство определённого текста и в 1915 году опубликовал работу «Лингвистические спектры». В 1916 году была напечатана статья А. А. Маркова [9], в которой он показал неубедительность выводов Н. А. Морозова (хотя и признал полезность математического подхода к таким задачам с привлечением более сложных моделей и увеличением объёмов выборок). Исследование литературных текстов математическими методами рассматривается в книге [10, § 5].

В начале прошлого века возник иной пример, оказавшийся цепью Маркова. Р. Росс (1857–1932) в 1902 году был удостоен Нобелевской премии за изучение распространения малярии москитами. Позднее он пытался построить математические модели перемещения москитов, чтобы предсказать плотность их популяции в районах заболевания малярией. В 1904 году Р. Росс написал письмо К. Пирсону — одному из создателей современной математической статистики, с просьбой помочь в этом деле. К. Пирсон отказался, сославшись на перегруженность работой, однако в 1905 написал статью в журнал «Nature», в которой ввёл стохастическую модель, связанную со случайными перемещениями в пространстве, под названием *случайное блуждание* (random walk). Точнее говоря, К. Пирсон описал модель и предложил читателям решить определённую задачу в рамках этой модели, но не упомянул письмо Р. Росса. В частности, модель случайного блуждания по целым точкам прямой такова: в момент времени n частица занимает положение $X_n = X_{n-1} + U_n$, $n \in \mathbb{N}$, где величины $X_0, U_1, U_2, \ldots$

независимы, X_0 (начальное положение) — случайная величина со значениями в множестве $\mathbb{Z}$ всех целых чисел прямой, а $\mathbb{P}(U_n = 1) = p$ и $\mathbb{P}(U_n = -1) = 1 - p$ для некоторого $p \in (0, 1)$. Другими словами, частица в момент $n \in \mathbb{N}$ смещается из положения X_{n-1} на $\mathbb{Z}$ с вероятностью p на единицу вправо или с вероятностью $1 - p$ на единицу влево, а в начальный момент частица находилась в случайной точке X_0. Такое случайное блуждание является цепью Маркова.

Последователи А. А. Маркова установили, что однородные цепи X_0, $X_1, \ldots$, в которых все величины X_j принимают значения в конечном или счётном множестве I, допускают полное описание. А именно, $X_n = f(X_{n-1}, U_n)$, где f — неслучайная функция двух переменных, семейство $X_0, U_1, U_2, \ldots$ состоит из независимых случайных величин, величины $U_1, U_2, \ldots$ одинаково распределены, $n \in \mathbb{N}$. Разумеется, соответствующая функция f принимает значения в множестве I, а при подстановке в f аргументов X_{n-1} и U_n должна получаться случайная величина. Мы не касаемся более общих формулировок, см., например, раздел 6.1 книги [11]. Таким образом, цепная зависимость (6) означает не только функциональную зависимость X_n от X_{n-1}, но и внесение на каждом шаге n некоторого случайного «обновления» U_n. В частности, упомянутое выше случайное блуждание получается при выборе $f(x, y) = x + y$, $x, y \in \mathbb{Z}$.

Интересный пример марковской цепи в 1907 году рассмотрели Пауль Эренфест и Татьяна Эренфест для объяснения второго закона термодинамики. Они предложили простую модель диффузии N частиц в параллелепипеде, разделённом проницаемой мембраной на две равные части A и B. Если обозначить $X(t)$ число частиц в контейнере A в момент $t = 0, 1, 2, \ldots$, то в данной модели $p_{i,i-1} = iq/N$ для $i = 1, \ldots, N$ (с вероятностью i/N в контейнере A выбирается одна из имеющихся i частиц, а затем с вероятностью q она перемещается в контейнер B), $p_{i,i+1} = (N - i)q/N$ для $i = 0, \ldots, N - 1$, $p_{i,i} = 1 - q$ при $i = 0, \ldots, N$ и $p_{i,j} = 0$ для остальных $i, j \in \{0, \ldots, N\}$, где $q \in (0, 1]$. Если в момент $t = 0$ в контейнере A было N частиц, то что можно сказать о числе частиц в этом контейнере, когда $t \to \infty$? Оказывается, в пределе для каждого $q \in (0, 1]$ вероятность иметь n частиц в контейнере A будет $2^{-N} C_N^n$, где

$$C_N^n := \frac{N!}{n!(N-n)!}, \quad N \in \mathbb{N}, \quad n = 0, \ldots, N$$

(напомним, что $n! := 1 \cdot \ldots \cdot n$ для $n \in \mathbb{N}$ и при этом $0! := 1$). Различные обобщения и приложения данной модели рассматриваются в статье [12].

В XX веке была развита общая теория случайных процессов и изучен класс марковских процессов $\{X_t, t \geqslant 0\}$ с непрерывным временем t, охватывающий цепи Маркова. Значения случайных величин X_t при каждом t уже не обязательно лежат в дискретном множестве. В 1900 году в диссертации Л. Башелье (1870–1946), в связи с изучением колебаний на бирже курсов ценных бумаг, был введён случайный процесс, который теперь называется винеровским процессом (или броуновским движением). В физике к процессу такого рода привело исследование диффузии, проведённое в 1905 году А. Эйнштейном (1879–1955) и М. Смолуховским (1872–1917). Позднее математически корректную модель этого процесса предложил Н. Винер (1894–1964), в дальнейшее изучение важный вклад внёс П. Леви (1886–1971). *Броуновское движение* и так называемый *пуассоновский процесс* являются важными примерами марковских процессов с непрерывным временем, им посвящены отдельные монографии. С помощью таких процессов и некоторых других удалось построить теорию (стохастического) интегрирования и теорию стохастических дифференциальных уравнений. Мы не даём определения общих «забывающих предысторию» марковских процессов с недискретным пространством состояний и непрерывным временем, чтобы не усложнять изложение привлечением аппарата условных математических ожиданий. В этой связи упомянем труды Е. Б. Дынкина (1924–2014). Построение любых марковских процессов по «переходной функции» и начальному распределению ещё не исследовано в полной общности.

2.3. Некоторые приложения цепей Маркова

Французский естествоиспытатель Ж.-Л. Леклерк, получивший титул графа де Бюффона (1707–1788), автор «Естественной истории» в 36 томах, положил начало методу статистического моделирования. В 1777 году он рассмотрел задачу о бросании иглы, имеющей длину l, на плоскость, расчерченную параллельными прямыми так, что расстояние между любыми соседними из них равно a, где $a > l$. Спрашивается, какова вероятность, что игла пересечёт одну из проведённых прямых. Вначале вводится модель, задающая механизм случайного бросания иглы. Мы не будем обсуждать эту классическую задачу. Приведём лишь ответ. Искомая вероятность $p(a, l) = 2l/(a\pi)$, где знаменитая константа $\pi = 3{,}1415\ldots$ Пусть в n независимых одинаковых опытах игла пересекла выделенные прямые m раз. Согласно закону больших чисел (даже усиленному), при больших n частота m/n будет близка к вероятности $p(a, l)$. Поэтому соотношение $m/n \approx 2l/(a\pi)$

при заданных a, l и всех достаточно больших n позволяет написать $\pi \approx 2ln/(am)$. Таким образом, с помощью случайных экспериментов оказалось возможным находить приближения для некоторых детерминированных величин.

В настоящее время во многих теоретических и прикладных задачах возникает необходимость приближённых вычислений сложных интегралов или сумм с очень большим числом слагаемых. Часто аналитические методы не помогают. Рассмотрим нахождение математического ожидания $\mathbb{E}\varphi(Z)$, где Z — случайная величина со значениями в множестве $V = \{x_1, \ldots, x_r\}$, $\varphi : V \to \mathbb{R}$. Классический метод Монте-Карло означает построение последовательности независимых одинаково распределённых случайных величин $Z_1, Z_2, \ldots$, имеющих то же распределение, что Z (рулетка в известном казино Монте-Карло может рассматриваться как простой датчик случайных чисел, что и объясняет происхождение названия данного метода). Пусть $p_j := \mathbb{P}(Z = v_j)$, $j = 1, \ldots, r$. Тогда, согласно усиленному закону больших чисел Колмогорова, с вероятностью единица

$$\frac{1}{n}\sum_{k=1}^{n}\varphi(Z_k) \to \mathbb{E}\varphi(Z) = \sum_{j=1}^{r}\varphi(x_j)p_j, \quad n \to \infty. \tag{8}$$

Поэтому, если z_k — значение случайной величины Z_k в k-м опыте, то используют приближённое равенство

$$\frac{1}{n}\sum_{k=1}^{n}\varphi(z_k) \approx \mathbb{E}\varphi(Z),$$

когда n достаточно велико.

Реализовать подход, основанный на (8), не получается, если p_j известны лишь с точностью до постоянного множителя (так как требуется знать все p_j, где $j = 1, \ldots, r$, чтобы построить величины $Z_1, \ldots, Z_n$). Так обычно бывает в сложных физических задачах. Разумеется, если $p_j = c\tilde{p}_j$, где $\tilde{p}_j$ известны для всех $j \in \{1, \ldots, r\}$, а константа c неизвестна, то

$$c = \frac{1}{\sum_{j=1}^{r}\tilde{p}_j}.$$

Однако $\tilde{p}_j$ могут иметь весьма сложный вид, а r быть огромным числом. Поэтому такой способ нахождения c не работает. Выход даёт следующая замечательная идея: вместо независимых величин $Z_1, Z_2, \ldots$ построить «хорошую» однородную марковскую цепь $X_0, X_1, \ldots$ со значениями в V такую, что соотношение (8) будет верно, если в него вместо Z_j подставить X_j, $j \in \mathbb{N}$. Далее отождествляем v_j с j и пишем $I = \{1, \ldots, r\}$ вместо V. Важно, что упомянутую цепь удаётся построить, когда величины p_j известны лишь с точностью до умножения на некоторую константу.

Рассмотрим цепь Маркова, для которой строка $\tau = (p_1, \ldots, p_r)$ задаёт инвариантное распределение вероятностей на I, т. е. сумма неотрицательных $p_1, \ldots, p_r$ равна 1 и верно равенство $\tau P = \tau$, где P — некоторая матрица переходных вероятностей. Иначе говоря,

$$(\tau P)_j = \sum_{i \in I} p_i p_{i,j} = p_j, \quad j \in I.$$

При этом начальное распределение марковской цепи можно взять произвольным (например, в начальный момент времени $X_0 = i_0$ с вероятностью единица, где $i_0 \in I$). Матрица $P = (p_{i,j})$ выбирается удовлетворяющей условию

$$p_i p_{i,j} = p_j p_{j,i}, \quad i, j \in I. \tag{9}$$

При выполнении равенства (9) набор p_j, $j \in I$, задаёт инвариантное распределение вероятностей, поскольку

$$(\tau P)_j = \sum_{i \in I} p_i p_{i,j} = \sum_{i \in I} p_j p_{j,i} = p_j \sum_{i \in I} p_{j,i} = p_j$$

в силу (7).

Матрицу P, удовлетворяющую соотношению (9), можно выбирать многими способами. Широкую известность получил алгоритм Метрополиса, предложенный в 1953 году коллективом авторов, когда вначале берётся матрица $Q = (q_{i,j})$ переходных вероятностей, отвечающая неразложимой и апериодической цепи с пространством состояний $I = \{1, \ldots, r\}$. Неразложимость цепи означает, что для каждой пары состояний i, j найдутся целые неотрицательные числа m и n (вообще говоря, зависящие от i, j) такие, что $q_{i,j}(m) > 0$ и $q_{j,i}(n) > 0$, где $q_{i,j}(k) = (Q^k)_{i,j}$ представляет собой вероятность перехода из состояния i в состояние j за k шагов. Апериодичность состояния i означает, что наибольший общий делитель всех натуральных k, для которых $q_{i,i}(k) > 0$, равен единице. Апериодичность цепи означает апериодичность всех её состояний.

Эти понятия не были до конца разработаны А. А. Марковым (неприводимость он трактовал как невозможность определёнными преобразованиями привести матрицу переходных вероятностей Q к специальному блочному виду). Можно сказать, что для цепи с конечным пространством состояний I неприводимость и апериодичность эквивалентны тому, что для некоторого $m \in \mathbb{N}$ матрица Q^m (т. е. m-я степень матрицы Q) содержит только строго положительные элементы. Достаточно (при $m = 1$) взять стохастическую матрицу Q, все элементы которой строго больше нуля. Затем для $i \neq j$ $(i, j \in I)$ полагают

$$p_{i,j} = q_{i,j}\alpha_{i,j}, \quad \text{где} \quad \alpha_{i,j} = \min\left\{1, \frac{p_j q_{j,i}}{p_i q_{i,j}}\right\}, \tag{10}$$

$p_{i,i} = 1 - \sum_{j \neq i} p_{i,j}$ (считается, что $p_i > 0$ для $i \in I$). Легко проверить, что тогда справедливо (9), отдельно рассмотрев случаи

$$\frac{p_j q_{j,i}}{p_i q_{i,j}} \geqslant 1 \quad \text{и} \quad \frac{p_j q_{j,i}}{p_i q_{i,j}} < 1.$$

Подчеркнём, что этот алгоритм применим и тогда, когда p_i известны лишь с точностью до постоянного множителя, например, $p_i = e^{-U(i)}/K$, где функция $U : I \to \mathbb{R}$, а $K = \sum_{i \in I} e^{-U(i)}$ не поддаётся подсчёту ввиду очень большого числа элементов множества I. В этом случае p_j/p_i для $i, j \in I$ не зависит от величины K. Поэтому матрица P находится согласно (10), когда константа K неизвестна.

Почему было важно ввести однородную марковскую цепь с инвариантным распределением, задаваемым набором вероятностей τ? Положим $p_{i,j}(n) := \mathbb{P}(X_n = j \mid X_0 = i)$, $i, j \in I = \{1, \ldots, r\}$, $n \in \mathbb{N}$. Легко видеть, что $p_{i,j}(n)$ являются элементами матрицы P^n. При широких условиях $\lim_{n\to\infty} p_{i,j}(n) = p_j^*$ для любых $i, j \in I$, причём предел не зависит от выбора i (необходимо и достаточно, чтобы для некоторого $m \in \mathbb{N}$ все элементы матрицы P^m были строго положительны). Марковская цепь за большое время n как бы «забывает», из какого состояния i она стартовала. Сходимость $p_{i,j}(n)$ к p_j^* в нашем случае будет экспоненциально быстрой с ростом n, т. е. $|p_{i,j}(n) - p_j^*| \leqslant A\gamma^n$ для $i, j \in I$, где $A > 0$ и $0 < \gamma < 1$. Предельные значения $p_1^*, \ldots, p_r^*$ задают распределение вероятностей, именуемое *стационарным*, и $\mathbb{P}(X_n = j) \to p_j^*$ (экспоненциально быстро) при $n \to \infty$ для всех $j \in I$ при любом начальном распределении цепи. Теперь следует заметить, что для построенной цепи инвариантное распределение будет стационарным и $p_j^* = p_j$, $j \in I$. Таким образом, при больших k распределение X_k будет близко к распределению, задаваемому набором τ, т. е. к распределению Z_k. Это поясняет (нетрудно дать детальное доказательство), почему при любом начальном распределении в формулу (8) можно подставлять $X_1, X_2, \ldots$ вместо $Z_1, Z_2, \ldots$ и тем самым использовать приближённое равенство

$$\frac{1}{n}\sum_{k=1}^{n} \varphi(x_k) \approx \mathbb{E}\varphi(Z),$$

когда n достаточно велико, а x_k является значением случайной величины X_k, $k = 1, 2, \ldots, n$ (мы оперируем значениями $x_1, x_2, \ldots$, полученными как $X_1(\omega), X_2(\omega), \ldots$ для некоторого $\omega \in \Omega$ при наблюдении марковской цепи $X_1, X_2, \ldots$). Также достаточно сослаться на следующий вариант усиленного закона больших чисел для однородной марковской цепи $X_0, X_1, \ldots$, где X_n принимает значения в счётном (или конечном) множестве I при $n = 0, 1 \ldots$. Если цепь неприводима и имеет стационарное распределение, задаваемое последовательностью (или набором) $\tau = (\tau_i)_{i \in I}$, а функция $\varphi : I \to \mathbb{R}$ такова, что $\sum_{i \in I} |\varphi(i)| \tau_i < \infty$, то при любом начальном распределении

$$\frac{1}{n} \sum_{k=1}^{n} \varphi(X_k) \to \sum_{i \in I} \varphi(i) \tau_i$$

почти наверное при $n \to \infty$ (см. [13, с. 313]).

Заметим, что мы лишь затронули обширную область исследований, в которой имеется целый ряд известных алгоритмов выбора матрицы P, см., например, [14]. Данная область оказалась тесно связанной с математической (байесовской) статистикой. Об этом можно прочитать в [15].

Важные для приложений модели описываются с помощью марковских цепей с непрерывным временем (и конечным или счётным пространством состояний). Так, в § 19 главы VI из [6] даётся вывод классических формул Эрланга, установленных при анализе функционирования телефонных станций. Цепи Маркова используются и в биологии, например, известна модель Джукса — Кантора стохастической эволюции молекулы ДНК, в которой состояниями являются нуклеотиды A, C, T, G. Применения марковских цепей в стохастических моделях биологии — отдельная обширная область исследований.

Идея А. А. Маркова исследовать семейства цепным образом зависимых случайных величин оказалась очень плодотворной и получила дальнейшее развитие в теории случайных полей (случайным полем называется семейство случайных элементов $\{X_t, t \in T\}$, заданных на некотором вероятностном пространстве, где T — множество индексов). Если T — конечный граф и при каждом $t \in T$ величина X_t принимает значения в конечном множестве I, то понятие марковости может быть введено, в определённом смысле, аналогично формуле (6), а именно, требуется, чтобы для каждого $t \in T$ и любых $i_s \in I$, $s \in T$, выполнялось равенство

$$\mathbb{P}\big(X(t) = i_t \mid X(s) = i_s,\ s \in T \setminus \{t\}\big) = \mathbb{P}\big(X(t) \mid X(s),\ s \in N(t)\big),$$

если $\mathbb{P}(X(s) = i_s, s \in T \setminus \{t\}) \neq 0$, а $N(t)$ — множество соседей вершины t (т. е. тех точек s из T, которые соединены ребром с t). Имеются и другие обобщения понятия марковости.

В статистической физике широко используется модель Изинга, возникшая при изучении ферромагнетиков и связанная с гиббсовскими случайными полями. Если гиббсовское случайное поле задано на конечном множестве T и при каждом t величины X_t принимают значения в конечном множестве I, то

$$\mathbb{P}\Big(\bigcap_{t\in T}\{X_t = x_t\}\Big) = C\exp(-E(x_T)),$$

где «конфигурация» $x_T = (x_t)_{t\in T}$, $x_t \in I$ для $t \in T$, нормировочная константа $C > 0$, а действительная функция E имеет смысл энергии. Согласно теореме Аверинцева — Клиффорда — Хаммерсли марковское случайное поле, определённое на конечном графе T и имеющее строго положительное распределение вероятностей на пространстве конфигураций, это такое и только такое поле, которое является гиббсовским с потенциалом взаимодействия «ближайших соседей» (другими словами, энергия E выражается через потенциал специального вида). Введение в теорию случайных полей, включая доказательство упомянутой выше теоремы, содержится в [16].

Во второй половине XX века возник класс полумарковских процессов. Кроме того, важную роль для разнообразных приложений (в том числе для построения алгоритмов распознавания речи) играют скрытые марковские процессы, введённые в 1989 году Л. Р. Рабинером (род. в 1943), см., например, [17]. В XXI веке развивается теория так называемых «нелинейных марковских процессов». В дискретном случае основная идея заключается в том, что вероятность перехода из состояния i, в котором процесс находится в момент s, в состояние j в момент $s+1$ зависит не только от i, но и от распределения процесса в момент s. Этой тематике посвящена монография [18]. Оказалось, что процессы такого рода полезны для описания ряда сложных физических явлений. Они применяются, например, в теории плазмы. Понятие марковской цепи получило нетривиальное отражение в современной квантовой теории информации.

2.4. Асимптотическая нормальность

А. А. Марков внёс также существенный вклад в доказательство центральной предельной теоремы (так по предложению Д. Пойа (1887–1985) с 1920 года стали называть группу результатов о сходимости функций распределения нормированных сумм случайных величин к функции распределения гауссовского закона). Говорят, что последовательность случайных величин $Y_1, Y_2, \ldots$ сходится *по распределению* к случайной величине Y, если функции распределения Y_n сходятся к функции распределения Y во всех точках непрерывности предельной функции при $n \to \infty$. Тогда пишут $Y_n \xrightarrow{\mathscr{D}} Y$, $n \to \infty$. Напомним, что функция распределения случайной величины W задаётся формулой $F_W(x) := \mathbb{P}(W \leqslant x)$, $x \in \mathbb{R}$. Величина W имеет гауссовское распределение с параметрами $a \in \mathbb{R}$ и $\sigma^2 > 0$, если

$$F_W(x) = \frac{1}{\sigma\sqrt{2\pi}} \int\limits_{-\infty}^{x} \exp\Bigl(-\frac{(u-a)^2}{2\sigma^2}\Bigr)\, du, \quad x \in \mathbb{R}.$$

Используется также термин нормальное распределение, пишут $W \sim \sim N(a, \sigma^2)$. Пусть $X_1, X_2, \ldots$ — последовательность независимых случайных величин таких, что $\mathbb{E}X_k^2 < \infty$, $k \in \mathbb{N}$. Положим

$$S_n := X_1 + \ldots + X_n, \quad T_n := \frac{S_n - \mathbb{E}S_n}{\sqrt{\operatorname{var} S_n}},$$

где $\operatorname{var} S_n \neq 0$, $n \in \mathbb{N}$. Классическая центральная предельная теорема утверждает, что при определённых условиях $T_n \xrightarrow{\mathscr{D}} Z$, т. е.

$$F_{T_n}(x) \to \frac{1}{\sqrt{2\pi}} \int\limits_{-\infty}^{x} e^{-u^2/2}\, du = F_Z(x), \quad x \in \mathbb{R},\ n \to \infty, \tag{11}$$

где случайная величина $Z \sim N(0, 1)$. Поскольку функция распределения величины Z непрерывна в каждой точке $x \in \mathbb{R}$, указанная сходимость будет даже равномерной на всей действительной оси. Отметим, что А. А. Марков разработал метод подсчёта интегралов, фигурирующих в (11), с точностью до 16 знаков после запятой.

П. Л. Чебышёв (1821–1894) считал, что для доказательства центральной предельной теоремы достаточно убедиться, что для каждого натурального m имеет место соотношение $\mathbb{E}(T_n^m) \to \mathbb{E}(Z^m)$ при $n \to \infty$. Однако этот метод моментов, введённый П. Л. Чебышёвым, не был обоснован достаточно строго. А. А. Марков сумел закрыть пробелы в доказательстве своего учителя. А. М. Ляпунов (1857–1918) развил метод характеристических функций (аппарат преобразования Фурье) и получил простые достаточные условия выполнения центральной предельной теоремы без предположения конечности моментов любого порядка у независимых слагаемых. Им было доказано, что если

$X_1, X_2, \ldots$ являются независимыми величинами такими, что $\mathbb{E}|X_k|^s < \infty$ при некотором $s \in (2, 3]$ и всех $k \in \mathbb{N}$, то (11) выполнено, когда

$$L_n(s) := \frac{\sum_{k=1}^{n} \mathbb{E}|X_k|^s}{\left(\sum_{k=1}^{n} \operatorname{var} X_k\right)^{s/2}} \to 0, \quad n \to \infty. \tag{12}$$

Для случайной величины X с конечным числом различных значений, если $\mathbb{P}(X = x_k) = p_k$, где $k = 1, \ldots, r$, то $\mathbb{E}|X|^s = \sum_{j=1}^{r} |x_j|^s p_j < \infty$ при каждом $s > 0$.

У А. А. Маркова несколько лет работы ушло на то, чтобы предложить модификацию метода моментов, не предполагающую наличия всех конечных моментов у рассматриваемых величин (Андрей Андреевич писал о «методе математических ожиданий»). Для этого он ввёл новый приём «урезания» исходных величин, который изложил в своей книге [2]. Он писал: «С одной стороны, я нашёл, как надо дополнить метод математических ожиданий, чтобы охватить все случаи А. М. Ляпунова; а, с другой стороны, ряд моих работ показал, что тот же метод даёт довольно лёгкое средство для распространения предельной теоремы на связанные величины.» В серии своих работ А. А. Марков, используя аппарат производящих функций, доказал (при определённых условиях) центральную предельную теорему для введённых им цепей. Отметим, что он рассмотрел и неоднородные цепи, принимающие значения 0 или 1.

Более общее достаточное условие, чем (12), для справедливости центральной предельной теоремы было найдено Я. Линдебергом (1876–1932). Он рассмотрел последовательность независимых случайных величин $X_1, X_2, \ldots$, имеющих нулевое среднее и конечные вторые моменты (не предполагая, что $\mathbb{E}|X_k|^s < \infty$ для некоторого $s > 2$ и $k \in \mathbb{N}$). Теорема Линдеберга утверждает, что нормированные суммы T_n сходятся по распределению к $Z \sim N(0, 1)$, если

$$\frac{1}{\operatorname{var} S_n} \sum_{k=1}^{n} \mathbb{E}\left(X_k^2 \mathbb{I}\left\{|X_k| \geqslant \varepsilon \sqrt{\operatorname{var} S_n}\right\}\right) \to 0 \quad \text{при } n \to \infty$$

для каждого $\varepsilon > 0$. Здесь $\operatorname{var} S_n \neq 0$ при всех достаточно больших n, а индикаторная функция $\mathbb{I}\{A\}$ равна 1, когда $\omega \in A$, и равна нулю в противном случае. В. Феллер (1906–1970) доказал, что для рассматриваемой последовательности $X_1, X_2, \ldots$ условие Линдеберга является необходимым, если слагаемые «асимптотически малы» в том смысле, что

$$\frac{1}{\operatorname{var} S_n} \cdot \max_{k=1,\ldots,n} \operatorname{var} X_k \to 0, \quad n \to \infty. \tag{13}$$

Необходимые и достаточные условия справедливости центральной предельной теоремы для серий независимых случайных величин $X_{n,1}, \dots, X_{n,m_n}$, где n и m_n — натуральные числа, были получены (без дополнительных условий типа (13)) лишь во второй половине XX века в работах В. М. Золотарева (1931–2019) и В. И. Ротаря (род. в 1942). А именно, если для указанных величин

$$\mathbb{E}X_{n,k} = 0, \quad k = 1, \dots, m_n, \quad \sum_{k=1}^{m_n} \operatorname{var} X_{n,k} = 1, \tag{14}$$

то $S_n \xrightarrow{\mathscr{D}} Z$, где $S_n := \sum_{j=1}^{m_n} X_{n,j}$ (тогда $\mathbb{E}S_n = 0$ и $\operatorname{var} S_n = 1$) и $Z \sim N(0,1)$, в том и только том случае, когда при каждом $\varepsilon > 0$

$$\sum_{k=1}^{m_n} \int\limits_{|x|>\varepsilon} |x||F_{n,k}(x) - \Phi_{n,k}(x)|\,dx \to 0, \quad n \to \infty. \tag{15}$$

Здесь $F_{n,k}$ и $\Phi_{n,k}$ обозначают соответственно функции распределения величин $X_{n,k}$ и $Z_{n,k}$, причём $Z_{n,k} \sim N(0, \operatorname{var} X_{n,k})$, $k = 1, \dots, m_n$, а интеграл берётся по мере Лебега dx.

Отметим, что изучение серий случайных величин носит более общий характер, чем исследование последовательностей. Нормированные частные суммы $(S_n - \mathbb{E}S_n)/\sqrt{\operatorname{var} S_n}$ для последовательности случайных величин $X_1, X_2, \dots$ легко представить в виде $\sum_{j=1}^{n} X_{n,j}$, где

$$X_{n,j} := \frac{X_j - \mathbb{E}X_j}{\sqrt{\operatorname{var} S_n}}, \quad j = 1, \dots, n, \quad m_n = n, \quad n \in \mathbb{N}.$$

Теорема Линдеберга легко переформулируется для схемы серий независимых случайных величин $X_{n,1}, \dots, X_{n,m_n}$ таких, что выполнено (14). Соотношение

$$\sum_{k=1}^{m_n} X_{n,k} \xrightarrow{\mathscr{D}} Z \sim N(0,1)$$

верно, если

$$\mathscr{L}_n(\varepsilon) := \sum_{k=1}^{m_n} \mathbb{E}(X_{n,k}^2 \mathbb{I}\{|X_{n,k}| \geqslant \varepsilon\}) \to 0 \tag{16}$$

для каждого $\varepsilon > 0$ при $n \to \infty$. Заметим, что из условия Линдеберга (16) вытекает, что $\max_{k=1,\dots,m_n} \operatorname{var} X_{n,k} \to 0$ при $n \to \infty$, когда все $X_{n,k}$ центрированы. Наглядно смысл условия (15) заключается в следующем. В отличие от (16), оно означает (при выполнении условия (14)), что если слагаемые не являются «асимптотически малыми», то их распределениям надлежит быть, в должном смысле, близкими к определённым гауссовским законам.

Центральную предельную теорему можно рассматривать как некоторое уточнение закона больших чисел, приведённого в разделе 2.1. Пусть даны независимые одинаково распределённые случайные величины $X_1, X_2, \ldots$, для которых $\mathbb{E}|X_1| < \infty$. Тогда по усиленному закону больших чисел Колмогорова

$$\delta_n(\omega) := \frac{1}{n}\sum_{k=1}^{n} X_k(\omega) - \mathbb{E}X_1 \to 0 \quad \text{при } n \to \infty$$

для $\omega \in \Omega_*$, где $\mathbb{P}(\Omega_*) = 1$. Если предположить, что $\operatorname{var} X_1^2 = \sigma^2 \in (0, \infty)$, то для независимых одинаково распределённых случайных величин с конечной и ненулевой дисперсией получаем по центральной предельной теореме Леви (вытекающей из теоремы Линдеберга), что величины $\delta_n\sqrt{n}/\sigma$ сходятся по распределению к $Z \sim N(0,1)$, откуда при любых $-\infty < a < b < \infty$ имеем

$$\mathbb{P}\left(\frac{\sigma}{\sqrt{n}}a \leqslant \delta_n \leqslant \frac{\sigma}{\sqrt{n}}b\right) \to \frac{1}{\sqrt{2\pi}}\int_a^b e^{-x^2/2}\,dx, \quad n \to \infty. \tag{17}$$

Для каждого $\varepsilon \in (0,1)$ можно найти $b = b(\varepsilon) > 0$ такое, что

$$\frac{1}{\sqrt{2\pi}}\int_{-b}^{b} e^{-x^2/2}\,dx = 1 - \varepsilon.$$

Затем по заданному числу δ выбираем $N_0 = N_0(\varepsilon, \delta) \in \mathbb{N}$ так, что $\sigma b/\sqrt{N_0} < \delta$. Тогда $|\delta_n| \leqslant \delta$ с вероятностью, близкой к $1-\varepsilon$, при $n \geqslant N_0$.

С 40-х годов XX века появился целый ряд работ, посвящённых оценке скорости сходимости в соотношении (11). Поэтому ныне имеется возможность оценивать точность аппроксимации с помощью гауссовского закона в соотношении (17). В частности, доказано, что для величин, фигурирующих в (11), таких, что $\mathbb{E}|X_k|^s < \infty$ для некоторого $s \in (2,3]$ и $k \in \mathbb{N}$, при любом $x \in \mathbb{R}$ верно неравенство

$$|\mathbb{P}(T_n \leqslant x) - \mathbb{P}(Z \leqslant x)| \leqslant c_s L_n(s), \tag{18}$$

где L_n введено в (12), c_s — положительное число. И. Г. Шевцова в [19] привела результаты, согласно которым $c_3 \leqslant 0{,}4693$ для независимых одинаково распределённых величин, имеющих $\mathbb{E}|X_1|^3 < \infty$. Кроме того, известно, что константа c_3 не может быть меньше, чем 0,4090. В настоящее время не найдено наименьшее возможное c_s в неравенстве (18). Если независимые $X_1, X_2, \ldots$ имеют распределение Бернулли с вероятностью успеха $1/2$, то теперь мы можем утверждать, что

$\Delta_n(0{,}01) < 0{,}01$ при $n \geqslant 16\,900$ (Δ_n введено в (3)), поскольку для таких величин величина погрешности приближения (17) не превосходит $1/\sqrt{2\pi n}$, причём эта оценка оптимальна, как доказано в [20].

Отметим, что для независимых одинаково распределённых величин $X_1, X_2, \ldots$ должным образом нормированные частные суммы могут сходиться по распределению к гауссовскому пределу без предположения о конечности дисперсии слагаемых. В этом отношении интересно сопоставить теорему, установленную П. Леви, В. Феллером и А. Я. Хинчиным (см. теорему 4.17 в [21]) с результатом А. Я. Хинчина, упомянутым в разделе 2.1. Эти авторы доказали, что необходимым и достаточным условием сходимости по распределению величин вида $a_n \sum_{k=1}^n (X_k - b_n)$, где a_n и b_n — некоторые числа, к гауссовской величине $Z \sim N(0,1)$ при $n \to \infty$ будет условие: $L(x) := \mathbb{E}(X_1^2 \mathbb{I}\{|X| \leqslant x\})$ является функцией, медленно меняющейся на бесконечности в смысле Караматы, т. е. $\lim_{x\to\infty} L(cx)/L(x) = 1$ для каждого $c > 0$. Когда данное условие выполнено, можно выбрать $b_n = \mathbb{E}X_1$. Если упомянутая сходимость распределений имеет место при $a_n = 1/\sqrt{n}$ и $b_n = 0$, то $\mathbb{E}X_1 = 0$ и $\operatorname{var} X_1 = 1$.

Возвращаясь к деятельности А. А. Маркова, следует подчеркнуть, что им установлен вариант центральной предельной теоремы для зависимых случайных величин, образующих цепь. Р. Л. Добрушиным [22] была доказана центральная предельная теорема для неоднородных марковских цепей, которая далее неоднократно обобщалась. В [23, теорема 1.1] дано упрощённое доказательство этой теоремы, опирающееся на мартингальную технику (мартингалам, образующим важный класс случайных процессов, посвящена, например, глава IV книги [6]). Пусть для каждого $n \in \mathbb{N}$ величины $X_{n,1}, \ldots, X_{n,n}$, принимающие значения в борелевском пространстве $(S, \mathscr{B}(S))$, образуют цепь Маркова (для простоты можем считать, что S — счётное или конечное множество I, а $\mathscr{B}(I)$ — совокупность всех подмножеств I). Положим $S_n := \sum_{i=1}^n f_{n,i}(X_{n,i})$, где (измеримые) функции $f_{n,i}\colon S \to \mathbb{R}$, $i = 1, \ldots, n$, $n \in \mathbb{N}$. Пусть $\max_{i=1,\ldots,n} \sup_{x\in S} |f_{n,i}(x)| \leqslant C_n$ для некоторых положительных чисел C_n, $n \in \mathbb{N}$. Предположим, что

$$\lim_{n\to\infty} C_n^2 \alpha_n^{-3} \left(\sum_{i=1}^n \operatorname{var} f_{n,i}(X_{n,i}) \right)^{-1} = 0, \tag{19}$$

где α_n — некоторый «коэффициент эргодичности» (см. определение на с. 3 в [23]). Тогда

$$(\operatorname{var} S_n)^{-1/2}(S_n - \mathbb{E}S_n) \xrightarrow{\mathscr{D}} Z, \quad n \to \infty, \tag{20}$$

где $Z \sim N(0,1)$. Если условие (19) не имеет места, то (20) может не выполняться. Иного типа достаточные условия справедливости центральной предельной теоремы для однородных марковских цепей найдены в [24] и [25].

Начиная со второй половины XX века, центральная предельная теорема была доказана для различных семейств случайных величин (не только процессов, но и полей) при разнообразных условиях зависимости. Упомянем, например, исследование m-зависимых величин (понятие, введённое С. Н. Бернштейном), мартингалов, процессов и полей с перемешиванием, систем величин, обладающих свойством ассоциированности и родственным ему, см., например, [26]. Кроме того, была построена теория функциональных предельных теорем. Это — отдельная важная тема.

В XX веке уже после кончины А. А. Маркова исследователям удалось создать глубокую теорию, которая позволяет описывать классы предельных распределений (более общих, чем гауссовские), к которым могут сходиться должным образом нормированные суммы независимых случайных величин в схеме серий. Выдающуюся роль сыграла монография [27]. Далее выяснилось, что возникающие безгранично делимые и устойчивые распределения играют важную роль во многих стохастических моделях, например, в физике и в финансовой стохастической математике. В настоящее время в области предельных теорем теории вероятностей интенсивно развивается метод, предложенный Ч. Стейном в 1972 году. С помощью различных вариантов этого мощного метода в ряде случаев удалось получить неулучшаемые оценки аппроксимации изучаемых распределений (например, распределений должным образом нормированных сумм случайных величин) определённым предельным законом, см., например, [28]. Здесь же надо указать, что интересной темой является использование различных вероятностных метрик для анализа близости распределений.

2.5. Заключительные замечания

Мы не затронули ряд важных направлений теории марковских процессов, которые связаны с использованием сложного аналитического аппарата и требуют систематического изучения, однако считаем полезным дать некоторые ссылки. С общей теорией марковских цепей (в том числе с непрерывным временем) и её приложениями в теории массового обслуживания (теории очередей) имеет смысл ознакомиться, обратившись к [29]. Различные предельные теоремы для марковских цепей излагаются в [30, 31] и [32]. Разнообразные применения марковских цепей, в том числе в генетике и стохастической финансовой математике, даются в [33]. О приложениях в эпидемиологии см., например, [34]. Марковские ветвящиеся процессы со значениями в пространстве мер изучаются в [35]. Скрытым марковским и полумарковским моделям посвящены книги [17] и [36]. О марковских моделях и задачах распознавания образов можно прочитать в [37]. Методы оптимизации и управляемые марковские цепи рассматриваются в [38]. Некоторые аспекты теории марковских процессов и квантовой механики составляют содержание монографии [39].

Подводя итог, можно сказать, что глубокие идеи А. А. Маркова инициировали развитие целого ряда областей современной теории вероятностей. Сам автор стал классиком науки, его результаты и введённые им понятия вошли практически во все университетские учебники.

Список литературы

[1] *Делоне Б. Н.* Петербургская школа теории чисел. М.; Л.: Изд-во АН СССР, 1947.

[2] *Марков А. А.* Исчисление вероятностей / 3-е изд. СПБ.: С.-Петербургский университет, 1913.

[3] *Гродзенский С. Я.* Андрей Андреевич Марков / Под ред. Б. В. Гнеденко. М.: Наука, 1987.

[4] *Марков А. А.* Избранные труды. Теория чисел // Теория вероятностей / Под ред. Ю. В. Линника. М.; Л.: Изд-во АН СССР, 1951.

[5] *Basharin G. P., Langville A. N., Naumov V. A.* The life and work of A. A. Markov // Linear Algebra Appl. 2004. V. 386 P. 3–26.

[6] *Булинский А. В., Ширяев А. Н.* Теория случайных процессов. М.: Физматлит, 2005.

[7] *Марков А. А.* Распространение закона больших чисел на величины, зависящие друг от друга // Известия Физико-математического общества при Казанском университете. 1906. Сер. 15. Т. 94. С. 135–156.

[8] *Колмогоров А. Н.* Основные понятия теории вероятностей / 4-е изд. М.: УРСС, 2013.

[9] *Марков А. А.* Об одном применении статистического метода // Известия Императорской Академии наук. 1916. Сер. VI. Т. 4. С. 239–242.

[10] *Успенский В. А.* Труды по нематематике / 2-е изд. Кн. 4. Филология с приложением «Семиотических посланий А. Н. Колмогорова». М.: Объединённое гуманитарное издательство, Фонд «Математические этюды», 2012.

[11] *Brémaud P.* Probability Theory and Stochastic Processes. Cham: Springer, 2020.

[12] *Dharmaraja S., Di Crescenzo A., Giorno V., Nobile A. G.* A continuous-time Ehrenfest model with catastrophes and its jump-diffusion approximation // J. Stat. Phys. 2015. V. 161. P. 326–345.

[13] *Durrett R.* Probability. Theory and Examples / Fourth Ed. Cambridge: Cambridge University Press, 2010.

[14] *Robert C. P., Changye W.* Markov chain Monte Carlo methods, a survey with some frequent misunderstandings. https://arxiv.org/abs/2001.06249.

[15] *Reich B. J., Ghosh S. K.* Bayesian Statistical Methods. Boca Raton: CRC Press, 2019.

[16] *Bulinski A., Spodarev E.* Introduction to random fields // Stochastic Geometry, Spatial Statistics and Random Fields. Asymptotic Methods. E. Spodarev (Ed.). Berlin: Springer, 2013, P. 277–335.

[17] *Hoek J. van der., Elliott R. J.* Introduction to Hidden Semi-Markov Models. Cambridge: Cambridge University Press, 2018.

[18] *Kolokoltsov V. N.* Nonlinear Markov Processes and Kinetic Equations. Cambridge: Cambridge University Press, 2010.

[19] *Шевцова И. Г.* Об абсолютных константах в неравенстве Берри — Эссеена и его структурных и неравномерных уточнениях // Информатика и её применения. 2013. Т. 7, вып. 1. С. 124–125.

[20] *Hipp C., Mattner L.* On the normal approximation to symmetric binomial distributions // Теория вероятностей и её применения. 2007. Т. 52. С. 610–617.

[21] *Kallenberg O.* Foundations of Modern Probability. NY: Springer, 2002.

[22] *Добрушин Р. Л.* Центральная предельная теорема для неоднородных цепей Маркова I, II. Теория вероятностей и её применения. 1956. Т. 1, вып. 1. С. 72–89. (Вып. 4. С. 365–425.)

[23] *Sethuraman S., Varadhan S. R. S.* A martingale proof of Dobrushin's theorem for non-homogeneou Markov chains // Electron. J. Probab. 2005. V. 10, paper no. 36. P. 1221–1235.

[24] *Jones G. L.* On the Markov chain central limit theorem // Probab. Surv. 2004. V. 1. P. 299–320.

[25] *Peligrad M.* On the CLT for additive functionals of Markov chains // Electron. Commun. Probab. 2020. V. 25, paper no. 40. P. 1–10.

[26] *Булинский А. В., Шашкин А. П.* Предельные теоремы для ассоциированных случайных полей и родственных систем. М.: Физматлит, 2008.

[27] *Гнеденко Б. В., Колмогоров А. Н.* Предельные распределения для сумм независимых случайных величин. М.; Л.: Государственное издательство технико-теоретической литературы, 1949.

[28] *Arras B., Houdré C.* On Stein's Method for Infinitely Divisible Laws with Finite First Moment. Cham: Springer, 2019.

[29] *Brémaud P.* Markov Chains: Gibbs Fields, Monte Carlo Simulation and Queues / Second Ed. Cham: Springer Nature Switzerland AG, 2020.

[30] *Levin D. A., Peres Y.* (with contributions of E. L. Wilmer). Markov Chains and Mixing Times / Second Ed. Providence: AMS, 2017.

[31] *Douc R., Moulines E., Priouret P., Soulier P.* Markov Chains. Cham: Springer Nature Switzerland AG, 2018.

[32] *Vassiliou P.-C. G.* Non-homogeneous Markov Chains and Systems. Theory and Applications. Boca Raton: CRC Press, 2023.

[33] *Pardoux E.* Markov Processes and Applications: Algorithms, Networks, Genome and Finance. Chichester: J. Wiley and Sons Ltd., 2008.

[34] *Xu Z., Zhang H., Huang Z.* A Continuous Markov-Chain Model for the Simulation of COVID-19 Epidemic Dynamics // Biology. 2022. V. 11(2), paper no. 190. P. 1–22.

[35] *Li Z.* Measure-Valued Branching Markov Processes. Berlin: Springer, 2011.

[36] *Zucchini W., MacDonald I. L., Langrock R.* Hidden Markov Models for Time Series. An Introduction Using R / Second Ed. Boca Raton: CRC Press, 2016.

[37] *Fink G. A.* Markov Models for Pattern Recognition. From Theory to Applications / Second Ed. London: Springer, 2014.

[38] *Bonnans J. F.* Convex and Stochastic Optimization. Cham: Springer Nature Switzerland AG, 2019.

[39] *Nagasawa M.* Markov Processes and Quantum Theory. Cham: Birkhäuser, 2021.

Александр Вадимович Булинский, мехмат МГУ им. М. В. Ломоносова
alexander.bulinski@math.msu.ru